

l r f u

cea

saclay

- Retour d'expérience - Synchronisation et TDC sur KM3NET



Yassir MOUDDEN

En étroite collaboration avec :

- **IRFU** : Frédéric LOUIS, Hervé LE PROVOST, Frédéric CHATEAU, Bertrand VALLAGE, Shebli ANVAR, Eric ZONCA, Kevin MENAGER
- **NIKHEF** : Peter JANSWEIJER, Jan-Willem SCHMELLING, Sander MOS, Albert ZWART, Henk PEEK, Jelle HOGENBIRK
- **NESTOR** : Kostas MANOLOPOULOS
- **INFN** : Fabrizio AMELI

OPEN SOURCE of great inspiration :



Le projet de détecteur de neutrinos KM3NET

<http://www.km3net.org>

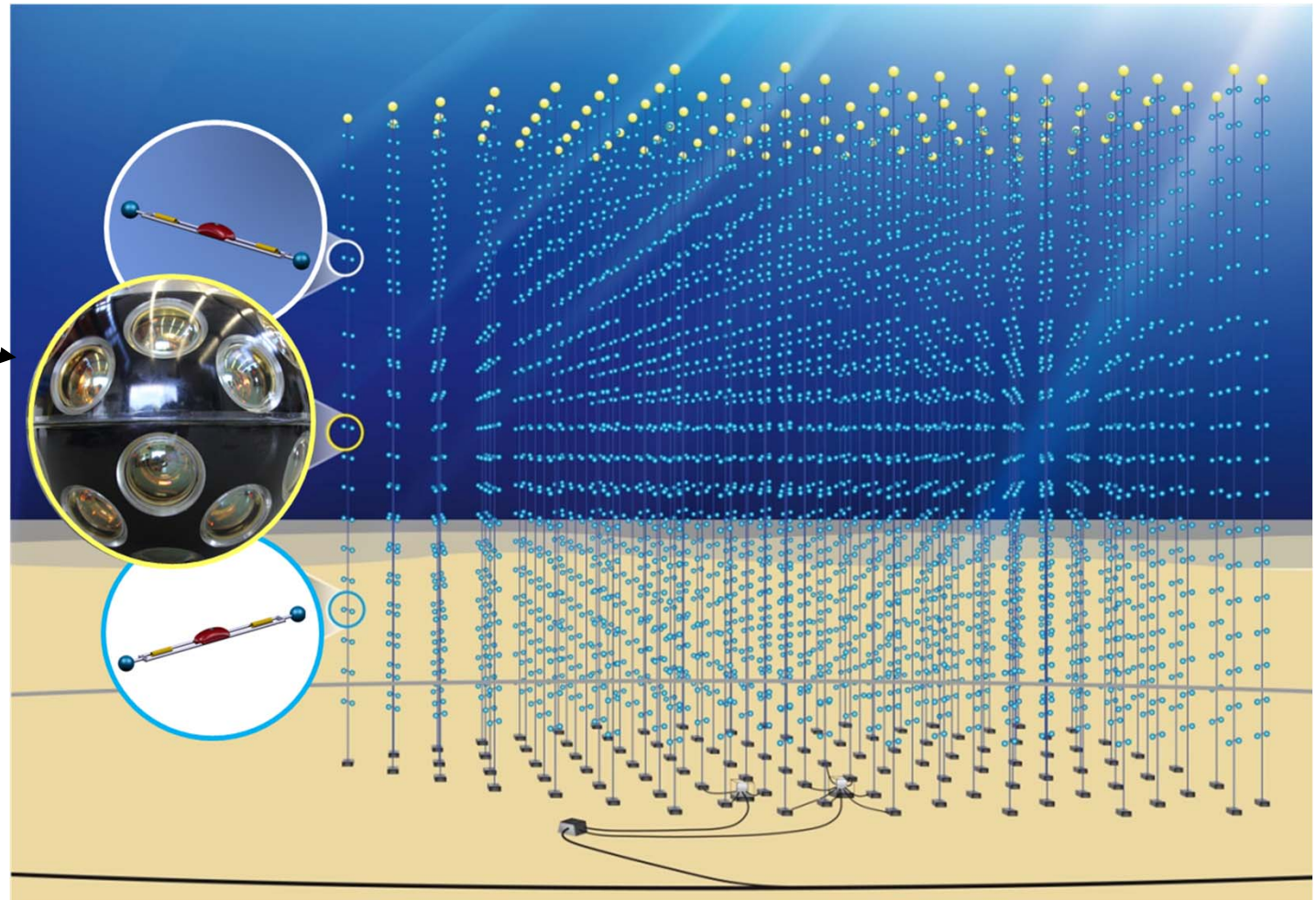
Technical Design Report (ISBN 978-90-6488-033-9)

Conceptual Design Report (ISBN 978-90-6488-031-5)

Digital Optical Module
(DOM) →

12800 DOMs

320 lignes

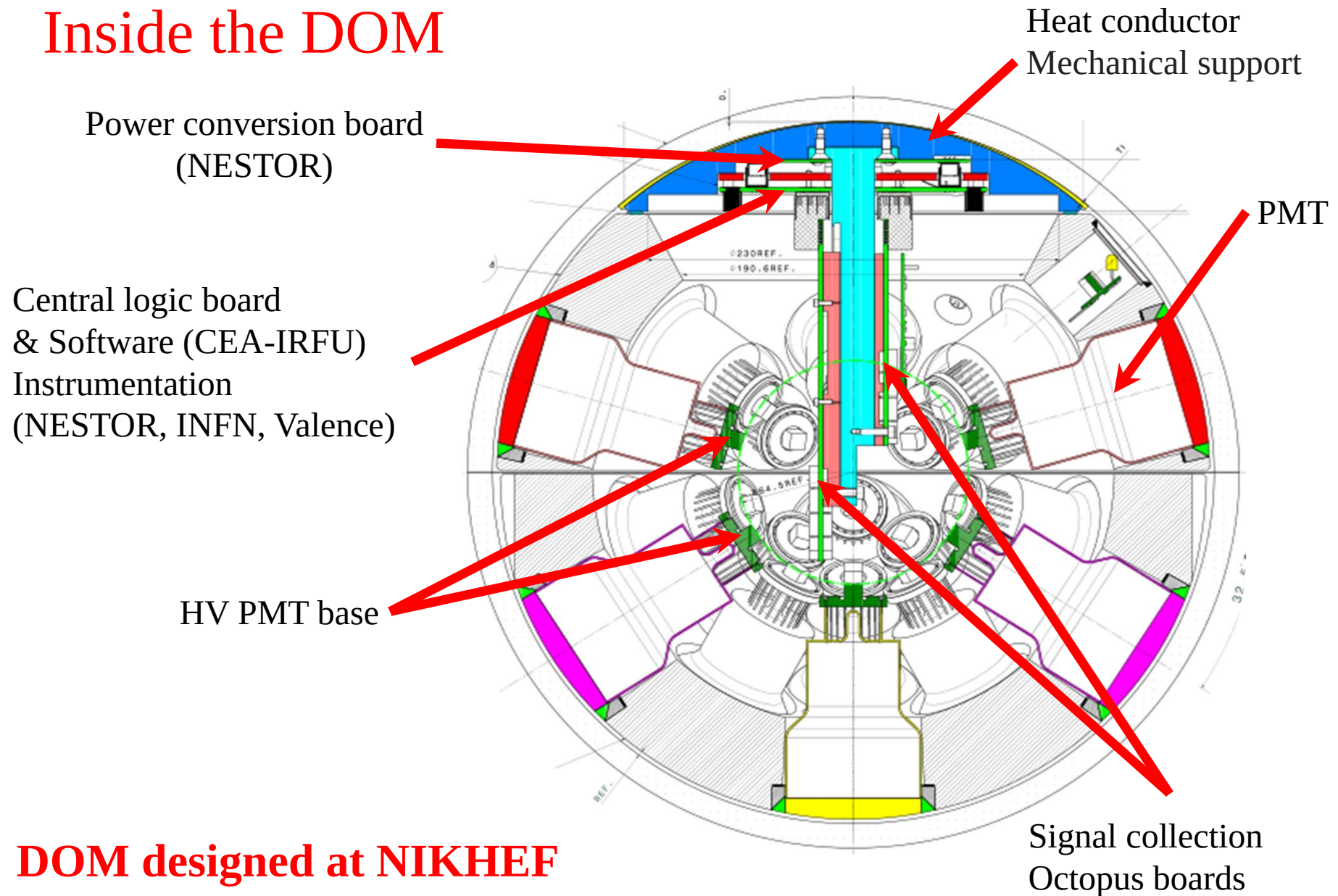


Digital Optical Module DOM



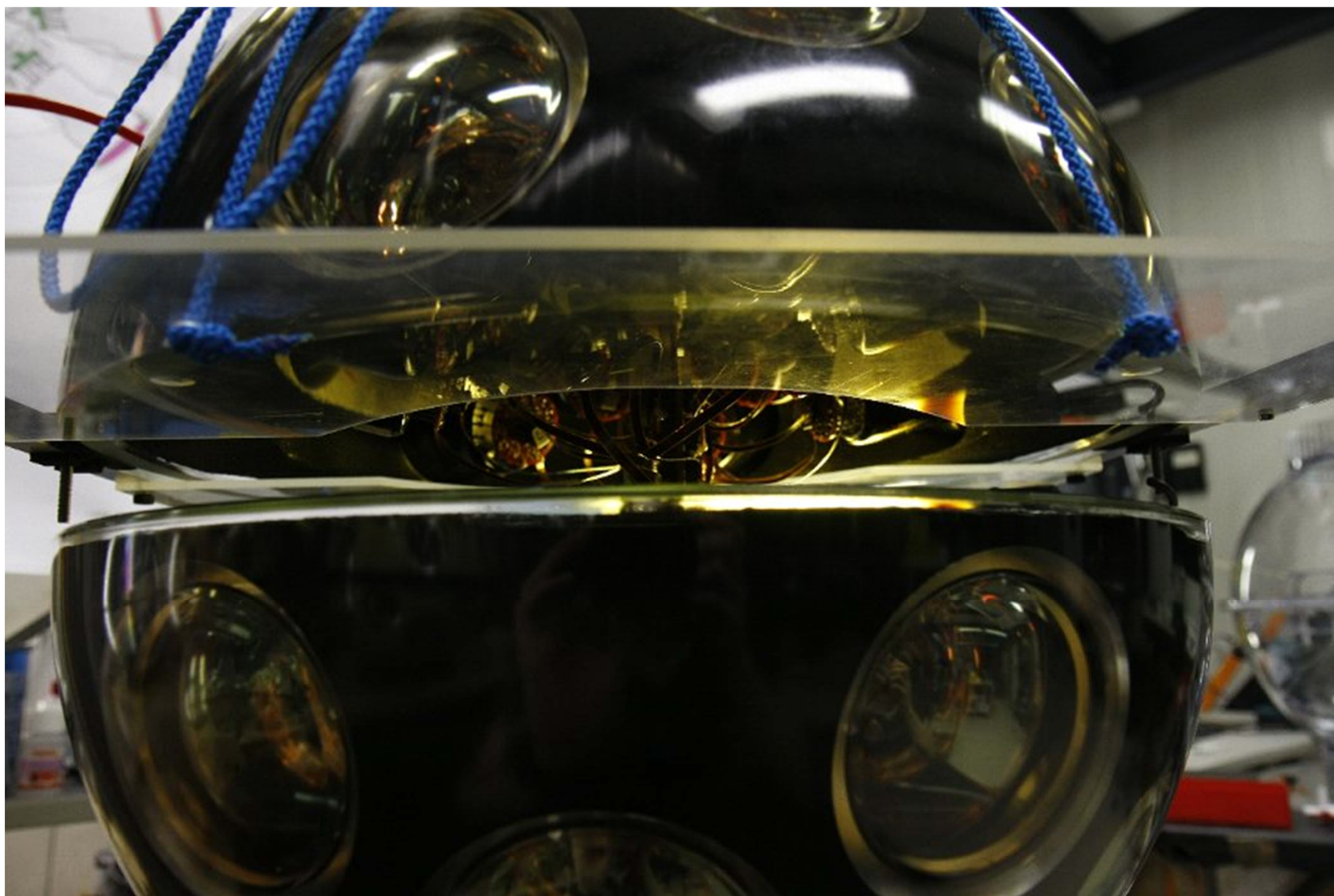
- Sphere de 17" de diametre
- 31 x 3 " PMT (19 Lower & 12 Upper)
- Read out des PMT – Connection gigabit ethernet optique vers la station onshore
- Slow control - Instrumentation et bases des PMT
- Alimentation 12V
- Read out system on chip (RSOC) sur FPGA Virtex 5 totalement intégré

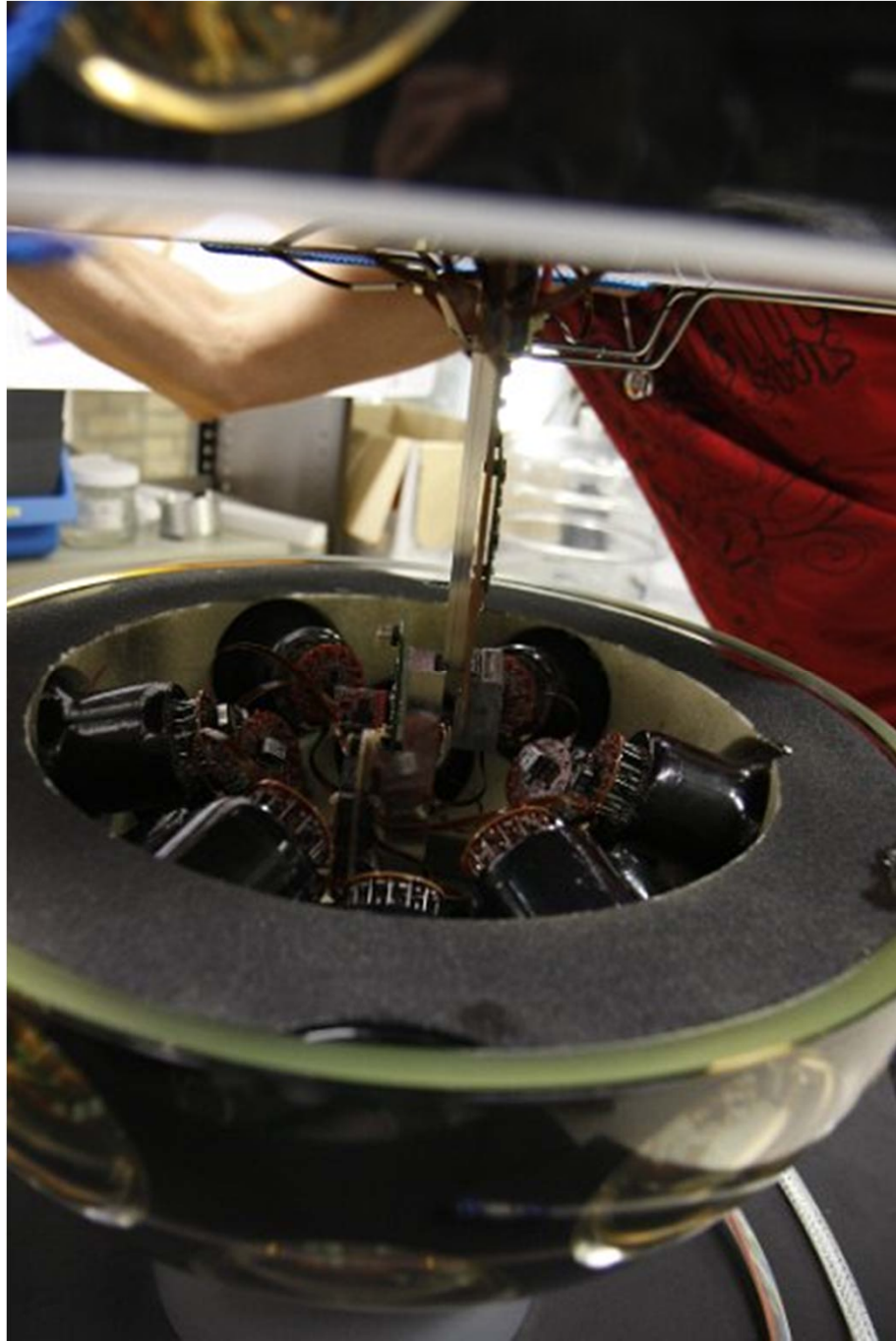
Inside the DOM



**DOM designed at NIKHEF
(Amsterdam, Netherlands)**

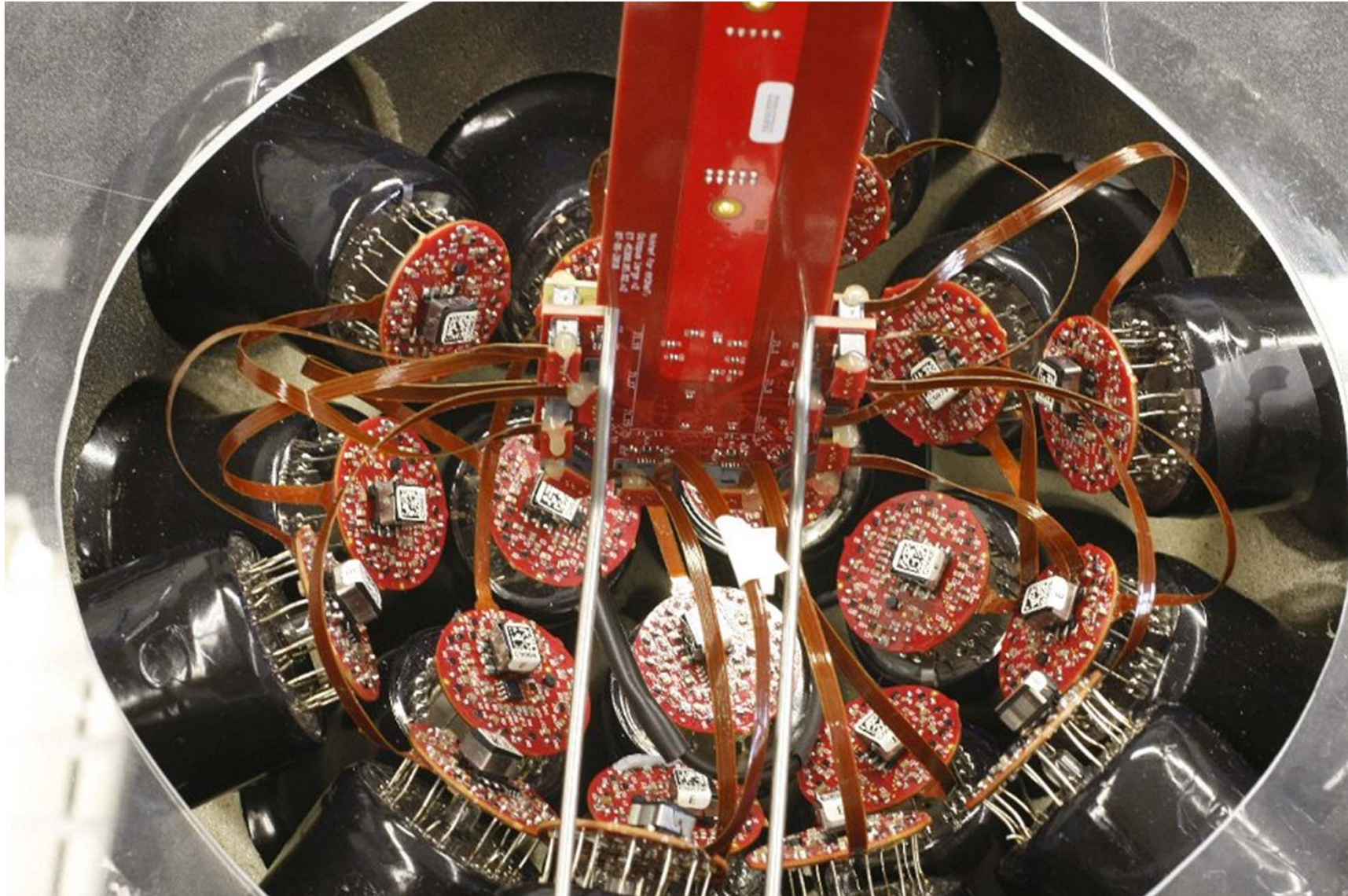


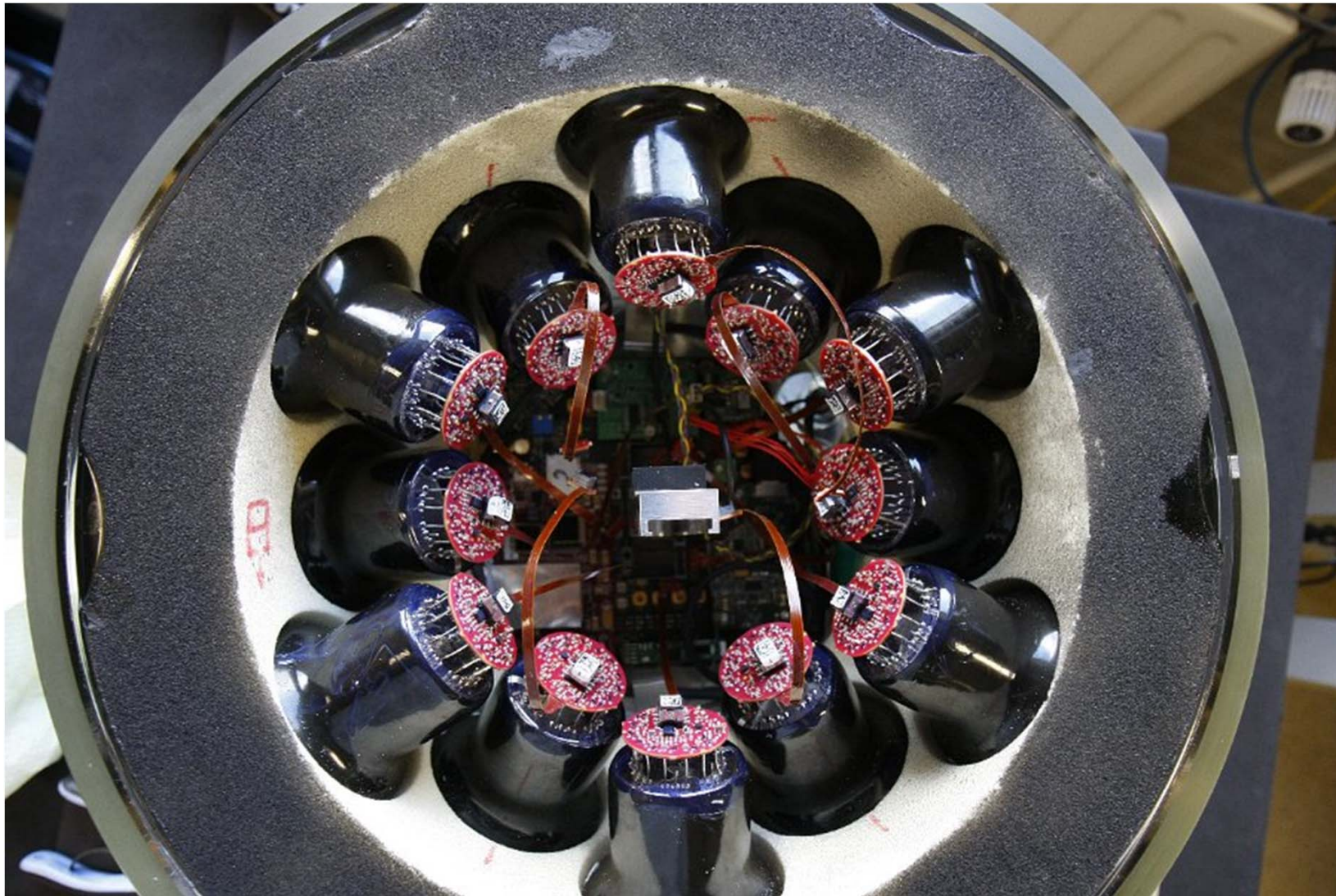




Octopus boards

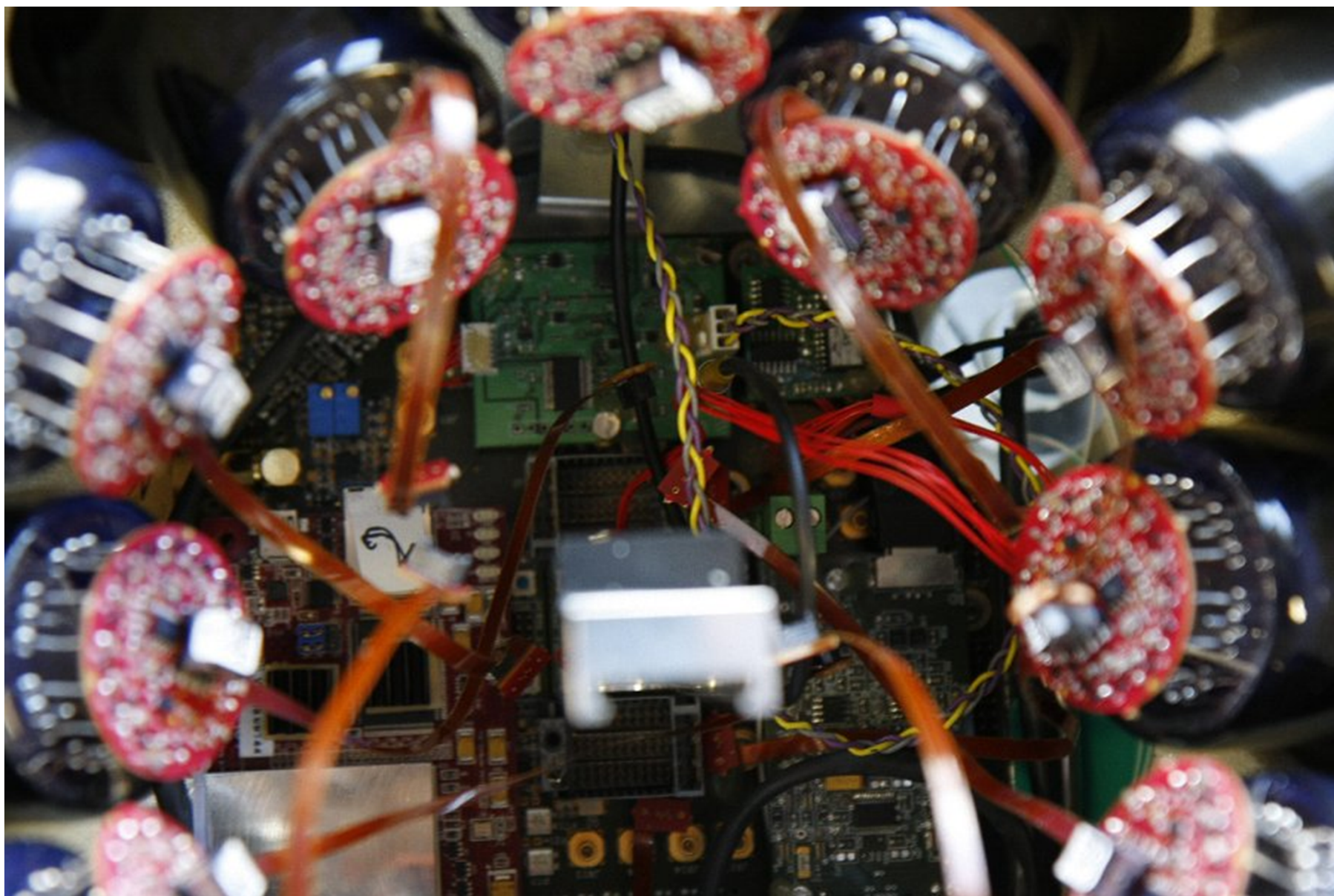
- Transfer signals from the PMT base to the Central Logic Board
- Provide and monitor power for individual PMT HV base

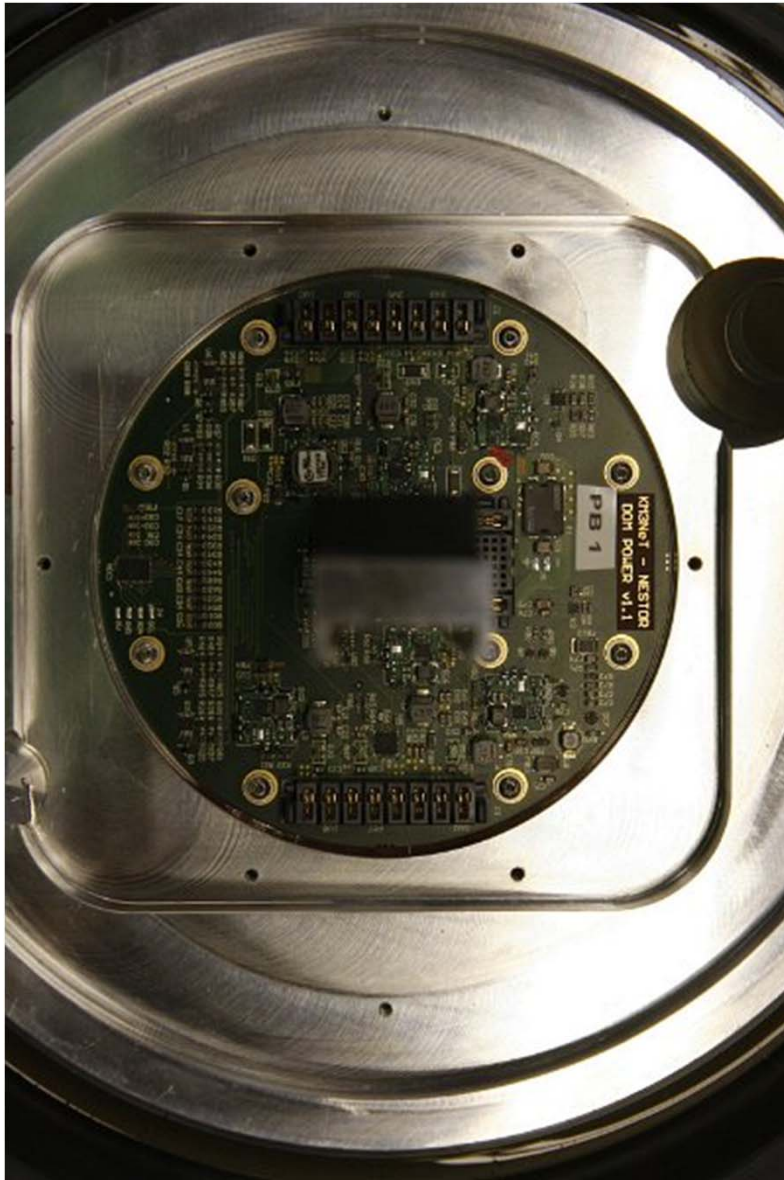




Photomultiplier base

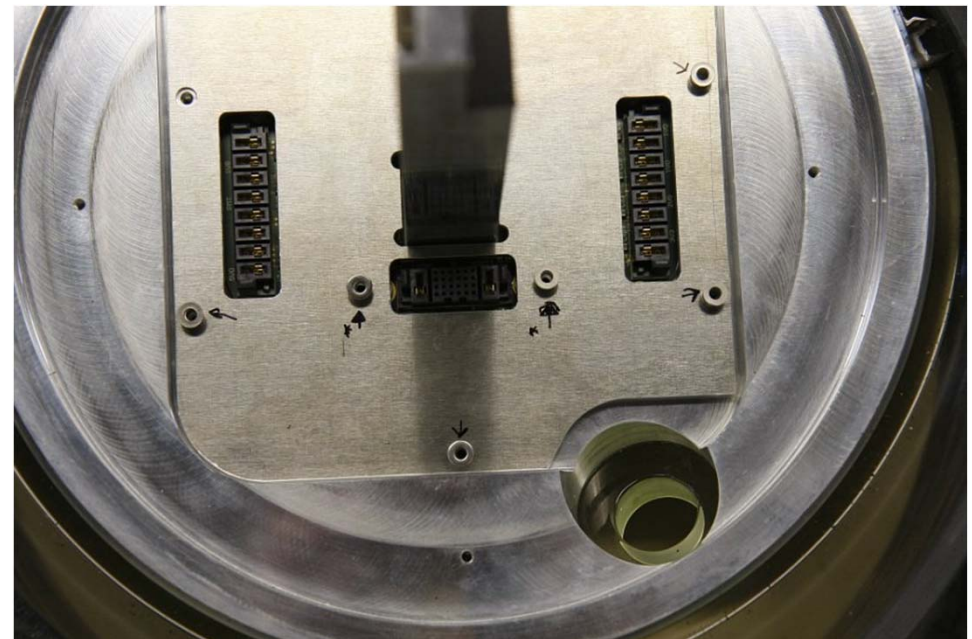
- Adjustable HV (800-1400V)
- Low power (4,5mW)
- ASIC with pre-amplifier and discriminator
- Connection to the signal collection boards done by flexible PCB
- Time over threshold (TOT) LVDS output signal





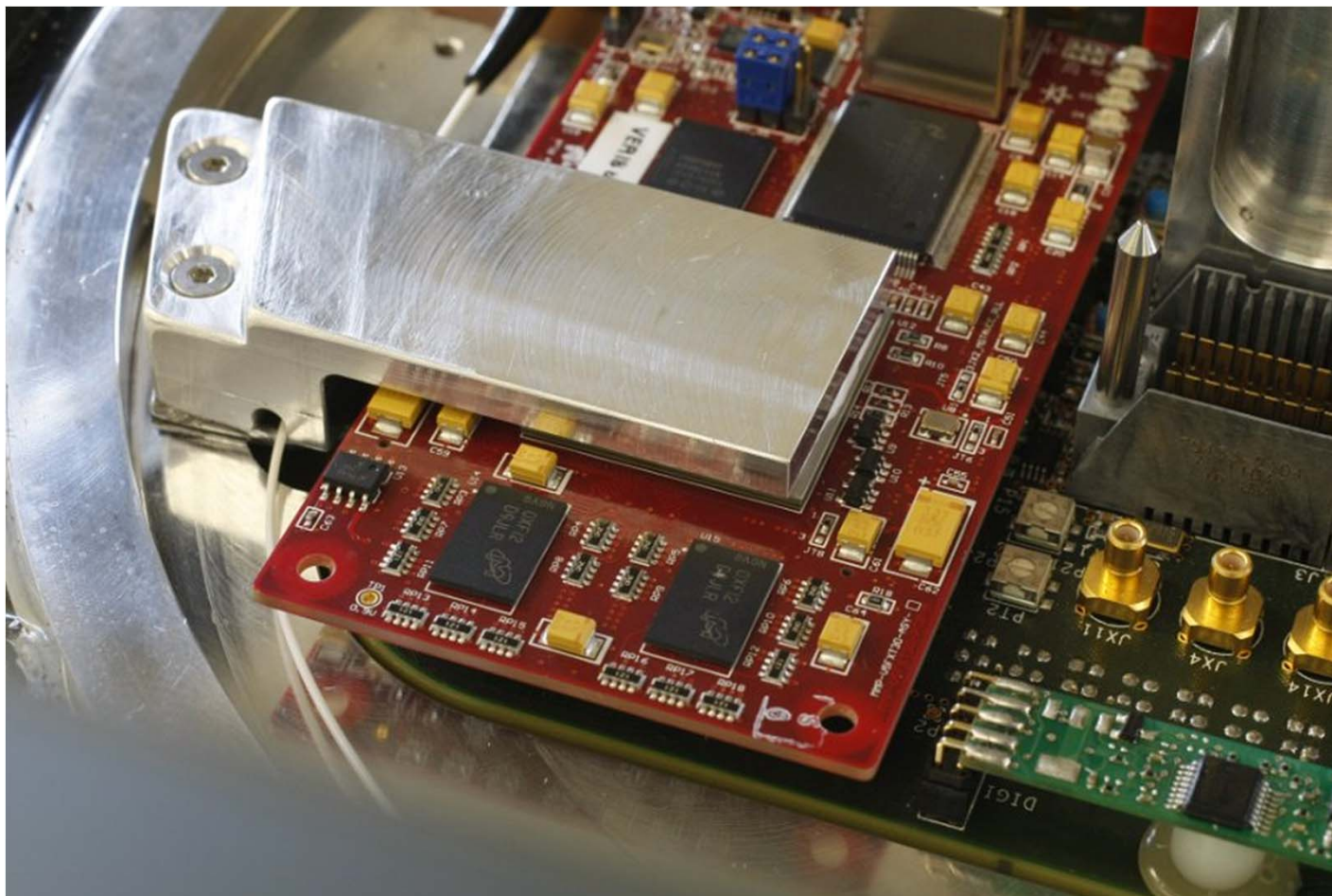
Power board

Convert 12V from the external break
Out Box to voltages for the DOM



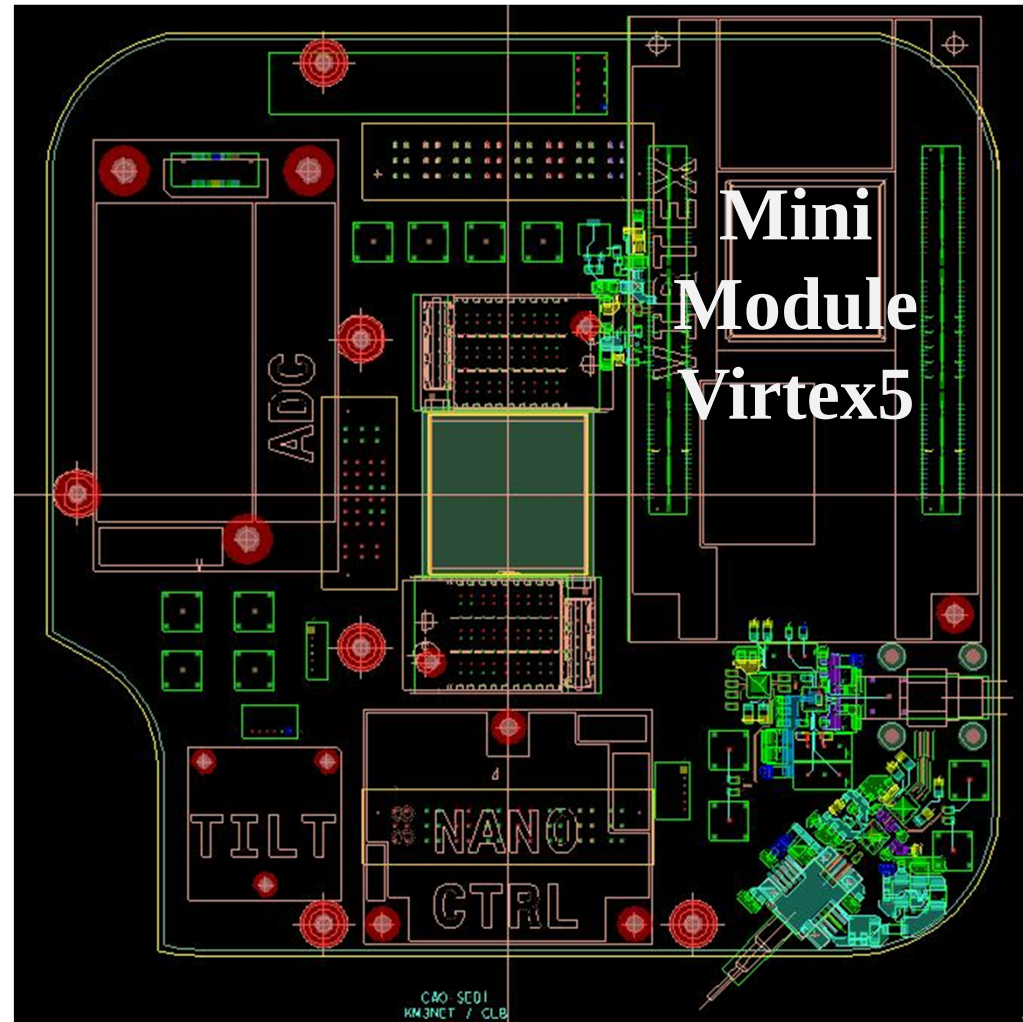
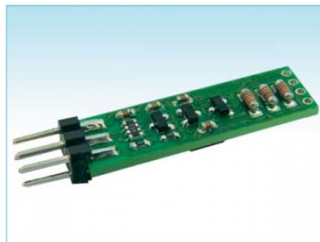
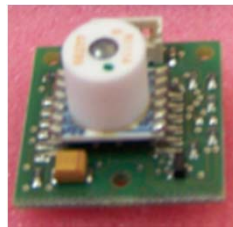
Central Logic Board

- Minimodule Virtex 5 - FX70T
- Instrumentation : Tilt meter, compass, digipico (temperature et **humidité**)
- Carte ADC des hydrophones



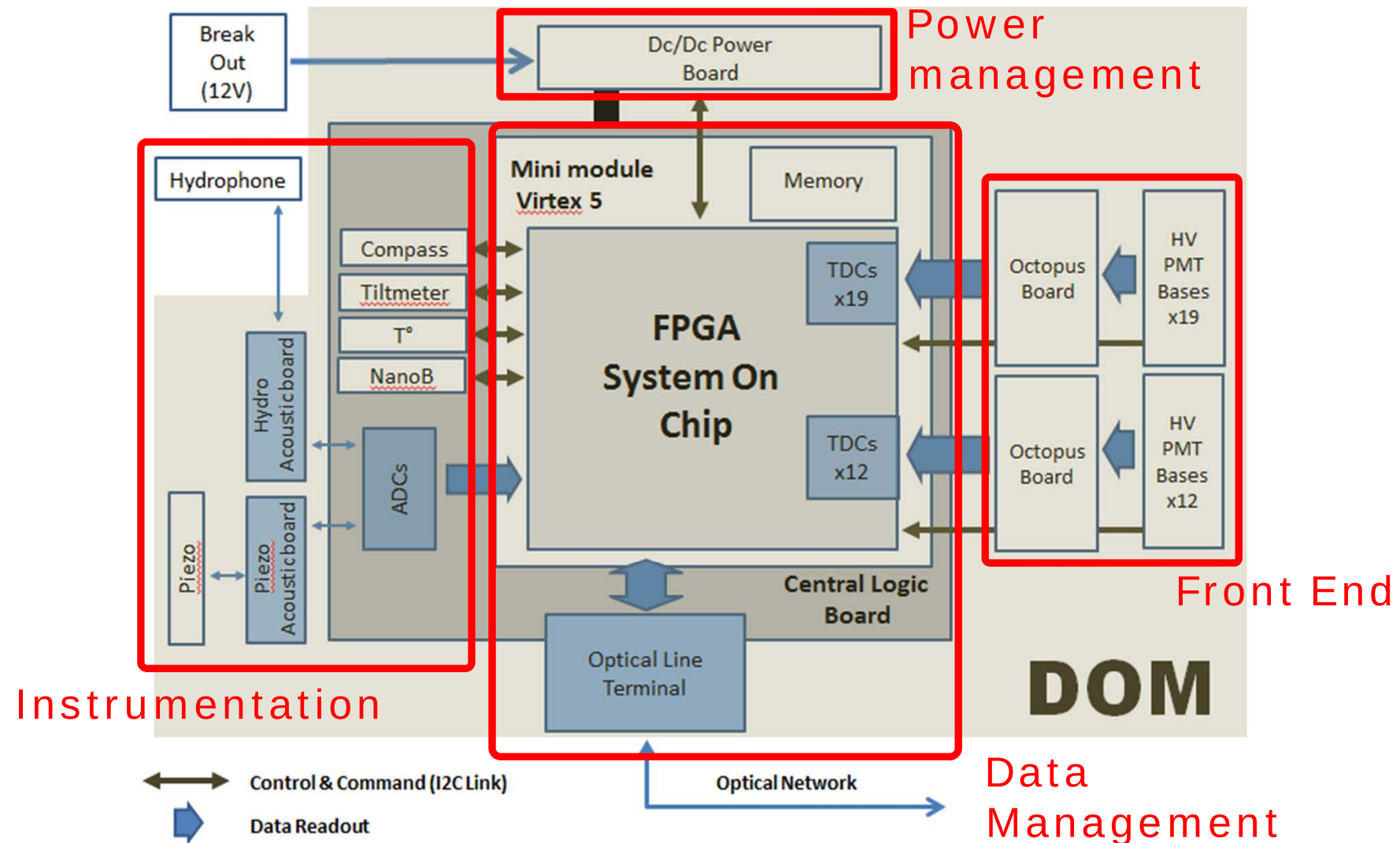
Central Logic Board

- Mini module with Virtex 5 System on Chip (SoC)
- Optical Line Terminal 1Gb/s Ethernet link
REAM Driver / Diode Pin
- Acoustic board (piezo + hydro)
- Tilt-meter board
- Compass
- Temperature sensor
- Nano Beacon control board

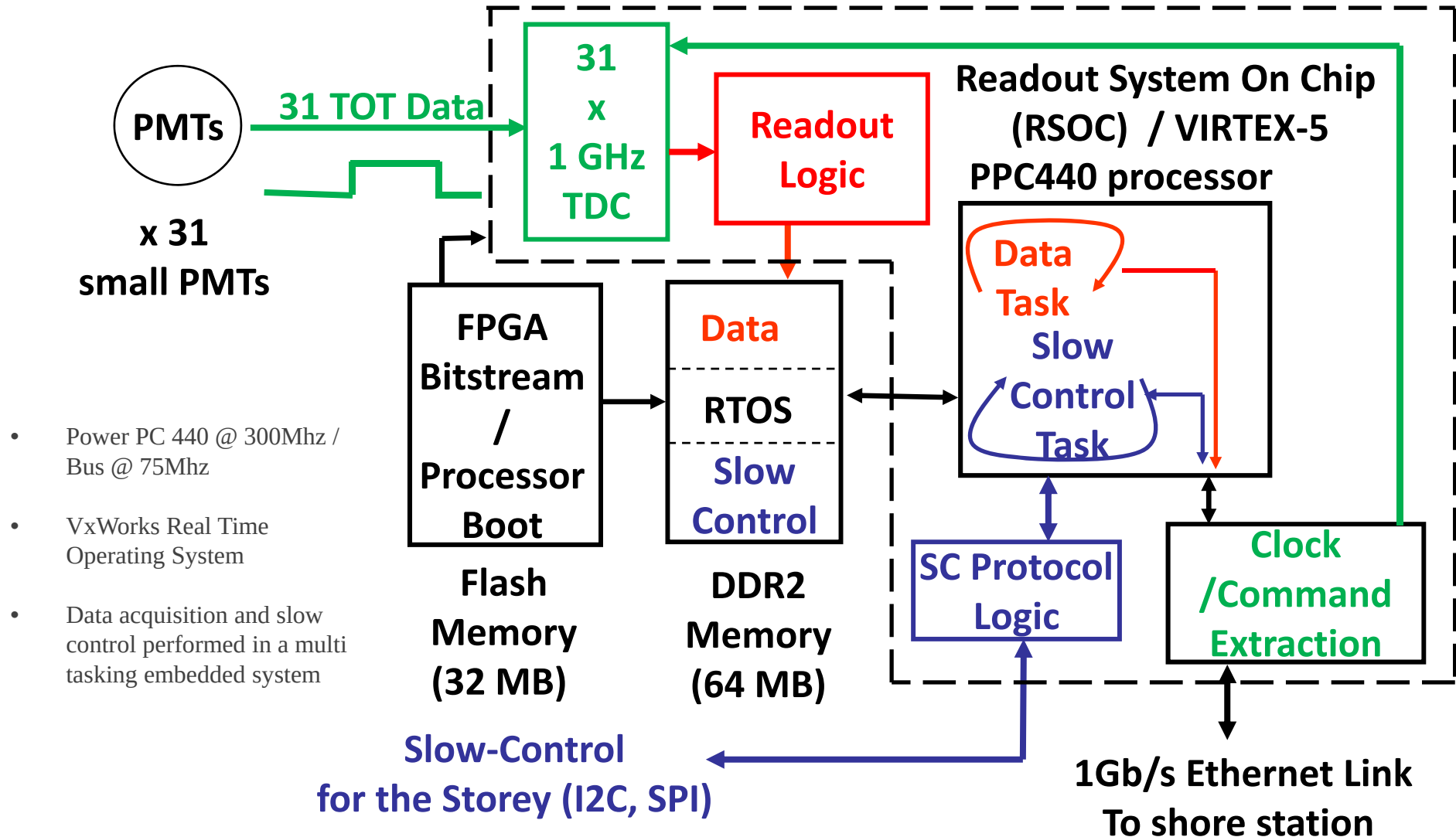


Optical Line
Terminal

DOM electronic description



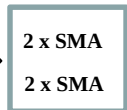
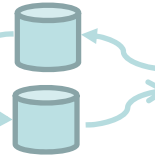
The KM3Net prototype Offshore Processor Board



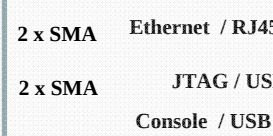
- Power PC 440 @ 300Mhz / Bus @ 75Mhz
- VxWorks Real Time Operating System
- Data acquisition and slow control performed in a multi tasking embedded system



Optical fiber



PPM DOM Shore Station



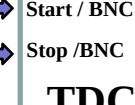
Ethernet / RJ45

JTAG / USB

Console / USB

Start / GPI-O

Stop / GPI-O



**TDC
SR620**

RS 232 / DB 9

USB

USB

**PC
Data
Control**

Ethernet
RJ45

USB

RS 232 / DB 9

**Ethernet
Switch**

ML605

1pps
/SMA

125 Mhz
/ 2x SMA

GPI-O GPI-O

PC
Link

GPS
Gate

RS 232 / DB 9

**RS232
Module**

**JTAG
Prog**



/ 2x SMA

/ Micro Port

/SMA

/SMA

MicroWire

3,3 V

GPI-O GPI-O

GPI-O

Power

ML507

GPI-O

JTAG / Pins

Console / DB 9

Ethernet / RJ45

1pps
/ BNC

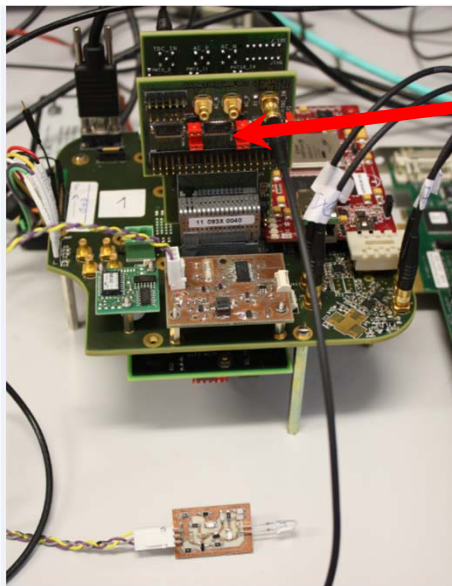
10 Mhz
/ BNC

NMEA / DB 9

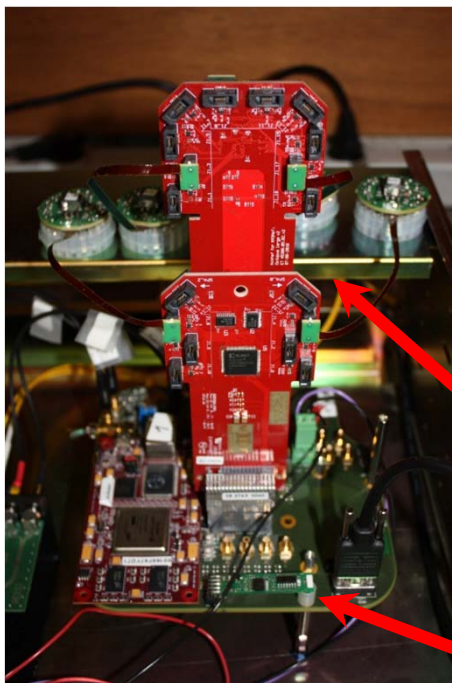
GPS

Ethernet / RJ45

Network



Pseudo Octopus



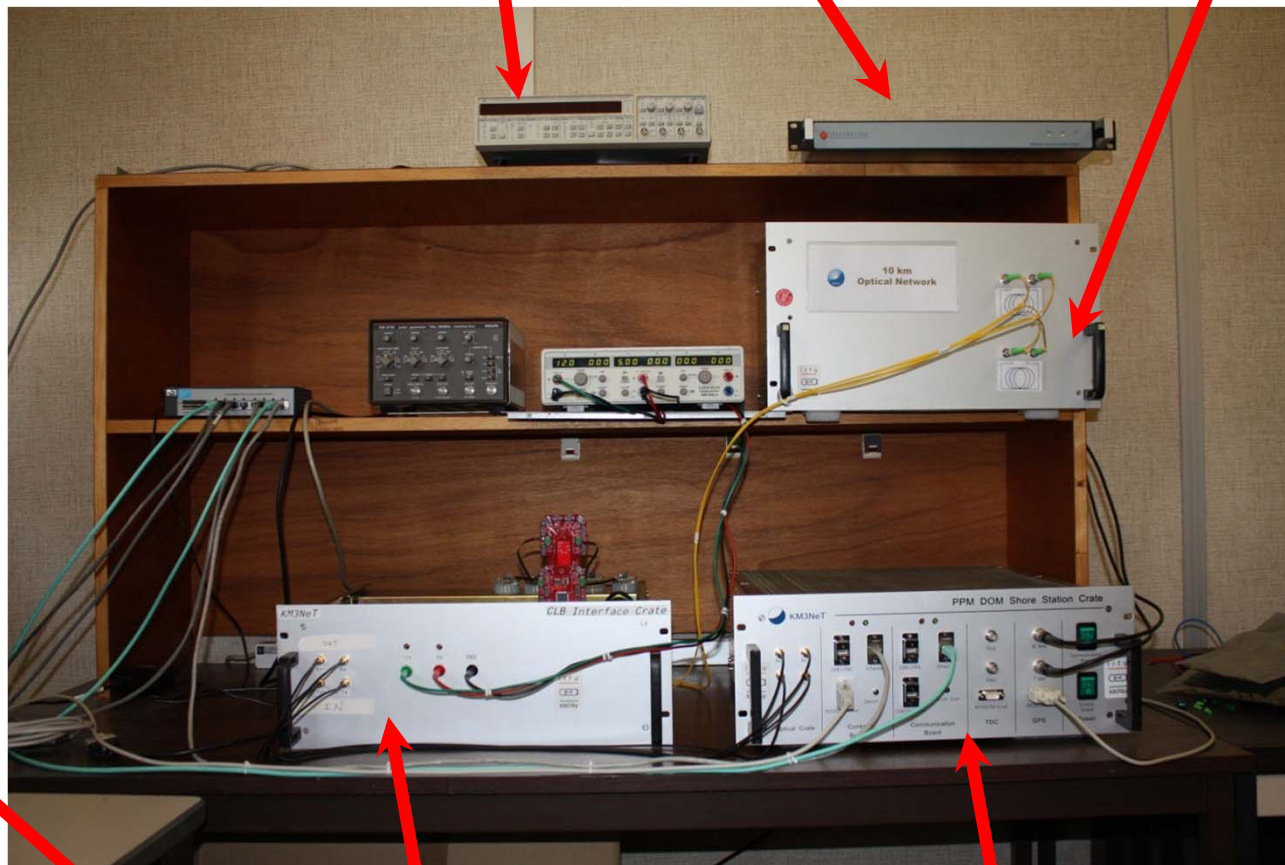
**Octopus + bases de PMT
tranceivers electro-optiques**

CLB + Minimodule Virtex 5

TDC SR620

GPS

20 km fibre optique



**Onshore Crate
Ethernet bridge**

Distribution d'horloge et Synchronisation

“Syntonization”

Distribution d'une fréquence issue d'une source unique (GPS, horloge atomique) d'un maitre *on-shore* vers un ou plusieurs esclaves *off-shore*.



Utiliser la fréquence *porteuse* des liens Gigabit codés 8b/10b.
(SONET/SDH, Synchronous Ethernet, White Rabbit)

“Synchronization”

Distribution d'une phase / d'une date d'un maitre vers un ou plusieurs esclaves à mieux que la nanoseconde pour les “*besoins*” de KM3NET.

- Latence fixe ou connue ?
- Mesurer la latence sur un Aller / Retour ?
- Pouvoir répartir cette latence entre Aller et Retour ?
- Latence liée à la propagation ?
- Latence dans le hardware ?
- Eléments aléatoires de latence à chaque *RESET* ou power on ?
- White Rabbit ? PTP ? Autre ?



Besoin de commandes synchrones (*top départ*) donc dans la couche physique, sans modifier les couches supérieures du protocole Ethernet.

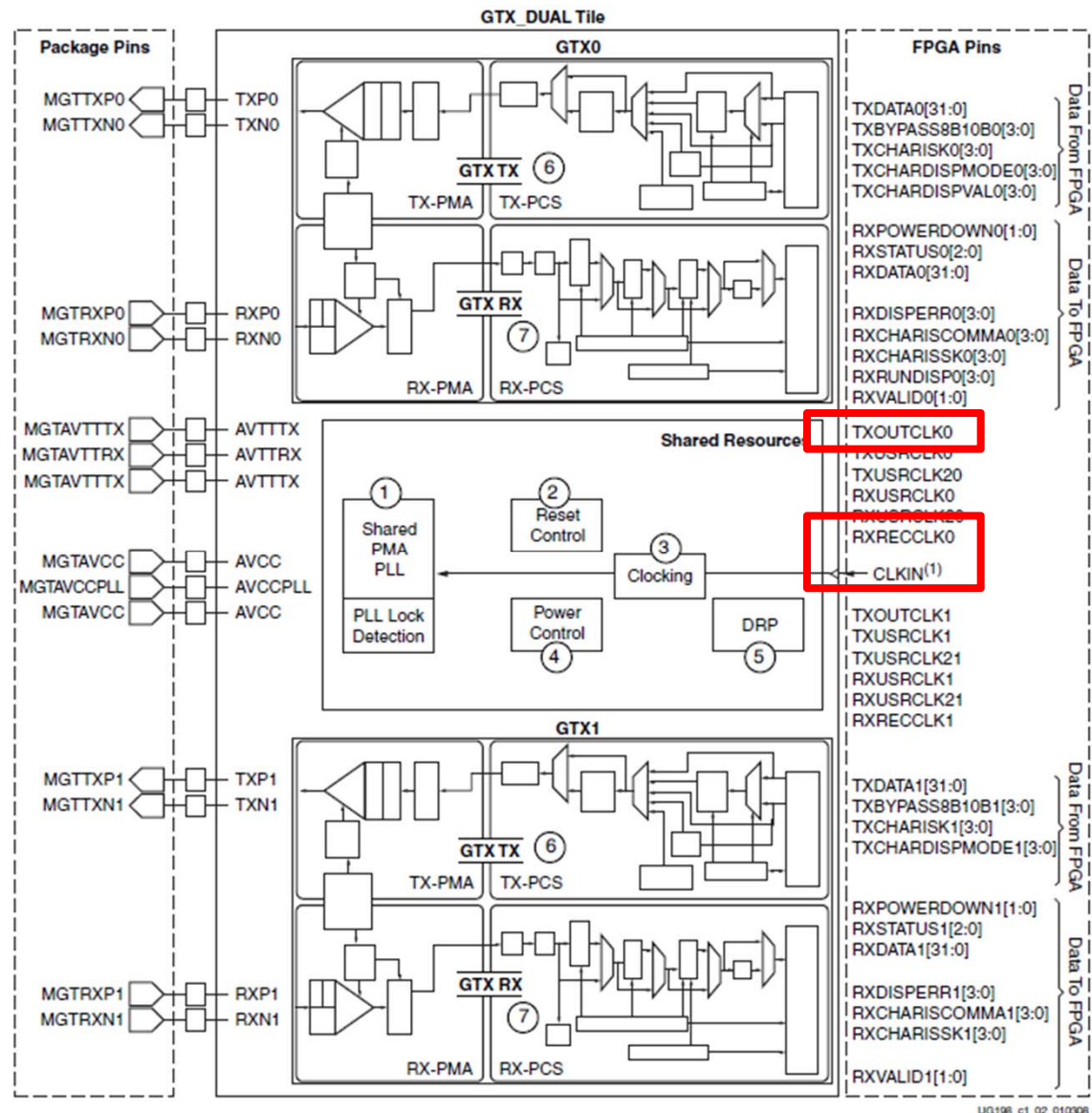
Virtex 5 GTX Tile

(Xilinx UG198 - 387 pages)

- Embedded hardware tile
- 8 to 24 GTX in Virtex5 FXT
- N parameters
- N' ports
- GTX Wizard to help designers
- 150Mbps to 6.5Gbps bit rate
- Shared PLL (1.5 to 3.25 GHz range)

Test en Loop back externe :

Latence fixe une fois le lien établi mais variation aléatoire a chaque reset.
Plusieurs valeurs différentes possibles séparées de 0.8 ns



Voir les publications de P. Jansweijer et al. , R. Giordano et al. et White Rabbit qui ont aussi développé sur Virtex 5/6 et Altera.

Table E-1: GTX TX Latency

Block Number	Block Name	Attribute Setting		Latency Contribution (TXUSRCLK cycles)	
				Min	Max
1	FPGA TX Interface	TXDATAWIDTH	0	0.5	0.5
			1	1	1
			2	2	2
2	8B/10B Encoder	TXENC8B10BUSE	0	0	0
			1	1	1
3	TX Buffer	TX_BUFFER_USE	FALSE	1	1
			TRUE	2	4
4+6+7+8+9	PMA + Interface	-	-	2	2

GTX Transmitter (TX)

Gamme d'accrochage de la PLL partagée : 1.5 à 3.25 GHz.

Existe un diviseur de fréquence par deux entre PMA et PCS qui introduit une variation de latence aléatoire d'une demi période.

Donc , preferer un bypass du buffer et le phase alignement.

(P. Jansweijer et H. Peek , Nikhef)

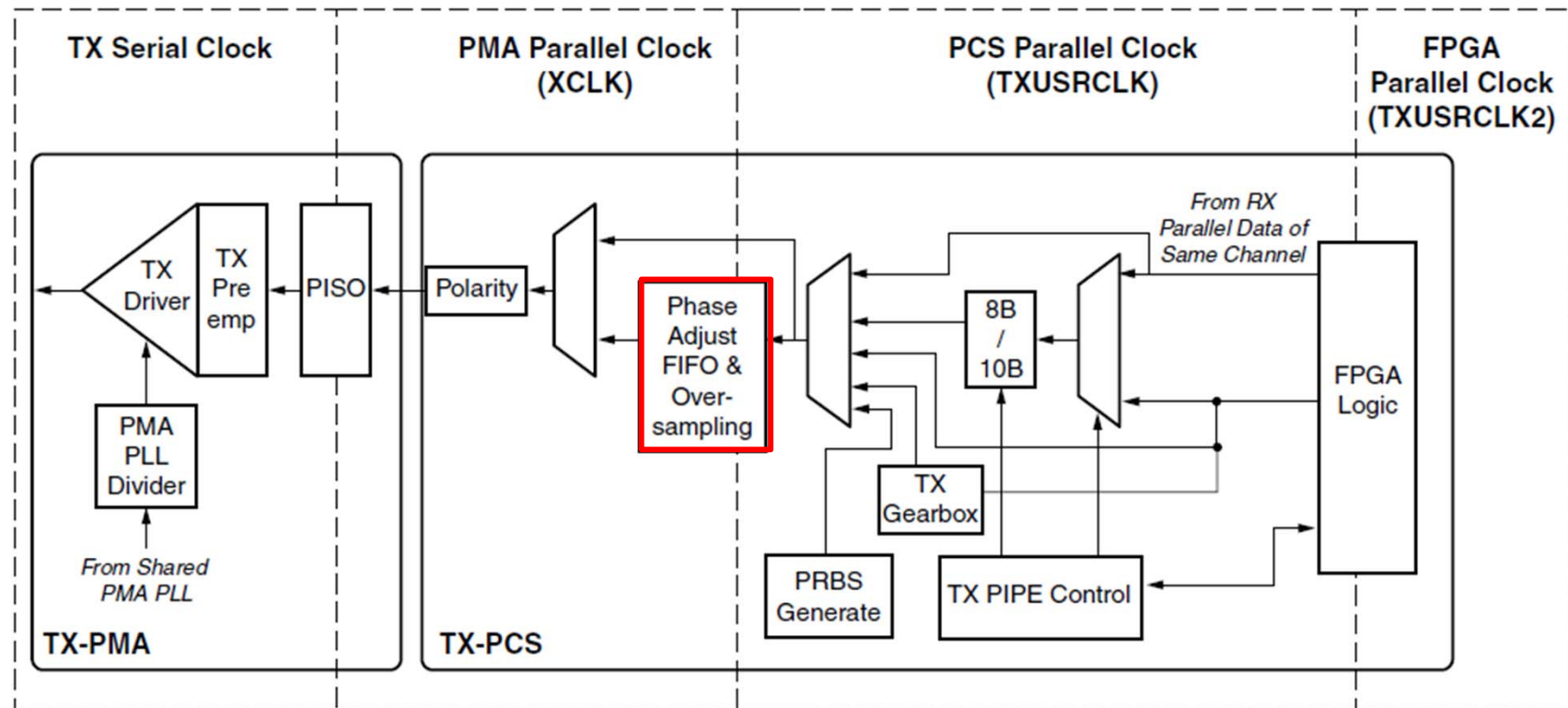
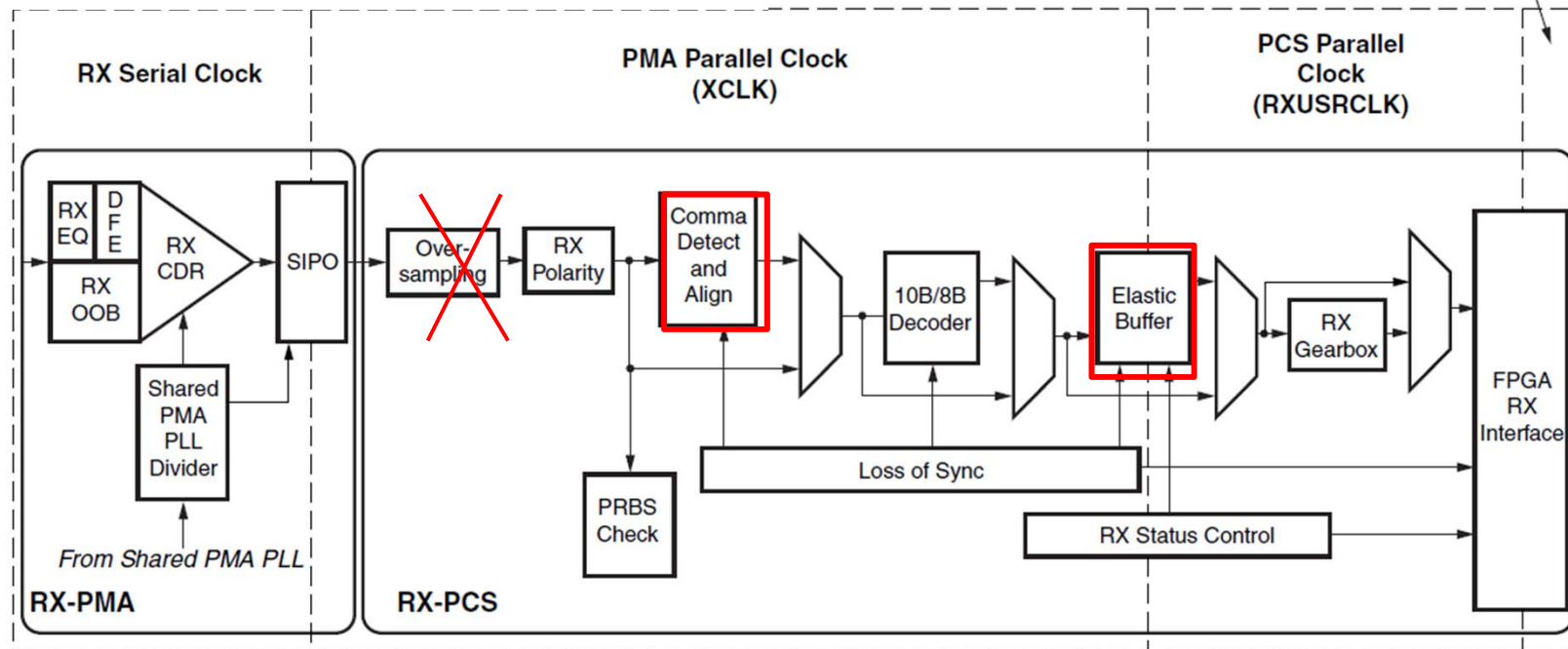


Table E-2: GTX RX Latency

Block Number	Block Name	Attribute Setting		Latency Contribution (RXUSRCLK cycles)	
				Min	Max
1+2+3+4	PMA + Interface	-	-	4	4
5+6	Over-sampling	OVERSAMPLE_MODE	FALSE	0	0
			TRUE		
9	Comma Alignment	RXCOMMADETUSE	0	1	1
			1	2.5	3.5
11	10B/8B Decoder	RXDEC8B10BUSE	0	0	0
			1	1	1
12	RX Elastic Buffer	RX_BUFFER_USE	FALSE	2	2
			TRUE	2 cycles + CLK_COR_MIN_LAT / 2	2 cycles + CLK_COR_MAX_LAT / 2
16	FPGA RX Interface	RXDATAWIDTH	0	1.5	1.5
			1	2	2
			2	3	3

GTX Receiver (RX)

- Detection de commas et alignement :
 - utiliser la logique interne en mode **manuel** ?
 - utiliser une logique **custom** externe au GTX ?
- Changement de domaine d'horloge PMA / PCS :
 - garder l'*elastic buffer* ?
 - ne pas le garder et utiliser *phase alignment*

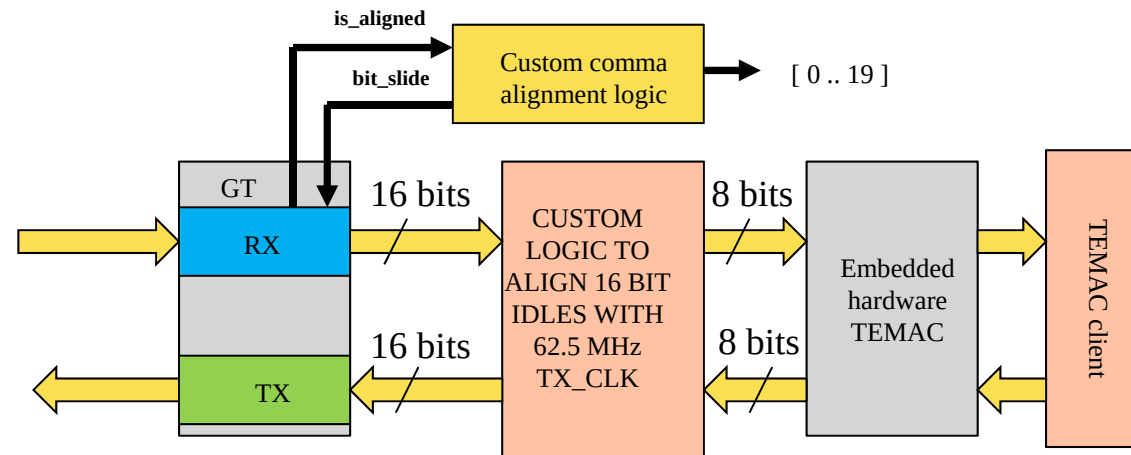


BITSLIDES et alignement manuel

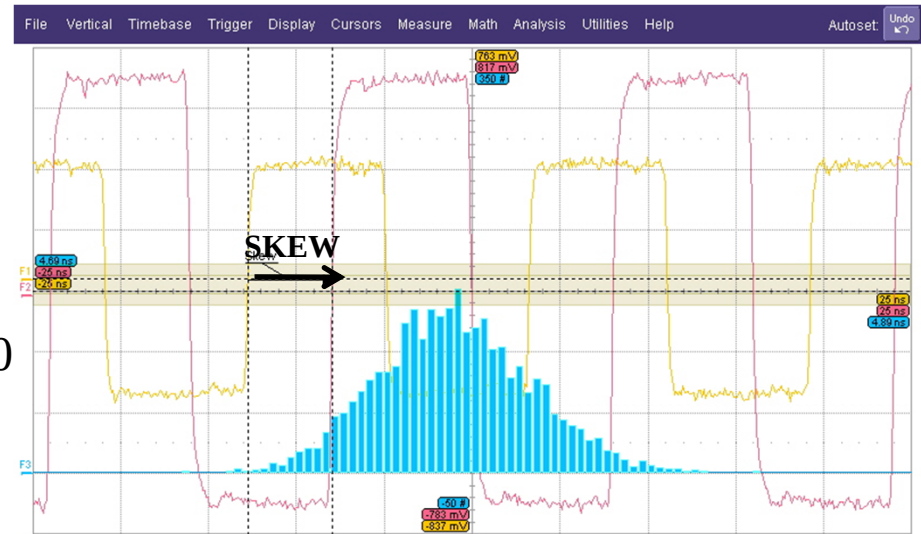
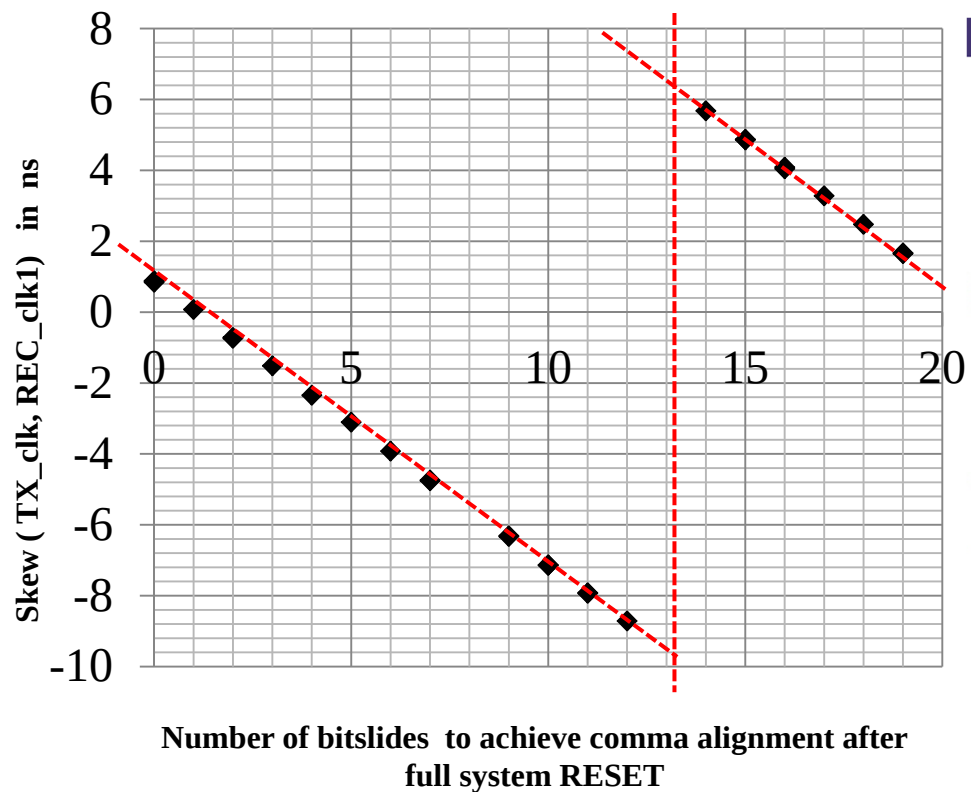
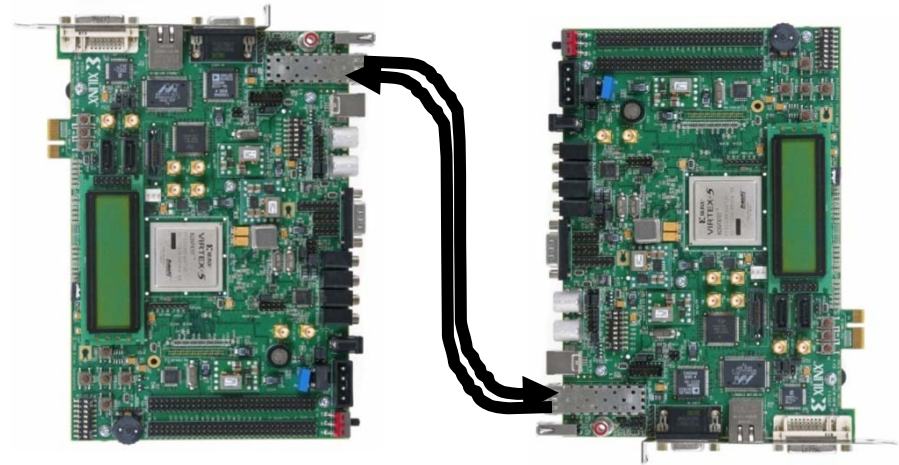
- Bus interne en 16 ou 20 bits de large. Interface utilisateur en 8, 16 ou 32 bits de large. Besoin de mots IDLES / COMMAS de 16 bits (20 bits après encodage) pour lever les indéterminations.
- Du fait de la serialisation / deserialisation de 20 bits, l'horloge reconstruite peut presenter 20 decalages possibles relativement a l'horloge propagée synchrone des données.
- Le GTX peut compenser ces décalages et resynchroniser les données parallèles et la RecClock.
- Désactiver l'alignement automatique. Trouver la bonne configuration de la logique d'alignement interne. Compter le nombre décalages de 0.8 ns nécessaires.
- GTX connecté au TEMAC qui génère par chance des IDLES 16 bits (K28.5+D5.6 ou K28.5+D16.2)
- Interface TEMAC coté GTX en 8 bits !
- Logique insérée entre TEMAC et GTX pour aligner correctement les IDLES sur l'horloge 62.5 MHz et pouvoir configurer le GTX en 16/20 bits. Necessary pour avoir une configuration valide de la logique interne d'alignement qui permette d'avoir une relation utilisable entre latence et nombre de Bitslides.

GTX alignment attributes

```
--Comma Detection and Alignment Attributes -----
ALIGN_COMMA_WORD_0    => 2,
COMMA_10B_ENABLE_0    => "1111111111",
COMMA_DOUBLE_0        => FALSE,
DEC_MCOMMA_DETECT_0    => TRUE,
DEC_PCOMMA_DETECT_0    => TRUE,
DEC_VALID_COMMA_ONLY_0 => TRUE,
MCOMMA_10B_VALUE_0    => "1010000011",
MCOMMA_DETECT_0       => TRUE,
PCOMMA_10B_VALUE_0    => "0101111100",
PCOMMA_DETECT_0       => TRUE,
RX_SLIDE_MODE_0       => "PCS ", -- PMA
-----
RXDATAWIDTH1          => "00",
TXDATAWIDTH0          => "00",
INTDATAWIDTH          => '1',
```



BITSLIDES et alignement manuel - Résultats

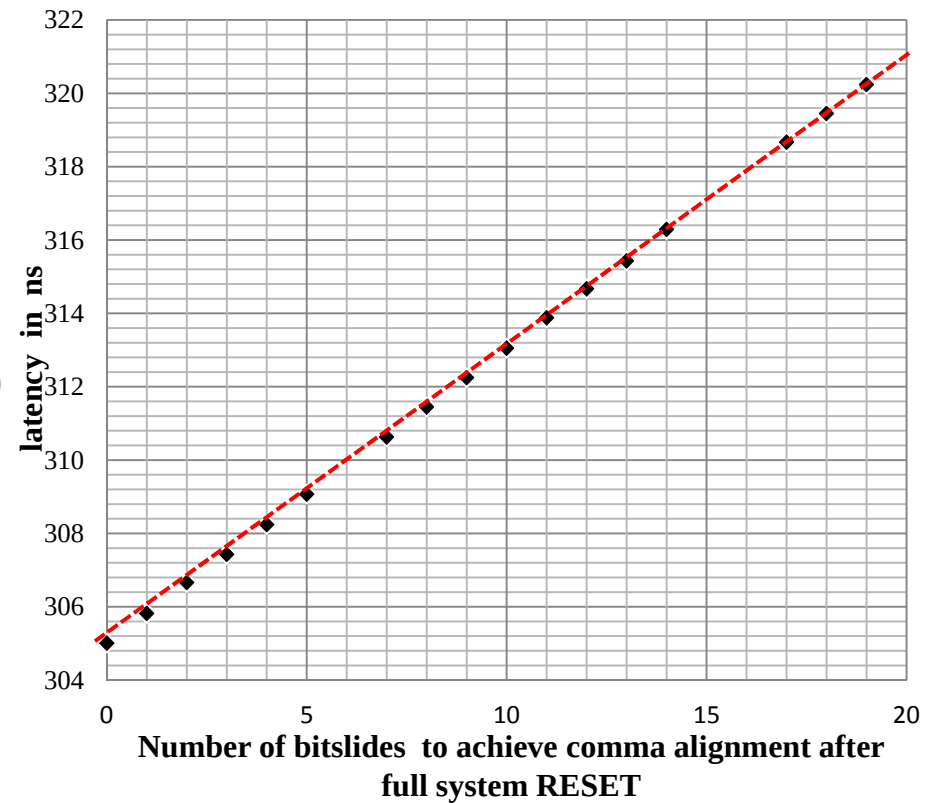
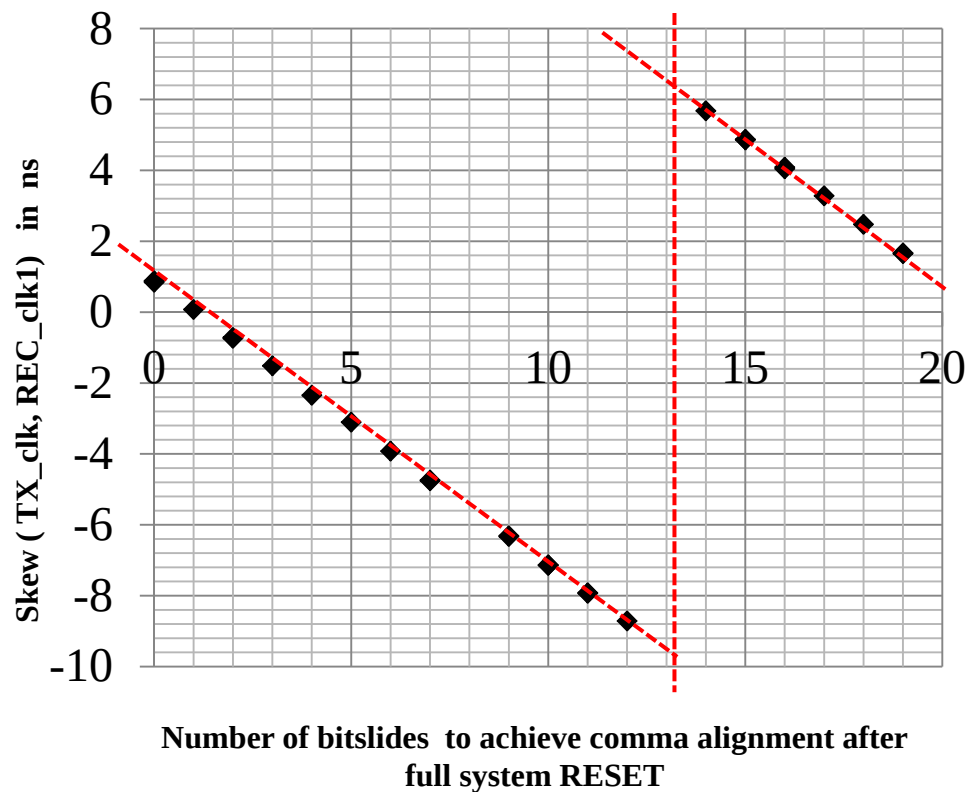
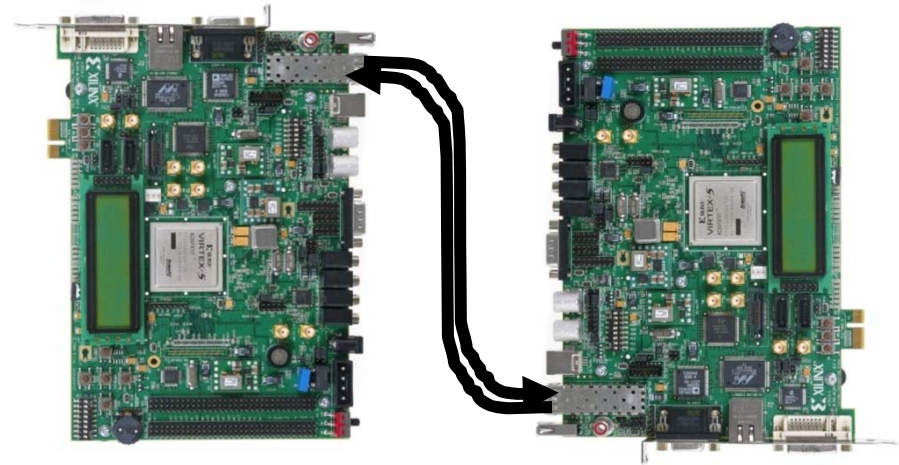


Histogram range < 200 ps

Histogram standard deviation < 25 ps

Sampling at 10 GS/s

BITSLIDES et alignement manuel : résultats



Transmitting back from offshore to shore using REC_clk1

Utiliser une seule horloge pour l'Aller et le Retour ?

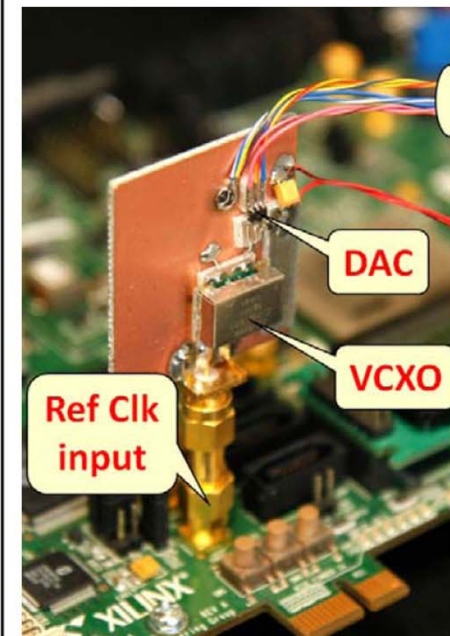
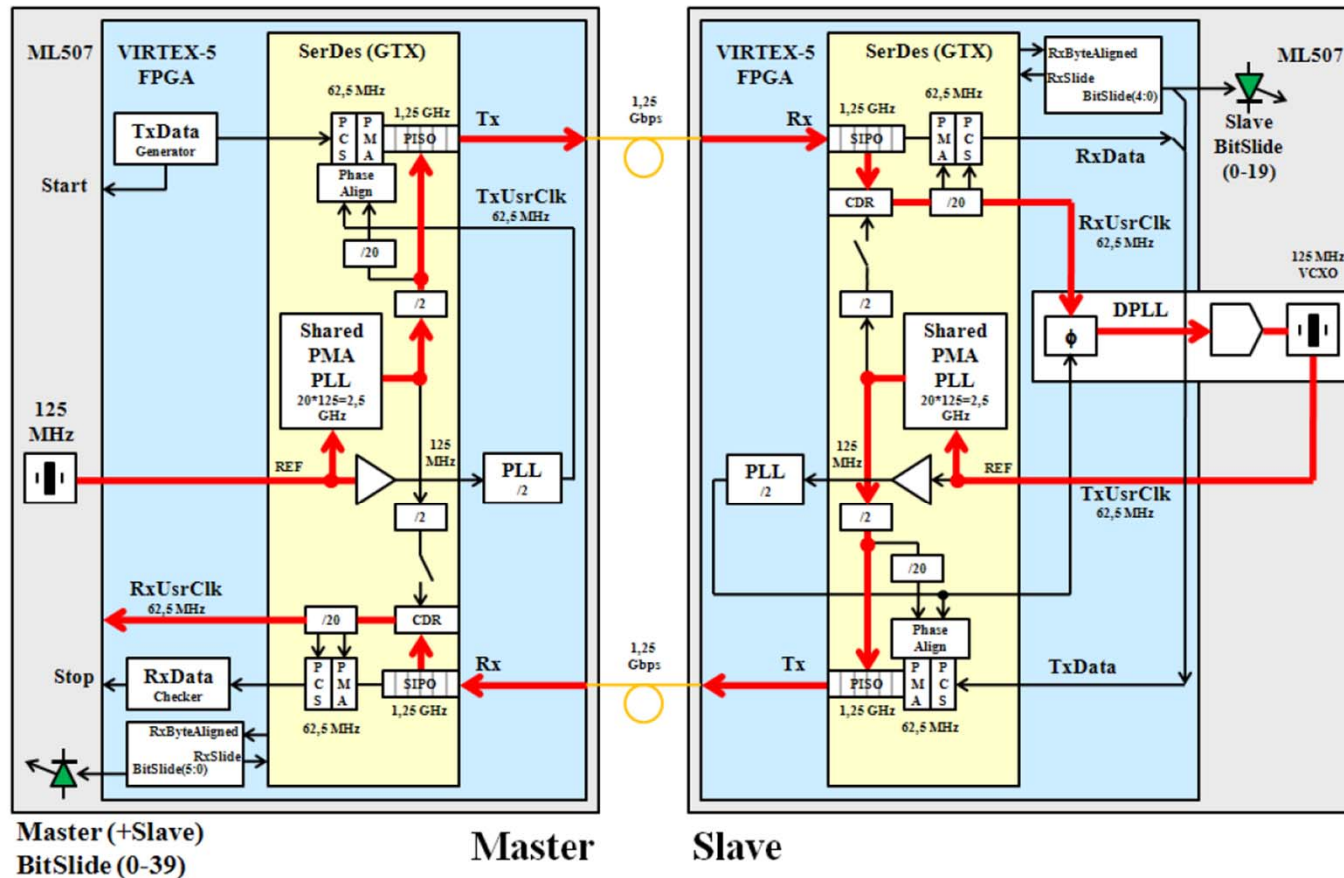
➔ Solution 1 :

Measuring propagation delay over a 1.25 Gbps bidirectional data link

P.P.M. Jansweijer,^a H.Z. Peck,^a

^aNikhef, Science Park 105, 1098 XG Amsterdam, Netherlands

+ White Rabbit



Transmitting back from offshore to shore using REC_clk1

Utiliser une seule horloge pour l'Aller et le Retour ?



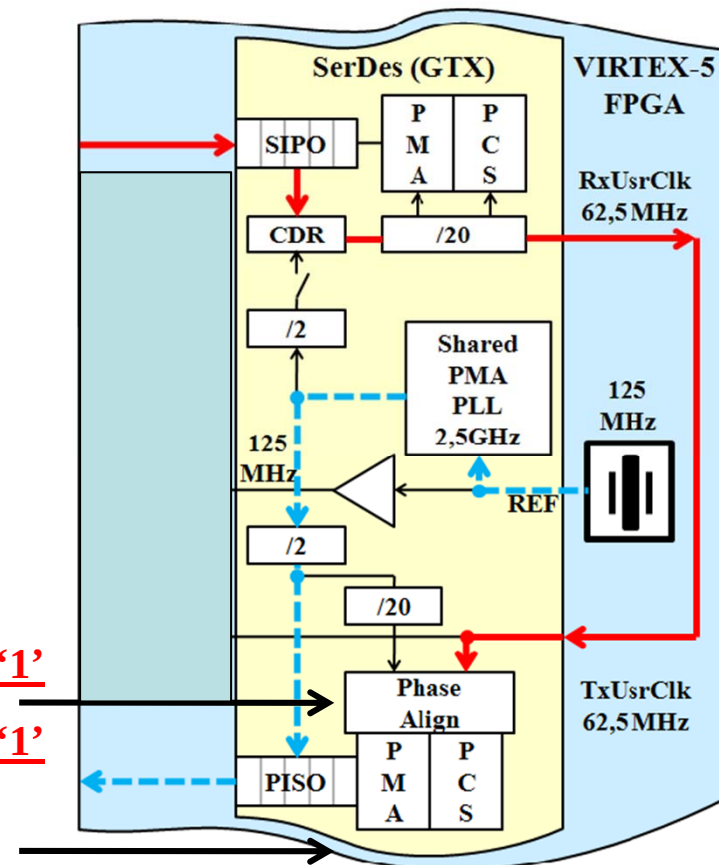
Solution 2 :

IBM Research Report

**Application Note : FPGA to IBM
Power Processor Interface Setup**

Ibrahim OUDA and Kai Schleupen

PMASETPHASE = '1'
ENPMAPHASEALIGN = '1'



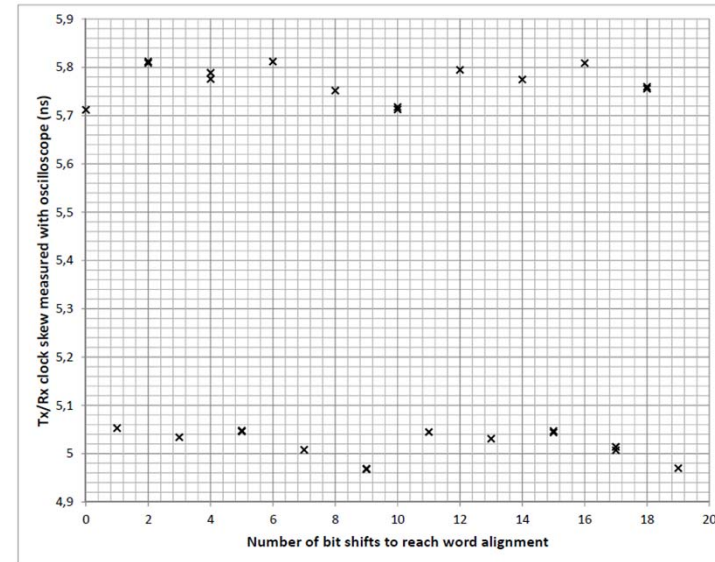
The slave node TxUsrClk directly fed by RxUsrClk



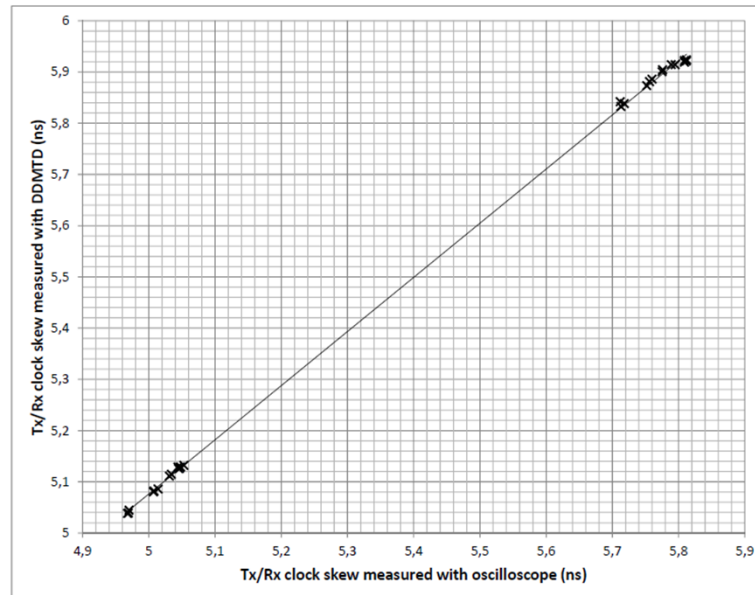
Le prix de la simplicité *pourrait être* une augmentation du jitter.

PCS vs PMA bitslides

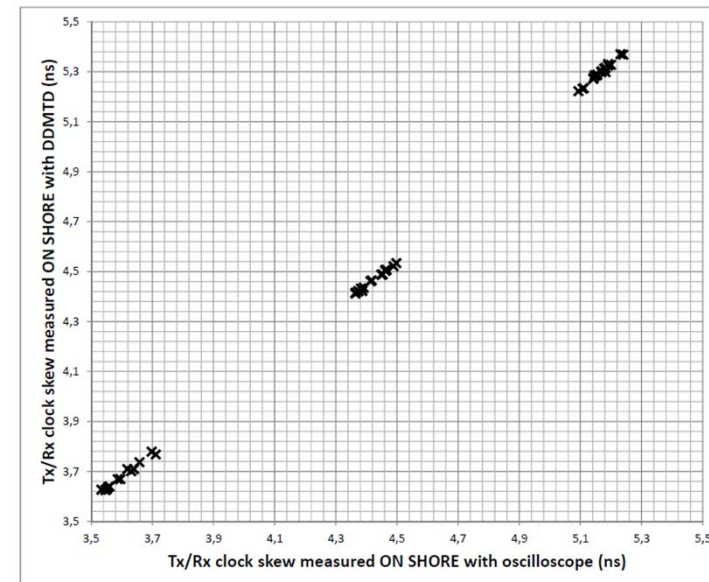
- The GTX can be configured to achieve bit sliding in the PCS or in the PMA.
- **PCS** : 20 different values of clock skew and latency associated to 20 different values of the number of bitslides (one way trip).
- **PMA** : 2 different values of clock skew and latency which map to even or odd numbers of bitslides. (R. Giordano *et al.*)



Aller simple

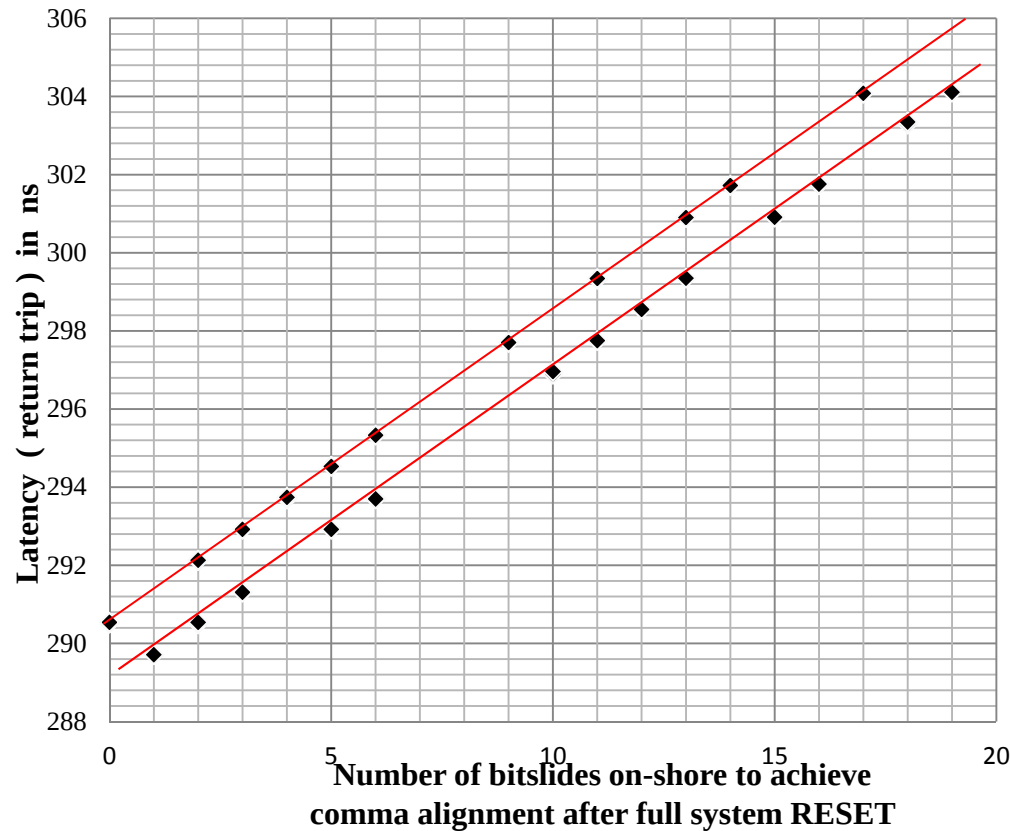


Aller simple



Aller et Retour

Unresolved random 1.6 ns latency on older Virtex 5



This issue appears to be hardware dependent : tests reveal that there is an unexplained problem on the RX channel of only one out of five ML507 boards. Appears to be correlated with **Virtex 5 date code**.

- Web case open but no clear answer : “could very well occur on all Virtex 5 , try extend temperaturatur range, voltage ..”
- Could be due to the RX Phase alignment block ... Needs some more investigating !

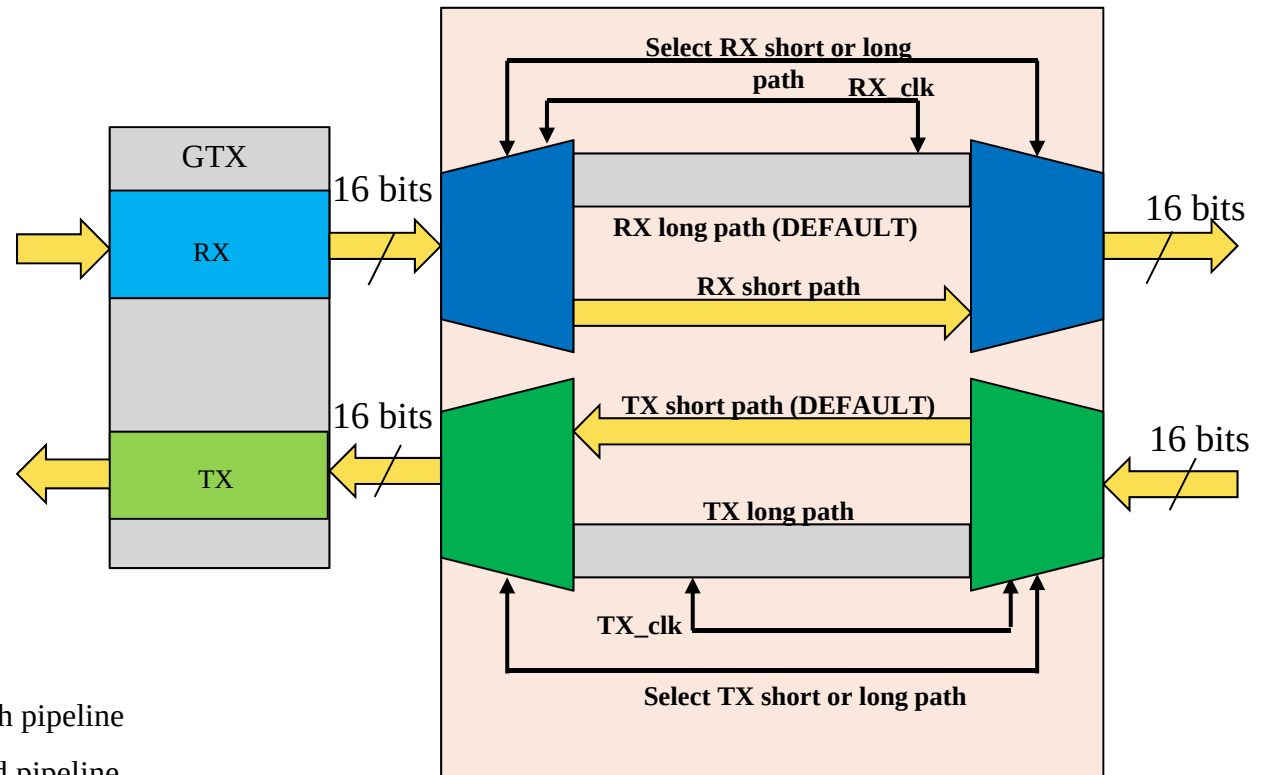
Synchronous Commands

Incoming commands :

- commands detected inside RX default long path pipeline. Could use a specific K-char to tag commands and second byte to define 256 commands.
- switch to short path , remove command from RX pipeline and replace it with 16 bit idle.
- switch back to default long path within inter frame gap, which is then one idle longer than initial gap.

Outgoing commands :

- command to be inserted is stored in long path pipeline
- switch to long path : move command out and pipeline Ethernet flow
- switch back to default short path within inter frame gap by removing one 16 bit idle.



N.B. The data bus is pipelined along with several other signals (running disparity, char_is_K, etc.)



Design is implemented in hardware for system_reset, tx_inhibit, start_of_run, reset_time_stamp, flash_led, etc.

Able to handle the necessary low rate of synchronous commands.

Commands inserted anywhere inside Ethernet flow without corrupting ethernet frames.

TEMAC

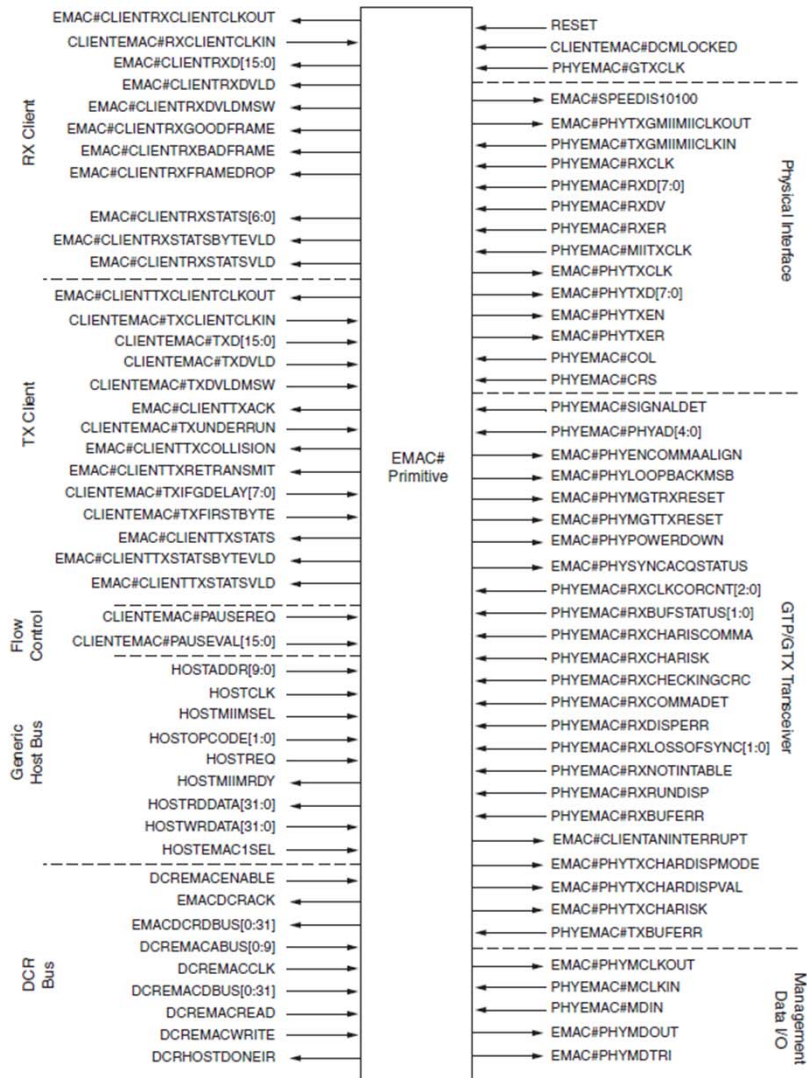
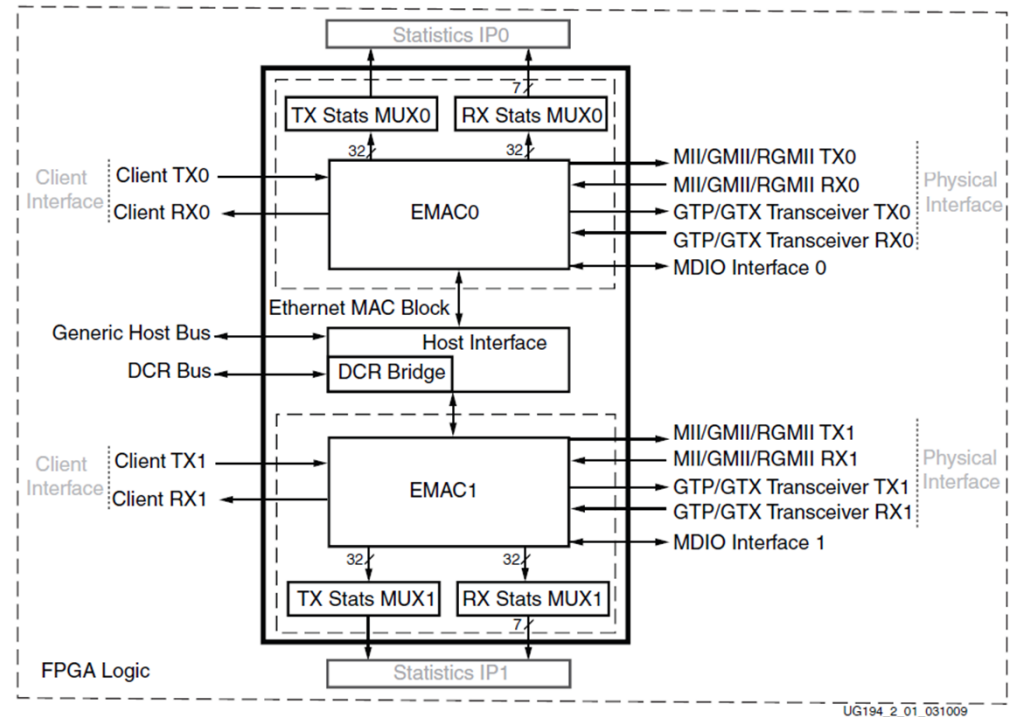
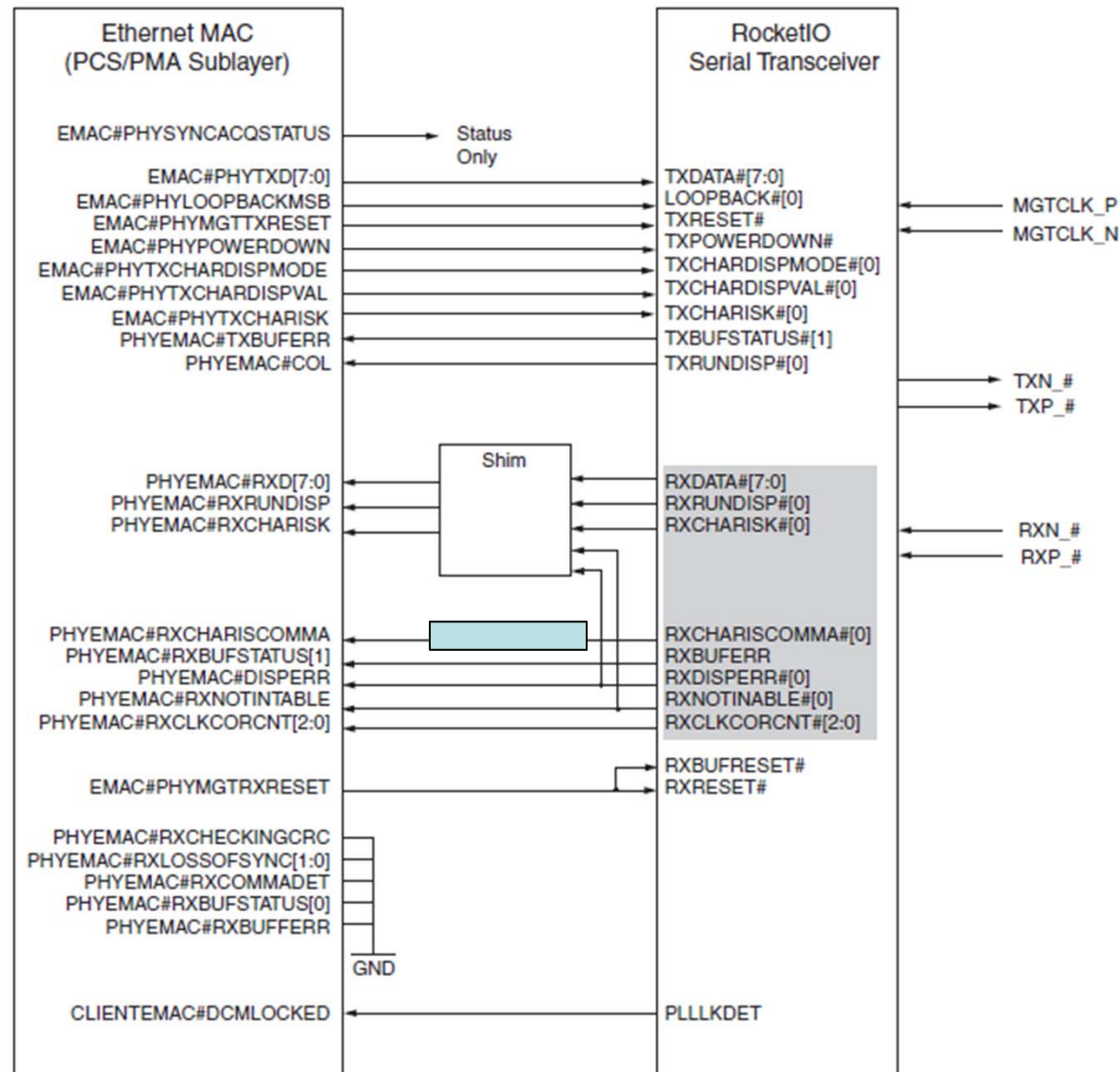


Figure 2-3: Ethernet MAC Primitive

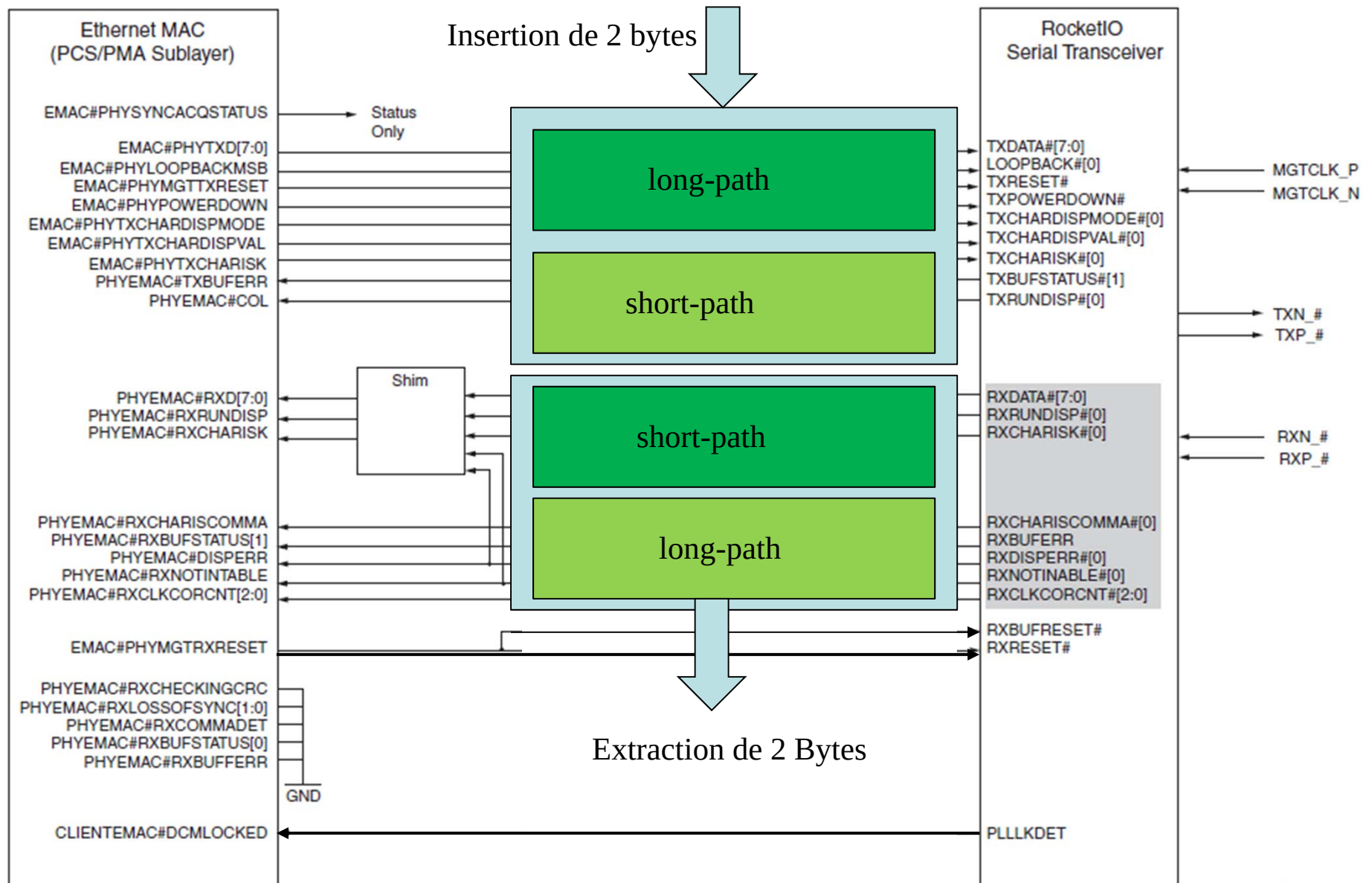


UG194_2_01_031009

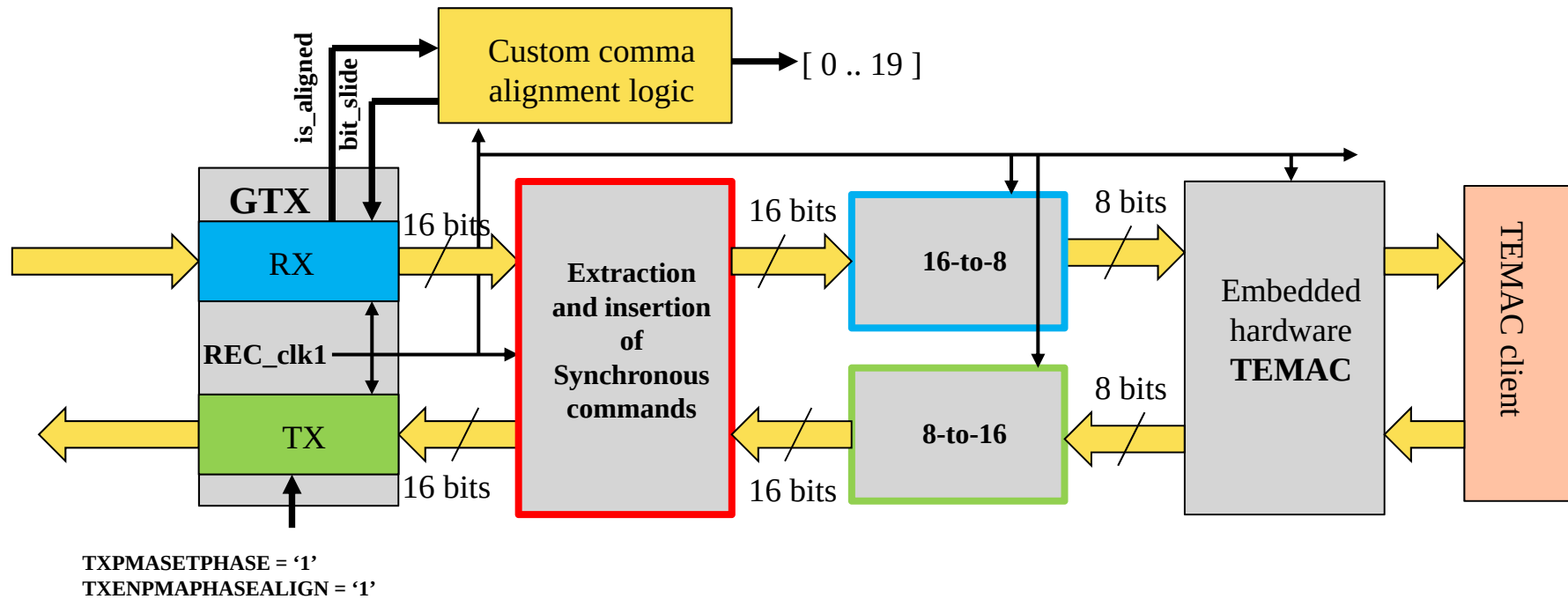
TEMAC to GTX connections in 1000 BaseX mode



TEMAC et Commandes Synchrones

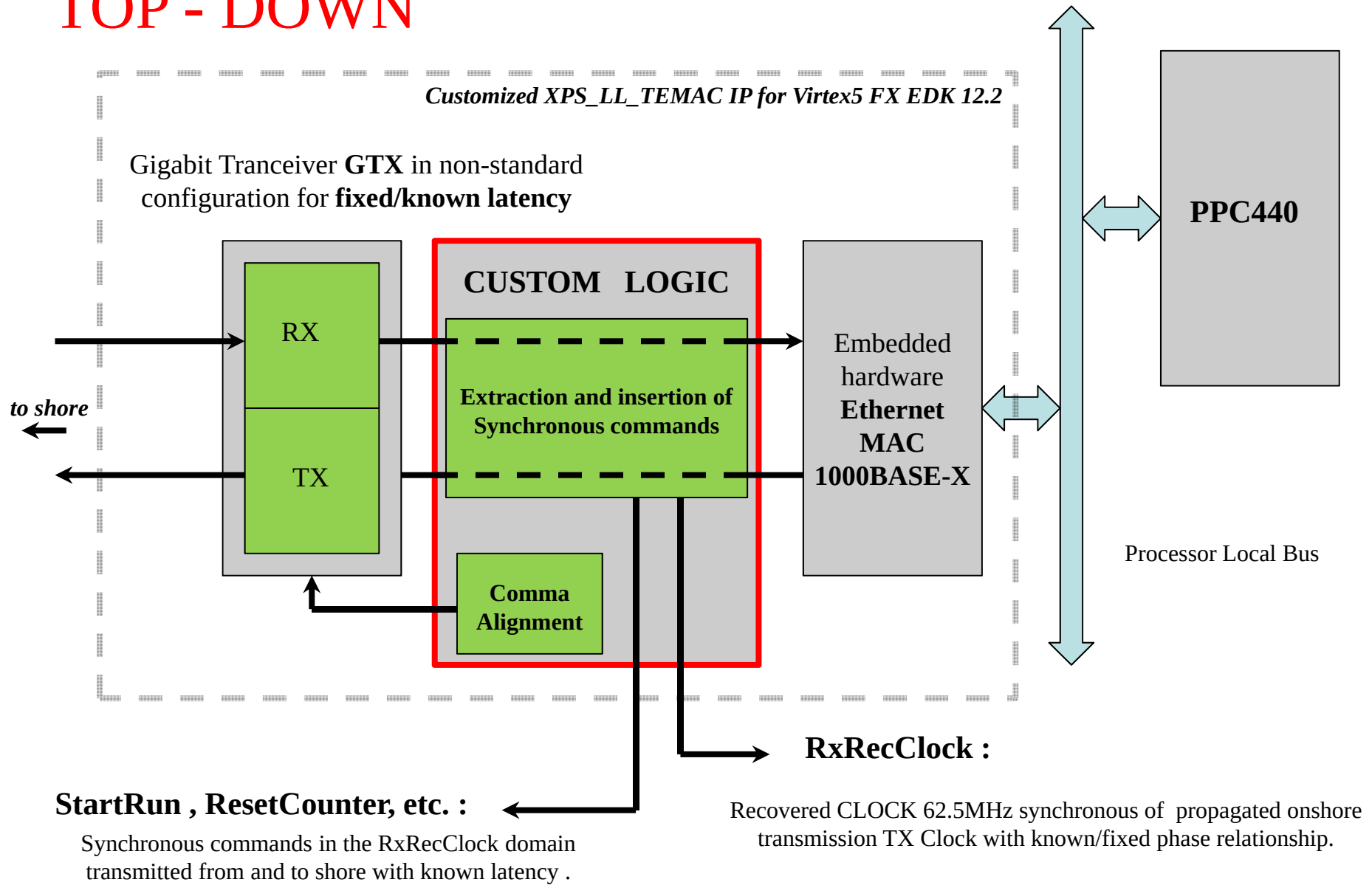


BOTTOM - UP



- TX and RX elastic buffers bypassed
- 16 bit idles are synchronized with the 62.5 MHz clock on TX path.

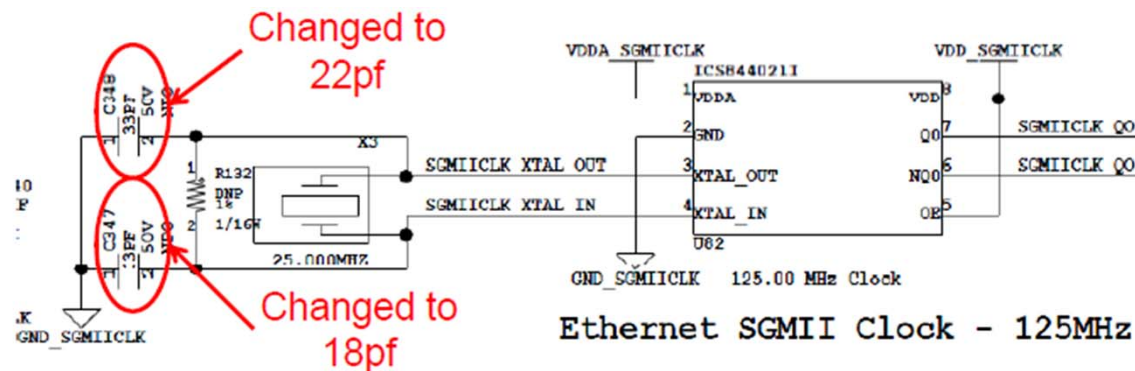
TOP - DOWN



- 1000BASE-X Ethernet drivers for VxWorks6.3
- logic insertion between TEMAC and GTX in the XPS_LL_TEMAC IP requires EDK12.2 (sources encrypted in prior versions!)

Onshore crate : Virtex 6 inside !

- Beaucoup de similarités entre GTX du V5 et GTX du V6
- GTX du V6 plus “friendly” : RX/ TX indépendants.
- Emission et reception dans deux domaines d’horloge différents mais synchrones
- Bridge entre deux domaines ethernet : deux LL_temac mis dos a dos.
- Coté “mer”, customized pour la latence fixe, coté “terre” standard
- insertion extraction de commandes
- Pour que **RxByteIsAligned** passe à 1, il faut corriger la valeur du parametre SHOW_REALIGN_COMMA donnée par défaut par le Core Generator comme suit :
SHOW_REALIGN_COMMA => (FALSE)
- L’horloge 125 MHz fournie sur la ML605 est hors specifications (P. Jansweijer, NIKHEF)



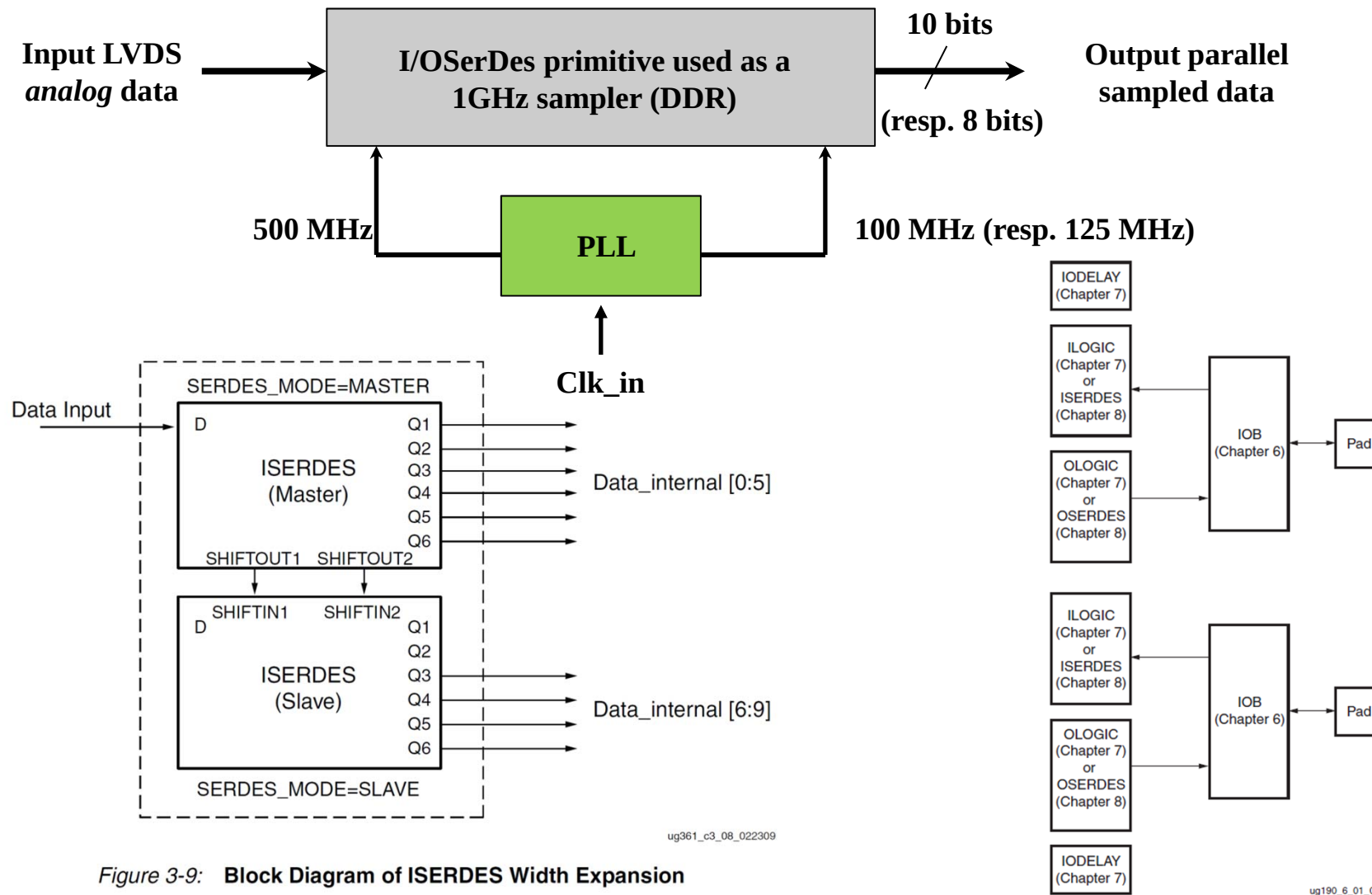
Frequency of the crystal ~81 ppm off target : SGMIICLK_Q0 = 124989880



Frequency measured after lowering C324 and C325: SFPCLK Q0 = 124998185 (= -15 ppm)

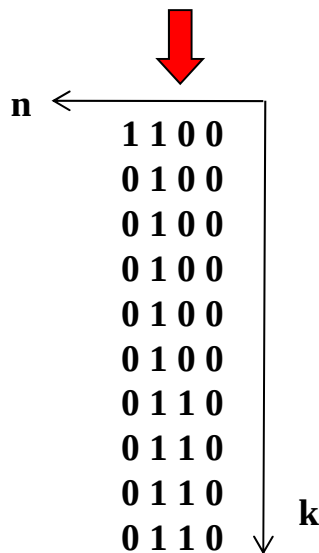
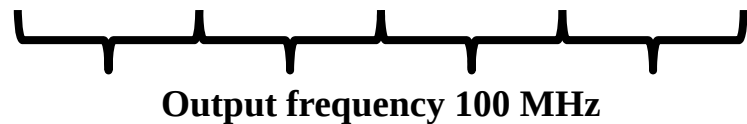
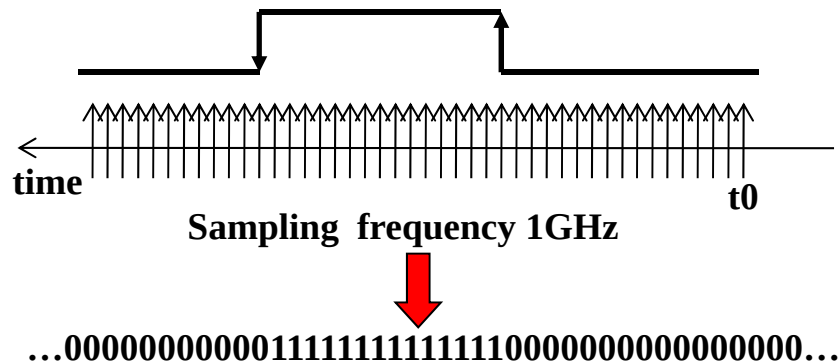
Single channel 1 GHz sampler as TDC

UG190 chapitre 8
UG361 chapitre 3

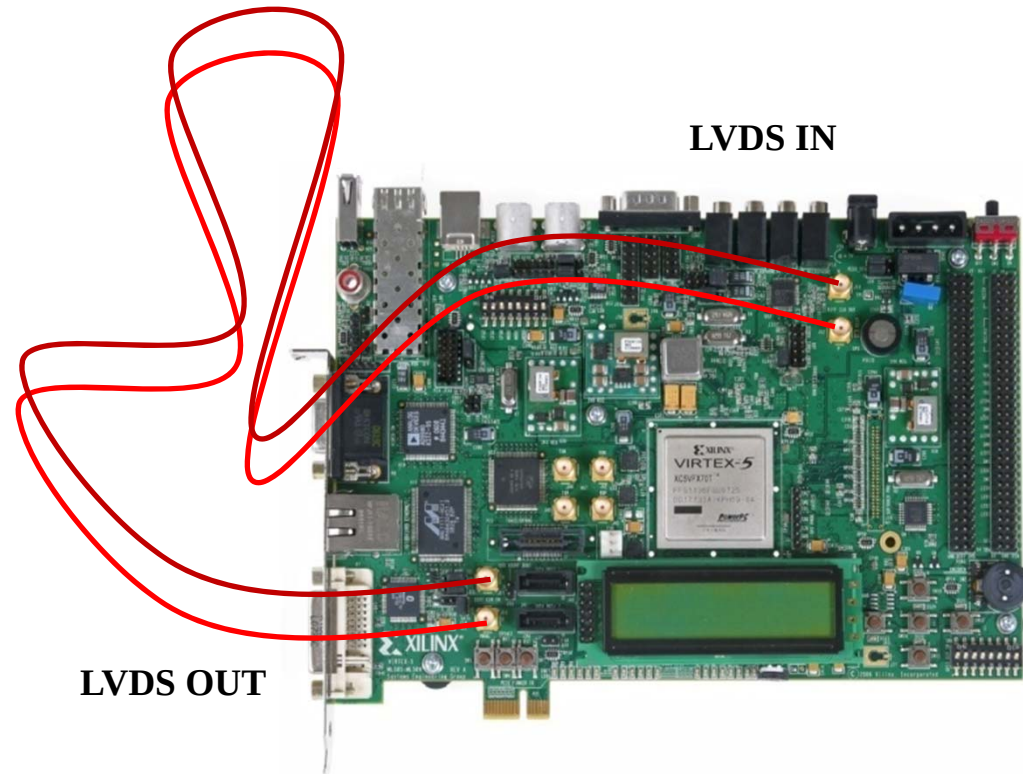


[Similar design by Albert ZWART (NIKHEF) / small PMTs test bench for ALTERA]

Single channel TDC

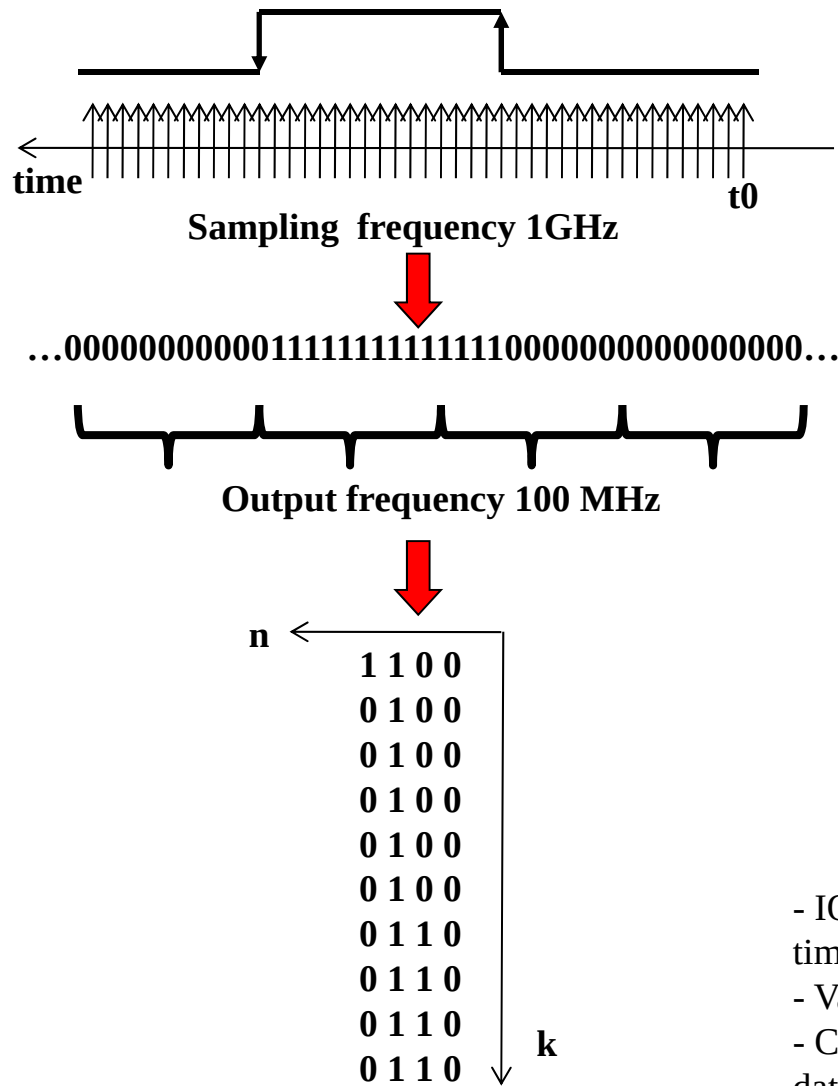


$$t = t_0 + n * 10ns + k * 1ns$$

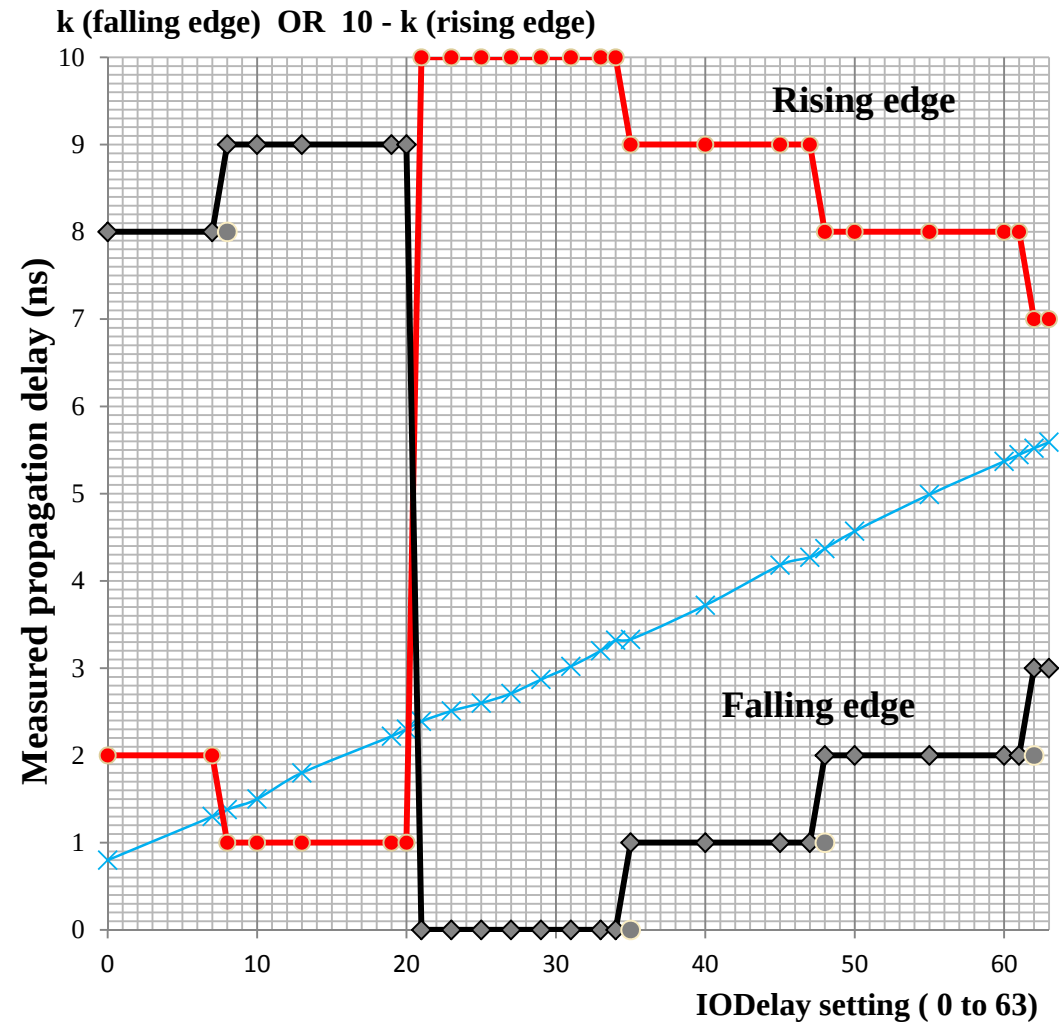


- IOdelay on the output path to insert a propagation delay 1 to 64 times 78 picoseconds.
- Validation of proposed single channel TDC architecture.
- Check for metastability : detect 010 or 101 sequences in the parallel data. None observed.

Single channel TDC

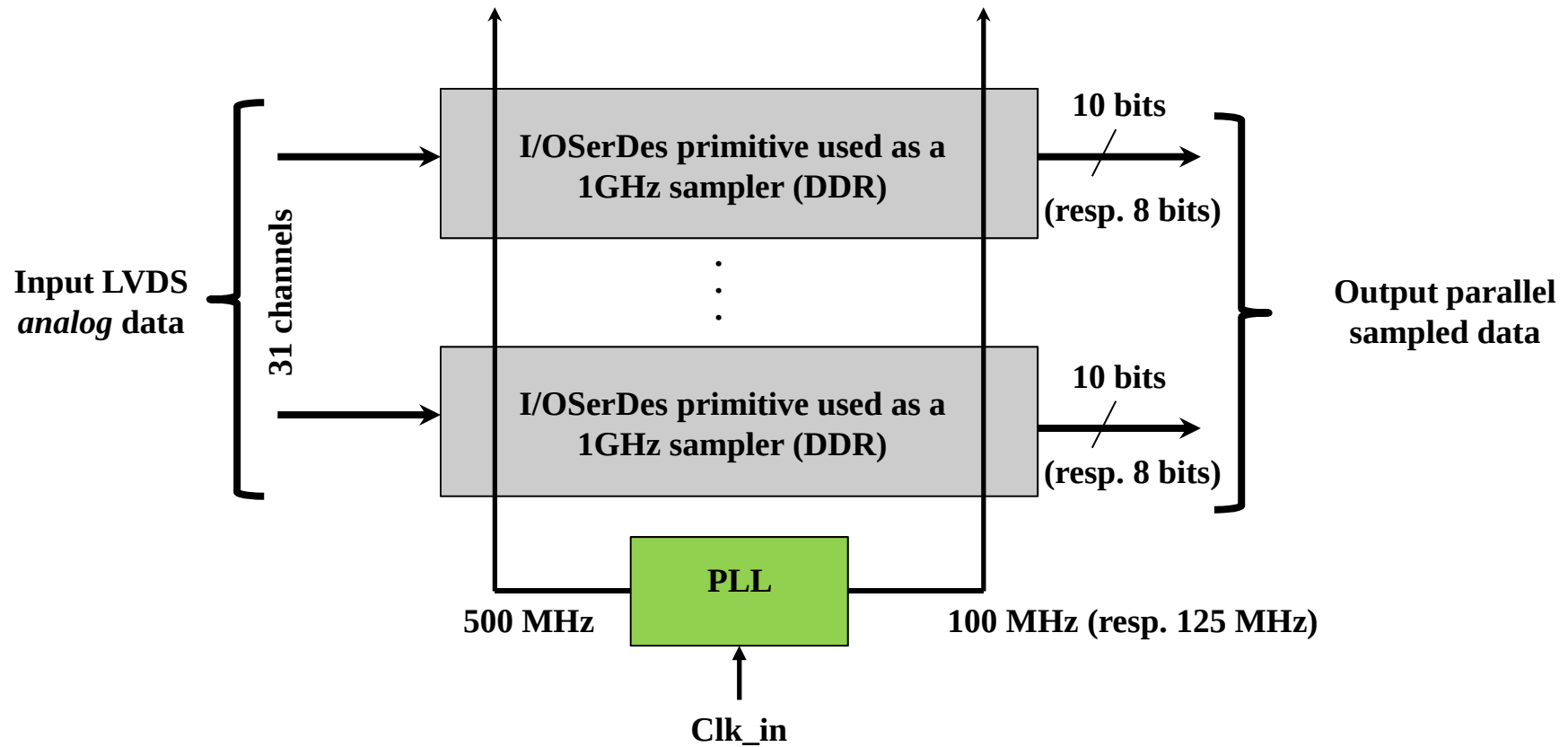


$$t = t_0 + n * 10ns + k * 1ns$$



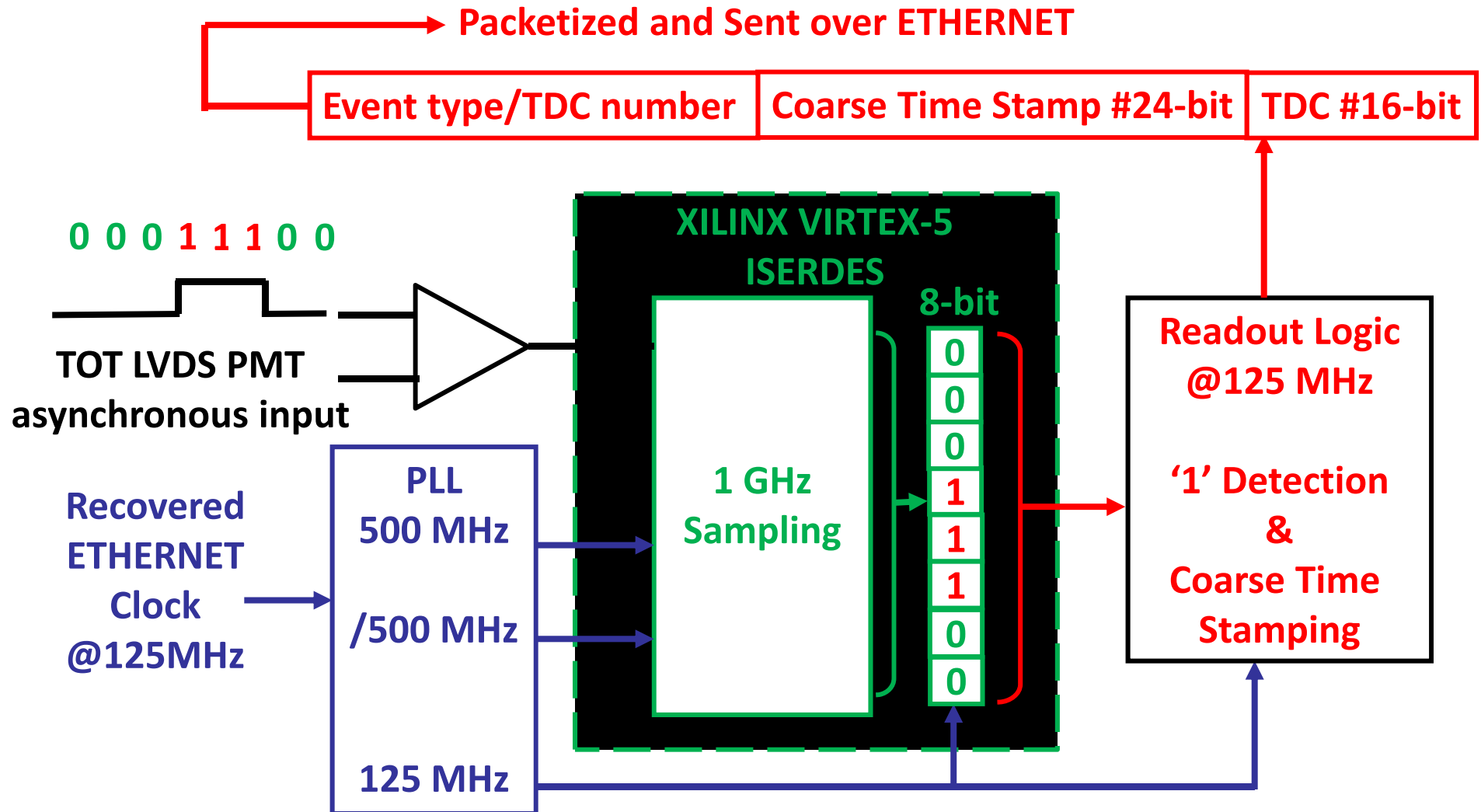
- IODelay on the output path to insert a propagation delay 1 to 64 times 78 picoseconds.
- Validation of proposed single channel TDC architecture.
- Check for metastability : detect 010 or 101 sequences in the parallel data. None observed.

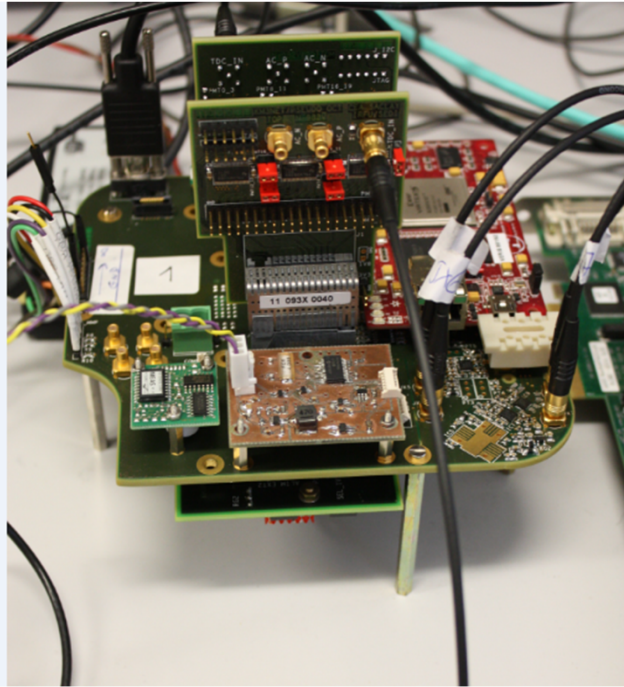
TDC 31 VOIES



- Synthesized, mapped and routed in a Virtex 5 FXT70.
- Xilinx routing tools show a 20 ps dispersion in the fast clock arrival times to the 31 samplers.

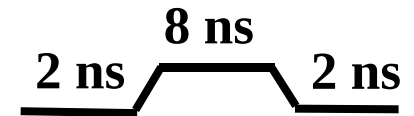
DOM 31 PMT readout





TDC : résultats quantitatifs (1)

Générateur d'impulsions :



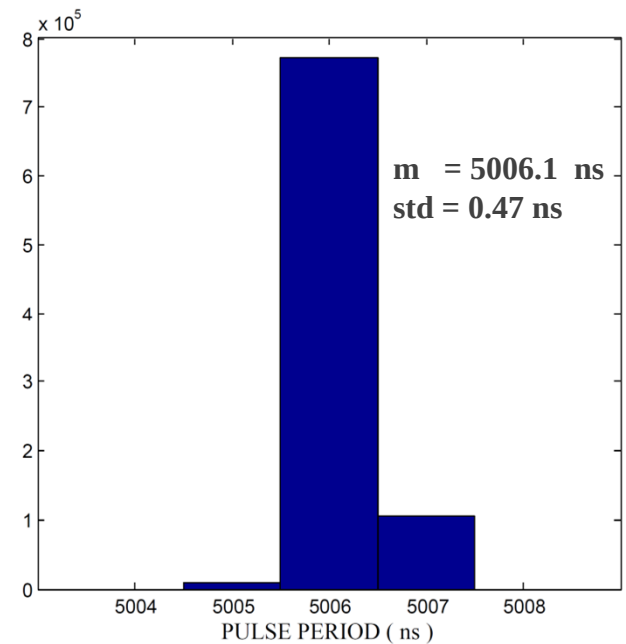
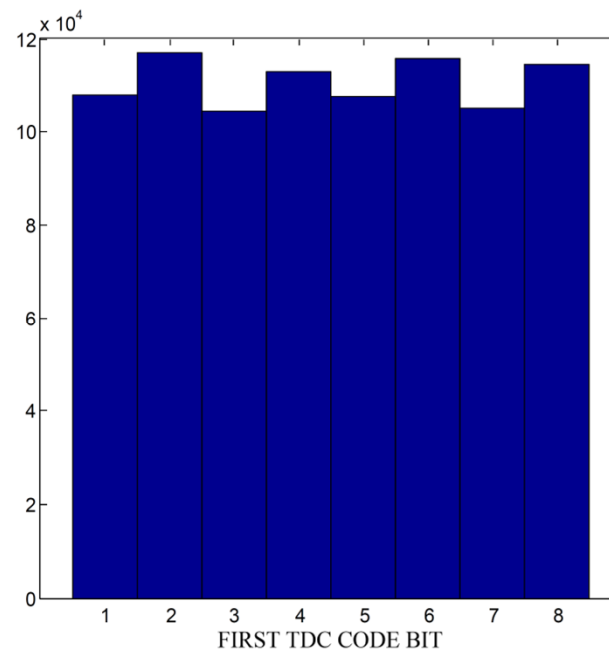
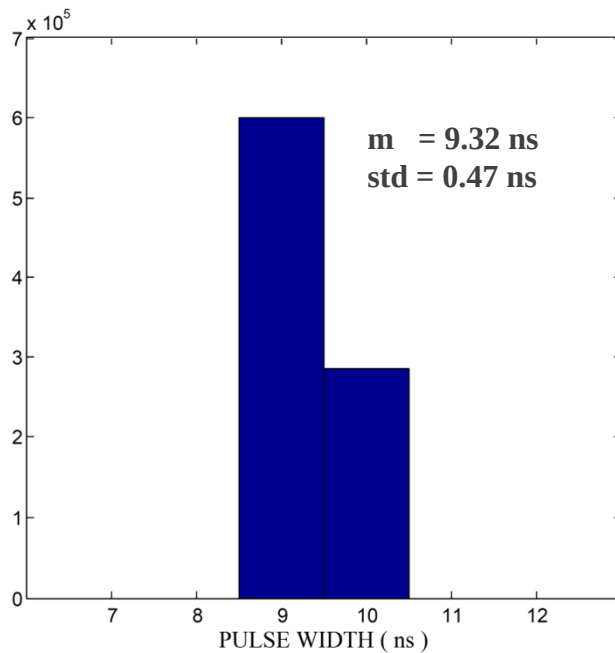
Fréquence : 200 kHz

(LeCroy 64Xi 600 MHz donne std = 220 ps)

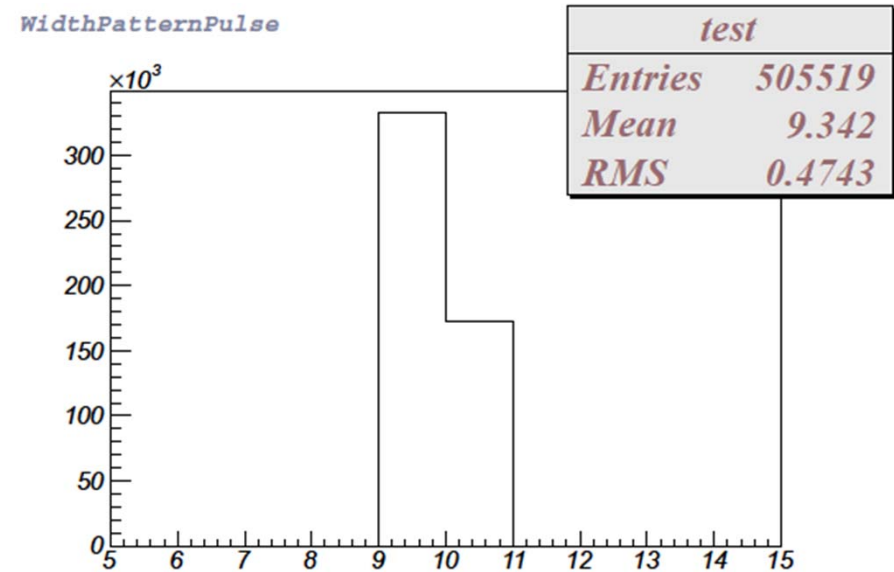
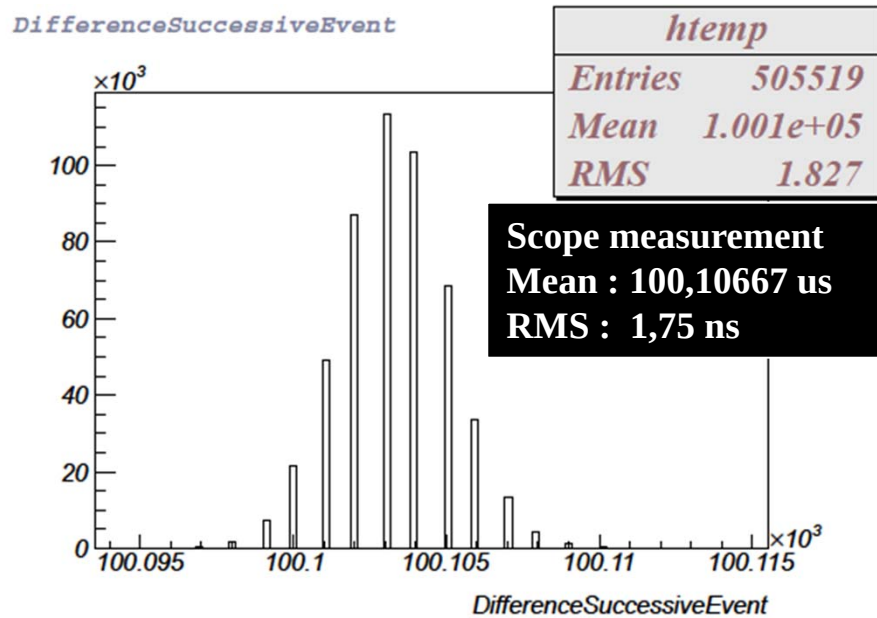


Les cartes pseudo Octopus réalisent un fan out sur les 31 voies LVDS.

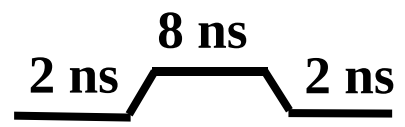
Analyse des mesures sur une voie :



TDC : résultats quantitatifs (2)

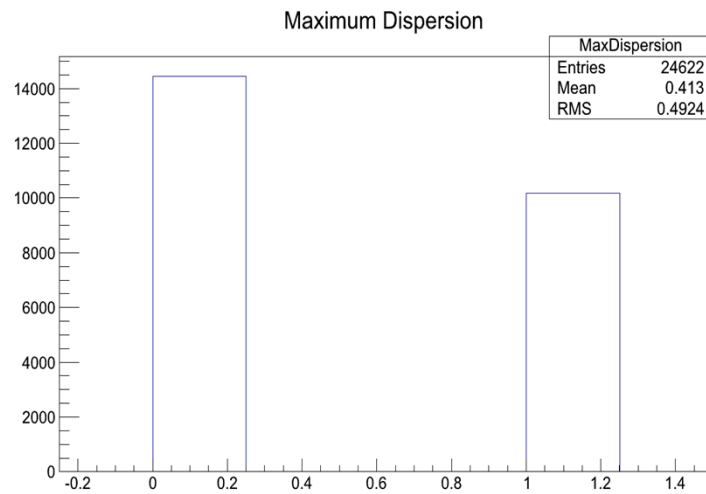


Pulse @ f=10 kHz

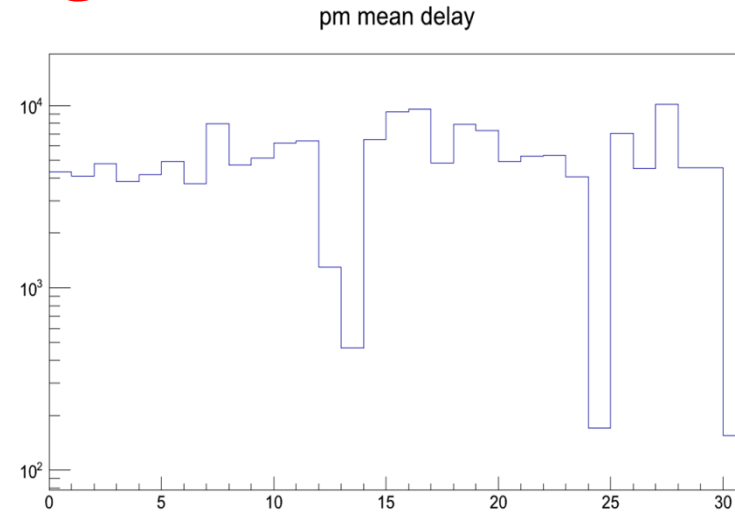


A diagram of a pulse waveform. It shows a series of horizontal lines representing the pulse. The first segment is labeled '2 ns', followed by a rising slope, a flat top labeled '8 ns', a falling slope, and a final segment labeled '2 ns'.

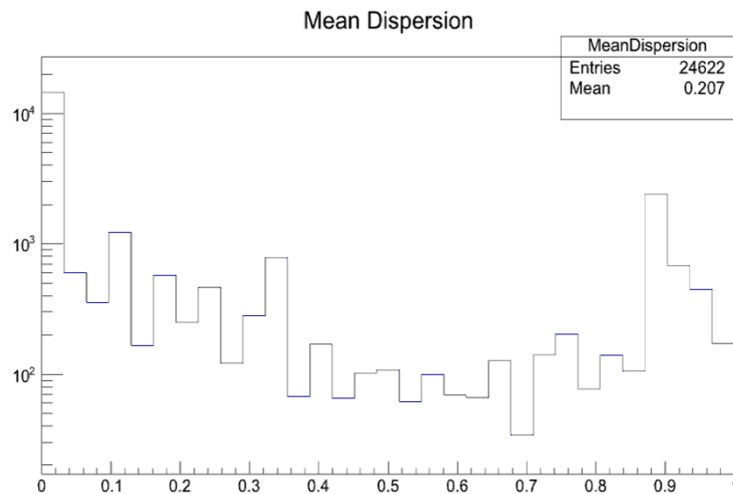
Caractérisation du ... TDC ? générateur ? ...



Différence des temps d'arrivée entre la première voie et la dernière voie touchées.



Nombre de fois qu'une voie est touchée en même temps que la première voie touchée.



Nombre de voies touchées une nanoseconde après la première voie touchée.



Permet aussi de valider la chaîne d'acquisition dans son ensemble : pas de trous dans les données, Jumbo frames, mesures de débit, etc.

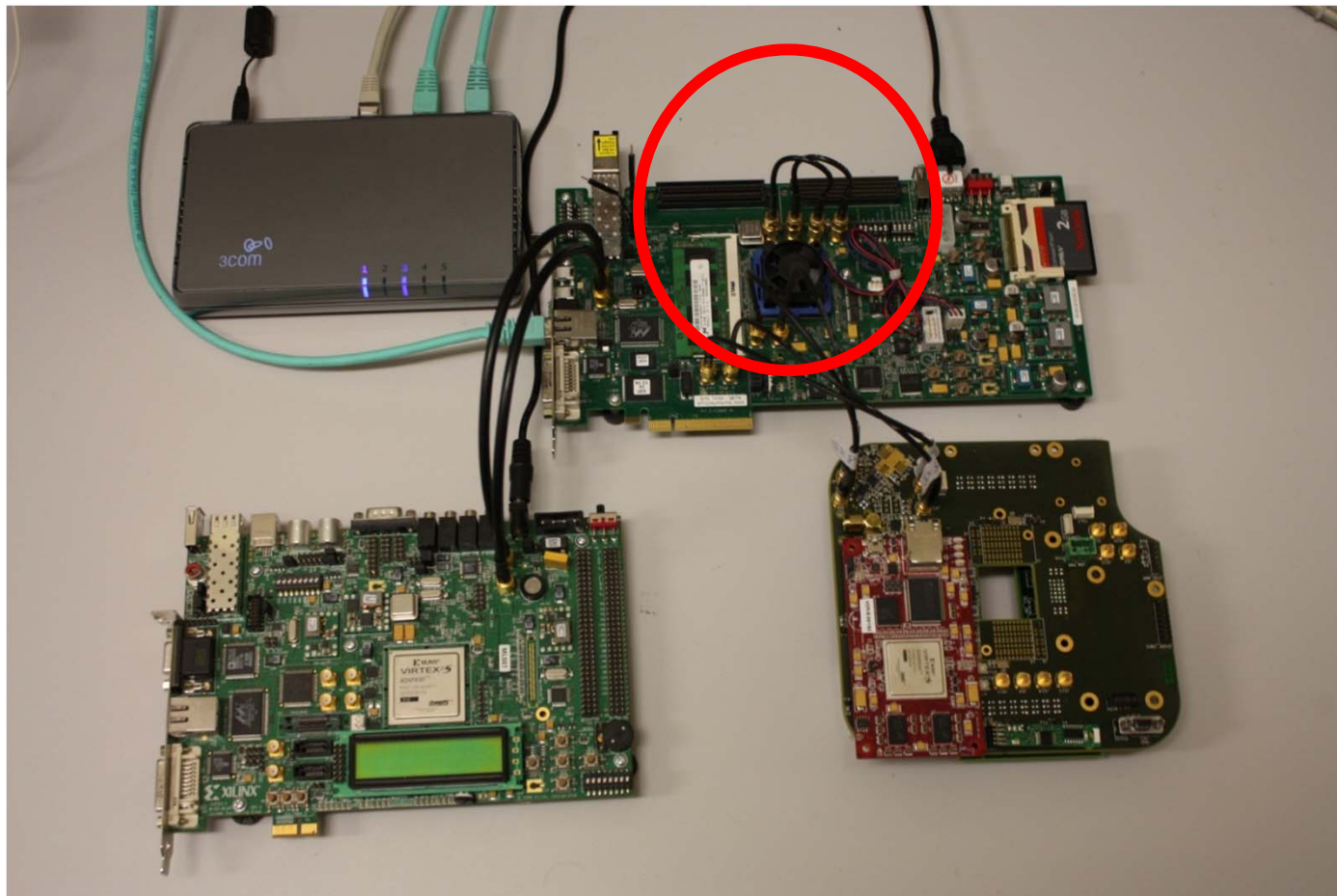
KDAQ avec ZeroCopyBuffer et MTU = 7000 : débit = 437 Mbit/s
TTCP avec ZeroCopyBuffer et MTU = 7000 : débit = 708 Mbit/s
TTCP sans ZeroCopyBuffer et MTU = 7000 : débit = 185 Mbit/s

TDC “ON SHORE” pour mesure de latence AR

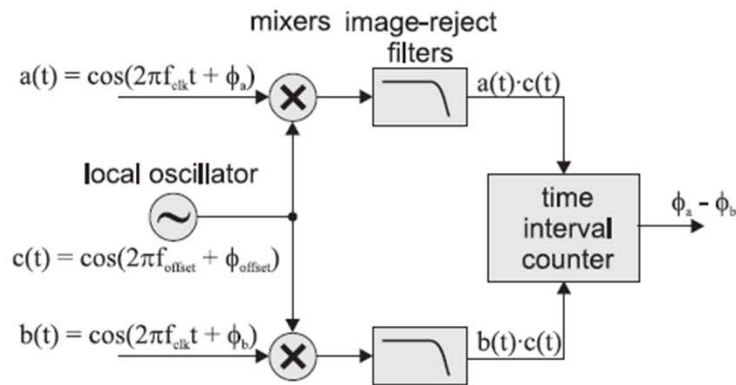
Commande synchrone émise vers le DOM => “START” synchrone de TXCLK

Commande synchrone reçue (echo) du DOM => “STOP” synchrone de RXCLK

Ces impulsions sont combinées et sorties du Virtex 6 sur un seul lien LVDS vers l’entrée différentielle d’un TDC implémenté sur ce meme FPGA.



Dual Mixing Time Difference (DMTD)



$$\begin{aligned} a(t) \cdot c(t) &= \cos(2\pi t f_{clk} + \phi_a) \cdot \cos(2\pi t f_{offset} + \phi_{offset}) \\ &= \frac{1}{2} \cos(2\pi t (f_{clk} + f_{offset}) + \phi_a + \phi_{offset}) \\ &\quad + \frac{1}{2} \cos(2\pi t (f_{clk} - f_{offset}) + \phi_a - \phi_{offset}) \end{aligned}$$

(from T. Wlotowski's thesis)

Digital Dual Mixing Time Difference (DDMTD)

(from T. Wlotowski's thesis)

- Mesure de Phase par comptage
- Mélange / translation en fréquence par échantillonnage
- White Rabbit préfère une PLL externe
- Risqué de cascader les PLL internes du Virtex 6 ?
- Compromis temps d'intégration / précision de la mesure / proximité des fréquences
- Filtrage des "glitches"

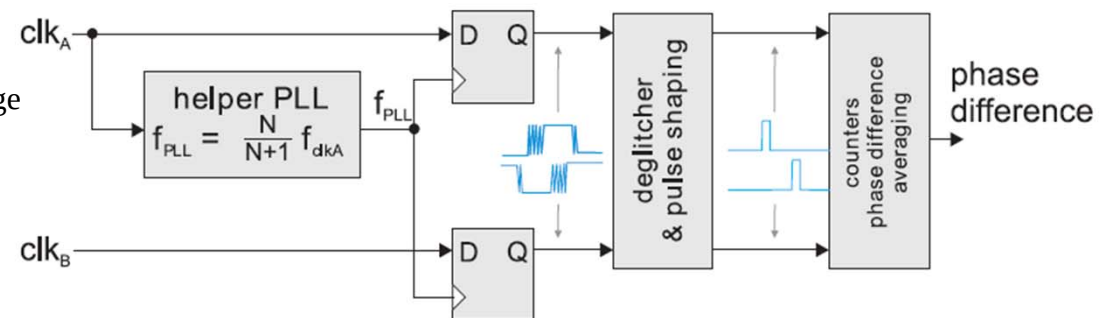
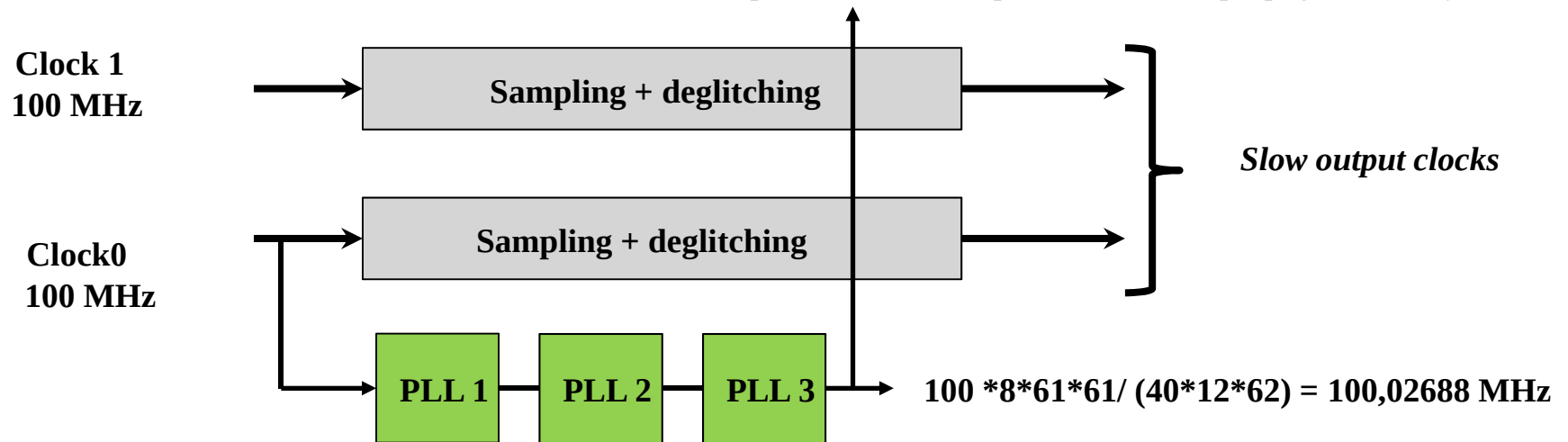


Figure 3.19. Structure of a digital DMTD phase detector

DDMTD

Using the same ML507 testbench, Clock1 is generated as a delayed (out of phase) Clock0 with a programmable delay multiple (0 to 63) of 78ps in addition to propagation delay.

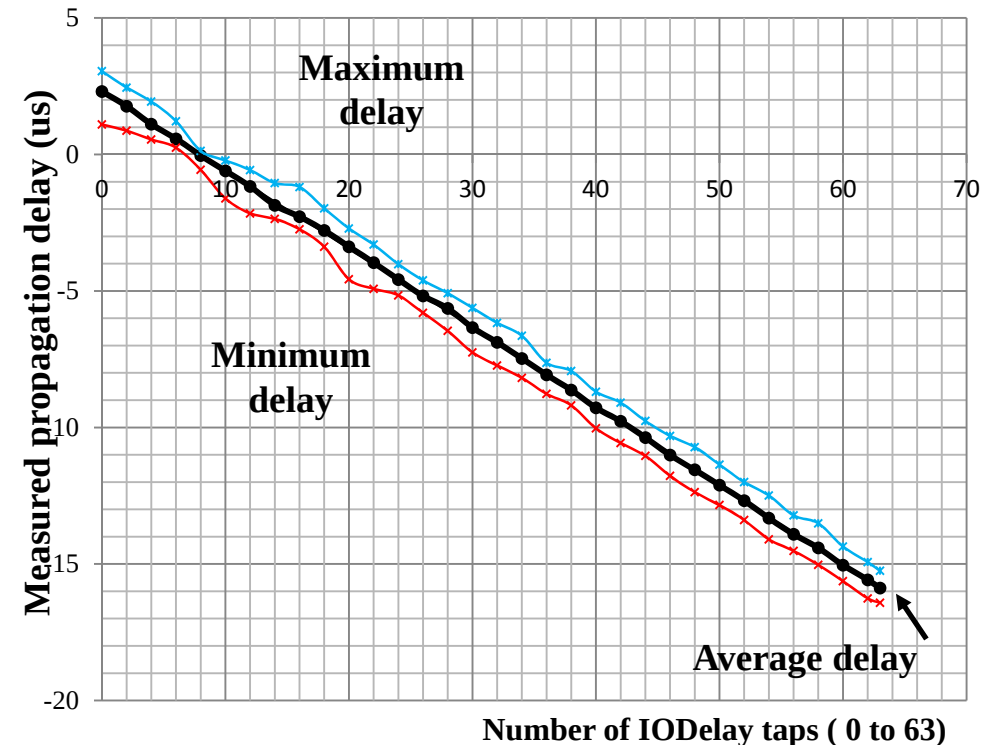


Delays amplified by a factor $1/0,0002688$ are more conveniently measured on output.

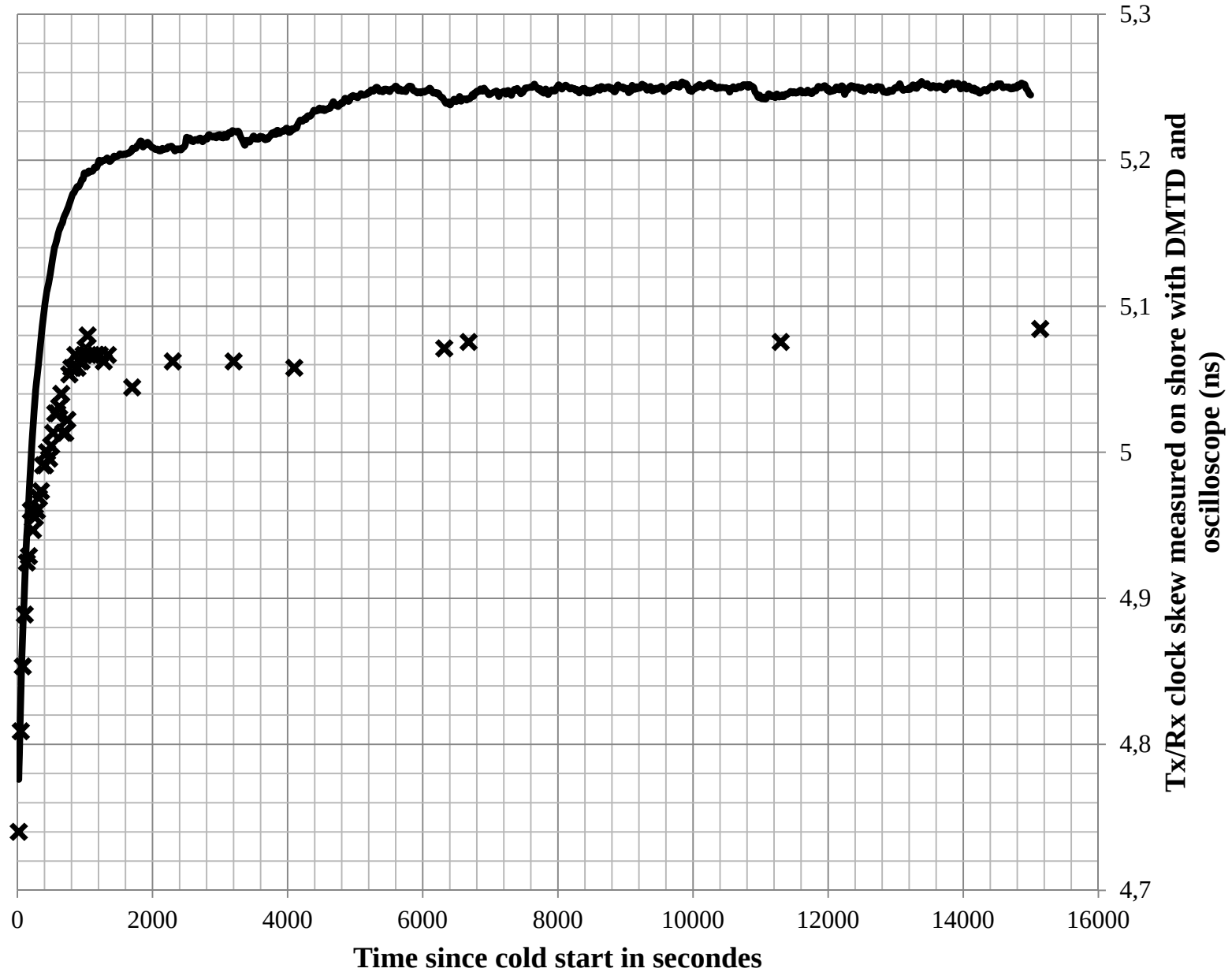
Estimated slope is 0,2888 us per delay tap which is coherent after conversion with 78 ps per tap : $0,2888 * 0,0002688 * 1000000 = 77,65 \text{ ps}$.

To Do :

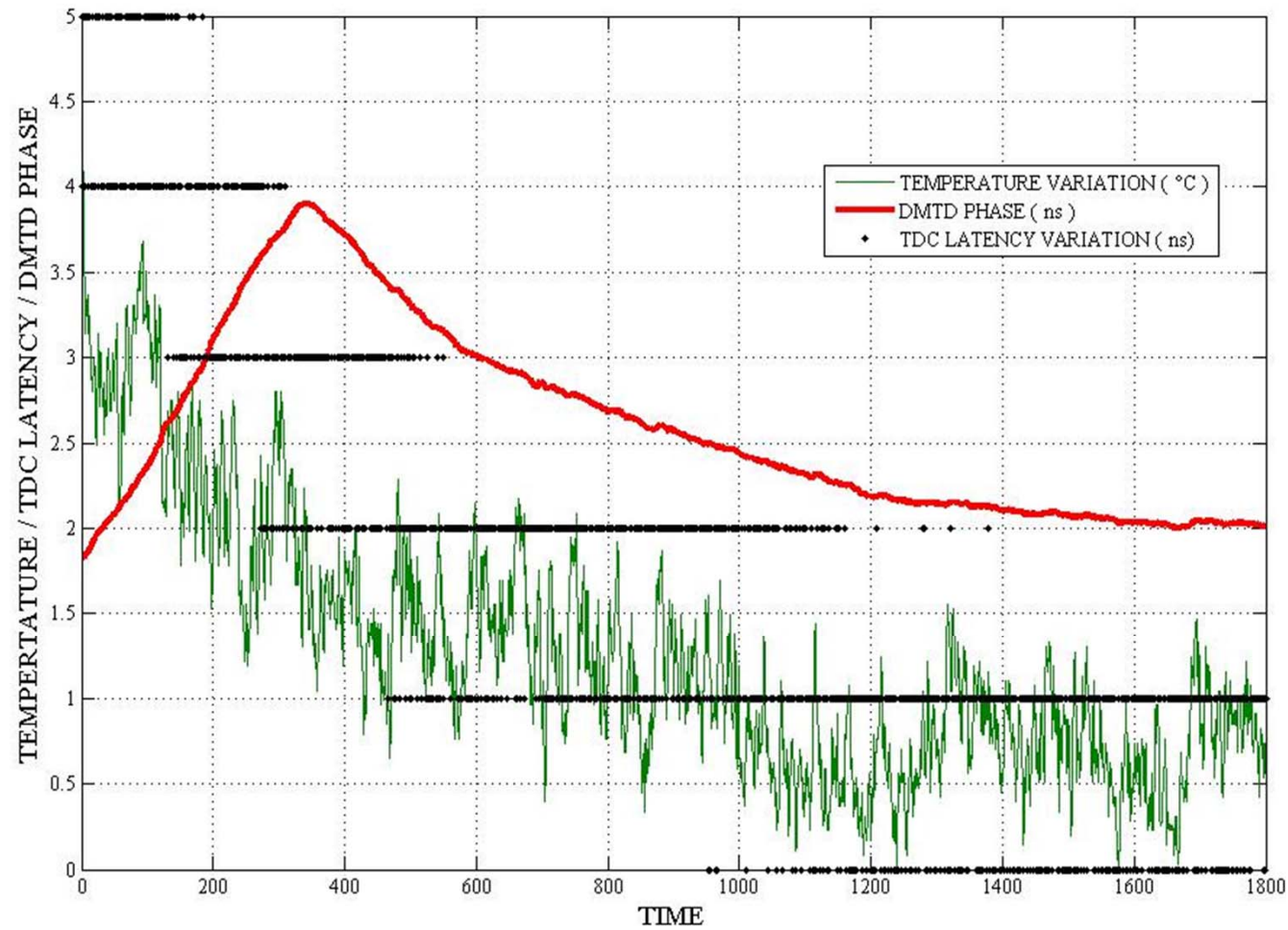
- refine measurements and statistical analysis of measurements.
- digital measurement of the long delay in the FPGA fabric.
- adapt PLL settings for the 62,5MHz clocks.
- include in the *onshore* clock distribution IP.



DDMTD



Mesures de phase, latence et température

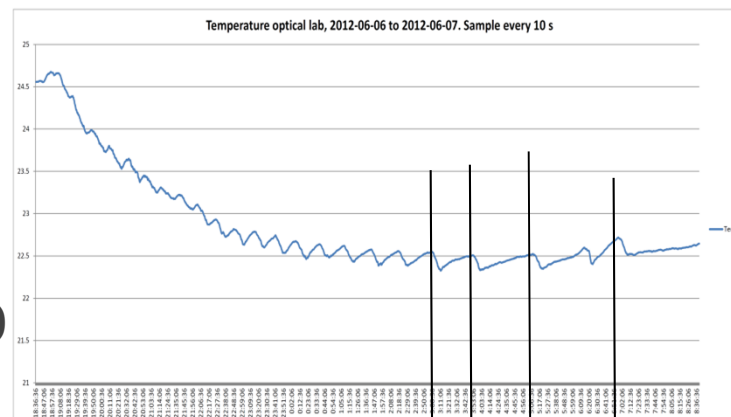


~ 10 h

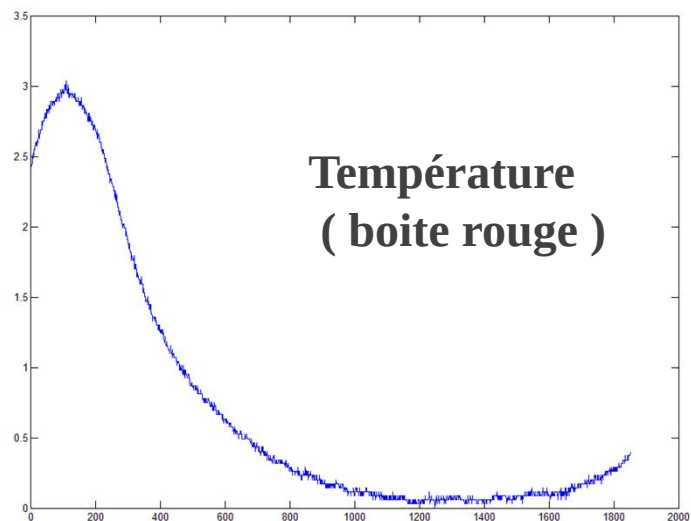
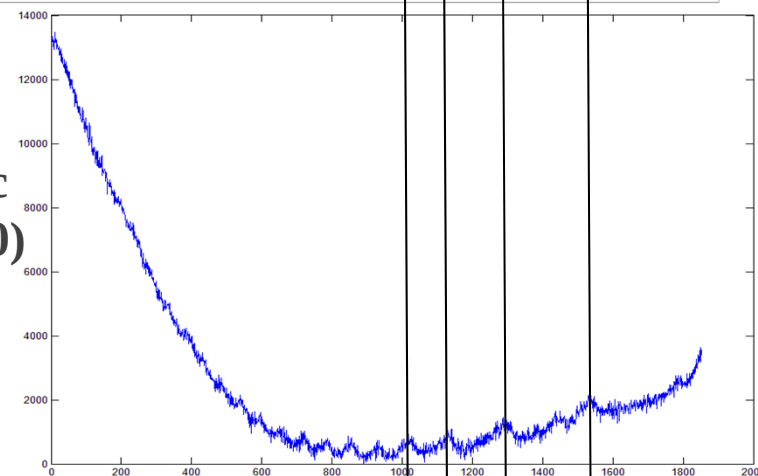
Nuit de mesures à Nikhef



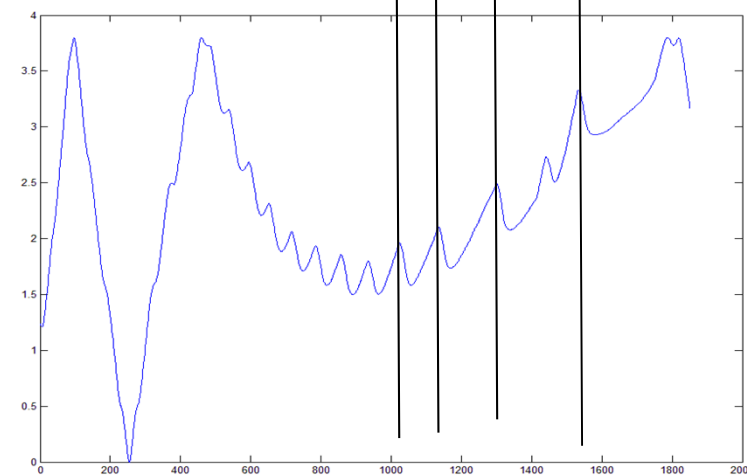
Température (labo)



Latence mesurée avec
TDC Externe (SR620)

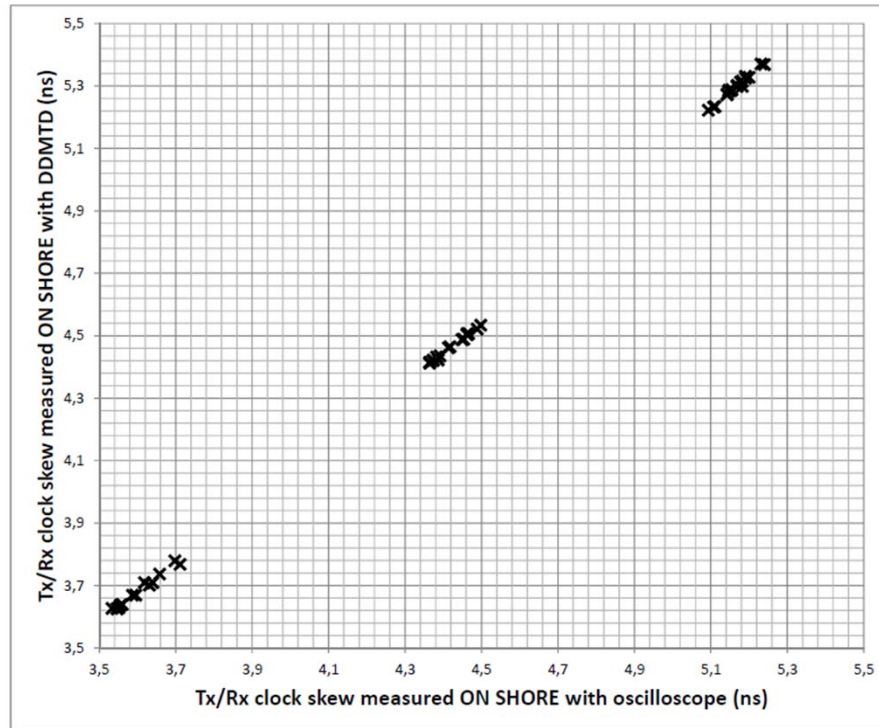


DDMTD



Mesures A/R et Calibration

- Probleme : attribuer a l'aller et au retour sa part de latence variable due à l'alignement.



Mesures sur un aller et retour

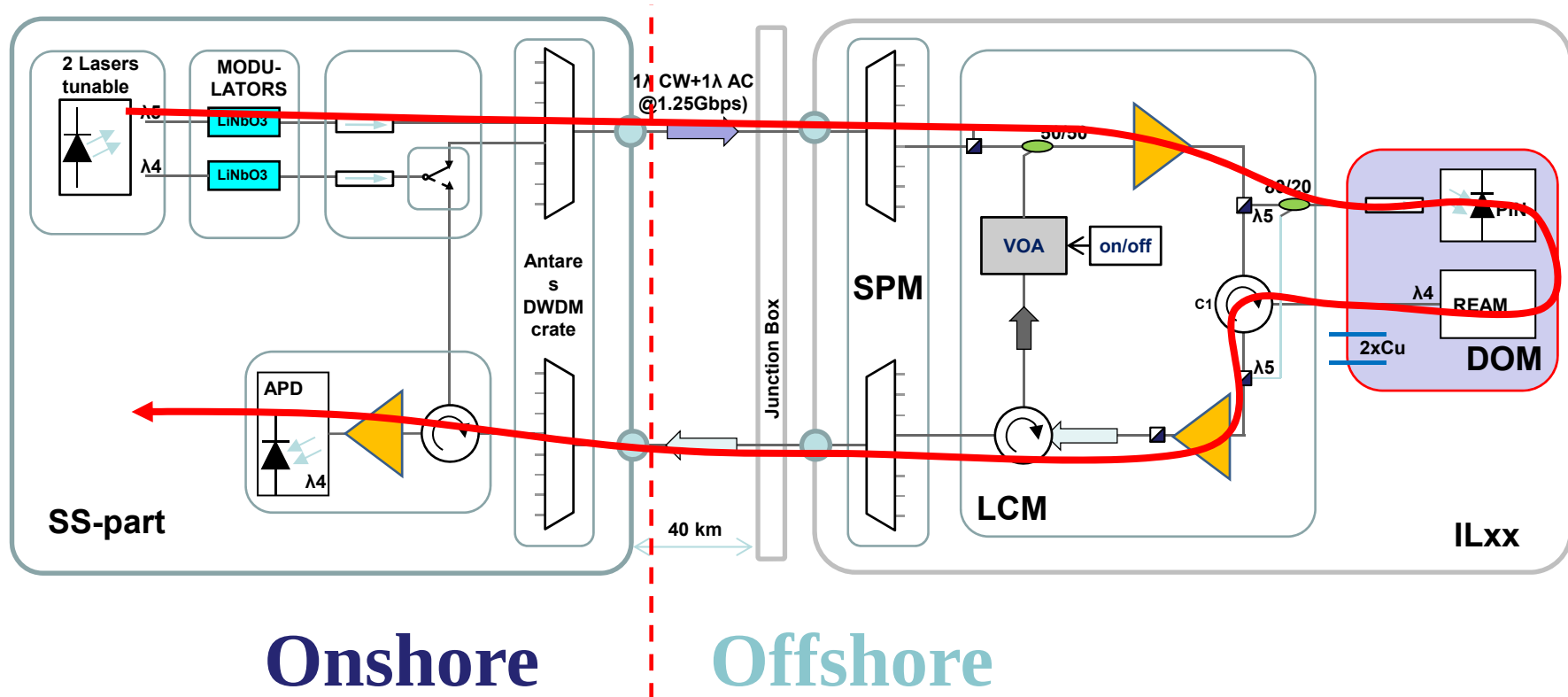
```
MM = zeros(39,1);
DD = zeros(39,1);
AA = zeros(20,19);

for ii=1:20
    A_on = 1 * ( BITSLIDES_on== (ii -1) ) ;
    MM(ii) = A_on * DMTD';
    DD(ii) = A_on * A_on';
    for jj = 1:19
        A_off = 1 * (BITSLIDES_off == (jj) );
        AA(ii,jj) = A_on * A_off';
        MM(jj + 20) = A_off * DMTD';
        DD(jj + 20) = A_off * A_off';
    end
end

BB = diag(DD);
CC = [[zeros(20,20), AA] ; [AA', zeros(19,19)]];
vv = inv(BB+CC) * MM;
```

**Inversion d'un systeme lineaire au
sens des moindres carrés**

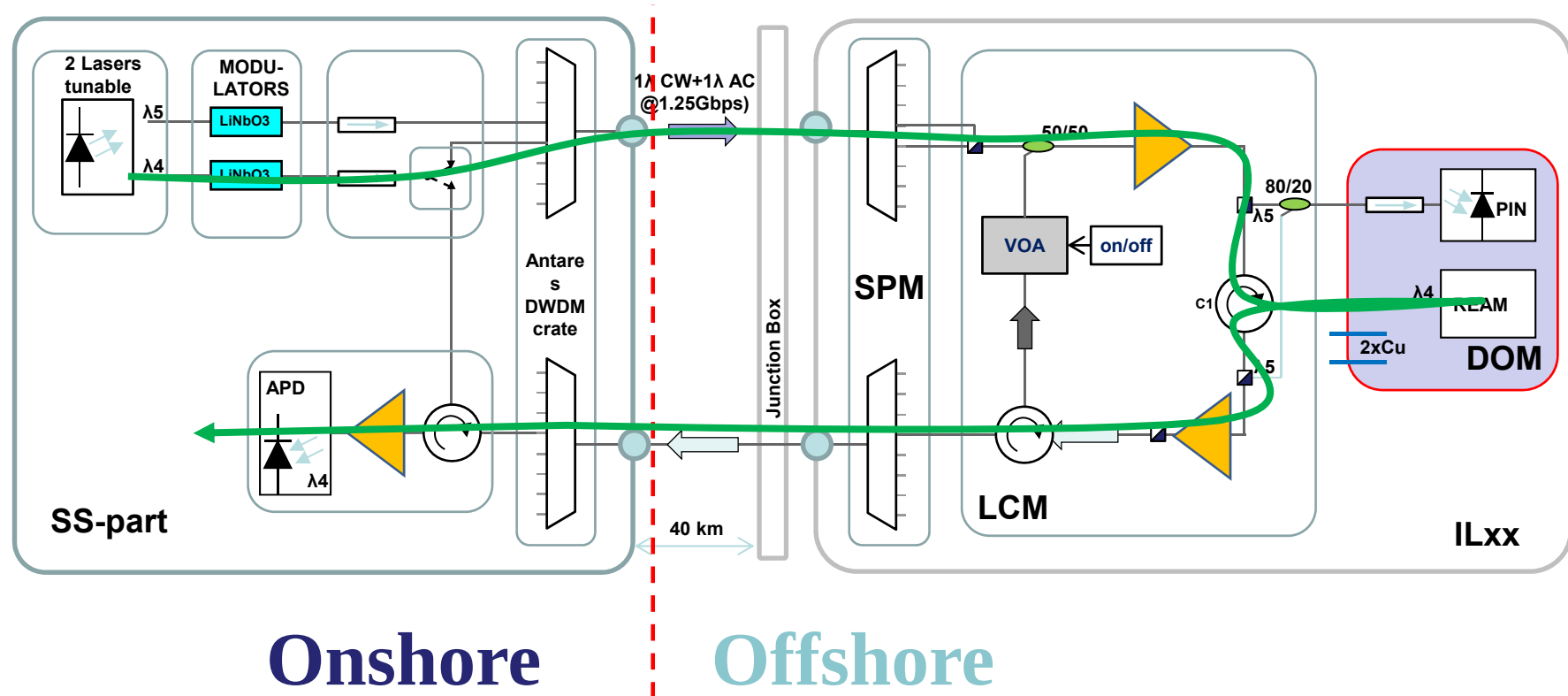
Mesures A/R et Calibration



- Calibration par commande synchrone Ethernet
- Calibration optique Aller + Retour
- Calibration optique Aller + Aller

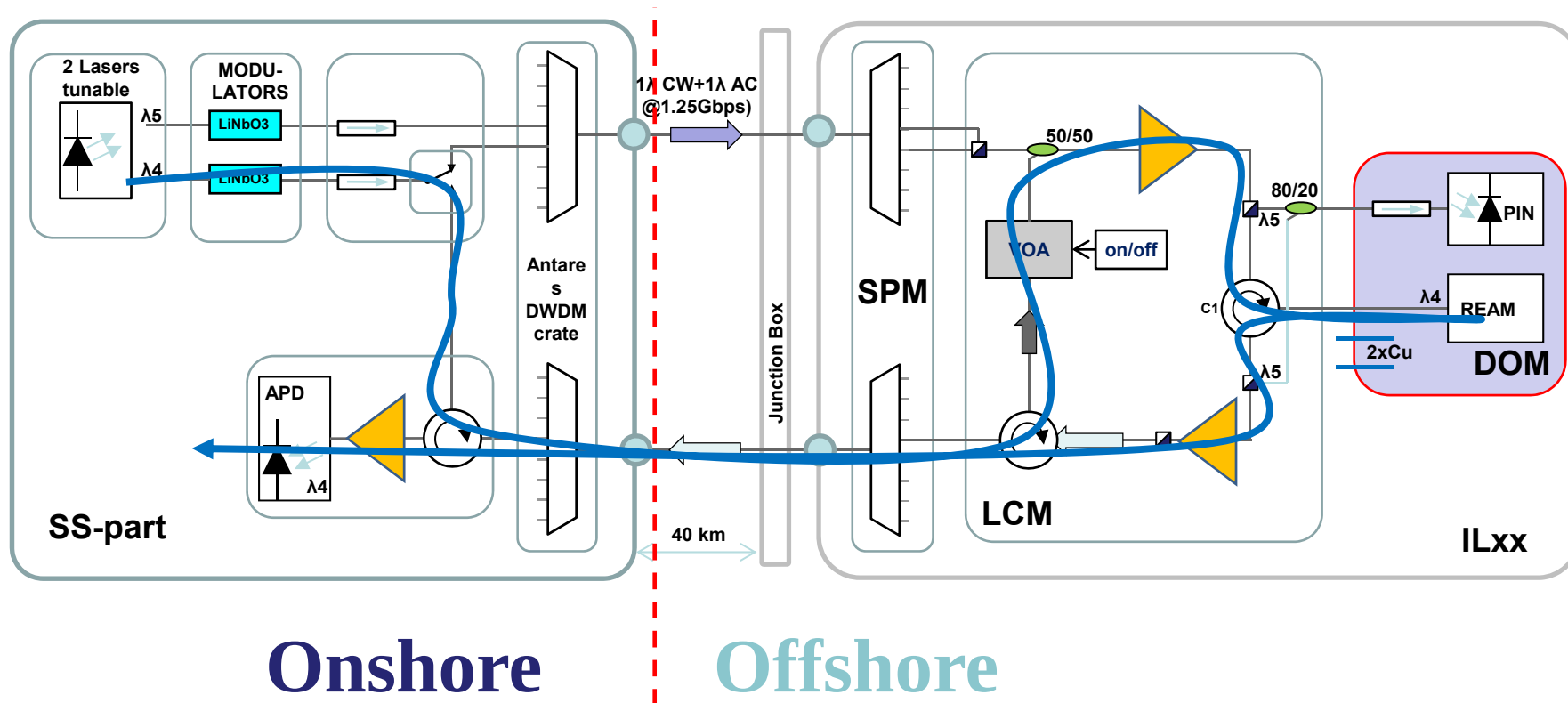
(conception NIKHEF)

Mesures A/R et Calibration



(conception NIKHEF)

Mesures A/R et Calibration



(conception NIKHEF)

Hopes for the future

The DOM integrates all common Readout functions in a single component (RSOC):

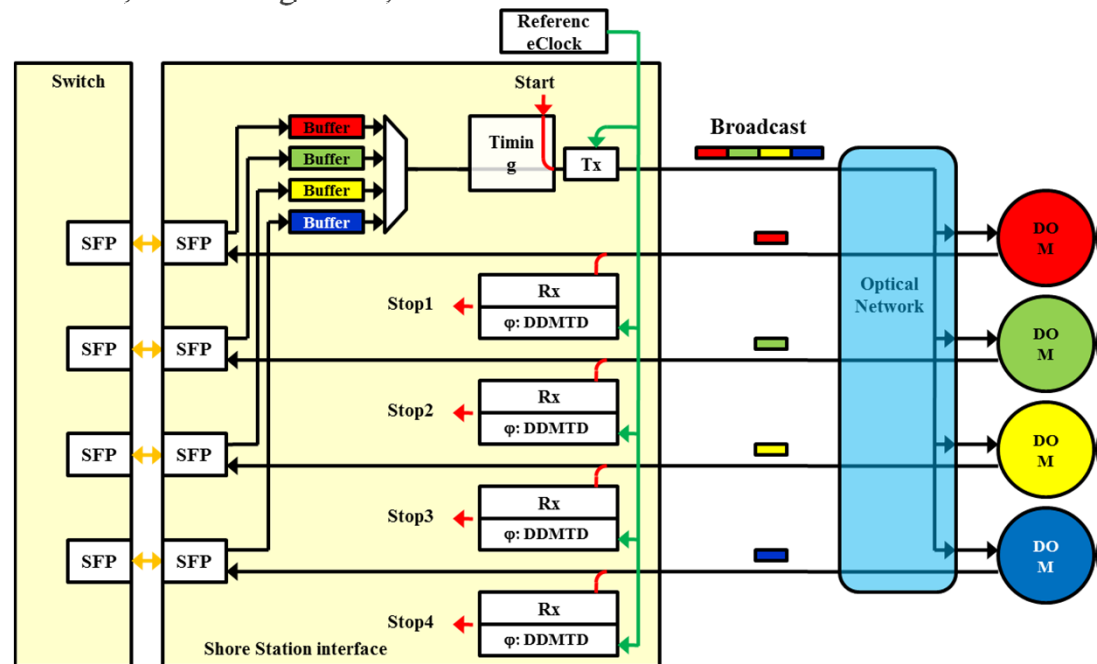
- Event Time stamping @ 1 GHz
- Clock and command distribution
- Slow-control and data acquisition performed in a RTOS multi-tasking embedded system
- Embedded Instrumentation for monitoring and calibration
- Optical Gigabit Ethernet link
- 31x 3" photomultipliers

Currently 1 DOM completed, deployment on Antares expected soon.

Final calibration on site using Laser and nano beacon, K disintegration, ...

Hopefully three more DOMs on the way ...

Hopefully a new shore station for Clock and Synchronous command distribution to several targets and additional great challenges ...



(image from P. Jansweijer, NIKHEF)

Références

A readout system-on-chip for a cubic kilometer submarine neutrino telescope

H. Le Provost *et al.*, TWEPP 2011

Time calibration of the ANTARES neutrino telescope

J.A. Aguilar *et al.*, Astroparticle Physics, 34, 2011

Measuring propagation delay over a 1.25 Gbps bidirectionnal data link

P. Jansweijer and H. Peek, Atlas note, 2010

PET system synchronization and timing resolution using high speed data links

R. Aliaga *et al.*, RT2010, Lisbonne.

Fixed latency, multigigabit serial links with Xilinx FPGAs

R. Giordano *et al.*, IEEE Nuclear Science, 2011

High speed serial IO made simple – A designer's guide with FPGA applications

A. Athavale *et al.*, Xilinx

Precise time and frequency transfer in a White Rabbit network

Tomasz Wlostowski, Master Thesis

White Rabbit homepage : <http://ohwr.org/projects/white-rabbit/documents>

Highly accurate timestamping for ethernet based clock synchronization

P. Loschmidt *et al.*, Journal of Computer networks and Communication, 2012

FPGA to IBM Power Processor Interface Setup

I. OUDA *et al.*, Application Note, IBM Research Report