Real-Time Fault Diagnosis of Node and Link Failures for Industrial Controller Redundancy

Kasra Ekrad[1,2], Sebastian Leclerc[1,2], Bjarne Johansson[2,1],
Inés Alvarez Vadillo[2], Saad Mubeen[1], Mohammad Ashjaei[1]
[1] Mälardalen University, Västerås, Sweden [2] ABB AB, Västerås, Sweden
{kasra.ekrad, sebastian.leclerc, mohammad.ashjaei, saad.mubeen}@mdu.se,
{bjarne.johansson, ines.alvarez-vadillo}@se.abb.com,

***Abstract*— In an Industrial Control System (ICS), link failures causing partitioning between the redundant controller pair can prevent synchronization. This disruption could potentially trigger incorrect redundancy role-switching procedures. Distinguishing link and node failures from each controller's perspective can effectively mitigate these dependability threats. We introduce an algorithm capable of detecting and differentiating link failures from node failures within a bounded time. The proposed algorithm leverages periodic messages generated by industrial protocols operational in network devices. Our preliminary evaluation indicates the feasibility of the proposed algorithm in terms of meeting the real-time requirements of the ICSs.**

***Index Terms*—Node and Link Failure, Industrial Control System, Redundant Controller Pair**

I. Introduction

In the era of Industry 4.0, Industrial Control System (ICS) will serve as the backbone of industrial systems [1]. Currently, these systems are structured in a segmented infrastructure, with different technologies tailored to meet the stringent requirements of industrial processes. This segmented architecture limits the implementation of novel applications such as those envisioned by Industry 4.0. For this reason Ethernet has gained a lot of attention from both academia and industry. Ethernet advantages drive this shift as it provides higher bandwidth, is more cost-effective, and supports scalability, enabling a network-centric architecture envisioned by Industry 4.0 [2]. However, Ethernet was designed for data communications and cannot independently support this transition.

To increase reliability in an Ethernet-based ICS, such as the one depicted in Fig. 1, fault tolerance techniques and mechanisms should be considered [3]. Employing redundancy which is a fundamental aspect of many fault tolerance strategies is essential for enhancing the reliability of ICSs. These systems consist of various components, namely Supervisory Control and Data Acquisition (SCADA), field, and controller level where Distributed Control Nodes (DCNs) reside. DCNs control and monitor critical field-level device processes to meet the industries' operational requirements. In a redundant DCN pair use case, where both an active and a backup DCN exist, link failures resulting in network partitioning, prevent the DCN from exchanging information [4]. This can lead to an overall failure of the system, as the active and backup controllers could simultaneously become active and send contradictory instructions to the I/Os. To prevent this from happening, the active and backup DCNs should be aware of the type of failure i.e. node or link, to perform the role-switching mechanism correctly if necessary within a bounded time [5].

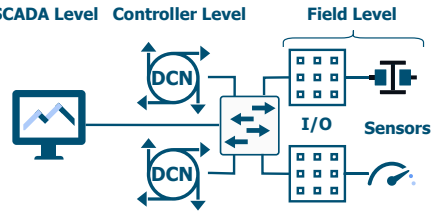

Fig. 1. Industrial Control System components and hierarchical structure.

Contributions: The current techniques to prevent dual active DCN in the event of link failures include the deployment of heartbeat signals. From now on, we will refer to these as node heartbeat (NHB). Nevertheless, existing strategies often fail to meet the real-time (RT) demands of an ICS and do not account for network partitioning. To bridge this gap, in this paper, we propose to use a more reliable source of periodic messages generated by network devices as network heartbeat (NWHB) to distinguish the types of failures. The solution is implemented on both active and backup DCNs as a distributed algorithm to listen to incoming NHBs and NWHBs. Further, the backup DCN performs a role-switch in the absence of NHBs from active DCN while it receives the NWHBs. On the other hand, the active DCN decides to switch roles in the absence of NWHBs. In this work, we propose to use the Media Redundancy Protocol (MRP) as the source of periodic messages for its extensive use in industrial networks. This approach leverages MRP beyond its conventional applications in conjunction with NHB signals, to enhance the reliability of ICSs. We also show the feasibility of the solution by implementing a proof-of-concept, yet further evaluation is an ongoing work.

II. Related Work

Many works have provided fault tolerance to Ethernet to aid the industrial transition process by leveraging protocols and standards, as surveyed in [3]. There is also literature focused on link error detection using Software-Defined Networking,

mostly focused on large-scale non-industrial networks, with the target on setting alternative routes after failures in the network [6], [7]. Unfortunately, most of these works rely on non-real-time layer 3 protocols, which cannot be supported in most existing industrial layer 2 devices.

Other works focus on detecting link errors but do not provide any RT guarantees on the detection, e.g., works based on Ping and Syslog. Specifically, Syslog-based error detection in [8] is based on periodic log aggregation and analysis, which can lead to delays that cannot be tolerated in the industrial use case addressed here. For this reason, we focused on the widely used industrial protocol MRP, which supports error recovery with a worst-case failover time of 200 milliseconds for a ring network [9]. Unfortunately, MRP cannot designate the failure location or distinguish between node and link failures.

The problem in this paper was inspired by the study in [5], which assesses dual active DCN problem in an industrial context. They proposed the network reference point to properly choose the active DCN if network partitioning occurred during link failures. However, their fault diagnosis method relies solely on heartbeats and does not include mechanisms for locating the fault. Other solutions are available to choose the active DCN in a network such as the bully leader election algorithm evaluated in [10]. However, this algorithm elects one leader per each network partition which can lead to dual active DCN.

Therefore, this work proposes a solution for diagnosing and handling faults that affect the redundant DCN in a layer 2 industrial Ethernet network considering real-time requirements.

III. Distinguishing Node and Link Failure

To increase the reliability of the DCNs, we can use fault prevention and fault tolerance mechanisms [11]. One way of applying fault tolerance is employing spatial redundancy. In the use case of a switched layer 2 Ethernet redundant DCN pair depicted in Fig. 2, spatial redundancy is achieved by duplicating the DCN, resulting in an active DCN and a hot standby backup DCN. These systems typically employ a bidirectional NHB signal to switch roles when unforeseen failures occur, e.g. when the active node fails. However, NHB alone is not enough for the DCNs to decide on whether the backup DCN should become active or not. In fact, when a link failure partitions the network, i.e., the DCNs are not connected to each other, yet they are connected to the I/O, NHB can result in dual active DCNs.

To detect and differentiate faults alongside the NHB, we can use another periodic message originating from the network devices, namely the NWHB. To meet the RT requirements of industrial operations, NWHB are transmitted periodically and within a bounded time. Additionally, NWHBs should be sent towards both DCNs from the network device. Some existing protocols including MRP and Bidirectional Forwarding Detection can support these requirements. However, since the scope of the study is limited to a layer 2 Ethernet switched network, MRP protocol is the selected candidate.

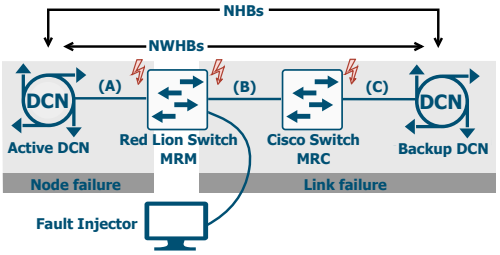

Fig. 2. Experimental test bed of redundant DCN use case employing NHB originated from DCNs and NWHBs from switch device.

MRP, defined in IEC 62439-2, is a layer 2 redundancy protocol designed for maintaining ring topology integrity within industrial Ethernet networks in the event of link failures [9]. MRP verifies the integrity of the ring utilizing two primary devices: A Media Redundancy Manager (MRM) which deterministically sends MRP_Test frames through its ring ports at specific intervals, and one or multiple Media Redundancy Clients (MRCs) that forward these MRP_Test frames through both ring ports. We can utilize the MRP_Test frames as the previously mentioned periodic messages for detecting and differentiating faults. The solution will be evaluated in the use case depicted in Fig. 2 where an active DCN is connected to the backup DCN through a line topology. Nonetheless, our solution can be employed in any topology that contains two DCNs connected over a line of n switches. Note that even though MRP provides space redundancy by means of a ring topology, MRP can operate in a line topology.

A. Overview of the Model

Let us assume that active and backup DCNs only notify each other's status by providing bidirectional NHB to keep the integrity of the DCN redundancy, as shown in Fig. 2. In the event of a link failure at locations (A), (B), (C), or an active DCN failure, both the active and backup DCNs stop receiving NHB. Consequently, it becomes ambiguous for both the backup and active to determine the location of link failure and if the disruption is due to a link or node failure. This issue may arise across various topologies, network devices in between, and active/backup nodes [4]. It is assumed that when links fail, they cannot transmit frames in any direction and we only cover Fail-Silent semantics.

To distinguish a node failure from a link failure, the device generating the NWHB (Red Lion switch) is designed to be adjacent to the active DCN as depicted in Fig. 2. By this design, from the **backup DCN perspective**, if only the NHBs are missing, there might be a node/link failure designated as (A) since NWHBs are being received. The absence of NHB and NWHB will be identified as link failure in locations designated as (B) and (C). Note that the solution cannot distinguish node and link failure in (A), as a result, the failure at this location is considered node failure. Also differentiating (B) and (C) is not considered since by only differentiating node failure from link failure, the role-switching mechanism can correctly be executed. **From the active DCN's perspective**, the role-switching procedure is triggered only if it stops

receiving NWHBs, indicating a node failure. Role-switching, in this case, can be assumed as a $FailSilent$ state to cease all operations and also stop sending NHBs. If NHBs transmitted from the backup are not received, this will indicate a backup failure. In this scenario the active can still drive the process to keep the system available, therefore backup failure is not time-critical and can be reported to the SCADA level. As a result, from the perspectives of both DCNs, a node failure prompts the role-switching procedure, while link failures do not lead to any change in roles. However, the type of decisions after differentiation is up to the system's requirements. This solution is applicable in larger networks, Fig. 2 is provided as an example.

TABLE I
UNIQUENESS OF CATEGORIZED FAILURE FROM EACH DCN PERSPECTIVE WHILE LINK FAILURES OCCUR AT A, B, AND C OR DCN LOCATIONS.

Failure location	Perspective	NWHB Received	NHB Received	Diagnosed Failure
A	Active	No	No	Node
B	Active	Yes	No	Link
C	Active	Yes	No	Link
Backup	Active	Yes	No	Link
A or Active	Backup	Yes	No	Node
B	Backup	No	No	Link
C	Backup	No	No	Link

Table I presents the unique differentiation between node and link failure scenarios from each DCN's perspective. Note that, if the role switching relied solely on NHBs, both DCNs would stop receiving NHBs in any failure scenario.

Based on the stated logic and scenarios in Table I, we constructed an algorithm to distinguish between node and link failures. The algorithm is designed in a decentralized manner, while it is presented as a unified algorithm in Algorithm 1. The logic behind distributing the algorithm across both DCNs is twofold. It prevents scenarios where redundant DCNs are unaware of each other's status and it can passively monitor the network without interfering with the critical traffic. The algorithm receives NHB and NWHB messages and is designed to handle only one fault in the system at a time. Upon detecting a single NHB or NWHB miss in $EvaluateSingleFrameMiss$, it starts to sequentially evaluate further misses, prioritizing NHBs due to their higher frequency and critical role in timing evaluations as will be discussed in Sections III-B. It identifies the failure type, i.e., node or link failure, and initiates the appropriate role-switching procedure for both active and backup DCNs, depending on their current roles. Note that the algorithm only bases its decisions on the presence or absence of NHB and NWHB. Identifying the root causes leading to the failure (e.g., switch or port) is left for future work.

Algorithm 1 Node and link failure differentiation

```
1: procedure FAULTHANDLER
2:     EvaluateSingleFrameMiss()
3:     if no NHB received then
4:         if no NWHB received then
5:             if Node = Active then NodeFailureDetected()
6:                 FailSilent()
7:             end if
8:             if Node = Backup then LinkFailureDetected()
9:             end if
10:        end if
11:        if NWHB received then
12:            if Node = Active then LinkFailureDetected()
13:            end if
14:            if Node = Backup then NodeFailureDetected()
15:                SwitchRoles()
16:            end if
17:        end if
18:    end if
19: end procedure
```

B. Realization of the model

Layer 2 Ethernet switches configured with MRP protocol were used to evaluate the model. MRP protocol served as the source of NWHB messages. Any MRP-capable switch can be used which in our case, the Cisco IE-2000-8TC-L running IOS Version 15.2(7)E2 and Red Lion NT-4008-000-PN-M running Firmware 1.0.9 were utilized. MRP_test frame explained in Section III is the realization of the NWHBs. These messages are configured to be transmitted every 20 milliseconds from the leftmost switch. The leftmost Red Lion switch is configured as MRM and the rightmost Cisco switch is configured as MRC. This study utilized non-RT Lenovo M720q7 devices running Ubuntu 22.04 to emulate two DCNs and a fault injector instead of actual controllers as a proof of concept. These devices connected by the aforementioned switches, formed the test bed as illustrated in Fig. 2.

The role-switching of the active node in this setup is not dependent on NHBs transmitting from the backup node. Thus, to prove the concept, only a unidirectional NHB was implemented to transmit every millisecond with negligible jitter from the active to the backup node and its one millisecond interval enables the timing evaluation of the deployed algorithm at millisecond precision. This method eliminates the need to deploy time synchronization protocols. It implies that the time evaluation of each event in the system will be relative to the first NHB missed time from the backup perspective. Furthermore, the topology is considered as known by the algorithm. However, the topology does not affect the differentiation of node and link failures since the active DCN is always located adjacent to the MRM node.

The algorithm is developed in C++ as distributed software across active and backup nodes. It can passively monitor and evaluate NHBs and NWHBs i.e. MRP_Test frames. The algorithm starts to identify specific types of packets in $EvaluateSingleFrameMiss$ by their unique Ethernet headers, i.e. UDP-based NHBs and MRP frames. It records the last received times for these frames and employs timeout logic to detect delays in their inter-arrival time, with specific thresholds set for NHBs (1 millisecond) and NWHBs (20 milliseconds). An additional 0.5 millisecond for the NHBs and 1 millisecond for NWHBs was added to the configuration to compensate for the execution time and jitter.

Suppose NHBs are missed beyond the allowed jitter-adjusted period. In that case, the system evaluates NWHBs to

decide on a role-switching procedure based on the explained algorithm in Section III-A. The implemented algorithm on the active node relies on the absence of NWHBs to make decisions, without requiring to determine the fault's location or receive NHBs from the backup node even though Algorithm 1 presents the general idea.

Finally, the fault injector was built to turn the lightning bolt-marked interfaces up and down periodically utilizing commands over Telnet. During the experiment, the fault injector changed the interface status every 15 seconds, and in total 3000 tests were executed at either location (A), (B), and (C). During the tests, role-switching times were measured from both active and backup perspectives which will be presented in Section IV.

IV. Results and Discussion

We have conducted 3000 tests by injecting link failures using the injector. We observed that the algorithm could differentiate the node and link failures correctly with no false positives. From the active perspective, the differentiation led to role-switching taking a minimum of 43.203 milliseconds and a maximum of 43.54 milliseconds. Role-switching time on the active perspective comprises two NWHB periods plus algorithm execution time. The minimum and maximum role-switching times when the backup differentiated node and link failures were 72.943 milliseconds and 73.213 milliseconds, respectively. Role-switching time on the backup constituted three NWHB plus six NHB in addition to the algorithm's execution time.

The backup DCN initiates the role-switching procedure after three NWHB intervals, whereas the active DCN executes this task after two intervals. Originally, if the active DCN had been set to switch roles after only one frame interval, the algorithm might have mistakenly interpreted a single missed NWHB, possibly due to a mere temporary delay, as a permanent fault. This misinterpretation could have resulted in unnecessary role-switching. To address this, an additional NWHB interval was incorporated into the active DCN's role-switching timeframe, extending it to two intervals. Additionally, to prevent the occurrence of dual active DCNs simultaneously, the role-switching for the backup DCN is deliberately delayed to occur one NWHB interval after the active DCN has performed its role-switching mechanism. Assuming that the role-switching mechanism executes instantly, the backup switches roles at three MRP_Test frame intervals.

Furthermore, jitter can affect the inter-arrival time of NHBs and NWHBs. This jitter may result from the systems' non-RT nature or the duration of the algorithm's evaluations. The results would be enhanced by utilizing the RT environment. To account for the non-RT characteristics of our evaluation system, we employ an NHB tolerance of six. This means that a permanent failure is only detected after six consecutive NHB misses, which helps to prevent temporary failures, such as single packet delays, from being classified as permanent failures.

V. Conclusion and future works

This study aimed at proposing a solution to enhance the reliability of redundant DCNs within ICNs in the event of failures. The proposed solution distinguishes between node and link failures in a line topology of a layer 2 industrial Ethernet switched network within a bounded time. This supports the role-switching mechanism of the DCNs and prevents the dual active DCN problem. The solution is applicable to any dual redundant nodes in any communication environment. The maximum role-switching procedure was executed in 73.213 milliseconds after the occurrence of an actual failure.

Further investigations are to explore other protocols that can deliver periodic messages in an industrial environment and propagate the source of the problem in an RT manner to exactly localize the failure. Additionally, this solution should be evaluated in more complex topologies and assess its behaviour in an actual RT environment.

Acknowledgements

This work is supported by the Swedish Agency for Innovation Systems and the Knowledge Foundation via the ISECURE and SEINE projects.

References

[1] I. Álvarez, D. Bujosa, B. Johansson, M. Ashjaei, and S. Mubeen, "Centralised architecture for the automatic self-configuration of industrial networks," in *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation*, 2023.

[2] B. Leander, B. Johansson, T. Lindström, O. Holmgren, T. Nolte, and A. V. Papadopoulos, "Dependability and security aspects of network-centric control," in *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation*, 2023.

[3] I. Álvarez, A. Ballesteros, M. Barranco, D. Gessner, S. Djerasevic, and J. Proenza, "Fault tolerance in highly reliable ethernet-based industrial systems," *Proceedings of the IEEE*, vol. 107, no. 6, pp. 977–1010, 2019.

[4] S. Gilbert and N. Lynch, "Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services," *Acm Sigact News*, vol. 33, no. 2, pp. 51–59, 2002.

[5] B. Johansson, M. Rågberger, A. V. Papadopoulos, and T. Nolte, "Consistency before availability: Network reference point based failure detection for controller redundancy," in *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation*, 2023.

[6] H. Gao, L. Zhao, H. Wang, Z. Tian, L. Nie, and K. Li, "Xshot: Light-weight link failure localization using crossed probing cycles in sdn," in *Proceedings of the 49th International Conference on Parallel Processing*, ser. ICPP '20. New York, NY, USA: Association for Computing Machinery, 2020.

[7] U. C. Kozat, G. Liang, K. Kokten, and J. Tapolcai, "On optimal topology verification and failure localization for software defined networks," *IEEE/ACM Transactions on Networking*, vol. 24, no. 5, pp. 2899–2912, 2016.

[8] A. M. Vargas-Arcila, J. C. Corrales, A. Sanchis, and Á. R. Gallón, "Peripheral Diagnosis for Propagated Network Faults," *Journal of Network and Systems Management*, vol. 29, no. 2, p. 14, Jan. 2021.

[9] "Industrial communication networks - High availability automation networks - Part 2: Media Redundancy Protocol (MRP)," Standard IEC 62439-2:2015, 2015.

[10] M. Numan, F. Subhan, W. Z. Khan, B. Assiri, and N. Armi, "Well-organized bully leader election algorithm for distributed system," in *2018 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications*, 2018.

[11] C. Vitucci, D. Sundmark, M. Jägemar, J. Danielsson, A. Larsson, and T. Nolte, "A reliability-oriented faults taxonomy and a recovery-oriented methodological approach for systems resilience," in *2022 IEEE 46th Annual Computers, Software, and Applications Conference*, 2022.