

Performance of Position Based Fast Fault Recovery Protocol for Industrial Bridged Ring Networks

산업용 브리지 망을 위한 위치 기반의 신속한 망 장애 복구 절차의 성능분석

Ju Sang Seo, Chong Ho Yoon, Hong Soon Park, Jin Uk Kim
서주상, 윤종호, 박홍순, 김진욱

Abstract

With the proposal-agreement procedure, RSTP can reduce the network recovery time to 400 ms or less in the case of 40 bridges. While the legacy RSTP reverts the previous agreement at the bridge with the alternate port role in the ring during the fault recovery, a new position based fast fault recovery procedure is proposed in this paper to guarantee a single proposal-agreement transaction which can provide more faster recovery. By knowing the relative position of the faulty link or bridge in hops, the bridge on the middle of the ring can complete the recovery procedure without revert. The performance of proposed procedure is numerically calculated and verified by simulation and the result shows that the recovery time can be reduced up to 100 ms, which is 1/4 times of the legacy RSTP.

RSTP는 링크 상태에 반응하는 제안-동의 절차를 통해 포트의 상태 천이 시간을 감소시킴으로써 노드 수 40개로 구성된 링 망에서 장애 복구 시간을 400 ms 이하로 단축시킬 수 있다. 본 논문에서는 기존 RSTP가 장애 위치 및 종류에 따라 이웃 브리지 간에 제안-동의 절차가 대체 포트를 갖는 브리지에서 반복되는 현상에 유의하여, 장애 발생 시 오직 한 번의 제안-동의절차가 수행되게 함으로써 장애 복구 시간을 최대 1/4로 감소시킬 수 있는 새로운 위치 기반의 장애 복구 절차를 제안하고 성능을 분석하였다. 제안된 절차는 링 형 망에서 각 브리지가 루트 브리지에 대한 상대적인 위치, 즉 홉 수를 사전에 인지하도록 하고, 이를 기반으로 루트 브리지의 장애 또는 링크 장애 발생 시 이를 감지한 이웃 브리지에 의해 개시된 제안-동의 절차가 양 방향의 링에서 전파될 때 장애 감지 브리지를 기준으로 항상 링의 중간 브리지에서 제안-동의 절차가 종료되도록 함으로써 신속한 망 장애 복구 절차가 가능하다. 수치 분석 및 모의실험을 통하여 제안된 절차가 기존 400ms의 RSTP에 비하여 1/4인 100 ms 이내의 복구 시간으로 단축될 수 있음을 검증하였다.

Keywords: Rapid Spanning Tree Protocol, RSTP, Ring Network, Position Based, Optimization, Simplifying Proposal and Agreement Procedure

I. Introduction

산업 IT의 핵심 중 하나인 산업용 제어 통신망 기술이 기존 직렬 통신 기반에서 이더넷 기반으로 발전하고 있는 이유 중 하나는 이더넷으로 저렴하게 고속 전송할 수 있어 실시간성 확보가 용이하기 때문이다 [1]. 특히 산업용 이더넷은 네트워크 이중화 또는 신속한 망 장애 복구 절차를 통해 장애 발생 시 허용하는 고장 감내 시간 내에 복구할 수 있다.

네트워크 이중화를 통한 복구 절차 중 하나인 IEC61850 표준인 HSR (Highly available Seamless Redundancy) 기술은 0 ms의 복구 시간을 지원할 수 있어 스마트 변전소용으로 사용된다 [2]. 하지만 이러한 이중화 방식은 효율성과 비용 문제를 함께 갖는다.

반면에 장애 발생 시 가급적 신속하게 망을 재설정하는 망

복구 방식으로는 RSTP (Rapid Spanning Tree Protocol), MRP (Media Redundancy Protocol), Turbo Ring 등이 있으며, 상대적으로 저렴하다 [3]-[5]. 최근에는 RRP [6]가 IEC62439-7 표준 규격으로 추가 등록되었다.

이러한 이중화 기술들은 고가용성 측면에서 링 형태의 망을 선호한다. 메쉬 망에서의 복구 시간을 줄이기 위한 접근법도 있지만 [7] 링 형태의 망은 이중화에서 다른 형태의 망보다 복잡성과 비용 그리고 보수 유지 측면에서 우수하다 [8]. 특히 항공용 네트워크 시스템 규격인 AFDX와 Aeroring 기술도 링 형태의 망을 기반으로 한다 [9][10].

IEEE802.1 표준의 RSTP는 기존 STP가 망 재구성에 따른 포트 상태 천이 시 타이머에 의존함으로써 장시간 소요되는 단점을 극복하기 위하여 이웃 브리지 간에 새로운 제안-동의 절차를 사용

Article Information

Manuscript Received April 17, 2019, Accepted May 3, 2020, Published online September 30, 2020

The authors are with School of Avionics and Information, Korea Aerospace University, 76 Hanggongdaehak-ro Gyeonggi-do 10540, Republic of Korea.

Correspondence Author: Ju Sang Seo (donworry1012@kau.kr)

ORCID: 0000-0002-5949-1401 (J. S. Seo); 0000-0002-6012-4078 (C. H. Yoon); 0000-0003-2774-4846 (H. S. Park); 0000-0002-3123-7401 (J. U. Kim)



This paper is an open access article licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0>. This paper, color print of one or more figures in this paper, and/or supplementary information are available at <http://journal.kepcoco.kr>.

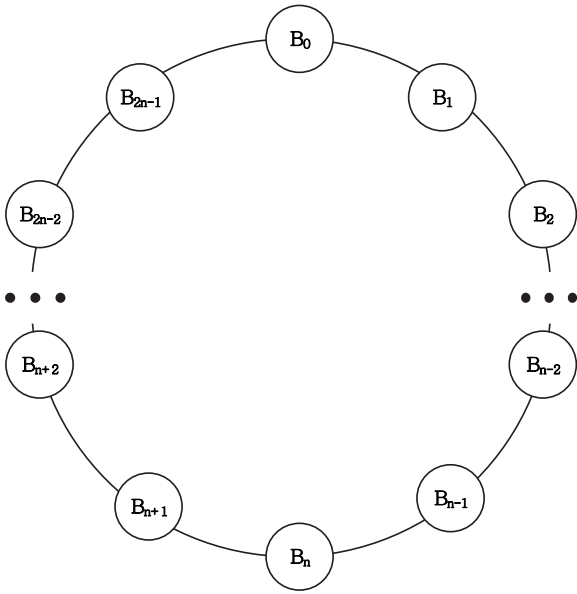


Fig. 1. 링 형태의 브리지 망.

하여 망 재구성 시간을 크게 단축하였다 [3]. 하지만 오래된 기술이므로 망 복구 성능이 좋지 않다고 알려져 있다. 그 이유는 STP가 메쉬 형태의 망을 전제로 루프 발생이 없는 트리 망으로 설정하기 위해 고안된 기술이기 때문이다.

하지만 망의 구조를 링으로 제한하는 경우, RSTP는 산업용 통신망에 적용될 수 있는 정도의 성능을 제공할 수 있다. 그 예로 eRSTP는 RSTP를 링 형태의 망에서 최적화함으로써 40개의 브리지로 구성된 링 망에서 망 복구 시간을 약 400 ms 이내로 보장한다 [11].

본 논문에서는 RSTP의 장애 절차의 문제점을 제시하고, 이를 해결하여 복구 시간을 단축할 수 있는 새로운 위치 기반의 장애 복구 절차를 제안하고 성능을 분석하였다.

기존 RSTP의 복구 시간은 장애 위치 및 종류에 따라 이웃 브리지 간에 제안-동의 절차가 반복될 수 있어 장애 복구 완료 시간이 증가한다. 그 이유는 대체 포트를 가지는 브리지에서 제안-동의 절차에 의해 직전에 갱신된 루트 브리지 정보가 링의 다른 포트로 수신된 새로운 제안-동의 절차를 통해 반복되면서 재 전파되기 때문이다. 그 뿐만 아니라 발생 장애 위치와 종류에 따라 복구 완료 시간이 변동될 수 있다.

이러한 문제점을 해결하기 위하여, 본 논문에서는 장애 발생 시 오직 한 번의 제안-동의 절차가 수행됨으로써 장애 복구 시간을 감소시킬 수 있는 새로운 위치 기반의 장애 복구 절차를 제안하고 성능을 분석하였다.

제안된 절차는 링 형태의 망에서 각 브리지가 루트 브리지에 대한 상대적인 위치, 즉 홉 수를 사전에 인지하도록 하고, 이를 기반으로 루트 브리지의 장애 또는 링크 장애 발생 시 이를 감지한 이웃 브리지에 의해 개시된 제안-동의 절차가 양 방향의 링에서 전파될 때 장애 감지 브리지를 기준으로 항상 링의 중간에 위치한 브리지에서 제안-동의 절차가 마감되도록 함으로써 신속한 망 복구 절차를 보장한다. 제안된 방식에 대한 수식 분석 및 모의실험을 통하여 제안된 절차가 기존 400 ms의 RSTP에 비하여 100 ms 이내로 복구 시간을 75% 단축할 수 있음을 검증하였다.

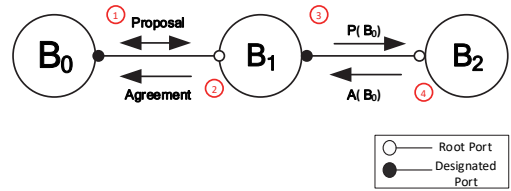


Fig. 2. 제안-동의 절차.

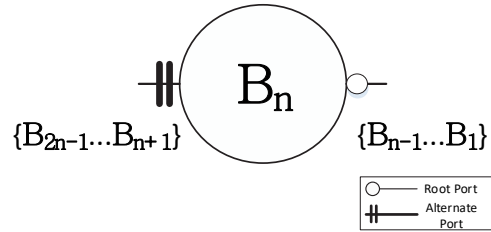


Fig. 3. 대체 포트를 갖는 브리지.

본 논문의 구성은 다음과 같다. 본 서론에 이어 II장과 III장에서는 기존 RSTP의 제안-동의 절차에 따른 장애 복구 과정과 문제점을 도출한다. IV장에서는 링 형태의 망에서 RSTP의 복구 절차를 간소화할 수 있는 새로운 위치 기반의 고속 장애 복구 절차를 제시하였다. V장에서는 제안된 절차에 대한 수치적인 분석과 시뮬레이션을 통하여 제안된 절차가 기존 RSTP에 비하여 최대 1/4배 정도로 복구 시간을 단축시킴을 검증하였다. 마지막으로 VI장에서 결론을 맺는다.

II. Rapid Spanning Tree Protocol (RSTP)

A. 개요

RSTP의 기본 기능은 MAC 주소에 대한 자동 학습을 통한 Filtering Database (FDB) 작성과 프레임 중계 및 필터링을 위해 다수의 브리지가 임의의 토폴로지로 연결된 브리지 망에서의 루프를 방지할 수 있는 스패닝 트리를 설정하는 것이다 [3]. 정상 동작 시 경로상의 브리지의 고장이나 링크의 단절 시 새로운 경로를 설정함으로써 장애 복구 능력을 제공한다.

기존 STP를 폐기하고 RSTP 기능으로 보완된 IEEE 802.1D-2004 [3] 규격은 기존 STP에 비해 장애에 대한 반응 속도를 향상시키고 브리지 망에 참여하지 않는 포트에 대해 엣지 포트로 설정하는 기능을 추가하였다. 또한 망 장애 복구를 위하여 대체 포트 (Alternate Port) 및 백업 포트(Back up port) 등도 추가되었으며 기존 STP에서 포트의 차단(Blocking), 청취(Listening), 비활성(Disabled) 상태는 RSTP에서 폐기(Discarding) 상태로 통합되었다. 가장 큰 개선점은 타이머에 의존하던 포트 천이 기능 대신에 이웃 브리지 간의 제안-동의(Proposal-Agreement) 절차를 통해 포트의 상태를 즉시 포워딩 상태로 전환 가능하도록 함으로써 망 재구성 시간을 크게 줄인 것이다.

본 절에서는 RSTP의 가장 핵심이 되는 제안-동의 절차를 링 형태의 망에 대하여 망 구성 과정, 링크 장애와 루트 브리지 장애 시 복구 과정을 분석한다.

Fig. 1은 본 논문에서 사용하는 링 형태의 망으로써 다음과

같이 정의된다.

- 1) B_n 은 n 번째 브리지이다.
- 2) 총 브리지의 개수 $N=2n$ 으로 짝수 개다.
- 3) 브리지 간 비용(path cost)은 동일하다.
- 4) B_0 가 가장 우선순위가 높은 루트 브리지이다.

B. 제안-동의(Proposal and Agreement) 절차

초기 망 구성 시, 각 브리지는 자신을 루트 브리지로 설정하고 루트 브리지 선출 과정에 참여한다. 이때 이웃 브리지에게 자신이 루트 브리지임을 제안하는 제안(Proposal)비트를 1로 설정한 BPDU를 송신한다. 이 과정은 Fig 2에 도시되어 있다. 여기서 $P(B_n)$ 와 $A(B_n)$ 는 각각 브리지 B_n 에 의한 제안(Proposal) 및 동의(Agreement) BPDU를 뜻한다.

- 1) B_0 와 B_1 : 각각 자신의 제안 BPDU를 전송한다.
- 2) B_1 : 우선순위가 B_0 보다 낮으므로 B_1 은 B_0 를 루트 브리지로 인정하고 수신 포트의 역할과 포트 상태를 각각 루트 포트(Root Port, RP)와 포워딩 상태로 변경하고 B_0 에게 동의(Agreement) BPDU로 응답한다.
 B_0 : 동의 BPDU를 수신하면, 해당 포트의 역할과 상태를 각각 지정 포트(Designated Port, DP)와 포워딩 상태로 변경한다.
- 3) 이후, B_2 가 망 구성 요소로 추가되면 B_1 과 B_2 또한 제안-동의 절차를 통해서 포트의 역할과 상태를 계산한다. 이때 B_1 은 B_0 를 루트 브리지로 인정한 상태이므로 제안 BPDU의 Root ID 영역에 자신의 루트 브리지인 B_0 의 정보를 수납하여 B_2 에게 전송한다.
- 4) B_2 와 B_1 은 각각 다음과 같이 동작한다.

B_2 : 자신의 우선순위가 B_0 보다 낮으므로 수신 포트의 역할과 상태를 각각 RP와 포워딩 상태로 천이시 키고 B_1 에게 동의 BPDU를 전송한다.

B_1 : 동의 BPDU를 수신하면 포트의 상태와 역할을 각각 DP 및 포워딩 상태로 천이한다. 만약 B_0 보다 B_2 의 우선순위가 더 높다면 B_1 은 B_2 를 루트 브리지로 인정하고 새로운 루트 브리지를 알리기 위한 갱신된 제안-동의 절차를 진행한다.

C. 스페닝 트리 설정 절차

브리지는 여러 포트에서 동일한 루트 브리지 정보가 수납된 BPDU의 수신 시 루프 방지를 위한 특정 포트의 역할을 하는 대체 포트(Alternate Port, AP)를 설정하고, 해당 포트의 상태를 블로킹(폐기) 상태로 설정하여, 프레임 중계에 참여하지 못하도록 함으로써 루프를 방지한다.

예를 들어 Fig. 1과 같은 링 망에 있는 B_n 은 이웃한 브리지인 B_{n-1} 과 B_{n+1} 모두로부터 B_0 가 루트 브리지임을 명시한 제안 BPDU를 수신한다. 이 경우, 두 포트의 B_0 로 향하는 비용이 동일하므로 추가로 B_{n-1} 과 B_{n+1} 의 우선순위를 비교한다. B_{n-1} 이 우선순위가 더 높다면 B_{n-1} 을 향하는 포트를 루트 포트로 설정하고 B_{n+1} 로 향하는 포트를 대체 포트로 설정한다. 이것은 Fig. 3과 같으며 B_{n+1} 이 우선순위가 더 높을 경우 전자의 경우와 반대로 포트를 설정한다. 여기서 bubble 기호는 해당 포트가 RP 역할임을 표시하고, 이중선은

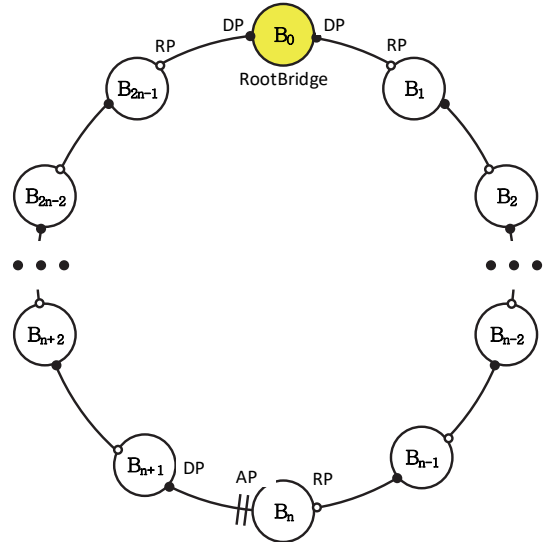


Fig. 4. 정상상태인 링 형태의 브리지 망.

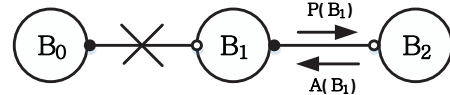


Fig. 5. 링크 장애 복구 과정 1.

AP 역할로써 블로킹 상태임을 표시한다.

Fig. 4는 앞서 설명한 절차에 따라 우선순위가 가장 높은 루트 브리지 B_0 와 B_0 로부터 가장 멀리 떨어져 있으며 루프 방지용 대체 포트를 갖는 브리지 B_n 으로 구성 완료된 링 형 브리지망의 예이다. 이후 링크 단절 또는 루트 브리지의 장애와 같은 망 장애 감지 시 루프방지 목적이었던 AP상태는 DP 또는 RP 등으로 천이할 수 있다.

III. RSTP 장애복구 절차와 문제점 분석

A. 링크 장애 시 복구 절차

RSTP에서는 링크 장애 시 대체 포트(AP)를 갖는 브리지가 AP를 RP 또는 DP로 역할을 전환하여 망을 복구한다. 이 과정에서 TC 비트가 1로 설정된 TC (Topology Change) BPDU의 전파를 통해 브리지의 장애가 발생한 지점을 향하는 포트의 FDB를 초기화한다.

Fig. 5는 Fig. 4의 B_0 와 B_1 간의 링크가 단절된 경우에 대한 것이다. B_1 은 B_0 로부터의 Hello BPDU를 3회 이상 수신하지 못하거나, 이더넷의 물리적인 링크 단절 시그널링에 의해 장애를 감지한다. 장애를 감지한 B_1 은 루트 브리지의 장애인지 링크 단절인지를 판단할 수 없으므로 자신을 루트 브리지로 설정하고 새로운 루트 브리지 선출 과정에 참여한다. B_1 과 B_2 는 다음 절차를 진행한다.

- 1) B_1 은 RP 포트에 장애가 발생했음을 감지한다.
- 2) B_1 은 모든 포트의 역할과 상태를 각각 지정 포트 (DP) 및 폐기(Discarding) 상태로 설정한다.

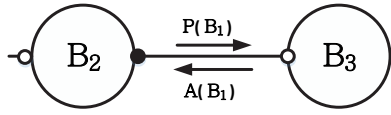


Fig. 6. 링크 장애 복구 과정 2.

- 3) B_1 은 자신을 루트 브리지로 판단하고 B_2 에게 제안 BPDU인 $P(B_1)$ 을 전송한다.
- 4) B_2 는 루트 포트(RP)로부터 새로운 루트 브리지의 정보가 수신되면 B_1 을 새로운 루트 브리지로 설정한다.
- 5) B_1 이 B_2 보다 우선순위가 높으므로 B_1 에게 동의 BPDU인 $A(B_2)$ 로 응답한다.

이어 새로운 루트 브리지 정보에 동의한 B_2 는 Fig. 6과같이 자신의 또 다른 이웃인 B_3 과 제안-동의 절차를 진행한다.

이러한 새로운 루트 브리지 선출을 위한 제안-동의 절차는 장애를 감지한 B_1 에 의해 개시되어 B_{n-1} 까지 진행된다. B_{n-1} 이 B_n 에게 전송하는 제안 BPDU를 수신한 B_n 은 다음과 같은 절차를 진행한다.

- 1) 루트 포트(RP)로부터 $P(B_1)$ 을 수신하지만 B_n 은 대체포트(AP)를 가지므로 루트 브리지 선출 과정에 참여하지 않고 기존 AP를 {루트 포트/포워딩} 상태로 전환한다.
- 2) 제안 BPDU를 수신한 포트를 {지정 포트/폐기} 상태로 전환하고 기존 루트 브리지 B_0 를 주장하는 $P(B_0)$ 를 전송한다. B_{n-1} 으로부터 $A(B_0)$ 를 수신하면 {지정 포트/포워딩} 상태로 전환한다.
- 3) 수신한 포트를 제외한 나머지 포트로 토폴로지변화를 알리는 TC BPDU를 전송한다.

B_{n-1} 은 B_n 으로부터 B_0 를 루트 브리지로 하는 제안 BPDU를 수신 후 다음과 같은 절차를 진행한다.

- 1) 새로운 루트 브리지인 B_1 보다 기존 루트 브리지 B_0 의 우선순위가 높으므로 B_n 을 향하는 포트를 {루트 포트/포워딩}상태로 전환하고 B_n 에게 $A(B_n)$ 을 전송한다.
- 2) B_{n-2} 으로 향하는 포트를 {지정/폐기}로 전환하고 $P(B_0)$ 를 송신한다.
- 3) B_{n-2} 으로부터 동의 BPDU를 받은 후 {지정/포워딩}상태로 전환한다.

B_0 를 루트 브리지로 주장하는 제안-동의 절차는 B_1 까지 연쇄적으로 이어지며 $\{B_{n-1} \dots B_1\}$ 에 속하는 브리지들은 포트의 상태와 역할을 재계산하고 B_0 를 루트 브리지로 인정한다. B_n 이 전송하는 TC BPDU를 수신한 B_{n+1} 은 다음과 같은 절차를 진행한다.

- 1) TC BPDU를 수신한 포트를 제외하고 모든 포트의 FDB를 초기화 시키고 학습(Learning)상태로 천이한다.
- 2) 수신 포트를 제외한 포트로 TC BPDU를 flooding 한다.

위와 같은 절차는 $\{B_{n+1} \dots B_0\}$ 의 장애 지점을 향하는 포트의 FDB를 초기화 시킨다.

결과적으로 $B_0: B_1$ 사이의 링크 장애 시 복구되는 망은 Fig. 7 과 같다. 앞에서 보인 과정들은 $B_0: B_{2n-1}$ 의 링크 장애 시에도 동일

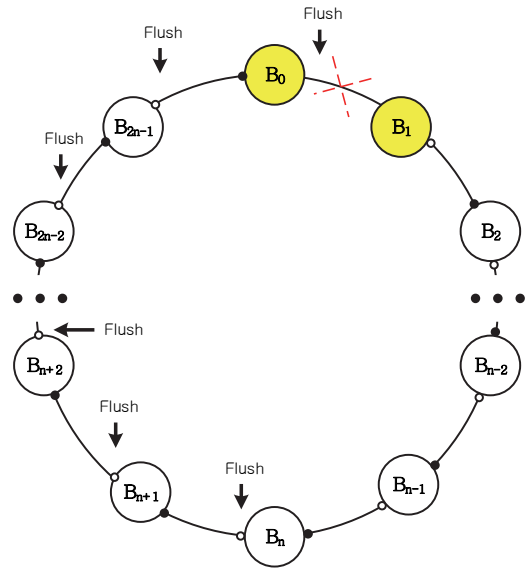


Fig. 7. 링크 장애 시 망 복구 결과.

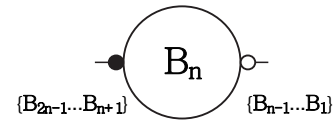


Fig. 8. 대체 포트에서 먼저 수신한 경우의 포트 전환.

하다. 이 경우 B_n 의 대체 포트는 지정 포트로 전환되고 루트 포트는 전환되지 않는다.

B. 루트 브리지 장애 시 복구 절차

루트 브리지 장애가 발생하는 경우 B_1 과 B_{2n-1} 의 장애 감지로 시작하여 새로운 루트 브리지 선출 과정이 $\{B_1 \dots B_{n-1}\}, \{B_{2n-1} \dots B_{n+1}\}$ 에서 이루어진다. B_n 은 대체 포트에서 $P(B_{2n-1})$ 을, 루트 포트 쪽에서는 $P(B_1)$ 을 수신한다.

B_0, B_1, B_{2n-1} 순으로 우선순위가 높을 때, B_n 의 BPDU 처리 순서에 따라 망 복구 시간이 링크 장애의 경우와 유사하게 또는 2배 가까이 차이 날 수 있다. Fig 8은 대체 포트 쪽에서 먼저 제안 BPDU를 수신한 경우의 전환된 포트의 상태이며, B_n 은 다음과 같은 절차를 진행한다.

- 1) Fig. 8과 같이 포트를 전환 시키고 링크 장애 의 경우와 같이 $P(B_0)$ 가 B_{2n-1} 까지 전파된다.
- 2) 루트 포트로 $P(B_1)$ 을 수신한다.
- 3) B_n 은 1)의 과정을 거치고 대체 포트가 없으므로 새로운 루트 브리지 선출 과정에 참여하며 B_1 이 루트 브리지임에 동의한다.
- 4) B_{n-1} 에게 $A(B_1)$ 을 전송하고 B_{n+1} 에게 $P(B_1)$ 을 전송한다.
- 5) $P(B_1)$ 은 B_{2n-1} 까지 전파되어 모든 브리지는 B_1 이 새로운 루트 브리지인 것에 동의한다.

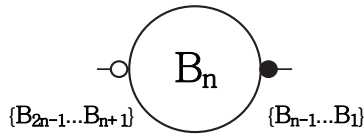


Fig. 9. 루트 포트에서 먼저 수신한 경우의 포트 역할 전환.

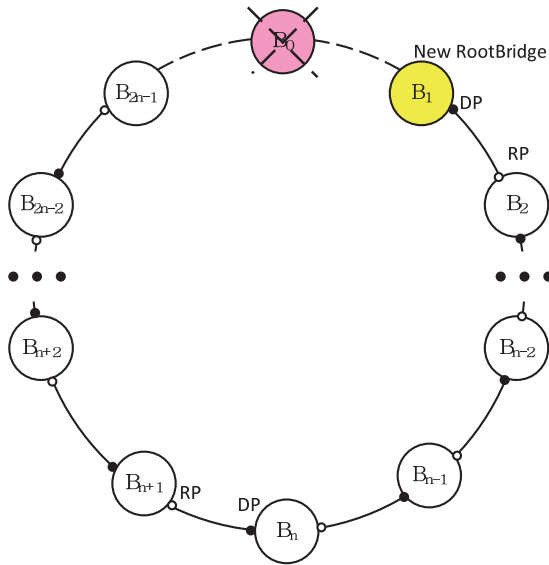


Fig. 10. 루트 브리지 장애의 망 복구 결과.

Fig. 9는 B_n 이 루트 포트 쪽에서 먼저 제안 BPDU를 수신할 경우의 전환된 포트 상태이며 다음과 같은 절차를 진행한다.

- 1) B_n 은 Fig. 9와 같이 포트를 전환시키고 B_{n-1} 에게 $P(B_0)$ 를 전송한다.
- 2) 새로운 루트 포트로부터 $P(B_{2n-1})$ 을 수신한다.
- 3) 대체 포트가 없기 때문에 자신의 정보보다 우선순위가 높은 $P(B_{2n-1})$ 에 동의하며 B_{n-1} 에게 $P(B_{2n-1})$ 을 전송한다.

B_n 을 제외한 나머지 브리지는 다음과 같은 절차를 진행한다.

- 4) $\{B_{n-1} \dots B_1\}$ 은 $P(B_0)$ 에 동의하고 B_n 을 향하여 루트 포트를 설정한다.
- 5) $\{B_{n-1} \dots B_2\}$ 는 루트 포트에서 3)에서 B_n 이 전송한 $P(B_{2n-1})$ 을 수신하여 $P(B_{2n-1})$ 에 동의한다. B_2 는 B_1 에게 $P(B_{2n-1})$ 을 전송한다.
- 6) B_1 은 B_{2n-1} 보다 우선순위가 높으므로 B_2 에게 거절의 의미인 $P(B_1)$ 을 전송한다. B_2 는 B_1 이 우선순위가 더 높으므로 $A(B_1)$ 을 B_1 에게 전송하고 B_3 에게 $P(B_1)$ 을 전송한다.
- 7) $P(B_1)$ 은 B_{2n-1} 까지 전파되어 모든 브리지는 B_1 이 루트 브리지인 것에 동의한다.

루트 브리지의 장애로 복구 과정에서 모든 브리지들의 FDB는 초기화된다. Fig 10은 포트 천이와 학습이 완료된 Active

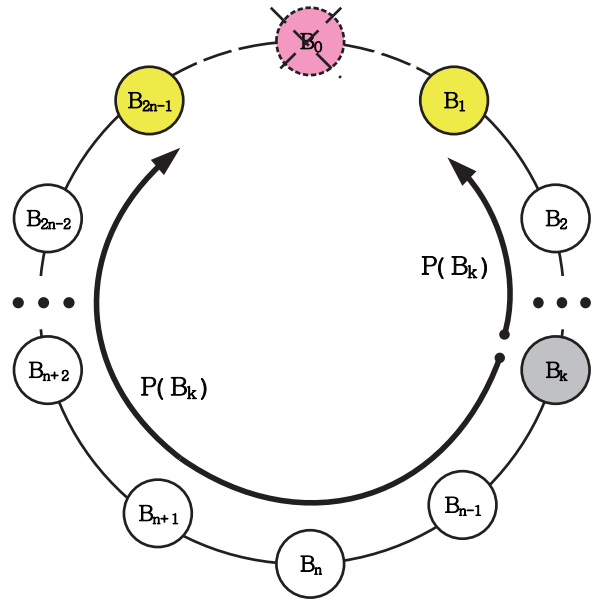


Fig. 11. B_k 에 의한 제안 반복 절차.

Topology 망이다.

C. 문제점과 해결방안

앞에서 기술된 것처럼 루트 브리지 장애 시 B_n 이 수행하는 BPDU의 처리 순서에 따라 복구 과정이 달라진다. 그 이유는 B_n 이 B_{2n-1} 를 루트 브리지로 주장하는 제안-동의 절차를 진행하면서 이 제안-동의 절차에 반복하는(B_1 을 루트 브리지로 주장하는) 절차가 다시 B_{2n-1} 까지 진행되기 때문이다.

Fig. 11은 다음 루트 브리지인 B_k 로부터 진행되는 $P(B_{2n-1})$ 에 대한 반복 절차의 진행 과정이며 다음과 같은 조건과 수식을 통해서 다음 루트 브리지의 위치에 따라 망 복구에 소요되는 제안-동의 절차의 개수를 분석할 수 있다.

차기 루트 브리지인 B_k 가 $1 \leq k \leq n$ 을 만족할 때, 필요한 제안-동의 절차 수(N_p)는 다음과 같다.

$$N_p = (4n - 2k - 3) \tag{1}$$

Eq. (1)에서 $k=1$ 일 때, 최대 $4n-5$ 의 제안-동의 절차가 필요하다. $k=n-1$ 의 경우 최대 $2n-1$ 의 제안-동의 절차가 필요하다. $k=n$ 일 때 $2n-3$ 의 제안-동의 절차가 필요하다.

수식을 통해 B_1 이 다음 루트 브리지일 때 최악의 경우가 발생하며 브리지 수가 많아질수록 복구에 필요한 절차 수는 $2N(4n)$ 에 가까워진다.

이러한 불필요한 절차 수를 최대한 감소시키는 방법 중 하나는 B_n 을 명시적으로 차기 루트 브리지로 사전에 지정하는 것이다. 이렇게 함으로써 B_n 이 루트 브리지 선출 과정에 참여할 때, B_n 이 가장 우선순위가 높기 때문에 기존과 같이 $P(B_n)$ 을 반복하는 절차가 진행되는 것을 방지할 수 있다.

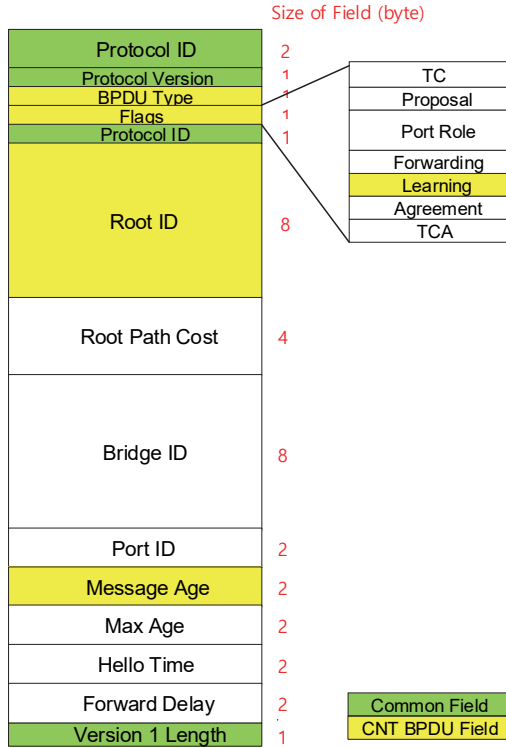


Fig. 12. CNT BPDU의 구성.

IV. Position Based RSTP

A. 개요

최악의 경우를 방지하기 위해 B_n 을 다음 루트 브리지 후보로 지정함으로써 불필요한 절차를 방지할 수 있지만 기존 RSTP의 방식으로는 링 형태의 망에서 최적화에 한계가 있다.

메쉬 형태에서 루프 방지를 위해 고안된 RSTP를 링 형태에서 복구 절차를 줄이기 위해서는 복구 과정을 최적화할 필요가 있다. B_n 을 차기 루트 브리지로 사전에 지정한다면 링 형태에서의 RSTP는 다음과 같은 특징을 갖는다.

- 1) 링 형태의 망을 구성하는 대부분의 브리지는 하나의 루트 포트와 하나의 지정 포트를 갖는다. 단 B_n 은 루트 포트와 AP를 갖는다.
- 2) 링크 장애의 경우: 루트 브리지와 경로가 단절된 브리지들의 새로운 루트 포트는 B_n 을 경유하여 B_0 에 연결된다.
- 3) 루트 브리지 장애의 경우: 브리지들의 루트 포트는 다음 루트 브리지인 B_n 을 향한다.

B_n 을 다음 루트 브리지로 지정하였을 때, 링크 장애와 루트 브리지 장애의 경우 재계산되는 브리지의 포트 천이 결과는 루트 포트와 지정 포트의 자리바꿈과 같다.

이 점을 활용하기 위하여 루트 브리지로 향하는 경로가 단절된 브리지들의 기존 루트 브리지의 생존에 관한 암묵적인 동의가 필요하다. 암묵적 동의는 새로운 루트 브리지 선출 과정에 참여하는 브리지들의 포트 재계산을 위한 제안-동의 절차를 장애 전파를 통한 포트의 자리바꿈으로 대체하는 것을 가능하도록 한다.

또한 링크 장애의 경우 대체 포트를 가지는 B_n 에 의해 진행

되는 기존 루트 브리지에 관한 제안-동의 절차를 생략할 수 있으며 루트 브리지 장애의 경우에는 새로운 루트 브리지인 B_n 에 관한 제안-동의 절차가 진행되기 전에 브리지들의 자리바꿈에 의한 즉각적인 포트 천이로 데이터 프레임의 경로를 보장할 수 있다.

이것이 가능 하려면 각 브리지는 브리지 망 내부의 브리지의 개수(N)와 루트 브리지에 대한 자신의 상대적 위치를 인지할 필요가 있다. 브리지들이 루트 브리지를 기준으로 하는 자신의 위치를 인지하고 있으면 장애 감지 시 장애 위치에 관한 정보를 제공할 수 있다.

이를 통하여 장애를 감지한 2개의 이웃 브리지에 의해 TC 메시지가 링의 양 방향으로 전파될 때 $N/2$ 홉에 해당하는 전파만 될 수 있도록 하고, 이 과정에서 각 브리지는 장애 위치에 대한 정보를 기반으로 신속한 포트 천이와 FDB 초기화를 수행할 수 있다. 링 형태의 망에서 장애 복구 과정을 간소화하기 위한 요구 사항은 다음과 같다.

- 1) 대체 포트(AP)를 갖는 브리지는 차기 루트 브리지로 선정되도록 함.
- 2) 각 브리지는 상대적 위치를 인지하고 망을 구성하는 브리지의 개수를 전파하기 위한 CNT (COUNT) BPDU 사용
- 3) 장애를 감지한 2개의 브리지로부터 양방향으로 전파되어 FDB 초기화와 포트 천이를 가능케 하는 TCR (Topology Change Request) BPDU 사용

AP를 갖는 브리지를 다음 루트 브리지로 선정하기 위한 방법은 간단하다. 대체 포트를 가지는 브리지 B_n 이 루트 브리지의 장애를 감지할 때 B_n 의 우선순위를 가장 높게 설정하면 된다. CNT BPDU와 TCR BPDU의 정의는 다음 절에서 다룬다.

B. CNT (Count) BPDU 정의

본 절에서는 브리지의 수(N)을 파악하고 전파하기 위한 브로드캐스팅 BPDU인 CNT BPDU를 제안한다. CNT BPDU는 Type 값이 0x03으로 기존 RSTP BPDU의 Message Age 필드, Flags 필드의 Learning bit, Root ID 필드를 사용한다. 루트 브리지는 망 초기 설정 시 또는 주기적으로 해당 메시지를 시계방향으로 송신한다. 초기 탐색 완료 후 루트 브리지는 링 상의 총 브리지의 개수 N 을 시계방향으로 전파하는 절차를 이어 수행한다. 이를 통하여 나머지 브리지는 루트 브리지를 기준으로 자신의 위치와 브리지의 개수를 알 수 있다.

Fig. 12에서 Common Field는 CNT BPDU가 기존 RSTP BPDU와 공통으로 사용하는 필드이며 CNT BPDU에서 각 영역의 기능은 다음과 같다.

- 1) Message Age: CNT BPDU가 거친 홉 수 또는 전체 브리지 수
- 2) Learning: 브리지 수 조사 및 브리지 수 전파 기능 선택
- 3) Root ID: 전파하는 루트 브리지 ID

CNT BPDU의 전파 과정은 다음과 같다.

- 1) 루트 브리지는 BPDU Type과 ROOT ID를 설정한다. Learning bit를 1로 설정하고 망 구성 요소로 자신을 포함하는 의미의 1을 Message Age에 저장하여 지정 포트 중 우선순위가 가장 높은 포트로 CNT BPDU를 송신한다.

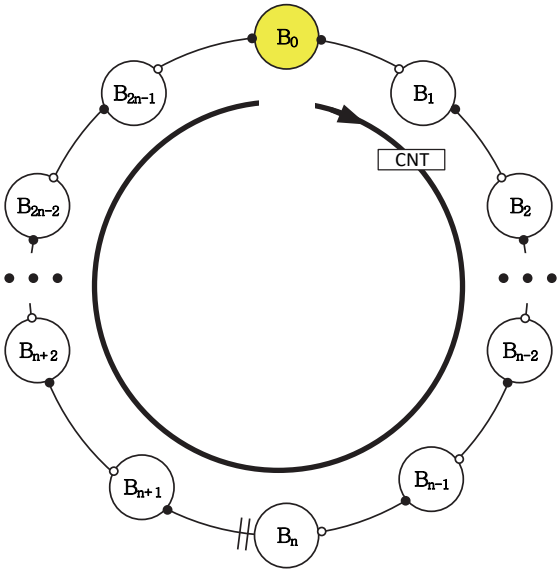


Fig. 13. 시계방향으로 순서대로 정해지는 브리지들의 위치.

- 2) CNT BPDU를 수신한 이웃 브리지는 BPDU의 Message Age 필드 값을 루트 브리지를 기준으로 하는 자신의 위치로 판단하고 Age에 1을 더하여 DP(또는 AP 포트)로 증계한다.
- 3) 링을 순환하여 루트 브리지가 이를 다른 포트에 수신하면 Root ID가 같음으로 모든 브리지를 거친 BPDU라고 판단하여 조사를 완료한다.
- 4) 수집된 브리지의 개수 정보를 전파하기 위해 1)번 절차에서 Learning bit를 0으로 설정하고, Message Age 영역에 브리지 수 N 을 저장하여 전송한다.
- 5) BPDU를 수신한 브리지는 해당 정보를 기록한 후 DP를 통해 BPDU를 다음 브리지에 증계한다.

위 과정을 통해 각 브리지는 Fig. 13과 같이 루트 브리지를 기준으로 시계방향으로 형성되는 자신의 위치를 알 수 있다.

C. TCR (Topology Change Request) BPDU

각 브리지는 이웃한 브리지의 장애를 감지할 수 있다. 즉 물리적인 방법을 통하여 장애를 감지하거나 일정 횟수 이상의 Hello BPDU를 수신하지 못할 경우 해당 브리지 또는 링크 장애로 판단한다.

제안하는 TCR (Topology Change Request) BPDU는 장애를 감지한 브리지가 전송하는 장애 감지 알림 BPDU이다. BPDU type이 0x04인 BPDU로 ROOT ID와 Message Age 필드와 Max Age 필드를 사용한다. Fig. 14에서 Common Field는 TCR BPDU가 기존 RSTP BPDU와 공통으로 사용하는 필드이며 TCR BPDU에서 각 영역의 기능은 다음과 같다.

- 1) ROOT ID: CNT BPDU를 통해 인지한 자신의 위치 정보로 장애를 감지한 브리지를 의미한다.
- 2) Message Age: TCR BPDU가 거친 홉 수를 의미한다
- 3) Max Age: TCR BPDU가 전파될 수 있는 최대 홉 수를 의

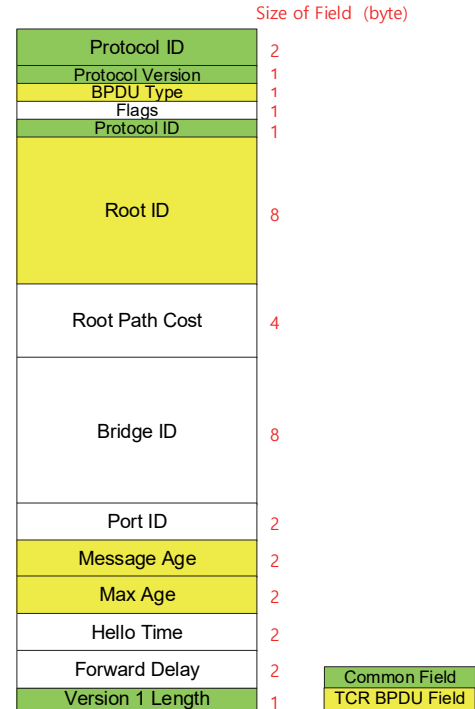


Fig. 14. TCR BPDU의 구성.

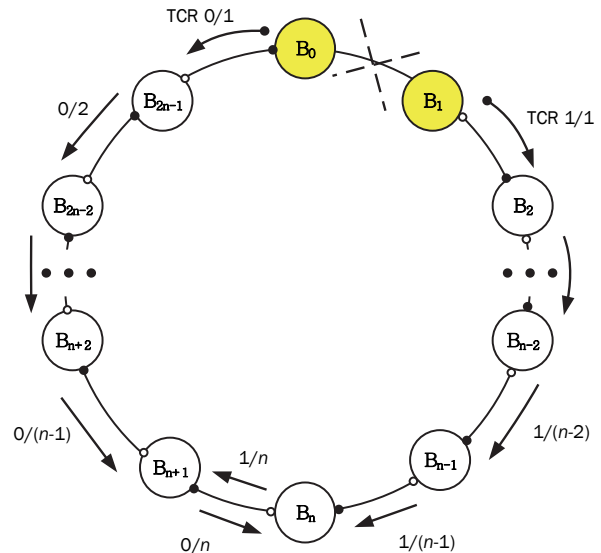


Fig. 15. 링크 장애 시 TCR BPDU전파.

미하며 N 이 짝수일 때 $N/2$, 홀수일 때 $(N-1)/2$ 이다.

D. 장애발생시의 동작절차

Fig. 15는 브리지 B_0 와 B_1 사이에 링크 장애의 경우이다. B_0 와 B_1 은 장애를 감지하고 장애 감지 포트를 제외한 포트로 TCR BPDU를 전송한다. 그림에서 K/hop 은 TCR BPDU의 내용을 축약한 것으로 K 는 장애를 감지한 브리지의 위치이고 hop 은 거친 홉 수를 의미한다.

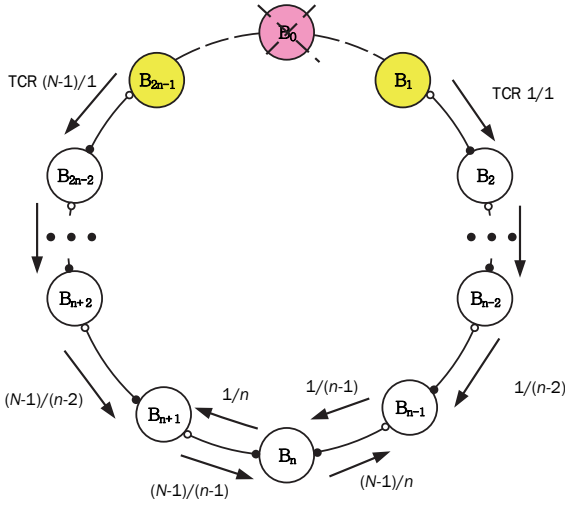


Fig. 16. 루트 브리지 장애 시 TCR BPDU 전파.

B_1 과 B_0 가 전송한 TCR BPDU는 $N/2$ 홉만 전파되어 B_n 과 B_{n+1} 까지 전파된다. B_n 은 양방향에서 TCR BPDU를 수신하고 기존 RSTP와 동일한 방법으로 루트 브리지가 될 것인지 대체 포트를 전환시킬 것인지를 결정한다. 이 경우는 링크 장애로 판단하여 포트를 전환한다.

Fig. 16은 루트 브리지에 장애가 발생하여 B_{2n-1} 과 B_1 이 장애를 감지한 경우이다. 루트 브리지 장애 발생 시 B_n 은 양 방향에서 TCR BPDU를 수신하고 자신을 루트 브리지로 전환하여 양 방향으로 $P(B_n)$ 을 전송한다. 이때 이미 모든 브리지의 모든 포트는 포워딩 되어있는 상태이다.

TABLE 1의 Pseudo 코드는 브리지가 TCR BPDU를 수신하였을 때 진행되는 포트 천이 과정을 보여준다. 수신 브리지의 위치 정보와 포트의 역할, 장애 감지 브리지의 위치정보를 기반으로 포트 천이를 결정하거나 자신을 루트 브리지로 주장한다.

V. 성능 분석

RSTP에서는 제안-동의 절차를 통해 포트의 상태를 즉시 포워딩으로 전환한다. 이러한 제안-동의 절차는 초기 망 구성 뿐 아니라 망 장애 시 포트 재계산을 위해 브리지 사이에 연쇄적으로 진행된다.

RSTP의 복구 시간을 결정하는 가장 큰 요소는 브리지 간 주고받는 절차 수와 처리 시간이며 다양한 프로토콜을 사용하는 스위치나 브리지는 내부 사정에 따라 처리 과정이 상이한 문제점이 발생한다.

본 장에서는 기존 RSTP와 제안하는 위치 기반의 RSTP 절차의 망 복구 완료 시까지의 소요 절차 수에 따른 복구 시간을 비교한다.

Fig. 17은 링크 장애의 경우의 절차를 간략하게 나타낸 것이다. 여기서 선로의 전파지연시간은 무시하고 망 복구에 따른 제안-동의 절차에 필요한 처리시간(T_p)만 고려한다.

여기서 기존 RSTP의 제안-동의 절차는 브리지의 구현 방식에 따라 절차 A와 절차 B로 나눌 수 있다. 절차 A는 처리 간격마다 제안 또는 동의용 1개의 BPDU만 전송 가능한 반면, 절차 B에

TABLE 1
수도(Pseudo) 코드

```

·B.N: Bridge Number 수신 브리지의 위치
·Arrival Port: 수신 포트 종류
·Alt port: 대체 포트
·DESIG: 지정 포트
·Root: 루트 포트
·BPDU.N: 장애 감지 브리지의 위치
·No Change: 포트 전환 없음

If( B.N == N/2 )
{
  If( Bridge has Alt port )
  {
    If( Arrival Port == Root )
    {
      If( BPDU.N == 0 )
      {
        Altport -> DESIG
      }
      Else
      {
        Altport -> Root
        Old root -> DESIG
      }
    }
    Else If( Arrival Port == Alt port )
    {
      If( BPDU.N == 0 )
      {
        Altport -> Root
        Old root -> DESIG
      }
      Else If( BPDU.N > N/2 )
      {
        Altport -> DESIG
      }
    }
  }
  Else If( Bridge has no Alt port )
  {
    If( Arrival Port == Root )
    {
      If( BPDU.N == 0 )
      {
        No change
      }
      Else
      {
        I'm Root -> All port DESIG
      }
    }
    Else If( Arrival Port == DESIG )
    {
      No change
    }
  }
}
Else If( B.N != N/2 )
{
  If( B.N < N/2 )
  {
    If( Arrival Port == Root )
    {
      If( BPDU.N == 0 or >N/2 )
      {
        No change
      }
      Else If( BPDU.N < N/2 )
      {
        Root -> DESIG
        Old DESIG -> root
      }
    }
    Else If( Arrival Port == DESIG )
    {
      No change
    }
  }
  Else If( B.N > N/2 )
  {
    If( Arrival Port == Root )
    {
      If( BPDU.N == 0 or <N/2 )
      {
        No change
      }
      Else if( BPDU.N > N/2 )
      {
        Root -> DESIG
        Old DESIG -> Root
      }
    }
    Else if( Arrival Port == DESIG )
    {
      No change
    }
  }
}
}
}

```

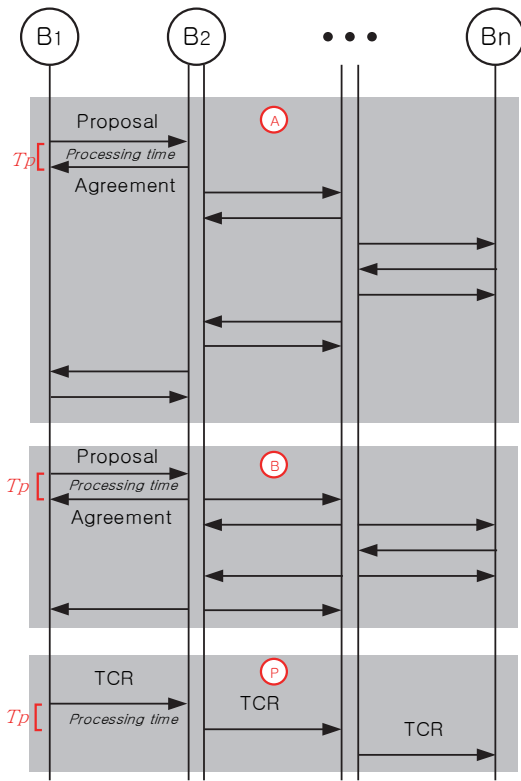



Fig. 17. 3 종류의 BPDU 처리절차.

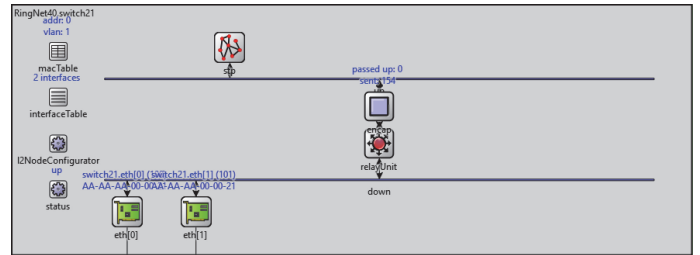


Fig. 18. 브리지 모듈의 구성.

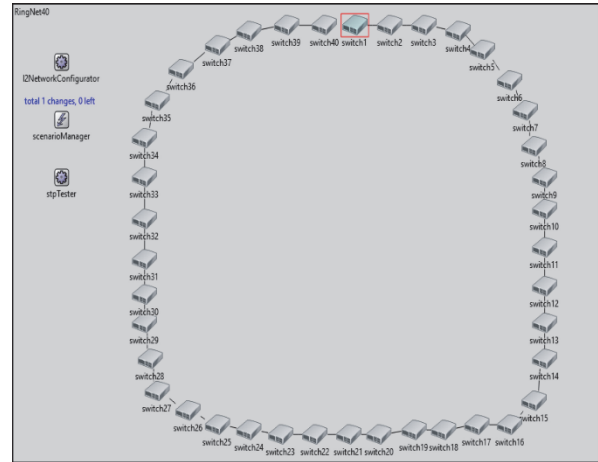


Fig. 19. 링크 장애 시물레이션의 경우.

TABLE 2
N이 짝수인 경우

링크 장애 시 복구 시간	루트 브리지 장애 시 복구 시간
A	$(2N-6) \times T_p$
B	$(N-1) \times T_p$
P	$(N/2-1) \times T_p$

TABLE 3
N이 홀수인 경우

링크 장애 시 복구 시간	루트 브리지 장애 시 복구 시간
A	$(2N-4) \times T_p$
B	$N \times T_p$
P	$(N-1)/2 \times T_p$

서는 제안 BPDU 수신 시 동의 BPDU로 응답과 동시에 다음 브리지에 제안 BPDU의 중계가 가능하다. 그리고 절차 P는 제안하는 위치 기반의 복구 절차이다.

링크와 루트 브리지 장애 시 수치적인 분석의 조건은 다음과 같다.

- 1) $N=2n$ 이다.
- 2) 물리적인 전파 지연 및 장애 감지 시간은 무시한다. 장애 감지 시간은 통신속도와 케이블의 종류에 따라 다르지만 대부분 microsec 단위로 무시 가능하다. [12]
- 3) T_p 는 브리지 내부에서의 BPDU 처리시간으로 [11]에서와 같은 2.5 ms이다
- 4) 수치분석은 최악의 경우(링크 장애의 경우 $B_0: B_1$ 또는

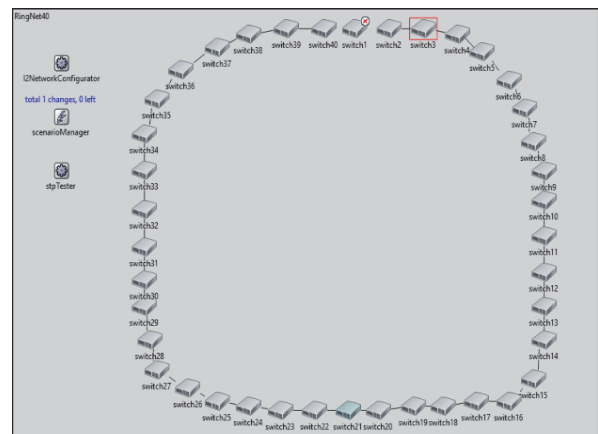


Fig. 20. 루트 브리지 장애 시물레이션의 경우.

$B_0: B_{2n-1}$, 루트 브리지 장애의 경우)를 가정한다.

- 5) 포트 천이와 새로운 루트 브리지에 동의하는 절차까지 포함하여 복구 시간을 측정한다.

TABLE 2와 TABLE 3은 각각 N이 짝수와 홀수인 경우, 링크 장애 및 루트 브리지 장애 시에 대한 최악의 경우의 장애 복구 시간을 비교한 것이다. A와 B는 기존 RSTP의 구현 방법에 따른 상이한 처리 과정이며 P는 제안하는 RSTP이다.

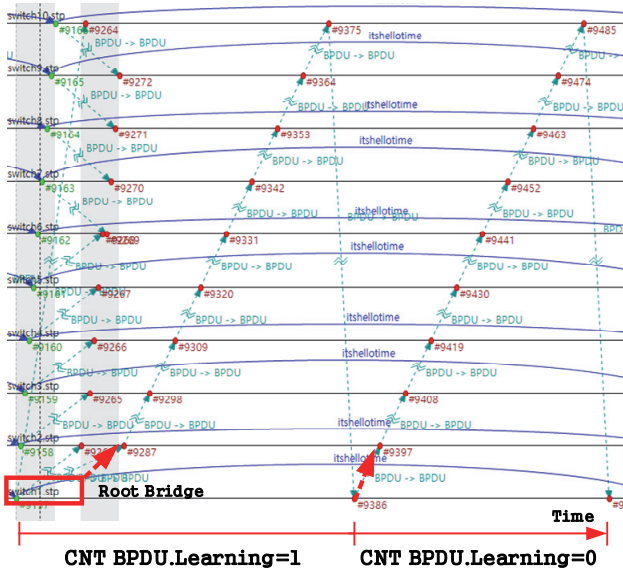


Fig. 21. CNT BPDUs의 전파과정.

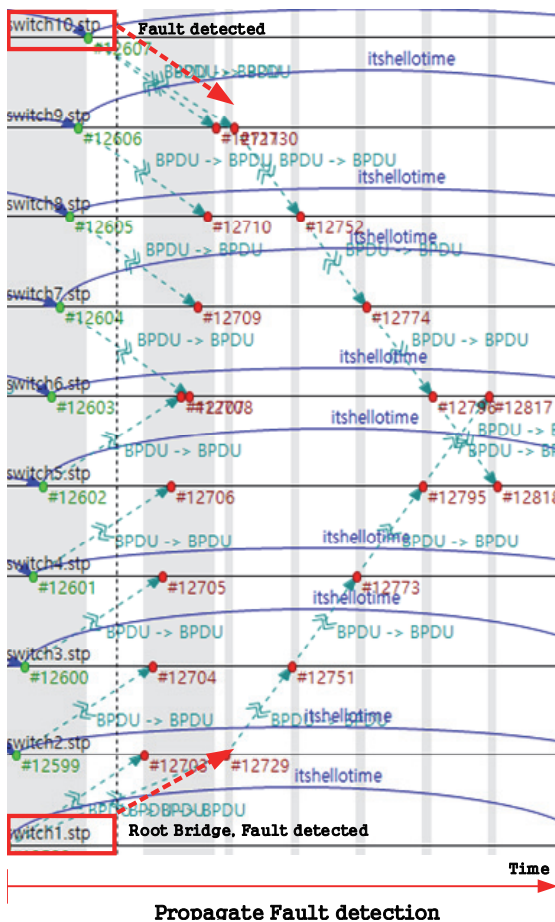


Fig. 22. B₁-B₁₀ 링크 장애 시 TCR BPDUs의 전파과정.

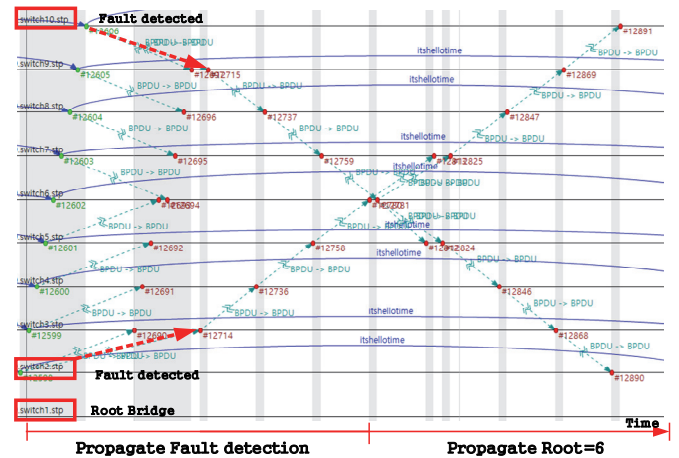
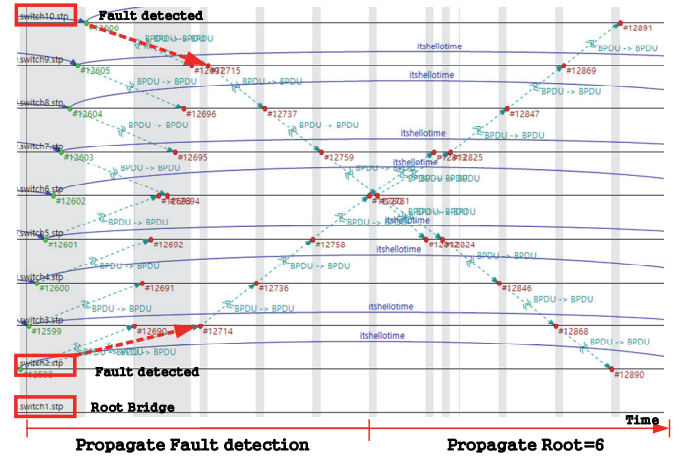


Fig. 23. 루트 브리지 장애 시 TCR BPDUs의 전파과정.

TABLE 4
복구 시간 최대치 비교 (N=브리지 개수)

N	기존 절차 A		기존 절차 B		제안된 절차	
	링크 장애	브리지 장애	링크 장애	브리지 장애	링크 장애	브리지 장애
10	35	75	25	42.5	10	20
15	65	130	35	70	17.5	35
20	85	175	50	92.5	22.5	45
25	115	230	60	120	30	60
30	135	275	75	142.5	35	70
35	165	330	85	170	42.5	85
40	185	375	100	192.5	47.5	95

Fig. 18부터 Fig. 20은 Omnet++ 시뮬레이터에서 제안하는 위치 기반의 복구 절차가 적용된 40개의 브리지로 링 형태의 망을 구성하였을 때의 시뮬레이션 모델이다.

링크 장애의 경우 B₁이 루트 브리지를 유지하고 루트 브리지 장애의 경우 B₂이 새로운 루트 브리지로 선출될 것을 확인할 수 있다. 또한 복구 시간이 TABLE 2의 수식 계산 값과 일치함을 확인하였다.

Fig. 21부터 Fig 23은 10개의 브리지로 망을 구성하였을 때 TCR, CNT BPDUs의 전파 과정을 도시한 예이다.

TABLE 4는 TABLE 2와 TABLE 3의 수식을 적용하여 10 ~ 40개의 노드에 대하여 B₀와 B₁간의 링크 장애와 루트 브리지 장애

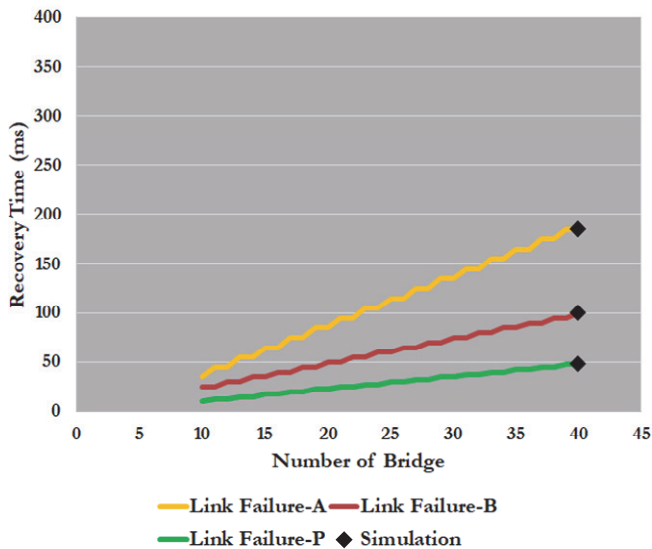


Fig. 24. 링크 장애 시 망 복구 시간의 비교.

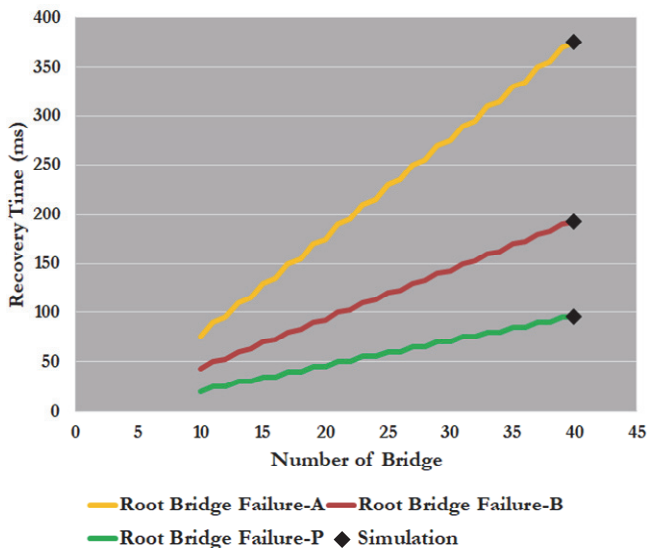


Fig. 25. 루트 브리지 장애 시 망 복구 시간의 비교.

시 복구 시간을 기존 RSTP의 2가지 방식 및 제안된 방식을 상호 비교한 것이다. 이 결과에서 알 수 있듯이 제안된 방식의 최악 복구 시간은 기존 절차 A에 비해 약 1/4로 단축되어 약 75%의 성능 개선이 가능하였다.

Fig. 24와 Fig. 25는 위의 결과를 브리지 수에 대하여 각각 링크 장애 및 루트 브리지 장애 시 소요되는 망 복구 시간을 비교한 것이다. 여기서 시뮬레이션 결과는 40개의 브리지에 대한 것으로서 수식적인 분석 결과와 일치하였다.

VI. Conclusion

본 논문에서는 장애 복구 절차를 간소화하여 링 형태의 망에서 RSTP를 최적화한 Position based RSTP를 제안하였다. 루트 브

리지를 기준으로 하는 브리지들의 위치정보를 활용하기 위해 2가지 새로운 BPDU로 CNT BPDU와 TCR BPDU를 정의하였으며 결과적으로 포트 천이, FDB 초기화를 한 번에 수행함으로써 루트 브리지 장애 시 기존 RSTP 복구 시간인 375 ms의 1/4인 95 ms, 링크 장애 시 기존의 185 ms (또는 100 ms)에 비하여 짧은 47.5 ms의 복구 시간이 소요되어 보다 신속한 망 복구가 가능함을 검증하였다. 물론 브리지 내부에서의 BPDU 처리시간이 단축되면 보다 고속으로 망의 복구가 가능할 것이다.

본 논문에서 제안된 Position based RSTP는 MOXA의 Turbo-ring, Siemens의 eRSTP처럼 장애 복구 능력을 요구하는 산업용 링 통신망에 활용될 수 있으며, 기존 RSTP에 대한 부정적인 인식을 바꿀 수 있을 것이다 [5][11].

Acknowledgment

This work was supported by Korea Electric Power Corporation under Grant R18XA06-10.

본 연구는 한국전력공사의 2019년 선정 기초연구개발 과제 연구비에 의해 지원되었음(과제번호: R18XA06-10)

References

- [1] ICT 중점기술 표준화 전략 맵 Ver. 2011, 한국정보통신 기술협회, Available at <https://www.tta.or.kr/> Accessed on Apr. 2020
- [2] IEC 62439-3, Industrial communication networks - High availability automation networks - Part 3: Parallel Redundancy Protocol (PRP) and High-availability Seamless Redundancy (HSR), DOI: 10.31030/2864600.
- [3] IEEE Standard for Local and metropolitan area networks: Media Access Control (MAC) Bridges, IEEE 802.1D-2004, DOI: 10.1109/IEEESTD.2004.94569.
- [4] IEC 62439-2, Industrial communication networks - High availability automation networks - Part 2: Media Redundancy Protocol (MRP), DOI: 10.31030/2833723
- [5] Vincent Liu, Redundancy in Automation, (MOXA White Paper), 2007, available at <https://www.moxa.com/>, Accessed on Mar. 2020
- [6] IEC 62439-7, Industrial communication networks - High availability automation networks - Part 7: Ring-based Redundancy Protocol (RRP) DOI: 10.31030/1895563.
- [7] Abishek Gopalan, Srinivasan Ramasubramanian, "Fast Recovery From Link Failures in Ethernet Networks," Reliability IEEE Transactions on, vol. 63, no. 2, pp. 412-426, 2014., DOI: 10.1109/TR.2014.2315957.
- [8] Mifdaoui, A., Amari, A., "Real-Time Ethernet Solutions supporting Ring topology from an Avionics Perspective: A Short Survey," Proceedings of the 22nd IEEE International Conference on Emerging Technologies and Factory Automation (IEEE ETFA'2017), Limassol, Cyprus, 12-15 Sep. 2017, DOI: 10.1109/ETFA.2017.8247608/
- [9] "Avionics Full-Duplex Switched Ethernet (AFDX) Network, ARINC Specification 664, Part 7, Aeronautical Radio."
- [10] Amari, A., Mifdaoui, A., Frances, F., Lacan, J., Rambaud, D., Urbain, L. "AeroRing: Avionics Full Duplex Ethernet Ring with High Availability and QoS Management," Proceedings of the European Congress on Embedded Real Time Software and Systems 2016, Toulouse, France, 27-29 Jan. 2016.
- [11] Michael Pustynnik, Mira Zafirovic-Vukotic, Roger Moore, "Performance of the Rapid Spanning Tree Protocol in Ring Network Topology 2007," Available at <http://siemens.com/Ruggedcom>, Accessed on Mar. 2020.
- [12] IEEE Standard for Ethernet, IEEE 802.3-2018, DOI: 10.1109/IEEESTD.2018.8457469.