White Rabbit: Sub-Nanaosecond Timing over Ethernet

P.P.M. Jansweijer[a], H.Z. Peek[a], E. de Wolf[b,a,*]

[a] *Nikhef, Science Park 105,1098 XG Amsterdam, The Netherlands*
[b] *University of Amsterdam, Science Park 904,1098 XH, Amsterdam, The Netherlands*

Abstract

The White Rabbit (WR) project is a multi-laboratory, multi-company effort to bring the best of the data transfer and the timing world together in a completely open design. WR is a fully deterministic Ethernet-based network for general purpose data transfer and synchronization. The aim is to enable the synchronization of a large number of nodes with sub-nanosecond accuracy and picosecond jitter over long lengths of fibre. The key technologies used are physical layer syntonization (clock recovery) and the Precision Time Protocol (IEEE 1588). WR generates sub-nanosecond synchronous precision timing in all nodes by continuous tracking and compensating the transmission delays. We give an overview of the WR project and describe the design goals and specifications of the project. The WR switch and the (user) node which are the central components of the WR system and real timing measurements of prototypes of WR hardware are presented.

Keywords: White Rabbit Synchronous Ethernet PTP Network Timing, IEEE1577
PACS: 95.55.-n, 95.85.Ry, 85.60.Ha

1. Introduction

White Rabbit (WR) is a project conveyed by particle physics laboratories and academic research labs in collaboration with the industry with the aim to develop a distributed timing and data network in which a large number of nodes are synchronized with an accuracy better than 1 ns relative to a master timing station. The data network should have a deterministic behaviour with low transmission latency. Currently, Ethernet is the most successful networking standard which allows for low costs for cabling and infrastructure, high bandwidth, efficient switching technology and interoperability. Hence, the challenge of White Rabbit is to accomplish high precision timing over Ethernet. In order to reach high precision, WR minimizes timing jitter by using "data carrier frequency transfer" at the physical layer, referred to as Synchronous Ethernet (SyncE) [1]. This ensures that master and slave node clocks are synchronous. Hardware assisted timestamping is used to synchronize the nodes with subnanosecond accuracy. Clock offset and link delay compensation is accomplished using the Precision Time Protocol (PTP) IEEE 1588-2008 [2] at the networking layer. The subnanosecond accuracy is only achieved when both sides of the data link implement the WR extension of PTP. Existing IEEE 1588 timing networks can be gradually and/or partially upgraded to WR PTP. A tree structure topology is used to distribute timing to a large number of nodes. The timing architecture is independent of the Ethernet networking. Each WR switch receives time information from a single master and distributes it transparently to all slaves. The WR project is an "open source" project. Both software and hardware sources are distributed via an Open Hardware Repository web portal [3].

2. White Rabbit Timing Distribution

WR timing distribution is based on three elements: Precision Time Protocol (PTP), Synchronous Ethernet (SyncE) in combination with clock loopback and phase measurement.

2.1. Precision Time Protocol

The Precision Time Protocol (PTP) constitutes a synchronization protocol for packet-switching networks. The protocol uses messages that carry precise timing information. Hardware timestamps at the physical layer are used to compute the point-to-point link delay. By using the bitslip trick that is described in [4], the time resolution is as small as a single symbol, which is defined as one bit-time (unit interval). Fig. 1 illustrates the clock offset measurement which relies on the propagation of PTP Sync, Follow_Up, Delay_Req, and Delay_Resp.

2.2. Synchronous Ethernet (SyncE)

Syntonization is the adjustment of two electronic circuits or devices in terms of frequency. Although syntonization exists in traditional Ethernet between the sending- and receiving-side on one link, the recovered clock is only

*Corrsponding author
Email address: e.dewolf@nikhef.nl (E. de Wolf)

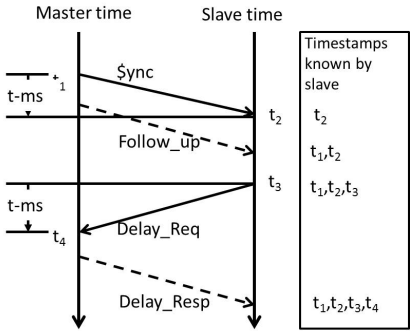

Figure 1: PTP clock synchronization process

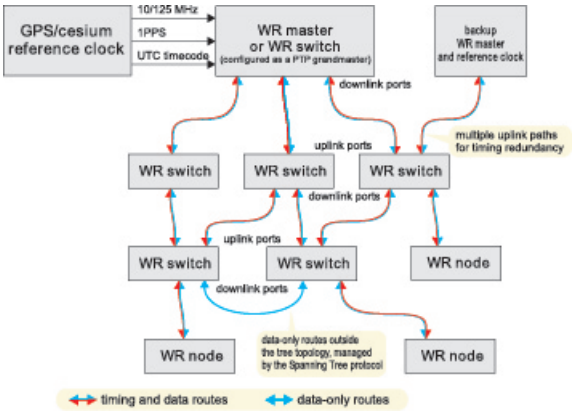

Figure 2: Schematic overview of a White Rabbit network.

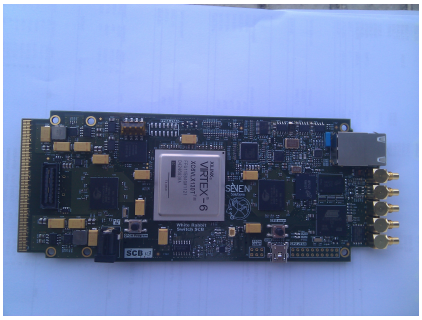

Figure 3: White Rabbit switch (v3); Switching Core Board

used to sample the incoming data. The slave receiving side does not propagate the clock further, so it is not possible to realize synchronous network structures. SyncE uses the data carrier frequency at the physical layer interface to pass frequency from master- to slave-node. Slave clocks are not free-running, but rather are locked and traceable to a primary reference clock as defined in ITU G.811 [5] and can be an external source such as GPS. SyncE provides a frequency transmission hierarchy formed on a link-by-link basis. Since clock recovery works on the physical layer, independent of data transmission the syntonization performance is immune to variations of traffic load and packet delay.

2.3. Clock loopback and phase measurement

When using SyncE, the slave node receiver recovered clock is locked to the master. A Phase Locked Loop (PLL) cleans the recovered clock, i.e. it removes jitter generated from the clock recovery circuitry. This PLL is based on a digital implementation of Dual Mixer Time Difference (DMTD) [6] technology. PLL optimization results in a jitter performance of 2.5 ps [7]. The digital implementation also enables easy control of the recovered clock phase offset. The cleaned clock is sent back to the originating master, creating a loop from master to slave and back. A DMTD phase measurement in the master measures the phase difference between the master reference clock sent to the slave and the recovered clock received back from the slave. The measured phase difference is sent to the slave node, using the correction field in the PTP message. The slave node corrects its phase offset such that the delay variations in the physical link are compensated. This results in a WR network that has the required sub-nanosecond accuracy.

3. White Rabbit Network

The timing architecture of the WR network is independent of the Ethernet networking. Network synchronization in WR is based on clock hierarchy with the high accuracy reference clock at the top. Slave clocks are synchronous to the recovered clock of the data link connected to the level above. The recovered clock is cleaned before it is further used in the clock distribution to the hierarchy levels below, creating a timing distribution tree. Additional links can be used as a backup to create a redundant timing distribution tree. Fig. 2 shows an overview of a WR network.

4. White Rabbit building blocks

The WR network consists of WR switches and WR nodes. WR switches provide all capabilities of normal Ethernet switches and in addition they provide high precision timing distribution as described above. Fig. 3 shows the current implementation of a WR switch [8], the Switching Core Board [9] in a Micro Telecommunications Computing Architecture (μTCA) [10] form-factor. The switch implements 2 uplink ports and 16 downlink ports. Fig. 4 shows the Simple PCIe FMC carrier (SPEC) [11] that is the current hardware implementation of a WR node. It contains a Xilinx Spartan-6 FPGA. The Firmware loaded into this FPGA holds the user application and the WR PTP Core (WRPC) [12]. The SPEC can hold one FMC [13] card and an SFP [14] module which is used to connect to the WR network. On the PCIe side it has a 4-lane interface. The FMC mezzanine slot uses a low-pin count connector. Most of the FMC cards designed within the "Open Hardware Repository" project [15] (e.g. ADC cards, Fine Delay) can be used. For boards requiring more

Figure 4: A WR node implementation: Simple PCIe FMC carrier (SPEC)

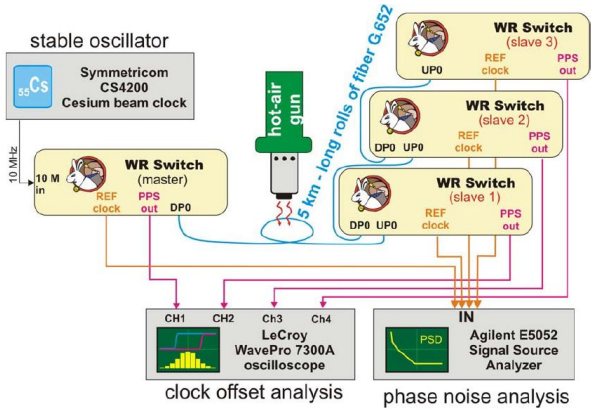

Figure 5: Time and frequency transfer test setup

facilities, the FMC PCIe Carrier [16] or its VME counter part [17] [18] can be used.

5. Measurements

In order to test the performance of time and frequency transfer, the test setup shown in Fig. 5 was assembled. It consists of a daisy chain of four switches connected with five kilometer fibre each (15 km in total). Varying operating conditions were simulated by heating the fiber with a hot air gun. A histogram of master-slave offsets is shown in Fig. 6. The accumulated skew is well below 1 ns. The standard deviation (sdev) is in the order of 6 ps.

6. Applications

In a WR network, the current time according to the primary reference clock, is distributed transparently to all WR nodes. This enables two main applications.

6.1. Digital to Time Converter

Since all WR nodes in the system agree with the precise current time, messages can be passed to generate events

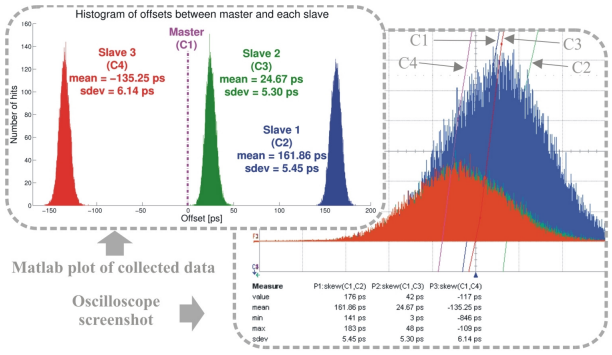

Figure 6: Example of timing measurements with the test setup.

at a certain time, specified in the messages. The WR network has deterministic behaviour with low transmission latency, still messages must be sent early enough to arrive in time at the WR nodes. This can be used to control large distributed systems such as accelerator facilities.

6.2. Time to Digital Converter

Events that occurs in a large distributed system can be timestamped with a high absolute accuracy due to the fact that all WR nodes in the system agree with the precise current time. This can be used in large volume particle detectors.

6.3. Control loops

Measured physical quantities can be processed to control distributed hardware. Loop stability is ensured by the maximum low-latency of the WR network.

7. Summary

The WR network pushes the frontiers of technology and is intentively based on well-established standards to ensure its long lifetime, wide support and commercial feasibility. By combining and extending existing technologies exceptional results are achieved that are still compatible with currently used technologies and standards. The WR PTP extension allows for increased accuracy and offers high reliability by supporting network redundancy. Compatibility with existing standards enables hybrid networks. WR PTP is used where high precision is required and standard PTP can be used where less precision is needed. WR is an open hardware and open software project that introduces new trends in cooperation between public institutions and commercial companies. The open hardware concept enables tendering for manufacturing WR devices. It reduces the risk of having a single source for WR device production. The WR applications extend beyond the developers community and many commercial companies are interested to provide WR capable instruments.

8. Acknowledgment

This work is part of the research programme of the 'Stichting voor Fundamenteel Onderzoek der Materie (FOM)', which is financially supported by the 'Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)'.

References

[1] ITU-T Study Group 15, "Timing and synchronization aspects in packet networks", International Telecommunications Union, Geneva, Switzerland, Apr. 2008, http://www.itu.int/rec/T-REC-G.8261-200804-I/en (2012)

[2] "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems", IEEE Std 1588-2008

[3] Open Hardware Repository, http://www.ohwr.org (2012)

[4] P.P.M. Jansweijer, H.Z. Peek, "Measuring propagation delay over a coded serial communication channel using FPGAs", Nucl. Instr. Meth. A, 626-627 (2011) S169

[5] ITU-T Study Group 13, "Timing characteristics of primary reference clocks", International Telecommunications Union, Geneva, Switzerland, Sep. 1997. URL: http://www.itu.int/rec/T-REC-G.811-199709-I/en (2012)

[6] D.W.Allan, H. Daams. "Picosecond time difference measurement system", Proc. 29th Annual Frequency Control Symposium, Atlantic City, USA, pp. 404-411, 1975

[7] Pedro Moreira, "PLL Optimisation for White Rabbit Applications", presentation at the Fifth White Rabbit Workshop, Geneva (Switzerland), 19 September 2011, http://www.ohwr.org/projects/white-rabbit/wiki/Sep2011Meeting (2012)

[8] White Rabbit Switch information, http://www.ohwr.org/projects/white-rabbit/wiki/WhiteRabbitSwitch (2012)

[9] White Rabbit Switch Technical Specification Version: 0.5 Date: September 9, 2011, URL: http://www.ohwr.org/documents/39 (2012)

[10] Micro Telecommunications Computing Architecture, http://www.picmg.org (2012)

[11] Simple PCIe FMC carrier (SPEC), http://www.ohwr.org/projects/spec/wiki (2012)

[12] White Rabbit PTP Core, http://www.ohwr.org/projects/wr-cores/wiki/Wrpc_core (2012)

[13] FPGA Mezzanine Card (FMC) Base Specification VITA 57.1, http://www.vita.com (2012)

[14] SFF Committee, INF-8074i Specification for SFP (Small Formfactor Pluggable) Transceiver, URL: ftp://ftp.seagate.com/sff/INF-8074.PDF (2012)

[15] Open Hardware Repository FMC projects, http://www.ohwr.org/projects/fmc-projects

[16] FMC PCIe carrier (PFC), http://www.ohwr.org/projects/fmc-pci-carrier/wiki (2012)

[17] Simple VME FMC carrier (SVEC), http://www.ohwr.org/projects/svec (2012)

[18] VME FMC CARRIER (VFC), http://www.ohwr.org/projects/show/fmc-vme-carrier (2012)

[19] M. Lipinski, T. Wlostowski, J. Serrano, P. Alvarez, "White rabbit: a PTP application for robust sub-nanosecond synchronization", International IEEE Symposium o Precision Clock Synchronization for Measurement Control and Communication (ISPCS), (2011) doi: 10.1109/ISPCS.2011.6070148