

BACKWARD COMPATIBLE UPDATE OF THE TIMING SYSTEM OF WEST

A. Barbuti, G. Caulier, Y. Moudden*, T. Poirier, B. Santraine, B. Vincent and the WEST Team†
CEA, IRFM, Saint Paul Lez Durance, France

Abstract

Between 2013 and 2016, the tokamak Tore Supra in operation at Cadarache (CEA-France) since 1988 underwent a major upgrade following which it was renamed WEST (Tungsten [W] Environment in Steady state Tokamak). The synchronization system however was not upgraded since 1999. At the time, a robust design was achieved based on AMD's TAXI chip: clock and events are distributed from a central emitter over a star shaped network of simplex optical links to electronic crates around the tokamak. Unfortunately, spare boards were not produced in sufficient quantities and the TAXI is no longer available. Designing replacement boards provides an opportunity to investigate new clock and data recovery (CDR) solutions and extended functionalities. However, backward compatibility is a major constraint given the lack of resources for a full upgrade of the synchronization network and electronics. This contribution reports on the implementation of a custom CDR in full firmware, using the IOSerDes of Xilinx FPGAs. Preliminary results on Xilinx development boards are provided.

INTRODUCTION

Making the most of information gathered or distributed over a large network of sensors and actuators for measurement and control requires some shared sense of simultaneity. In some cases, it is provided by sufficiently frequent external events simultaneously observable over the entire array and on which all measurements can be readjusted in time. Most often, a facility's timing system is in charge of generating and propagating complementary *internal* events to all its subsystems in order to achieve the specified rigidity of the shared time frame in terms of accuracy, precision and offset of the local time on remote nodes with respect to master time. As an example, the remarkable state of the art White Rabbit network [1] provides a generic scalable solution to synchronize hundreds of nodes with better than nanosecond accuracy over up to 10 km optical fibers. Many other different solutions have been engineered to closely match the synchronization needs of specific physics experiments with very different sizes, granularity and constraints ([2–7]). And the trend is set by the upcoming high luminosity runs at CERN which demand for even greater accuracy on the order of a few tens of picoseconds [8].

A common feature of many such synchronization solutions is the use of fast serialisers and deserialisers either as individual chips or as specific functional blocks inside modern FPGAs. Deserialisation is possible only if a clock

signal is propagated along with the data telling where to sample and read individual data bits. At higher transmission rates, a very tight phase relationship between data and clock is required and is obtained thanks to proper digital encoding of the data stream acting as a modulation of the carrier clock frequency. This serves well the purpose of synchronization: the carrier clock embedded in the high speed serial stream of data words is distributed and recovered in the physical layer on each remote node hence providing means to increment local counters at a common rate and to send specific synchronous command words for instance to reset these counters, with either known or measurable latency. Last upgraded in 1999 [9], the timing system of WEST (WTS) was designed based on these ideas and AMD's TAXI receiver and emitter chips [10] used for serial communications up to 17.5 Mbyte/s. Unfortunately, the TAXI chip was discontinued shortly after and spare boards had not been produced in sufficient quantities with respect to the long life span of the Tokamak. While a functionally compatible circuit is available from Cypress [11], the fact that serial communication standards have now reached multigigabit rates reasonably questions the future availability of such low rate SerDes. The obvious need to design replacement boards in order to ensure continued operation of the WTS during the next ten years, is also an opportunity to investigate new clock and data recovery (CDR) solutions and extend current board functionalities (e.g., error detection, measurements, etc.) using up-to-date FPGAs and embedded resources.

Backward compatibility is however a major design constraint given the span of the network, the lack of resources for a thorough upgrade and the fact that WEST is in operation. Hence state of the art timing solutions based on high speed multigigabit serializers [1] such as provided in latest generation FPGAs are not applicable. The following sections describe the current state of the timing network of WEST and plans for gradual conservative upgrades. Next, the implementation of a custom replacement CDR in full firmware is discussed. Finally, preliminary results obtained with prototype firmware and software on Xilinx FPGA development boards are presented.

WEST TIMING SYSTEM

Current Status

A sketch of the timing infrastructure of WEST (WTS) is given in Fig. 1: events and clock are distributed from a central TAXI emitter board over a star shape network of simplex optical links to remote receivers in electronic crates all around the tokamak for control, safety and data acquisition. The TAXI receiver includes a CDR circuit to recover the

* yassir.moudden@cea.fr

† <http://irfm.cea.fr/WESTteam/>

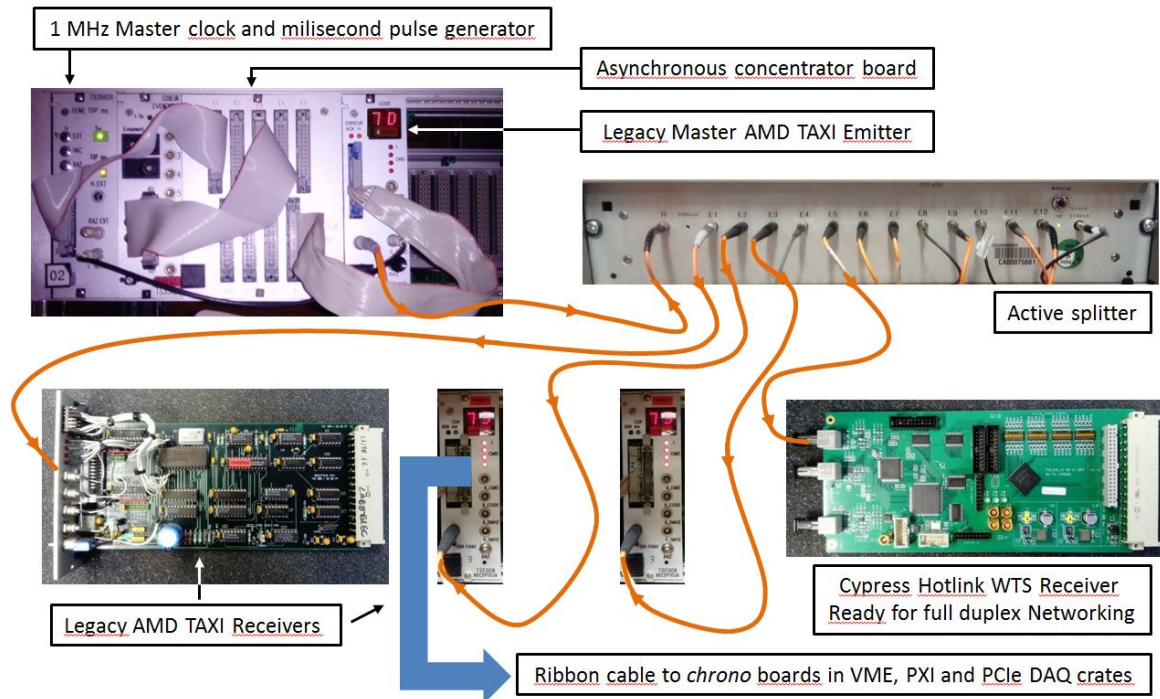


Figure 1: Sketch of the Timing System of WEST.

transmitted byte rate clock from the serial data flow. The master 10 MHz clock sets the byte rate; the resulting bit rate is 100 Mbps due to serialization and NRZi 4b/5b encoding. The 4b/5b encoding scheme provides special control and idle codes to recover word boundaries for proper deserialization inside the TAXI receiver chip. Synchronous events consist in individual bytes broadcasted over this network to sequence plasma phases and data acquisition. In addition to emitter and receiver boards, the WTS includes cascadable active Single In Multiple Out (SIMO) optical splitter boards and a Multiple In Single Out (MISO) concentrator board to funnel up to eight sources of events into the master emitter and down the network. Inside a typical cubicle, the deserialized codes and a divided 1 MHz clock are distributed electrically through ribbon and coaxial cables to dedicated *chrono* boards in either VME, PXI and PCIe crates which carry the required logic to interpret the received commands and drive proper actions related to data acquisition or control. A greater concentration of functions for enhanced signal integrity and system maintainability is a major concern in our plans for upgrade : a direct connection to the WTS optical network is now possible on the recently designed *chrono* board shown in Fig. 2. Figure 3 illustrates the dispersion of arrival times of synchronous commands on the TAXI receiver boards in a small subset of 6 cubicles located at most 20 meters apart : the maximum offset is at least 250 ns which translates to nearly 50 m in fiber length. Offset correction to enhance synchronization accuracy using the PTP algorithm [1, 12] requires monitoring the two way propagation latencies in real time on symmetric bidirectional links.

Towards a Full Duplex WTS

Fortunately, fibers were most often run in bundles of six so that the simplex links of the WEST timing network could easily be upgraded to full duplex using one of the spares allowing for an implementation of PTP or at least a periodic calibration of propagation latencies. Very high precision embedded measurements of time intervals between the moment a byte is sent to the nodes, received remotely and echoed back, and the moment the echoed bytes are received back on the emitter are possible in modern FPGAs using high resolution TDC (Time to Digital Conversion) and DDMTD (Digital Dual Mixer Time Difference) implementations [1, 7]. Transmission error monitoring is a further advantage of duplex links and of the latter *acknowledgment* mechanism with clear possibilities for increased robustness, hence all new hardware for the WTS will be full duplex by design. The legacy TAXI emitter and receiver boards have recently been combined in a new design based on the replacement part TAXI compatible Hotlink CY7C9689A delivered by Cypress which implements both transmit and receive functions on a single chip. This new board is gradually being installed on the WTS as a replacement for failing boards or at new nodes. The redesign of the MISO concentrator and SIMO splitter boards is also under way. The leading design idea is that a single MIMO board holding an adequate FPGA can be used for the two different functions to be implemented in firmware. In the present configuration, the concentrator board is connected to a master clock generator board and to the central TAXI emitter board. These could be integrated into the new design to avoid unnecessary ribbon cable lengths and to increase overall reliability thanks

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2019). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

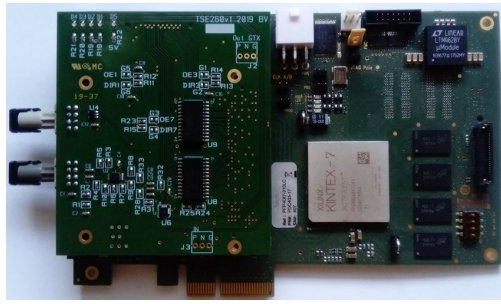


Figure 2: New *chrono* board based on Techway's PFP-KX7 board. A custom FMC was designed to connect to the WTS either through the standard ribbon cable (Fig. 1) or directly to the optical fiber.

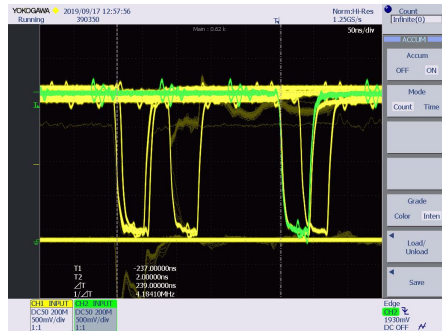


Figure 3: Dispersion of arrival times on different end points of the WTS. The green pulse is from a reference node. The yellow pulses were recorded on 5 other nodes. Each node was powercycled several times to check for latency variations.

to clock synchronous logic design targeting the embedded FPGA. Presently, the concentrator function cannot properly handle collisions of incoming code words too close in time, nor provide any monitoring information on their occurrence. The code rate on the inputs of the concentrator is rather low small, at most 2 KHz on the first of eight inputs which is connected to the periodic pulse generator of the master clock board : a deadtime-less firmware design is obvious.

The optical splitter boards of the timing network currently implement no more than an active replication of an input optical signal to 16 optical outputs. Here also including programmable logic resources in the design will allow for new functionalities. For instance, monitoring the traffic in and out on each port will provide helpful information to understand incidents on the timing network. In the meantime, two *spy* nodes are dedicated to monitoring the activity on the network : the byte transmission error rate is on the order of 10^{-7} based on the detection of *missing byte* events in a periodic sequence of specific bytes sent approximately every $512 \mu\text{s}$. The *spy* nodes also timestamp and log all bytes received which are greater than some threshold. A histogram of the timestamp differences recorded over several plasma discharges is given in Fig. 4 : the $\pm 1 \mu\text{s}$ range is readily explained by the fact that the recovered codes and 10 MHz clock are locally in phase on each node whereas the 1 MHz clock used for timestamping is derived from the

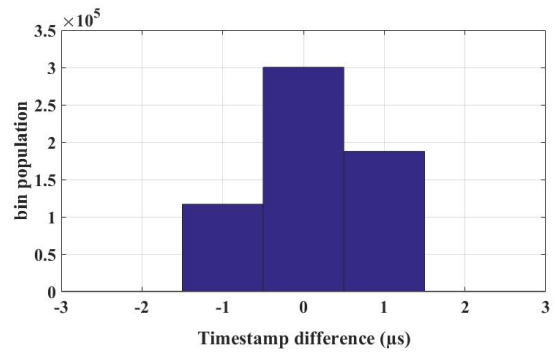


Figure 4: Histogram of timestamp differences for events recorded on 2 WTS spy nodes during several plasma runs.

10 MHz clock so that it is generally not edge aligned with the recovered byte codes, and the 1 MHz clocks on each node are generally not in phase because of propagation and clock division. Indeed, following a reset of the divider PLL, the 1 MHz will lock independently on each node on either of ten possible phase offsets in steps of 10 ns relative to the recovered 10 MHz master clock. It is worth mentioning that resetting or powercycling the TAXI receiver boards in different cubicles does not modify the measured offset of arrival times of synchronous commands (Fig. 3) : the TAXI receivers achieve fixed latency deserialization of the incoming 100 MHz serial bit stream.

The general lines of the proposed gradual and backward compatible upgrade of the WTS have been described with strong expectations in terms of timing accuracy, performance, reliability and maintainability. Nevertheless there is some risk in relying on the CY7C9689A chip, a reference already more than 15 years old. Hence the next section discusses the implementation of a custom CDR solution in firmware targeting up-to-date FPGAs.

FIRMWARE CLOCK AND DATA RECOVERY FOR WEST

The CDR solution described here, implemented and tested on a Xilinx Virtex 6 ML605 development board, is based on the IOSerDes primitive available in nearly all IO blocks of fairly recent Xilinx FPGA's [13]. They are commonly used for on board source synchronous communications with for instance fast memory components, etc. But they can also be used as high speed asynchronous samplers for high resolution edge detection or networking applications [7, 14]. Typically, the maximum double data rate achievable using this primitive on a speed grade -1 Virtex 6 is 1.1 Gb/s. This is more than ten times the bit rate of the WTS, a strong indication that handling it directly in firmware without resorting to an external CDR chip is possible.

Oversampling

IO blocks that pair to form a differential input also provide access to two cascadable deserializers. These can be configured into a 10 bit wide serial to parallel converter in

DDR mode : fed with two opposite 500 MHz clocks generated from a local 100 MHz the result is a high speed sampler operating at 1 GHz, with a ten bit parallel output at 100 MHz asynchronous relative to the WTS bit stream. The next steps of the CDR solution appear in Fig. 5. Locking is about detecting transitions in the oversampled data hence locating the data bits. But the bit limits drift : they were not generated remotely in the same clock domain as the one used locally for asynchronous sampling. Keeping track of the drifting data bits and transitions is the purpose of this second logic function. If the remote clock is slightly slower (resp. faster), the local bitstream will occasionally underflow (resp. overflow) meaning that there will occasionally be zero (resp. two) valid bits forwarded every clock cycle to the next module in charge of inverting the NRZ_i encoding. Alignment comes next : every clock cycle, either 0, 1 or 2 bits from the previous stage are shifted into a register from which this function is to recover and track the 10 bit word boundaries thanks to the redundant 4b/5b encoding and the reserved idle characters which are uniquely discernable by design. Finally, valid 10 bit words are delivered to the subsequent module in charge of inverting the 4b/5b encoding. There will most often be ten clock cycles separating two valid words, some times only nine and sometimes eleven depending on the frequency offset between local and remote clocks.

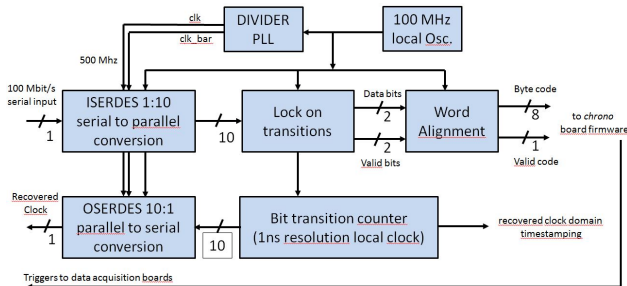


Figure 5: Block diagram of the full firmware clock and data recovery solution for the WTS.

Ten-fold oversampling is an overkill regarding data recovery. However, 1 ns resolution on bit transitions is useful for clock recovery. If needed by external equipment (e.g., acquisition boards), a clock signal (1 MHz or 100 KHz) can be generated using the OSerDes primitive configured to serialize 10 bits at 1 Gb/s by feeding it with proper transitions based on bit counts and boundary positions detected by the lock function. The 1 ns resolution of the ISerDes and OSerDes clearly adds jitter to the recovered clock. There are solutions to increase the sampling resolution up to 100 ps using ISerDes primitives, but none were found to use the OSerDesS at this rate.

Results obtained with the proposed CDR solution on a Xilinx ML605 development board connected to the WTS are shown in Figs. 6, 7, and 8.



Figure 6: Jitter on the recovered clock : The oscilloscope is set in infinite persistence to trigger on the rising edge of the recovered 1 MHz clock from a Taxi receiver, in purple. The green trace records the rising clock edges obtained using the proposed CDR solution. The total measured relative jitter is approximately 5 ns.

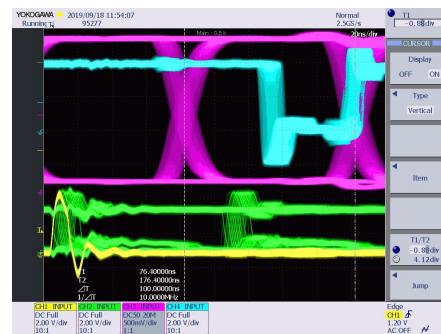


Figure 7: Testing for false positives : The oscilloscope is set to trigger on codes detected with the custom CDR solution (yellow) in infinite persistence. The blue trace records the codes detected with the TAXI receiver. Green and blue are as in Fig. 6.

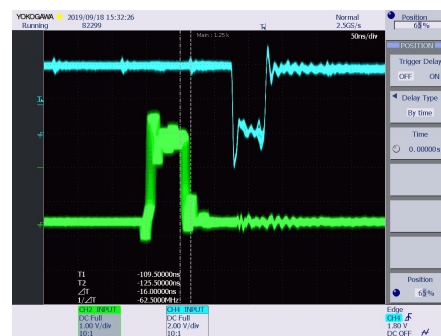


Figure 8: Testing for detection errors : The oscilloscope is set to trigger on codes detected with the TAXI receiver (blue) in infinite persistence. The green trace records the codes detected with the custom CDR receiver (due to jitter, the latter pulses had to be made longer compared to Fig. 7).

First, the CDR appears to lock durably to the WTS : no unlocking of the recovered clock with respect to the output clock from a TAXI receiver board was detected using an oscilloscope with infinite persistence for several hours. There

are assumptions made in the firmware which will impact locking : capture and lock ranges still need to be assessed. A second result, in a similar configuration, is that there were no observable missed detections nor false detections of the synchronous code bytes distributed over the WTS.

Ongoing Work

The above CDR module was easily ported from Virtex 6 to Kintex 7, and has been integrated into the firmware of a new PCIe *chrono* board design based on the PFP-KX7 PCIe board [15]. The custom FMC includes adequate optical transceivers allowing as an option feeding the 100 Mbit/s serial stream from the WTS directly to the FPGA. Integrated functional tests within the control and data acquisition system of WEST will be performed shortly. Also, the distributed synchronous commands can be timestamped within the recovered 10 MHz or a divide clock domain. The statistical analysis of the timestamps from this and other *chrono* boards also has to be performed.

Meanwhile, more FPGA primitives are being explored for other CDR solutions with possibly greater performance. The IOSerDes primitives used above are quite different from the multigigabit transceivers (e.g., GTP, GTX) available in the more expensive Xilinx FPGAs. They offer ever rising performance with bit rates up to 58 Gb/s on Xilinx's UltraScale+ GTM serial transceivers [13]. They have been extensively studied for synchronization purposes and they take a significant part in many state of the art solutions [1, 7, 16, 17] however with bit rates ten or more times faster than the 100 Mb/s rate of the WTS. In fact the datasheet of the Virtex 6 (resp. Kintex 7) specifies a minimum bit rate of 480 Mb/s (resp. 500 Mb/s). Nevertheless the GTX and all its constituents make up a complex black-box so that it is hard to tell what behaviour to expect when it is not used strictly within bounds. The GTX primitive includes a phase measurement block, a voltage controlled oscillator, a CDR circuit : it is currently being investigated as yet another alternative clock recovery option for WEST. Another feature of the WTS is that the serial bitstream consists almost exclusively of 4b/5b protocol so-called *SYNC (JK)* characters and synchronous user data codes are scarce. The resulting signal is nearly square and periodic : the use of FPGA embedded PLL primitives and external jitter cleaner chip in the design of CDR solutions for WEST is also being investigated.

CONCLUSION

Maintaining the timing network to ensure continued operation of WEST in the next years offers opportunities for a backwards compatible and gradual upgrade. Several actions have been discussed including a move towards a full duplex network to allow propagation offset correction, a higher concentration of functions for higher signal integrity and overall timing accuracy, and an increased use of FPGAs and SoCs for greater flexibility and independence. In fact, a full firmware CDR solution was developed and tests are ongoing.

The overall goal is to gradually enhance the accuracy of the WTS, its reliability and its maintainability.

ACKNOWLEDGEMENTS

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

REFERENCES

- [1] The White Rabbit Project, <https://white-rabbit.web.cern.ch/>
- [2] Z. Guzik, A. Chlopik, F. Formenti, A. Gianoli, and O. Vossnack, "Control and synchronization of the krypton calorimeter pipeline digitizer in na48 experiment at cern," *Nucl. Instrum. Methods*, vol. 427, pp. 574–582, 1999.
- [3] J. Aguilar *et al.*, "Time calibration of the antares neutrino telescope," *Astropart. Phys.*, vol. 34, pp. 539–549, 2011.
- [4] D. Abbott *et al.*, "A 250 mhz level 1 trigger and distribution system for the glux experiment," in *IEEE Real Time Conference*, 2009.
- [5] R. Aliaga *et al.*, "Pet system synchronization and timing resolution using high-speed data links," *IEEE Transactions on Nuclear Science*, vol. 58, pp. 1596–1605, 2011.
- [6] D. Calvet *et al.*, "The back-end electronics of the time projection chambers in the t2k experiment," *IEEE Trans. Nucl. Sci.*, vol. 58, pp. 1465–1471, 2011.
- [7] S. Anvar *et al.*, "Design and implementation of a nanosecond time-stamping readout system-on-chip for photo-detectors," *Nucl. Instrum. Methods*, vol. 735, pp. 587–595, 2014.
- [8] R. Rusack, "A precision pure clock distribution," presented at TWEPP'19, Santiago de Compostela, Spain, Sep. 2019, unpublished.
- [9] D. Moulin *et al.*, "Upgrade of the timing system for tore supra long pulses," *IEEE Trans. Nucl. Sci.*, vol. 47, pp. 119–122, 2000.
- [10] TAXIchip™ Integrated Circuits, http://hep.uchicago.edu/~thliu/projects/Pulsar/other_doc/TAXIchip.pdf
- [11] CY7B923, CY7B933: HOTLink® Transmitter/Receiver, <https://www.cypress.com/documentation/datasheets/cy7b923-cy7b933-hotlink-transmitterreceiver>
- [12] H. Weibel, "High precision clock synchronization according to iee 1588: Implementation and performance issues," in *Embedded World*, 2005.
- [13] XILINX, <http://www.xilinx.com>
- [14] M. Defossez, *Lvds 4x asynchronous oversampling using 7 series fpgas and zynq-7000 ap socs*, Xilinx Application Note 523, 2017.
- [15] Techway, <https://www.techway.fr/>
- [16] P. Jansweijer and H. Peek, "Measuring propagation delay over a coded serial communication channel using fpgas," *Nucl. Instrum. Methods*, vol. 626-627, pp. 169–172, 2011.
- [17] A. Aloisio, F. Cevenini, R. Giordano, and V. Izzo, "High-speed, fixed-latency serial links with fpgas for synchronous transfers," *IEEE Trans. Nucl. Sci.*, vol. 56, pp. 2864–2873, 2009.