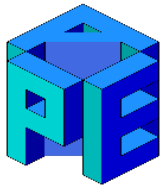


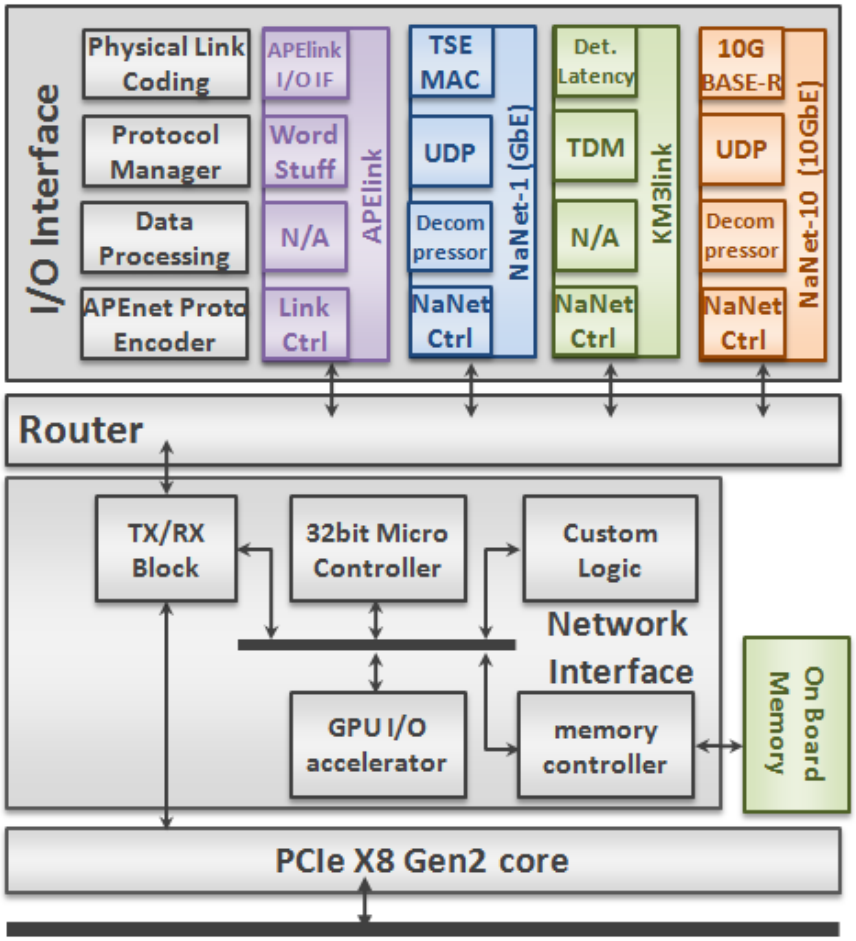
Sviluppo elettronica shore station
Stato della scheda NaNet^3

Piero Vicini
INFN – Sezione di Roma

NaNet^3

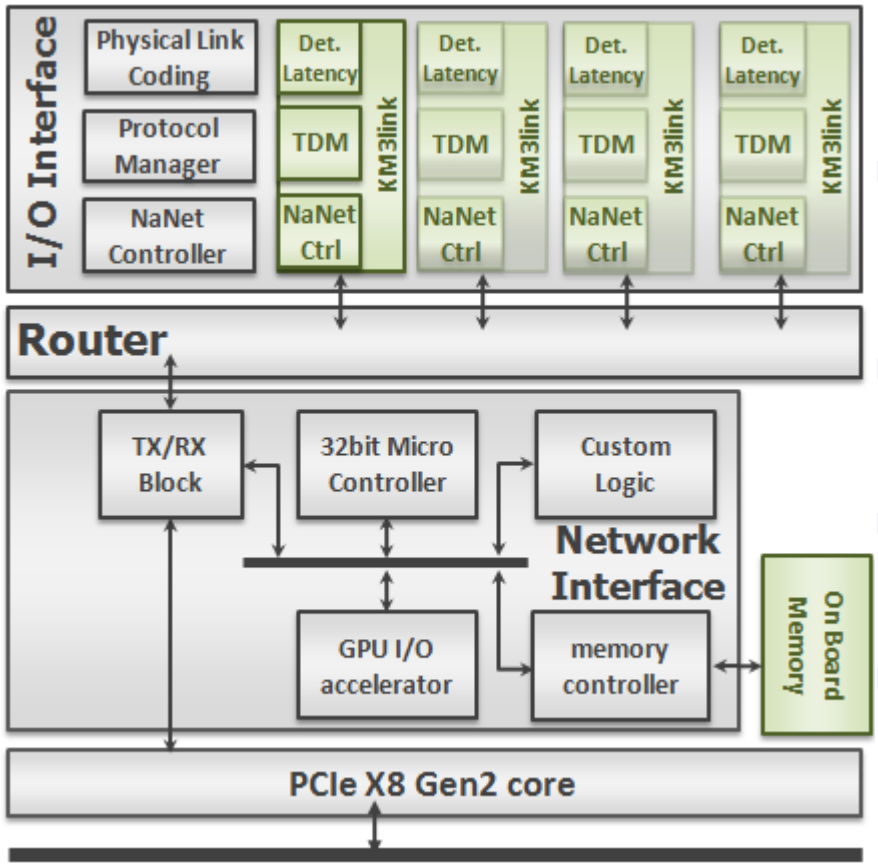
- Design and implementation of a family of network interface cards dedicated to real-time systems (for HEP).
  - NaNet^3 for KM3Net-IT implemented on TERASIC DE5-NET board
  - NaNet-1 (NaNet-10) twinned projects, targeted to NA62
- Low and stable communication latency.
- High bandwidth.
- Multiple link technologies and network protocols.
- Processing on data streams.
- Flexible.
- Extensible.
- Upgradable.
- Optimized communication with GPU accelerators

NaNet Modular Design



- I/O Interface
  - Multiple link.
  - Multiple network protocols.
    - Off-the-Shelf: 1-GbE, 10-GbE (work in progress).
    - Custom: APElink (34gbps/QSFP), KM3link (work in progress).
- Router
  - Dynamically interconnects I/O and NI ports.
- Network Interface
  - Manages packets TX/RX from and to CPU/GPU memory.
  - Nios II Microcontroller
    - Virtual memory management
- PCIe X8 Gen2 Core
  - CPU BW:
    - 2.8 GB/s Read ÷ 2.5 GB/s Write
  - GPU BW:
    - 2.5 GB/s Read & Write.

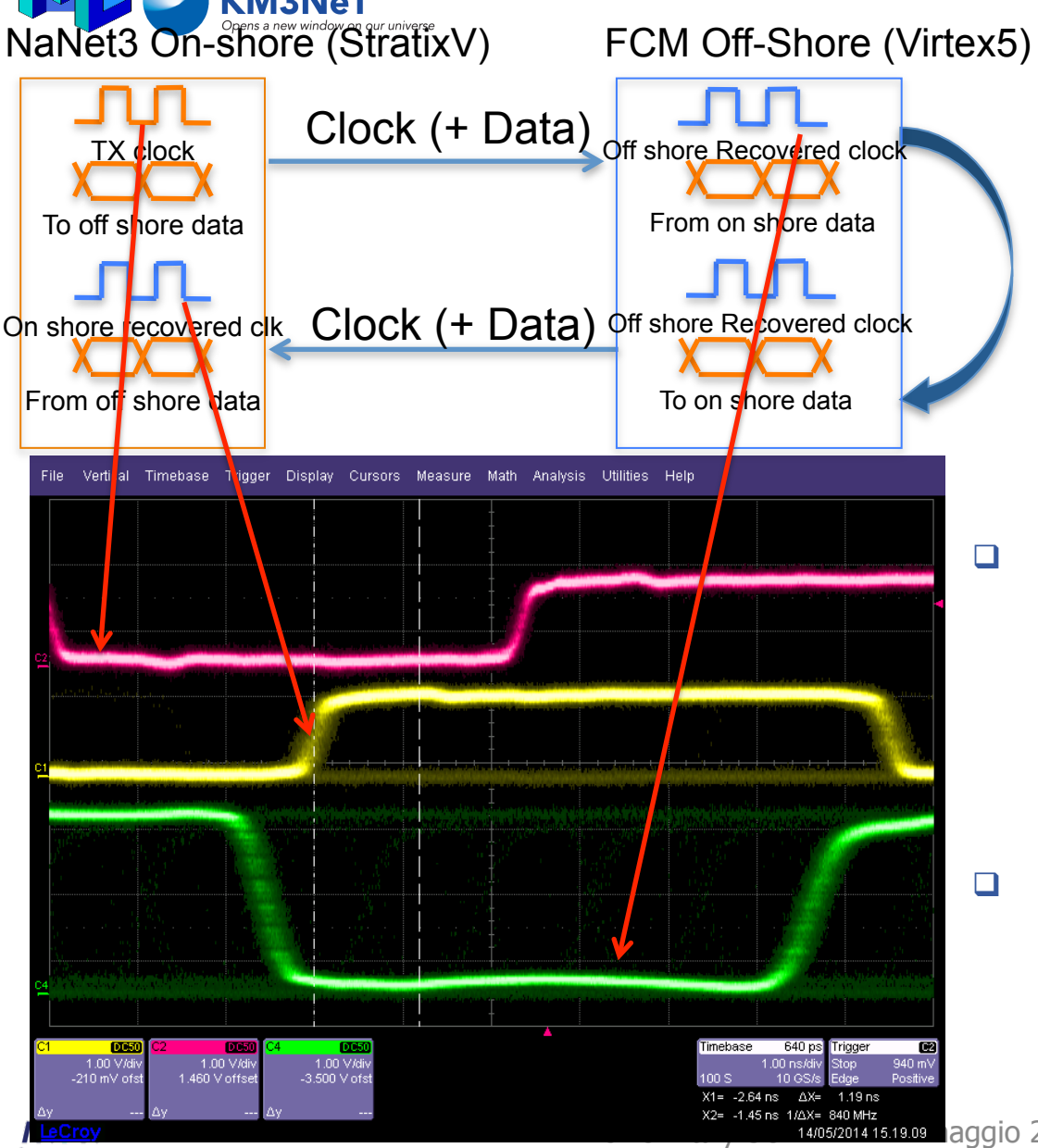
NaNet³ overview

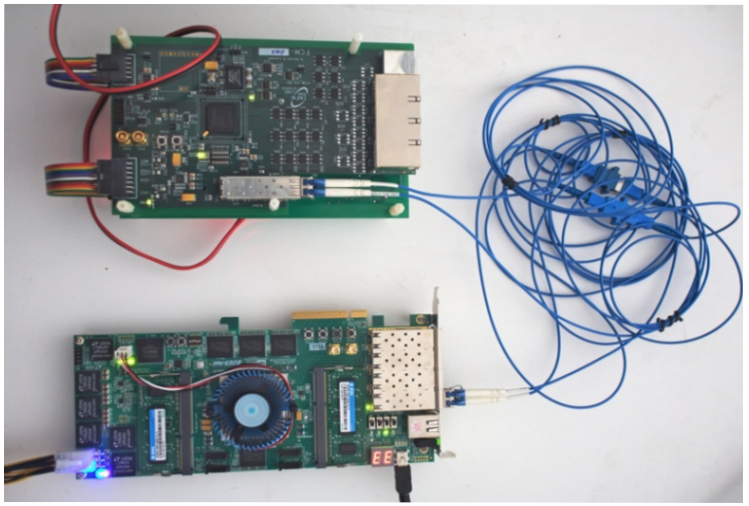


KM3link (1 link - 48 hours test passed)

- RX path: payload of different off-shore devices, multiplexed on continuous data stream at fixed time slot
  - (PCIe DMA transaction to CPU/GPU mem.)
- TX path: limited data rate per FCM (slow control)
  - (PCIe TARGET transaction from CPU/GPU mem.)
- Physical Layer: Altera Deterministic Latency Transceivers (8B10B encoding scheme)
- Data/Transport Layer: Time Division Multiplexing (TDM) data transmission protocol.
- **NaNet Ctrl:**
  - protocol translation; encapsulate TDM data stream in APEnet protocol
  - Virtual Memory management (CPU/μC offloading)
- **Virtual-to-Physical Translation**
  - Nios Implementation
  - HW acceleration: Translation Lookaside Buffer (TLB) based on associative memory

Deterministic latency link inter-operability





- Testbed: FCM vs Terasic DE5-Net
  - Custom hw mode for FCM Transceivers (Xilinx)
  - Latency deterministic mode for Stratix V Transceiver
  - 2mt copper and 2 mt long fiber
- Test:
  - 12 hours of periodic (~s) Tx clock reset to verify pll locking and rx word alignment

Time Division Multiplexing (TDM) support

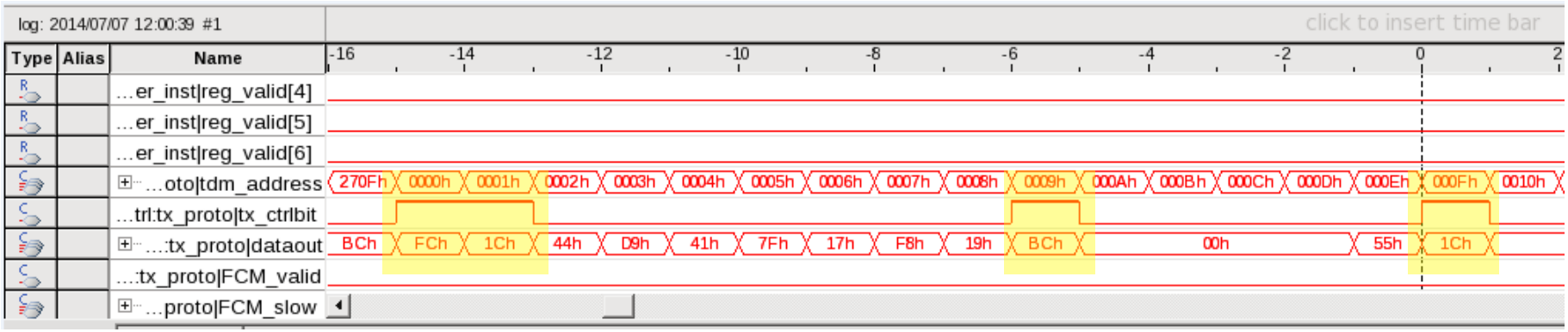
NaNet I/O interface Customization:

- TDM implementation: compliant with KM3Net-IT specification, echo test OK!

	K Code	Kout	D
IDLE_CODE	K28.5	1	0xBC
FRAME_CODE	K28.7	1	0xFC
DATA_CODE	K28.0	1	0x1C

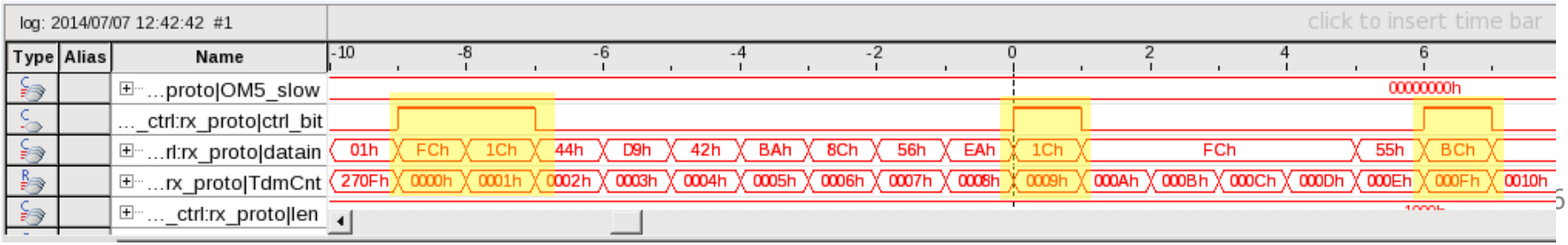
Field	Validity Code	Start Address	Stop Address	Length [byte]
Start Frame	FRAME_CODE	0x0000	0x0000	1
Frame Time	DATA_CODE	0x0001	0x0008	1+6+1
FCM Slow Control	DATA_CODE	0x0009	0x000E	1+4+1
OM0 Slow Control	DATA_CODE	0x000F	0x0014	1+4+1
OM1 Slow Control	DATA_CODE	0x0015	0x001A	1+4+1
OM2 Slow Control	DATA_CODE	0x001B	0x0020	1+4+1
OM3 Slow Control	DATA_CODE	0x0021	0x0026	1+4+1
OM4 Slow Control	DATA_CODE	0x0027	0x002C	1+4+1
OM5 Slow Control	DATA_CODE	0x002D	0x0032	1+4+1
Hydro data	No IDLE_CODE	0x0200	0x03FF	512
OM0 Phy Data	No IDLE_CODE	0x0400	0x07FF	1024
OM1 Phy Data	No IDLE_CODE	0x0800	0x0BFF	1024
OM2 Phy Data	No IDLE_CODE	0x0C00	0x0FFF	1024
OM3 Phy Data	No IDLE_CODE	0x1000	0x13FF	1024
OM4 Phy Data	No IDLE_CODE	0x1400	0x17FF	1024
OM5 Phy Data	No IDLE_CODE	0x1800	0x1BFF	1024

- NaNet^3 TX



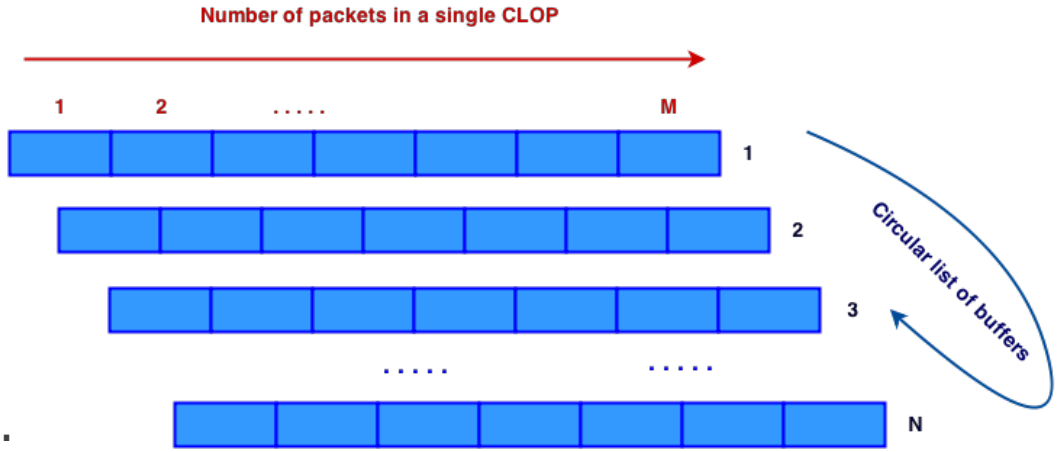
- NaNet^3 RX:

  Data flow: NaNet^3 → FCM board → NaNet^3



NaNet^3 Software

- Linux Kernel Driver
  - Status/Configuration registers.
  - TX registers interface.
  - Custom Event Queue management

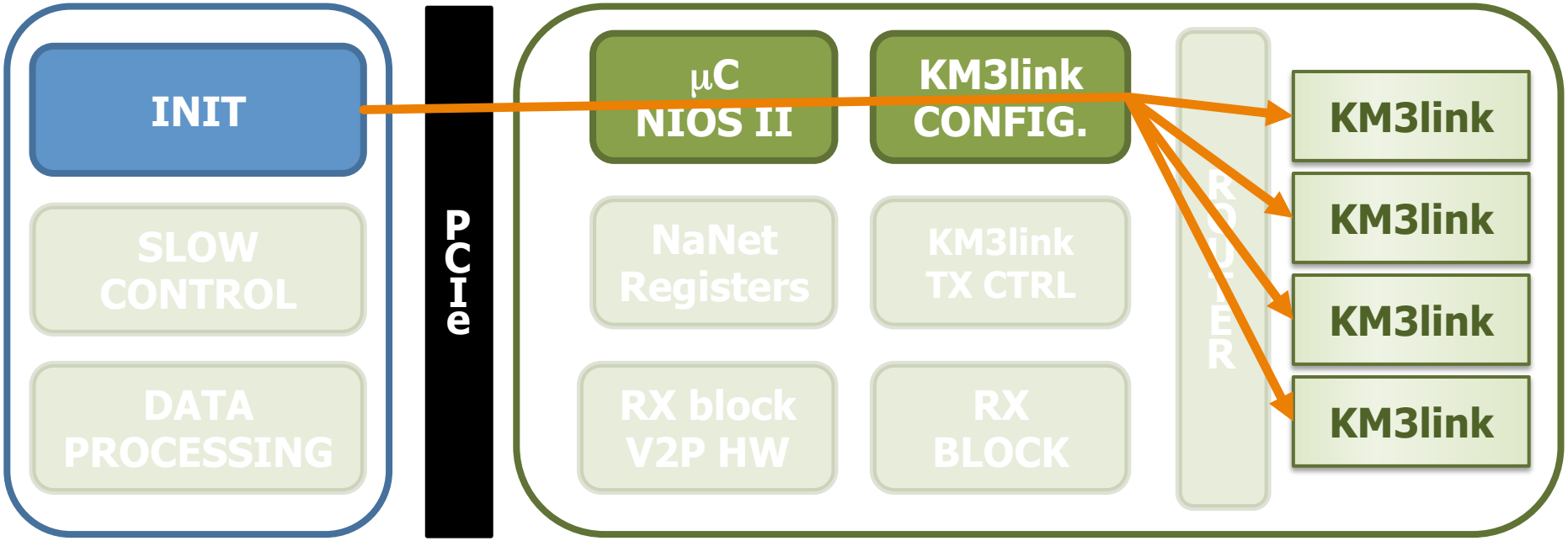


- NIOS II Firmware
  - New BSP for NaNet3 board.
  - Initialization of NaNet3 channels.
  - Management of 4 concurrent data streams

- Application Library
  - nan3_t nan3_open(int card_id); int nan3_close(nan3_t nan);
  - int nan3_register_clop(nan3_t nan, nan3_chan_id_t chan, u32 pkts, u32 items, u8 is_gpu);
  - int nan3_wait_event(nan3_t nan, nan3_event_recv_t* event);
  - u32 nan3_write_slow_data(nan3_t nan, u8 channel, u8 deviceId, u32 data);

- Application Template

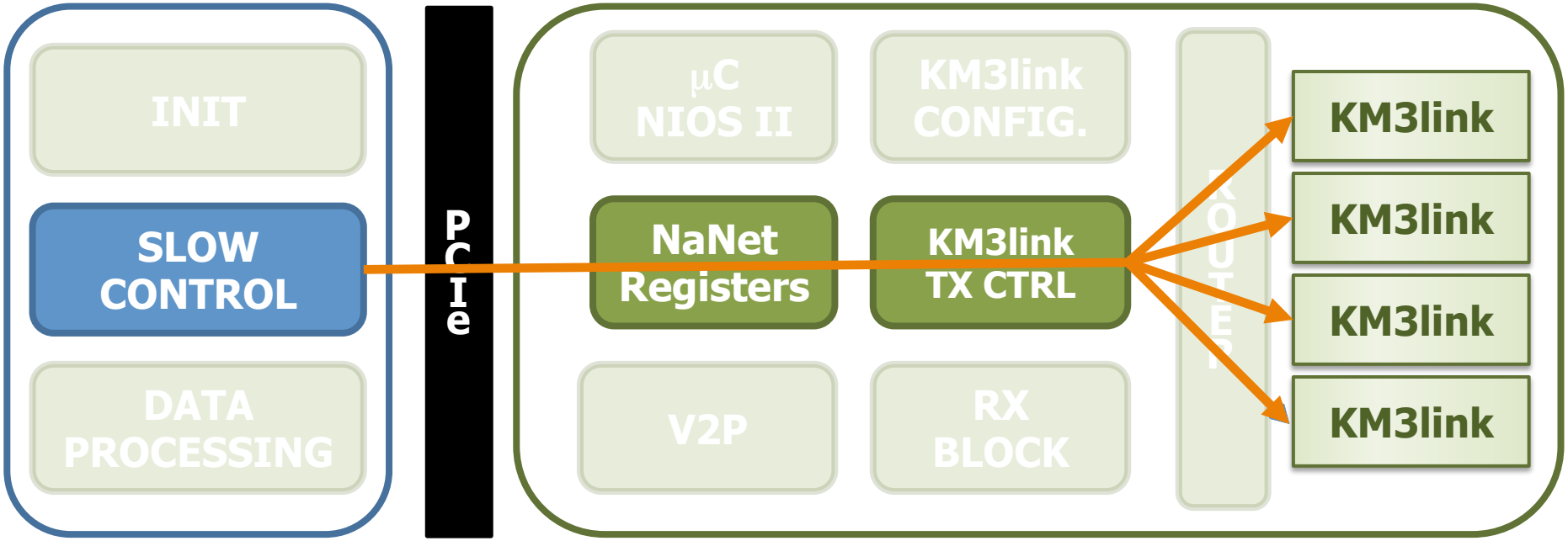
SW/HW Initialization



- CLOP buffer registration
- Pages pinning
- Virtual/Physical address
- Driver/µC communication

- Virtual/Physical address in Nios memory
- KM3link Configuration
- KM3link Alignment
- KM3link enabling

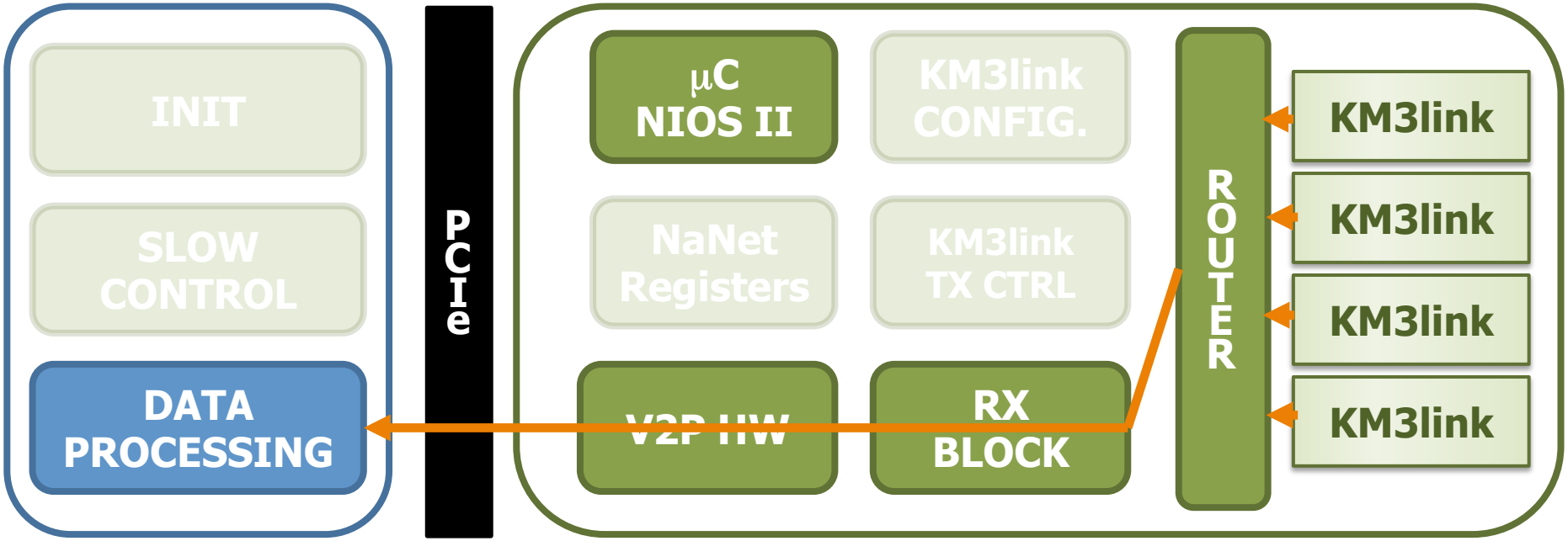
OM Slow Control Management



- TX registers interface
- Slow control buffer management

- Slow control messages distributed over the KM3links and data curruption avoidance
- Slow control encapsulated in TDM streams

Data Reception and Processing

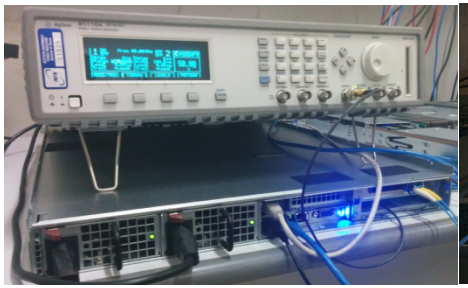


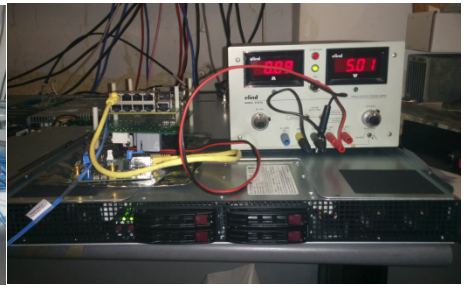
- Signalling to the application (FCM server) of reception of new frame.

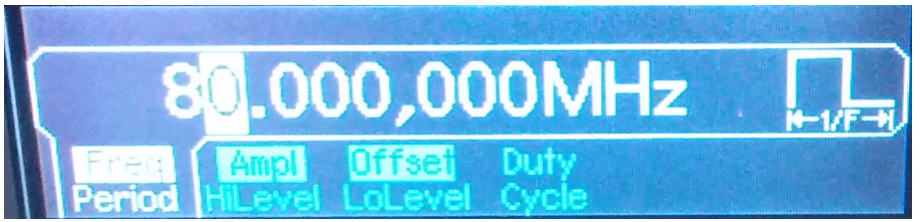
- Data encapsulated according to APEnet+ protocol
- Virtual Address generation
- Data Routing
- Virtual-to-physical translation
- PCIe DMA write process
- Interrupt generation (clop buffer)

NaNet^3 current status

- Testbed Environment
  - FCMserver
  - NaNet^3
  - FCM + FEM (off-shore systems)
- Echo test (single channel)
  - NaNet/FCM roundtrip
- Slow-control test
  - TX: from NaNet^3 to FCM
- Data Acquisition test
  - Frame ID in slow_control_0
  - Data integrity
  - Single channel 72 hours test passed

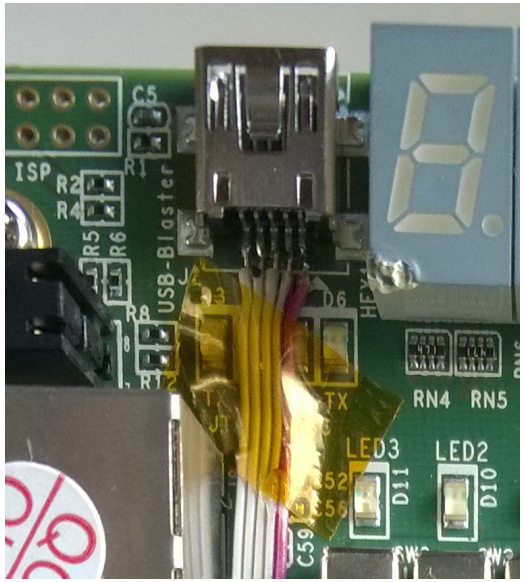


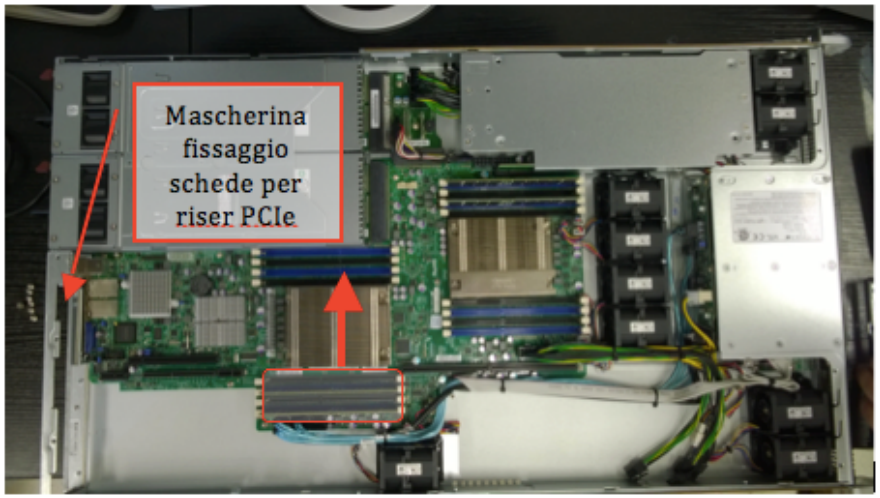




NaNet^3 current status

- Procedura di installazione HW/SW completata e documentata
  - Doc di installazione con patch da eseguire documentati con immagini…
    - FCMServerSetup.doc
      - patch per USB cable
      - procedura completa di assemblaggio FCMServer
      - …
  - Procedura automatica di configurazione OS/Driver delle FCMServer
    - Installazione via script di OS, driver e application software
    - One-click" setup: replicabile e testabile
    - Utilizzata per installare 4 server a LNS





NaNet^3 next steps

- Patch HW per segnali di calibrazione
  - Progetto completato, da implementare ASAP

- Versione 4 canali
  - Componenti HW/SW ready:
    - Re-design dell'interfaccia PCIe RDMA (nuova TLB per supporto x4 data throughput)
    - Re-design dell'architettura switch/clop-based
    - Supporto per 4 canali indipendenti
    - Driver software ottimizzato
    - … ulteriori piccoli ottimizzazioni e miglioramenti architetturali…
  - Under test
    - Effettuato test singolarmente per ogni canale
    - Pianificata la realizzazione nel nostro lab di Roma di un setup di test aggiuntivo a canale multiplo (FCMs, fibre, etc…)
    - Finalizzazione compatibile con l'installazione di > 2 torri

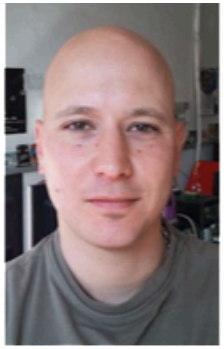
NaNet^3 at work....

- Di cosa abbiamo bisogno per il test del read-out di KM3Net-IT basato su NaNet^3 ?
  - Un piano della torre = 1 fibra.
  - Per ogni fibra utilizziamo un cage/connettore SFP+
  - NaNet design e' modulare:
    - L'hardware/software di una singola DE5-Net offre fino a 4 connessioni SFP+ indipendenti
    - Il design di NaNet supporta 1 Porta SFP+ (i.e 1 fibra, 1 piano, running da gennaio) fino a 4 Porte SFP+ (i.e. 4 piani, 4 porte, hw/sw ready, design under test)
  - Ogni torre necessita di 16(?) connessioni SFP+. Di conseguenza per il read-out di una torre:
    - 16 FCMServer con DE5-Net installata ognuna connessa ad un piano (via fibra su porta SFP+) (ORA!)
    - Appena la versione 4 canali verra' rilasciata in produzione una torre 4 FCMServer con DE5-Net installata
- Nel frattempo a LNS abbiamo:
  - 4 FCM server installati tramite la procedura automatizzata di installazione e configurazione del software necessario sulla FCMServer
  - Replica del testbed di Roma con FEM che emula traffico e struttura dei frames connessi a NaNet3
  - → Siamo pronti per connettere il sistema alla Torre

THANK YOU!!


Roberto Ammendola


Andrea Biagioni


Ottorino Frezza


Francesca Lo Cicero


Alessandro Lonardo


Michele Martinelli


Pier Stanislao Paolucci


Elena Pastorelli


Davide Rossetti


Francesco Simula


Laura Tosoratto


Piero Vicini

This work was partially supported by EU Framework Programme 7 EURETILE project, grant number 247846; Roberto Ammendola and Michele Martinelli were supported by MIUR (Italy) through INFN SUMA project.