RELIABILITY IN A WHITE RABBIT NETWORK

Maciej Lipiński, Javier Serrano, Tomasz Włostowski, CERN, Geneva, Switzerland
Cesar Prados, GSI, Darmstadt, Germany

Abstract

White Rabbit (WR) is a time-deterministic, low-latency Ethernet-based network which enables transparent, sub-ns accuracy timing distribution. It is being developed to replace the General Machine Timing (GMT) system currently used at CERN and will become the foundation for the control system of the Facility for Antiproton and Ion Research (FAIR) at GSI. High reliability is an important issue in WR's design, since unavailability of the accelerator's control system will directly translate into expensive downtime of the machine. A typical WR network is required to lose not more than a single message per year. Due to WR's complexity, the translation of this real-world-requirement into a reliability-requirement constitutes an interesting issue on its own – a WR network is considered functional only if it provides all its services to all its clients at any time. This paper defines reliability in WR and describes how it was addressed by dividing it into sub-domains: deterministic packet delivery, data resilience, topology redundancy and clock resilience. The studies show that the Mean Time Between Failure (MTBF) of the WR Network is the main factor affecting its reliability. Therefore, probability calculations for different topologies were performed using the "Fault Tree analysis" and analytic estimations. Results of the study show that the requirements of WR are demanding. Design changes might be needed and further in-depth studies required, e.g. Monte Carlo simulations. Therefore, a direction for further investigations is proposed.

INTRODUCTION

The WR project is a multi-laboratory, multi-company, international effort to create a universal fieldbus for control and timing systems to be used at CERN, GSI and possibly other such facilities. The rationale behind WR, the choice of the technologies and technical details of its functioning have been already described in a number of papers [1], [2], [3]. The resilience and robustness is one of the key features of any fieldbus. This article presents a study on the reliability of a White Rabbit Network (WRN) assuming a basic knowledge about WR.

Reliability is defined as the ability of a system to provide its services to clients under both routine and abnormal circumstances. It can be estimated by calculating the probability of the system's failure (P_f). The lesser the probability of WRN failure, the higher its reliability. Thus, in this article we identify critical services of a WRN based on the study of WR's requirements. Then, we analyze each critical service to identify possible reasons for their failure and propose targeted counter-measures to increase reliability. Finally, their impact on the overall system reliability is studied to identify the highest contributor and the focus for the further studies.

DEFINITION OF RELIABILITY IN A WRN

A WRN, consisting of White Rabbit Switches (switches) connected by fiber or copper, is meant to transport information among White Rabbit Nodes (nodes). We distinguish two types of information distributed over the WRN: (1) *Timing* (frequency and Coordinated Universal Time) and (2) *Data* (the Ethernet traffic). This translates into two types of services provided by the WRN which have their own requirements and can be handled separately. The requirements are defined by GSI and CERN as the prospective users of WR to control their accelerators.

Timing Distribution

Timing is distributed in the WRN from a switch/node called Timing Master (TM) to all the other nodes/switches in the network. All the devices in the WRN lock their frequency (syntonize) and adjust their local clocks (synchronize) to that of the TM. The deviation between the clock of the TM and that of any other node/switch is called **accuracy**. A stable and continuous synchronization of all the nodes with an appropriate accuracy is the key requirement of the Timing Distribution in the WRN.

Data Distribution

The critical data distributed over the WRN is the one carrying sets of commands (events) which are organized into Control Messages (CM). The CMs are sent by a privileged node (Data Master, DM) in the payload of the Ethernet frame(s). Therefore, the Data Distribution in the WRN is broken into (1) *Control Data (CD)* – the Ethernet frames carrying CMs, critical, and (2) *Standard Data (SD)* – the Ethernet frames which do not carry CMs, non-critical. The reliability of the WRN depends on the successful delivery of the CD to all the designated nodes. The CMs are always broadcast within a VLAN , which can span the entire network. The worst-case upper bound of their delivery latency from the DM to any node in the network, regardless of it's location (**maximum distance from the DM**), is required to be guaranteed by the network – this is **a determinism** requirement.

Reliability of the WRN

The reliability of the WRN relies on the **deterministic** delivery of the CD to all the designated nodes and their

sufficiently **accurate and stable synchronization**. This means that the WRN is considered non-functional if one or more of the following occur:

- A node is synchronized with insufficient accuracy.
- A designated node receives corrupted CD or no CD.
- The upper-bound delivery latency has been exceeded.

Unreliability is translated into the number of CMs considered lost (not delivered, delivered corrupted or in a non-deterministic way) in a given period of time. During this time, the synchronization must be always of the required quality. Quantitative requirements of the accelerator facilities are listed in Table 1.

Table 1: GSI's and CERN's Requirements Summary

Requirement	GSI	CERN
Max latency	100 µs	1000 µs
CM failure rate	$3.17 * 10^{-12}$	$3.17 * 10^{-11}$
CMs lost per year	1	1
d_{max} from DM	2 km	10 km
CM size	200–500 bytes	1200–5000 bytes
Accuracy	probably 8 ns	1 µs to 2 ns

FAILURE STUDY

One of the main possible reasons for WRN failure, which affects both Timing and Data Distribution, is a malfunction of its elements (switches or links). Since the distribution of information in the WRN is of one-to-all character (Data/Timing Master to all nodes), all the elements of the WRN are considered Single Points of Failure (SPoF)[4]. Malfunction of any SPoF results in failure of the entire system. SPoFs can be eliminated by introducing redundancy of the system components. Due to its special features (distribution of frequency over physical layer) and strict requirements (determinism, low data loss), the number of possible redundant topologies of the WRN is restricted, as explained in the following sections.

Imperfections of the physical medium as well as switching between redundant elements of the network (which takes time) can cause loss or corruption of data. The deterministic and mostly broadcast character of the data distribution in the WRN enforces application of the Forward Error Correction (FEC) – adding redundant information on transmission to enable recovery of lost or corrupted data on reception. This brings constant data overhead and the probability that the added redundancy is not sufficient to recover the data. However, it is the price to pay for ensuring low latency and determinism of data delivery in the WRN.

The delivery latency of an Ethernet frame varies with cable length and the number of hops (switches) it has to traverse to reach its destination, the traffic load on the way and the assigned Class of Service (CoS). Therefore, to ensure the required determinism of the CD delivery, we need to make sure that there is no congestion of Ethernet frames carrying CMs. Moreover, the number of hops (the latency introduced by them) needs to be sufficiently small, which can be done by restricting the topology.

The resilience of the Clock Distribution translates into continuous and stable synchronization of all the nodes and switches in the WRN (Table 1). Although, the network redundancy eliminates SPoFs, the switch-over between redundant elements might introduce instability and render the network unreliable despite the costly redundancy. Therefore, a seamless switch-over between redundant clock paths needs to be ensured. Another reason for the deterioration of the synchronization accuracy is the variation of external conditions (e.g. temperature) which needs to be compensated.

DETERMINISM

A carefully configured and properly used WRN offers deterministic Ethernet frame delivery thanks to the implementation of CoS and the fact that the delay introduced by the switch can be verified by analysis of **publicly available source code** [5]. Such analyses were performed to verify the worst-case upper bound delivery latency of a CM against the requirements listed in the Table 1. The results, presented in Table 2 (*Store-and-forward* column), take into account the fact that a CM is encoded into 4 Ethernet frames (as required by the FEC and described in the next Section), it is sent with the highest priority (CoS) and it always traverses 3 hops.

Table 2: Control Message(CM) Deliver Latency Estimations

CM size	CM deliver latency			
	Store-and-forward		Cut-through	
	GSI	CERN	GSI	CERN
500 bytes	221 µs	283 µs	76 µs	118 µs
1500 bytes	285 µs	325 µs	102 µs	142 µs
5000 bytes	324 µs	364 µs	162 µs	202 µs

The analysis revealed that GSI's requirements are not fulfilled: the upper-bound delivery latency for the required size of CM and max distance of 2km is greater then 100 µs.

The solution to decrease delivery latency is targeted into the CD only and takes advantage of its characteristics (broadcast within a VLAN, sent by privileged node). We propose to break the highest priority of the CoS into two (unicast and broadcast) and use the highest priority broadcast Ethernet traffic only for the CD. Moreover, this particular traffic shall be forwarded using the cut-through method (unlike the store-and-forward method used normally in the switch) which can be effectively fast for the broadcast traffic with a single source (DM). The results, presented in Table 2 (*Cut-through* column), show a significant improvement. The solution requires hardware-supported cut-through forwarding in the switch as described in [6].

DATA RESILIENCE

Forward Error Correction

The objective of the FEC scheme is to decrease the loss rate of the CMs, preferably, to less then one per year. WR uses as a physical medium Fiber Optic and CAT-5. The number of received corrupted bits compared to the total number of received bits is called Bit Error Rate (BER). The value of BER characterizes a physical medium and can be used to characterize the entire switched network. A WRN can be seen as a Packet Erasure Channel (PEC) or as a Binary Erasure Channel (BEC) depending on the effect of a bit error on the frame. If the frame is lost (e.g. dropped by the switch due to a corrupted header or lost during switch-over between redundant components), the WRN is a PEC. If the bit error happens in the link between a switch and node, a corrupted frame can be used (optional) to attempt frame recovery. In such case, the channel is called BEC. Each type of channel requires a different FEC solution. Therefore two concatenated FECs are used in WR. Reed-Solomon (R-S) coding is used for the PEC and allows to encode k original-frames into n encoded-frames ($n > k$). Reception of any k encoded-frames can be used to decode the original frames. Hamming coding with additional parity (SEC-DED) is used for the BEC and allows to detect up to two simultaneous bit errors and correct a single error. These two schemes (R-S and Hamming) are combined to encode each CM – it is split into two and encoded using R-S into four messages (two original and two with redundant data). Each of the four messages is then encoded using Hamming. Such encoded messages are sent in a burst of 4 Ethernet frames. Reception of any two of these frames enables to decode the original Control Messages. A systematic analysis, using the BER characteristic of the WRN, proves that the presented FEC scheme guarantees less than single CM lost per year due to physical medium imperfection, as can be seen from Table 3.

Table 3: GSI and CERN FEC Characteristics

Parameter	GSI	CERN
Control Message length	500 bytes	1500 bytes
Control Message per year	$3.145 * 10^{11}$	$3.145 * 10^{8}$
Max Bit Correct.	1	1
Payload Length	294 bytes	854 bytes
Num Encoded Frames	4	4
Needed Frames to Receiver	2	2
Probability of Loosing a CM	10^{-14}	10^{-13}

Rapid Spanning Tree Protocol (RSTP)

In an Ethernet network with redundant topology, the problem of loops (causing "broadcast storms") is handled by the Rapid Spanning Tree Protocol (RSTP) . It creates a loop-free logical topology by blocking appropriate ports in switches, and unblocks them in case of topology break (due to element failure).

The functionality provided by the RSTP is essential for the WRN. However, the convergence speed provided by the standard implementation of the RSTP (milliseconds at best) would cause many CMs to be lost during the process. This is not acceptable, we need a solution which is fast enough to prevent loosing the CMs at all. Since we know the size-range of the CMs (Table 1) and how they are FEC-encoded into Ethernet frames, we can calculate the maximum value of the convergence time: 3 µs. This time is smaller than the duration of transmitting a single frame with FEC-encoded CM – this ensures that no more than two frames with FEC-encoded CM are lost, thus the CM can be recovered.

In order to achieve a convergence time of 3 µs, the switch-over between active and backup connections needs to be performed in the hardware as soon as the link-down is detected. It can only be done if the alternative topology is known in advance. The knowledge of alternative topology is translated into an RSTP-assignment of alternative and backup roles of switch ports, i.e at least one port with alternative role must be identified in every switch (except the topology-root switch). If we ensure, by restricting the topology, that RSTP identifies the alternative links, we can use its data to feed the hardware, consequently achieving the required convergence time and staying standard-compatible: the hardware switch-over is just a faster RSTP-driven convergence. The required topology restrictions, described in [6], greatly overlap with these imposed by the Time Distribution.

CLOCK DISTRIBUTION

A seamless switch-over between redundant sources of timing (uplink ports) is heavily supported by the Clock Recovery System (CRS) [2] of the switch and the WR extension to PTP (WRPTP)[3].

Figure 1 presents an example where a switch (timing slave) is connected (by its uplinks 1 & 2) to two other switches (primary and secondary masters) – the sources of timing. On both uplinks the frequency is recovered from the signal and provided to the CRS. Similarly, WRPTP measures delay and offset on each of the links and provides this data to the CRS. The modified Best Master Clock (mBMC) algorithm [3] decides which of the timing masters is "better" and elects it the primary, the other is considered secondary (backup). The information from *uplink 1* (primary) is used to control the CRS and adjust the local time. However, at any time all the necessary information from the *uplink 2* is available and a seamless switch-over can be performed in case of primary master failure [2].

In addition to the switch-over-related synchronization instability, the variation of external temperature can cause an accuracy degradation. This problem, however, is solved by the PTP standard itself. By frequent link delay measurements, the fluctuation is compensated.

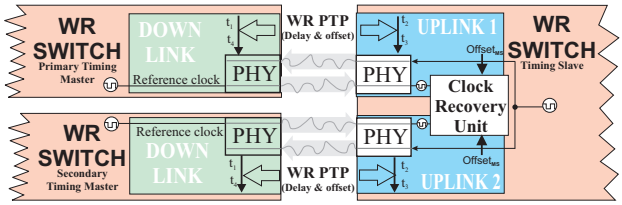

Figure 1: Seamless switch-over.

OVERALL RELIABILITY

The final equation of the WRN reliability is a sum of the data and clock distribution reliabilities. The clock distribution is assumed to be sufficiently accurate as long as there is a connection between the TM and all the nodes. The same applies to the CD distribution: as long as there is a valid connection, the FEC makes sure that the data is delivered with a sufficient reliability and the latency calculations prove it to be deterministic while the congestion is prevented by CoS and limited number of data sources (DM). Consequently, the overall reliability is strongly dependent on the WRN topology, which needs to be appropriate for the proposed solutions (SyncE, H/W-supported RSTP, upper-bound latency).

For the comparison of different network topologies, we consider the reliability of a network of switches. Each node is connected to such a network with M links (each to a separate switch). The value of M reflects the level of redundancy (M=1 for no redundancy, M=2 for double redundancy, etc).

In the calculations of the network reliability we used the idea of Mean Time Between Failure (MTBF) and its relation with the failure probability presented in [4] (a very simplified mathematical model). In order to calculate the MTBF of the entire network, we need the MTBFs of each network component: switches and links. Since the WR switches are still under development (no MTBF measured), we used representative values for CISCO switches (2, 10 and $100{*}10^4$[h]). Two estimation methods were used: "Fault Tree analysis" [7] and analytic. Both provide just rough estimations of the reliability. The former allowed to estimate two-terminal reliability (DM to single node) of simple non/double/triple-redundancy topologies (P_f). The most desired value is the all-terminal network reliability ($P_{f\_Network}$), where : $P_f < P_{f\_Network} < N_{nodes} * P_f$. Table 4 presents rough estimations of $P_{f\_Network}$ using analytic calculations for the three considered topologies ($MTBF_{Switch}$=200 000[h]). However, to meet the requirement of ≈2000 nodes and only three network layers (hops), the Data Master node is connected to more separate switches than the level of redundancy (M). The estimations show that a triple redundancy topology can barely satisfy the requirements by CERN (Table 1).

Table 4: WRN Topologies's Reliabilities

Redundancy	Switches	P_f	MTBF[h]
No	127	$2.08 * 10^{-3}$	$5.77 * 10^3$
Double	292	$4.71 * 10^{-7}$	$2.55 * 10^7$
Triple	495	$3.06 * 10^{-11}$	$4.08 * 10^{11}$

CONCLUSIONS

A WRN must be considered as an ordinary Ethernet network with extra optional built-in features which, when properly used, can make it robust and more reliable. This, however, comes at a price of topology restrictions and redundant elements (money). The reliability study described in this article and detailed in [6] presents areas which need to be addressed to increase the reliability of a WRN. The development of WR is an on-going effort and some of the suggested solutions have been already properly investigated or developed (FEC, clock distribution) while the others need further verification (RSTP, cut-through forwarding). Suggested solutions enable to fulfill the requirements set by CERN and GSI. However the costs might trigger double-checking and re-justifying of at least two of them: upper-bound latency by GSI and the number of CMs lost per year. The former requires additional development efforts to achieve the required 100 µs. The latter requires a high level of network redundancy (triple or more) which is very costly. Since the network topology and its reliability calculations turned out to be the greater factor in the overall system reliability, it is necessary to perform more precise calculations and simulations to verify the rough estimations. This might include different techniques (e.g. Monte Carlo simulations) but also more real-life use cases (i.e. of the network layout suggested in [8], which was not available at the time of described study). Especially, we need to take into account and include into calculations the fact that not all the nodes connected to the WRN are equally critical in real-life applications.

REFERENCES

[1] J. Serrano, P. Alvarez, M. Cattin, E. G. Cota *et al.*, "The White Rabbit Project," in *ICALEPCS*, Kobe, Japan, 2009.

[2] T. Wlostowski, "Precise time and frequency transfer in a White Rabbit network," Master's thesis, Warsaw University of Technology, may 2011.

[3] E. Cota, M. Lipinski, T. Wlostowski, E. Bij, and J. Serrano, "White Rabbit Specification: Draft for Comments," http://www.ohwr.org/documents/21, july 2011, v2.0.

[4] K. Dooley, *Designing Large-Scale LANs*. O'Reilly, 2002.

[5] White Rabbit. http://www.ohwr.org/projects/white-rabbit.

[6] C. Prados and M. Lipinski, "White Rabbit and Robustness," http://www.ohwr.org/documents/103, March 2011.

[7] "Reliability workbench, fault tree," www.isograph.com.

[8] J.-C. Bau and M. Lipinski, "White Rabbit CERN Control and Timing Network," http://www.ohwr.org/documents/85, July 2011.