

논문 2010-47TC-1-8

프로바이더 백본 브리지 망을 위한 백업링크 스페닝트리 기반 링크장애 복구기능과 구현

(A Link Protection Scheme with a Backup Link Spanning Tree for
Provider Backbone Bridged Networks and Implementation)

남 위 정*, 이 현 주*, 윤 중 호**, 홍 원 태***, 문 정 훈***

(Wie-Jung Nam, Hyun-Joo Lee, Chong-Ho Yoon, Won-Taek Hong, and Jeong-Hoon Moon)

요 약

본 논문에서는 브리지 기반의 백본망을 위한 프로바이더 백본 브리지(Provider Backbone Bridge) 망에서 각 브리지간 링크 장애가 발생한 경우 이를 효율적으로 보호할 수 있는 방법인 백업링크에 대한 스페닝트리 기반의 링크장애 복구기능을 제안하고 성능을 분석하였다. 제안된 방식은 PBB with Traffic Engineering (PBB-TE) 기반의 브리지망에서 대역예약된 마스터링크의 장애시 해당 링크의 트래픽을 루트브리지가 사전에 설정한 스페닝트리 기반의 백업링크로 우회시킴으로써 해당 트래픽의 QoS를 만족시킨다. 또한 백업링크 상으로 전달되는 프레임들은 스페닝트리 상에서 플러딩되도록 함으로써 주소학습과정에 따른 지연을 최소화시키고 루프백현상도 제거할 수 있는 장점이 있다. 제안된 방식의 동작 검증을 위하여 CCM메시지에 의한 링크보호절체기능과 MAC-in-MAC 수납방법이 모두 지원되는 PBB-TE 시스템을 리눅스 브리지 기반에서 구현하였고 관련 프로토콜 분석도구도 함께 구현하였다. 제안된 방식은 기존의 링크보호방법의 백업링크 예약에 의한 큰 대역손실이나 링크장애시 백업링크를 설정하는데 소요되는 신호절차지연시간이 큰 단점을 해결할 수 있어 QoS를 지원해야 하는 대규모 사업자용 백본 브리지 시스템 구현시 활용될 수 있다.

Abstract

In this paper, we propose an efficient link protection switching scheme for provider backbone bridge systems with a spanning tree for backup links exclusively, and evaluate its performance. The proposed scheme offers guaranteed QoS flows even when a link fault occurs in the primary link by flooding the flows over the profiled spanning tree. The flooding mechanism over the spanning tree can also provide low latency and remove the loopback flows. We also derive the efficiency of bandwidth usage for the normal flows and the number of lost frames during the link restoration. For evaluating its feasibility, we implement a prototype of PBB-TE systems based on the Linux bridge codes, which can support both link protection switching capability with CCM and MAC-in-MAC encapsulation. A related protocol analyzer is also developed. One can see that the proposed scheme and the prototype can be useful for developing carrier class Ethernet systems based on PBB-TE.

Keywords : Provider Backbone Bridge, Traffic Engineering, Link Protection, Connection Fault Management

I. 서 론

최근 정보통신 기술의 발전과 더불어 음성과 데이터

중심의 통신망은 방송서비스와의 융합을 시도하면서 IPTV(Internet Protocol Television), VoIP(Voice over Internet Protocol)등의 서비스가 이미 시작되었다. 이러한 멀티미디어 실시간 트래픽은 연속적인 데이터 스트림을 단방향 전송하며 지연 및 지터 성능에 따른 서비스 품질(Quality of Service, QoS)에 민감한 특징을 가진다. 이러한 실시간 트래픽에 대한 QoS를 망에서 제공하기 위해서는 수락제어 및 대역예약 절차뿐만 아니라

* 학생회원, ** 평생회원, 한국항공대학교 정보통신공학과 (Korea Aerospace University)

*** 정회원, 한국과학기술정보연구원 슈퍼컴퓨팅본부 (Korea Institute of Science and Technology Information)

접수일자: 2008년8월27일, 수정완료일: 2010년1월18일

라 망 장애 시 복구기능까지 필요로 한다.

최근 이러한 실시간 트래픽의 경제적인 백본 전달 망에서의 전달기능을 제공하기 위하여 기존 IP기반의 라우터 대신에 이더넷 브리지기반의 중계방식을 사용하는 캐리어 이더넷 기술이 등장하고 있다^[1~2]. 캐리어 이더넷 시스템은 확장성뿐만 아니라 기존 사무실 환경의 이더넷과 달리 종단간의 Ethernet virtual connection (EVC)별 대역할당절차와 고도의 신뢰성을 유지하기 위해 OAM절차를 따르는 링크보호 기능 및 장애관리 등을 제공해야 한다.

이러한 기능은 IEEE802.1의 Provider Bridge(PB), Provider Backbone Bridge(PBB), PBB-TE(Traffic Engineering)등의 표준화 기술에 의한 Q-in-Q 및 MAC-in-MAC, 그리고 링크의 연결성을 주기적으로 감시하고 관리하는 Connectivity Fault Management (CFM) 기술에 의하여 구체화되고 있으며, 이를 기반으로 End-to-End 이더넷을 궁극적인 목표로 하고 있다^[4~5].

캐리어 이더넷 망에서 서비스 기술 또는 서비스 장비 선택에 있어서 고가용성(high availability)은 매우 중요한 기준이 된다^[6]. 고가용성이란 서비스 장애 없이 얼마나 오랫동안 서비스를 안정적으로 이용할 수 있는가를 의미한다. 이를 위해서는 장애빈도가 매우 낮아야 하며 장애가 발생한 경우라도 그 복구시간이 매우 짧아야 한다.

높은 수준의 가용성을 제공하기 위하여 인터페이스 카드 및 사용 링크의 수 등의 redundancy를 제공함으로써 장애 발생율을 감소시키는 방안과 장애발생시 빠른 복구(restoration)를 위해 automatic protection switching(APS)방안을 고려할 수 있다.

이때 충분히 빠른 복구란 일반적으로 SONET/SDH가 제공하는 50ms 이하의 서비스복구지연시간 내이다. 이를 위한 기술 중 하나가 RPR-TE 기술에서 사용하는 CFM기술이다^[3~4].

ITU-T에서도 IEEE 802.3에서 정의한 표준 MAC 프로토콜을 그대로 이용하는 Ethernet protection switching(EPoS)기술에 대한 표준화 작업을 수행하여 G.8031/Y.1342를 작성하고 있다. ITU-T의 EPoS기술은 이더넷 전송망에서 점 대 점 VLAN 기반의 EVC에 대하여 경로와 대역이 예약된 백업링크(backup link) 및 백업 브리지를 사용하여 신속한 보호절체 기능을 제공하는 기술이다. 이것은 송신측이 동일한 프레임용 주

링크(primary link) 및 백업링크로 동시에 송신하고, 장애발생시 수신브리지가 프레임을 선택하는 방법으로 장애를 복구한다^[6].

이러한 보호절체방법들은 신속한 링크장애복구를 위하여 모든 브리지 간에 미리 백업링크를 예약해 두는 방법이므로 그만큼 일반 데이터 프레임 전달용 대역은 그만큼 감소하게 되는 문제가 발생한다. 또한 링크 장애 시 새로운 대역할당절차를 개시하는 것은 그 신호절체의 지연 때문에 바람직하지 않다.

본 논문은 이러한 문제를 해결하기 위하여 PBB-TE 기반의 캐리어이더넷 망에서의 링크장애를 대비한 고장 복구기능을 적용할 경우 QoS 트래픽을 전송하는 주 링크와 동일한 대역이 필요한 백업링크를 효율적으로 할당하여 일반 데이터 프레임용 대역을 많이 할당할 수 있으면서 동시에 짧은 보호절체지연이 보장되는 새로운 링크장애복구기능을 제안하고 이를 활용한 PBB-TE시스템의 구현에 대한 것이다.

제안된 방법은 루트브리지에서 모든 노드로 연결된 스페닝트리 형태의 백업링크를 미리 설정하고, 링크장애 시 QoS가 제공되어야 할 프레임에 대해서는 별도의 시그널링 절차 없이 트리상의 백업링크에 프레임을 전달하여 이들이 목적지까지 루프 없이 전달될 수 있도록 한다. 또한 스페닝트리에 참여하지 않는 링크에는 백업링크용 대역이 예약될 필요가 없으므로 일반 데이터 프레임 전달용 대역을 더 많이 제공할 수 있는 장점이 있다.

본 논문의 구성은 다음과 같다. 본 서론에 이어, 제 II장에서는 프로바이더 백본 브리지 기술의 특징과 링크보호방식을 분석한다. 제 III 장에서는 제안하는 백업링크 전용 스페닝트리 기반의 링크장애복구 기능의 동작과 이에 따른 일반 데이터 트래픽에 대한 성능 향상도를 비교 분석한다. 제 IV 장에서는 제안된 방식이 적용된 실험실 수준의 리눅스 기반의 PBB-TE 시스템의 구현을 다룬 후, 마지막으로 제 V장에서는 결론을 맺는다.

II. 프로바이더 백본 브리지와 링크보호절체 기술

1. Provider Backbone Bridge(PBB)-TE기술

가. Provider Backbone Bridge(PBB)

기존 브리지망 기술은 제어평면 관점에서 customer

망과 캐리어 백본망간의 구분이 없고, 4094개로 제한된 VID사용에 의한 확장성 문제 및 종단 간 QoS가 제공되는 않는 문제 등에 의해 캐리어 급 망에 부적합하다. 이러한 문제를 해결하기 위한 802.1ah Provider Backbone Bridge(PBB)기술은 MAC-in-MAC 기법을 사용하여 확장성을 제공 한다^[3]. 여기서 MAC-in-MAC 기법은 사용자 이더넷 MAC프레임 (C-MAC)을 새로운 백본용 MAC 프레임(B-MAC)의 페이로드에 수납하여 백본 에지 브리지(BEB: Backbone Edge Bridge)간에 전달하고, 이들을 연결하는 백본 코어 브리지(BCB: Backbone Core Bridge)는 B-MAC의 주소만 참조하여 다량의 프레임을 고속으로 중계한다.

또한 여러 개의 주요 종단 간(즉 data center, ISP gateway 등)연결을 가상적으로 구분하기 위한 Backbone VLAN ID(B-VID)을 추가 사용한다. 따라서 개별적인 PBB 트렁크는 {B-SA, B-DA, B-VID}의 조합으로 구분될 수 있다. 여기서 {B-SA, B-DA}는 각각 백본망에서의 송신지 및 목적지 브리지의 MAC주소이면서 해당 트렁크의 종단 식별자이다. 그리고 B-VID는 동일한 목적지에 대한 여러 개의 트렁크를 식별하는 값으로 사용된다.

이러한 특징뿐만 아니라 새로운 24비트 길이의 service identifier(I-SID)를 추가 사용하여 12비트에 불과한 B-VID값으로 제한된 기존 VLAN ID 확장성 문제를 해결하였다.

PBB망에 참여하는 각 브리지는 {VID}, {VID 및 DMAC}, 또는 {I-SID, VID, DMAC}를 참조하여 수신되는 프레임을 다음 브리지로 중계한다. 대부분의 중계 동작에서 각 브리지를 거칠 때 헤더는 변경되지 않지만 VID값은 변경될 수 있다. 하지만 여전히 PBB망에서의 링크장애 시에는 스페닝트리 알고리즘에 의한 망의 재설정 과정이 수행되므로 새로운 스페닝트리를 신속하게

형성하는 다양한 알고리즘에 대한 연구도 진행되고 있다^[7, 10].

나. PBB-TE(Traffic Engineering)

기존 PBB기술은 기존 브리지 망의 단점인 확장성 문제를 해결할 수 있었지만, 기본적으로 여전히 기존의 IEEE802.1D 브리지의 기능을 수행하는 장치이기 때문에 MAC 주소 학습에 따른 미등록 프레임에 대한 플러딩 동작과 Spanning Tree Protocol (STP) 또는 Rapid STP(RSTP)를 사용한 긴 고장복구 시간 등의 문제가 있다. 이러한 문제는 캐리어급 백본망에서는 심각한 문제를 초래한다.

이를 해결하기 위하여 동적인 FDB설정 및 STP에 의한 망의 재설정 기능 등을 제거하는 대신에, 관리자, 또는 SNMP 및 Path Computation Element (PCE)프로토콜에 의한 별도의 원격 경로설정절차를 사용하여 고정 FDB를 설정하는 이더넷 기반의 연결위주 서비스를 제공하고, IEEE 802.1ag connectivity fault management(CFM) 기능을 이용 하여 링크장애 시 백업링크를 사용하여 링크보호 (protection) 기능을 추가로 제공함으로써 PBB에 비하여 보다 예측가능하고 제어기가 용이하도록 한 방법이 IEEE 802.1Qay Provider Backbone Bridge -Traffic Engineering(PBB-TE)기술이다^[4].

참고로 PBB-TE에서는 고정 FDB가 사용되므로 자동주소학습과정이 없으며, 모든 입력되는 방송형 및 멀티캐스트 프레임은 폐기되거나 트렁크별 유니 캐스트 프레임에 수납된다. 플러딩을 예방하기 위하여 목적지가 FDB에 등록되어 있지 않은 프레임은 자동 폐기된다.

2. 링크보호절체기능

동작 중이던 링크의 단절 또는 노드의 고장 등의 장애가 발생한 경우 또 다른 링크를 제공하는 기능을 링크장애복구 또는 링크보호절체(protection switching) 기능이라고 하며 제공되는 링크의 수와 방법은 다음과 같다. 이것을 Automatic Protection Switching (APS)라고 부르기도 한다. 이러한 보호기능은 IEEE 802.1 뿐만 아니라 ITU-T에서도 G.8031/Y.1342표준에 명시되어 있다.

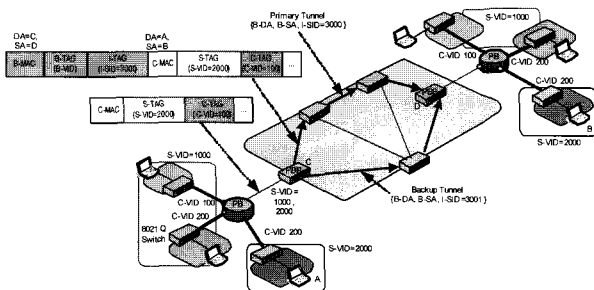


그림 1. PBB브리지망의 예
Fig. 1. Example of a PBB network.

가. 1+1 보호방식

이 방법은 “one-plus-one” protection방식으로, 데이터를 주 링크와 백업링크를 통하여 동시에 동일한 내용으로 송신하고 수신측에서는 FCS계산에 통과한 프레임 중 하나를 선택한다. 링크단절 뿐만 아니라 인터페이스 카드의 고장에도 대비할 수 있는 장점이 있다.

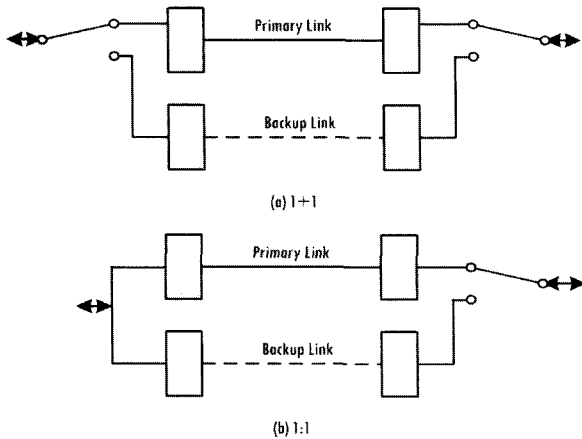


그림 2. 1+1와 1:1 APS방식
Fig. 2. 1+1 and 1:1 APS Mechanisms.

나. m:n 보호방식

이것은 m개의 백업링크를 사용하여 n개의 주 링크를 보호하며, “m-for-n” protection방식 이라고 한다. 이러한 방식은 모두 링크단절을 감지한 후 백업링크로의 전환중에 프레임의 손실이 발생할 수 있으며 또한 라인 카드의 고장에는 대비할 수 없는 단점이 있다. m의 값에 따라 다음과 같은 3가지 변형 형태가 가능하다.

- 1:1 : 하나의 주 링크에 대한 하나의 백업링크가 제공된다.
- n:1 : 하나의 주 링크에 대하여 n개의 백업링크가 제공된다.
- 1:n : n개의 주 링크에 대하여 1개의 백업링크가 보호를 하는 공유형태로써 대역활용 면에서 유리하다. 물론 다양한 목적지를 향하는 경로별로 이러한 백업링크를 설정해야 할 경우 대역활당이 복잡한 단점이 있다.

3. 링크보호절체과정과 백업링크 운용방법

가. 링크보호절체과정

링크장애보호절체과정은 장애감지로 부터 중단 간

서비스가 재개되는 데까지의 과정이다. 이 과정에서 소요되는 시간은 다음과 같다.

- Detection time(T_D): 고장발생시점에서 이를 감지하기까지의 소요시간.
- Hold-off Time(T_H): 장애감시시점에서 적절한 보호절체 기능을 적용할 때까지의 경과시간이다. 같은 네트워크의 둘 이상의 여러 계층에서 보호절체과정을 수행하려 할 때 유용하다. 예를 들어 상위 계층에서 Hold-off instant는 상위 계층의 protection이 동작하기 전에 하위 계층에서 protection을 수행할 기회를 줄 수 있다.
- Connectivity Restoration Time(T_R): 장애발생이후 중단간 전송이 가능하게 되기까지의 소요시간이다. 프레임 전달이 가능하지만 SLS가 요구하는 성능을 만족시키는지 알 수 없다.
- SLS Restoration Time(T_S): 장애발생이후에 사용자 트래픽이 원래의 성능을 만족하면서 중단간 전송이 이루어지기까지의 지연시간이다.

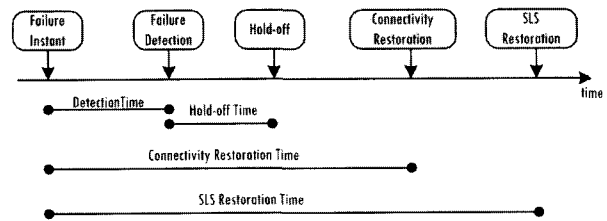


그림 3. 링크보호절체관련 소요시간의 종류
Fig. 3. APS Events and Times.

나. 백업링크 운용방법

백업링크를 운용방법으로 1:1(Cold standby), 1:1(Hot standby), shared redundancy방식이 있다. Cold standby는 경로를 미리 계산해 둔 상태에서 장애발생이후에 실제 대역활당과정을 수행하는 것인 반면, Hot standby는 백업링크에 대한 대역활당 등을 실제 링크에 설정하고 대기상태를 유지하는 것으로서 백업링크로 복사된 프레임이 이중으로 전달할 수도 있다. 물론 Hot standby방식이 자원을 낭비하는 단점을 가지지만 빠른 복구기능을 제공한다.

4. IEEE 802.1ag (connectivity fault management)기술

이더넷기반의 OAM기술은 오피스급 이더넷 기술을 캐리어급 망에서 사용할 수 있도록 하는 주요 기술 중의 하나로 목적지까지의 링크 또는 노드의 장애를 발견, 격리 및 통보하는 등 망을 관리할 수 있는 기능을 제공한다. 본 논문에서 다루는 이것의 핵심기능 중 하나인 connectivity fault management (CFM)으로써 연결 장애 감지 및 해당 링크의 고립기능을 수행한다.

CFM은 브리지간의 제어연결로인 maintenance associations(MA)을 설정하여 CFM 메시지를 상호 교환한다. 즉 MA는 Connectivity Check Message (CCM), TraceRoute, Loopback 메시지를 사용하는 CFM메시지를 생성하고 수신할 수 있는 2개의 maintenance association endpoint(MEP)간의 연결로서 경로상의 브리지 즉 maintenance association intermediate point (MIP)들도 이러한 CFM관련 메시지의 수신여부로부터 장애감지 및 고립기능을 수행한다.

특히 CFM의 장애감지기능은 continuity check message(CCM)메시지를 사용하여 MA간 연결장애 감지를 지원한다. 이를 위하여 개별 MEP는 connectivity check message(CCM)를 주기적으로 송신하고 MA상의 다른 MEP로부터의 CCM도 수신한다. 이러한 연결검사 기능은 프레임손실 및 지터 계산용으로도 사용된다.

참고로 장애확인 및 장애링크의 고립기능은 장애가 감지된 이후 수행하는 관리 동작으로서 루프백 메시지를 사용하여 링크의 장애를 확인한다. 또한 이러한 루프백메시지를 대량 송신하여 대역폭 및 신뢰성과 지터를 측정할 수도 있다. 그리고 link Traceroute 메시지를 사용하여 링크별 경로를 수집할 수 있다.

III. 백업링크 전용 스페닝트리 기반의 링크보호절체기능

1. 망의 구성과 동작

본 논문에서는 백업링크 전용 스페닝트리를 사용한 새로운 링크보호절체 방법을 제안한다. 동작원리의 설명을 위하여 그림 3-1과 같은 NxN의 그리드 형태로 구성된 PBB-TE브리지망을 가정하였다. 브리지를 연결하는 모든 링크의 대역은 C이며, QoS가 제공되어야 할 트래픽 전달용 주 링크용 대역 Cp 가 모두 예약되어 있으며, 남은 대역은 백업링크용 대역(Cb) 및 일반 데이터 프레임용 대역(Cr)으로 활용된다.

이러한 망에서 백업링크 기능을 제공하기 위하여 대

역을 각 브리지 상호간을 풀메쉬 형태로 1:1로 보호방법을 사용하는 것이 가장 바람직하지만 이것은 대역활용 면에 문제가 있다. 또한 링크장애감지 이후에 대역 할당절차를 개시하는 것도 그 신호절차의 지연 때문에 바람직하지 않다.

이를 해결하기 위하여 루트브리지에서 모든 노드로 연결된 스페닝트리 형태의 사전에 예약된 백업링크를 설정하고, 링크장애시 QoS가 제공되어야 할 EVC에 대한 프레임에 대해서는 별도의 시그널링절차없이 트리상의 백업링크에 프레임을 전달하여 이들이 목적지까지 루프없이 전달될 수 있도록 한다.

여기서 $N_{1,1}$ 은 백업링크전용 스페닝트리의 종단에 참여한 노드로서 한 개의 링크만 스페닝트리에 소속된 반면 $N_{2,2}$ 는 모든 링크가 스페닝트리에 참여되고 있다. 그리고 $N_{3,2}$ 는 스페닝트리의 루트노드이다.

이렇게 스페닝트리를 사용하는 이유는 브리지와 브리지간의 링크가 단절된 경우 이를 감지한 브리지는 별도의 시그널링 절차없이 QoS트래픽을 모두 백업용 스페닝트리에 연결된 링크상으로 송신하고 이것을 수신한 이웃노드는 자신의 FDB에 기록되어 있지 않은 프레임에 대해서는 나머지 모든 포트로 플러딩할 수 있도록 한다. 이 플러딩된 프레임은 망내부에서 루프없이 목적지 노드까지 전달될 수 있다. 만약 모든 노드간에 백업링크를 설정한다면 링크장애시 각 노드별로 장애시 대처할 수 있는 FDB를 미리 작성해 두어야 하는 문제가

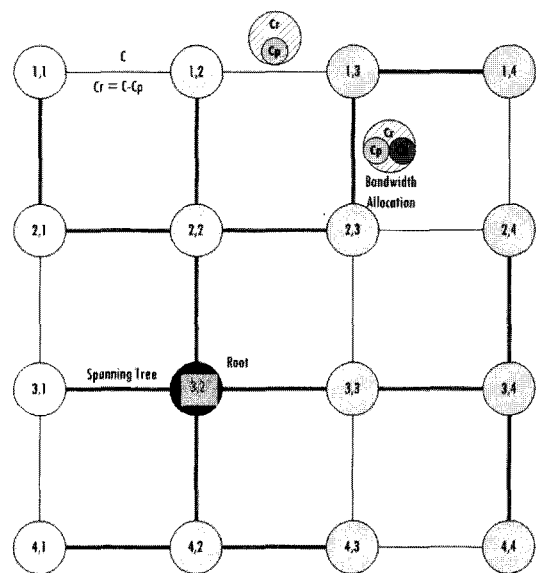


그림 4. 백업링크용 스페닝트리가 설정된 망의 예
Fig. 4. An example network with the spanning tree for backup links.

있기 때문이다.

제안된 방식에서 사용한 백업링크 사용방법은 hot standby방식의 shared redundancy방식으로 분류되며 local 링크 및 노드 장애시 이를 신속히 우회하는 local path detour 방식인 Aggregated Link and Node Protection (ALNP)개념이 적용되었다.

즉 백업링크는 루트브리지를 중심으로 모든 브리지를 루프없이 연결하는 스패닝트리형태로 미리 설정되며, 해당 예약된 대역은 모든 주 링크중에서 최대의 예약된 대역과 동일하며, 이 예약된 대역은 장애를 대비하여 대기상태를 유지한다. 물론 이러한 hot standby방식이 자원을 낭비하는 단점을 가지지만, 신속한 복구기능을 제공하기 위함이다.

참고로 제안하는 방법은 고속도로의 {전용차선, 일반차선, 갓길}과 {QoS링크, 일반데이터링크, 링크보호용링크}로 유사점을 이해할 수 있다.

2. 링크장애복구절차

가. 스패닝트리에 참여하지 않는 링크 단절시

스패닝트리에 참여하지 않는 링크가 단절된 경우 해당 브리지는 스패닝트리에 접속된 다른 인터페이스의 백업링크로 QoS트래픽을 우회시킨다. 이 트래픽은 스패닝트리에 참여한 노드들에 의해 플러딩되면서 모든 목적지로 전달된다. 이러한 플러딩 동작시 QoS프레임은 스패닝트리에 의해 루프백되지 않는 특징이 있다.

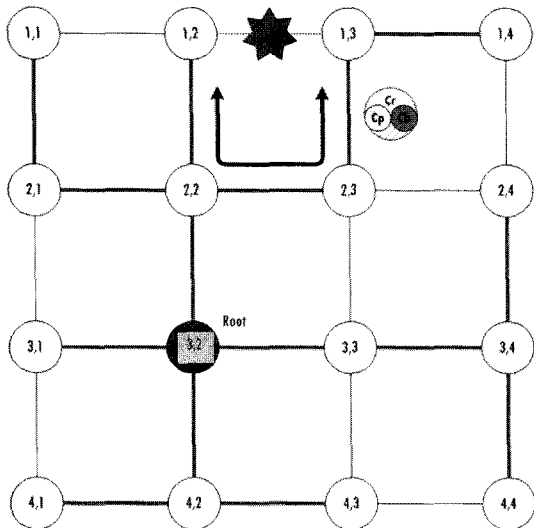


그림 5. 스패닝트리에 참여하지 않는 링크 단절시
Fig. 5. On failure event of the link not belonging to the spanning tree.

또한 일반 데이터프레임은 사전에 설정된 경로정보에 의해 다음 브리지로 전달한다.

나. 스패닝트리에 참여한 링크가 단절된 경우

그림 6은 스패닝트리에 참여한 링크가 단절된 경우의 예이다. 이 경우 하위의 노드가 스스로 이웃노드와의 이음동작(Splicing)을 통하여 신속하게 국부적인 스패닝트리 변경을 수행한다. 이러한 이음동작은 이웃노드에 대해서만 수행되므로 신호방식절차치연이 최소화되는 장점이 있다. 이 경우에도 일반 데이터프레임은 사전에 설정된 경로정보에 의해 다음 브리지로 전달된다.

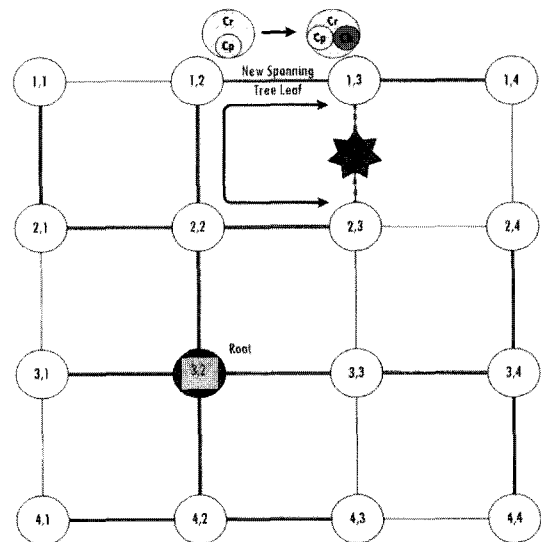


그림 6. 스패닝트리에 참여한 링크가 단절된 경우
Fig. 6. On failure event of the link belonging to the spanning tree.

다. 알고리즘

이러한 과정은 다음과 같은 절차로 요약된다.

```

Backup Link Step with Spanning Tree(ST) Algorithm
for Bridge  $i$ , ( $1 \leq i \leq N$ ).
begin
1. During the ST setup process, assigns a port as
   the root port.
2. Assigns another one port as a candidate root port.
3.1 while (true) do
3.2 if (link failure detected){
3.3 if (the link is not a part of the ST)
3.4 makes the link as the blocked port;
3.7 else
3.8 the candidate root port tries to splice a STP
   leaf link to its upstream bridge;
3.5 forward all QoS frames to the backup virtual
link
    
```

```

of the root port;
3.6 forward another frames to the root port;
3.7 }else then
3.8 forward frames;
3.9 }
end
    
```

3. 스페닝트리 이음절차

스페닝트리에 참여하지 않는 링크로 접속된 노드간에 새로운 가치를 설정하여 연결하기 위하여 기존 표준에서 사용되는 CCM메시지 헤더의 플래그 영역(현재 Remote Defect Indicator 만 사용됨)에 Splicing 플래그를 추가하여 상위 브리지에게 자신과의 스페닝트리 연결을 요청할 수 있도록 한다. 이러한 기능은 ITU-T의 G.8031/Y.1342규격에 정의된 APS OAM메시지를 변형하여 사용할 수도 있다.

4. 성능분석

가. 일반 데이터 프레임용 링크 효율

각 링크별 전체대역 C 에 대하여 주 링크용 대역과 백업링크용 대역을 각각 C_P, C_B 라고 하면, 일반 데이터 프레임에 주어지는 가용 대역 C_R 은 $C - (C_P + C_B)$ 이다. 완벽한 백업링크의 동작을 지원하는 경우에 $C_P = C_B$ 이어야 하므로 $C_R = C - 2C_P$ 이다. 예를 들어, $C=100\text{Mbps}$ 이고, $C_P=C_B=30\text{Mbps}$ 일 때 $C_R=40\text{Mbps}$ 이다.

정상적인 경우, 스페닝트리에 참여하지 않는 링크에 연결된 노드간의 일반 프레임 전송용 대역 ($C_R = C - C_P$)은 스페닝트리에 참여하는 링크에 비하여 더 많은 대역을 제공할 수 있는 장점이 있다. 즉 스페닝트리에 참여하지 않는 링크의 수는 $N \times N$ 망에서 필요한 총 링크 수 $(N-1)N + N(N-1)$ 의 $1/3$ 이므로, 망 전체의 관점에서 일반 데이터 프레임 전달용으로 제공할 수 있는 총 대역 C_R 은 $(2(N-1)N) \times 1/3 * (C - C_P) + (2(N-1)N) * 2/3 * (C - (C_P + C_B))$ 이다.

만약 모든 링크에 백업링크가 설정된다면 일반 데이터 프레임용으로 제공하는 총대역 C_R 은 $(2(N-1)N) * (C - (C_P + C_B))$ 로 주어지며 이것은 제안된 스페닝트리 방법에 비하여 작은 대역이다.

물론 스페닝트리에 참여한 링크가 단절된 경우, 스페닝트리에 참여하지 않고 있던 링크가 새롭게 스페닝트리에 참여하면서 $C_R = C - C_P$ 에서 $C_R = C - (C_P + C_B)$ 로 일반 프레임 전송용 대역이 감소하게 된다. 이

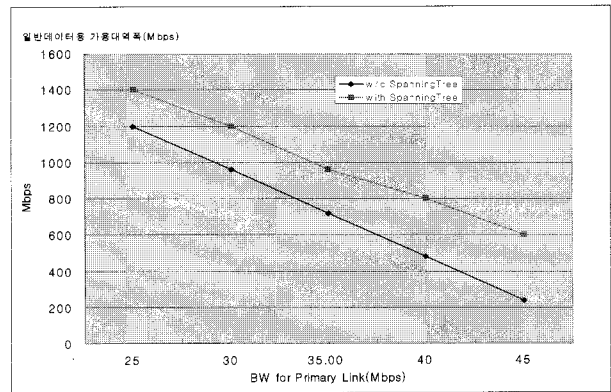


그림 7. 일반 데이터용 대역폭 비교
Fig. 7. Available bandwidth comparison for normal data traffics.

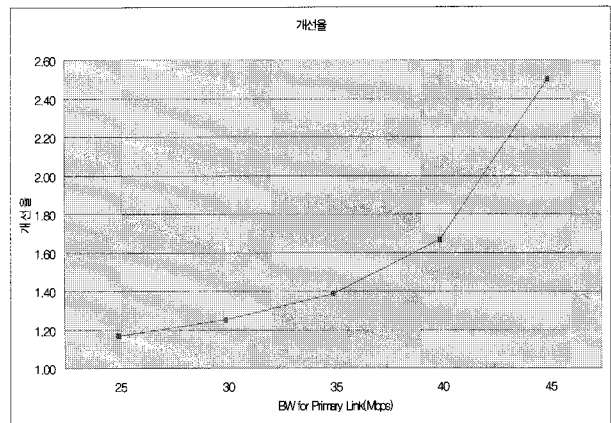


그림 8. 일반 데이터용 대역폭의 성능향상을 비교
Fig. 8. Available bandwidth efficiency comparison for normal data traffics.

러한 대역감소는 비 QoS트래픽에 대한 것이므로 큰 문제가 되지 않는다.

그림 7은 모든 링크가 백업링크로 설정된 경우와 스페닝트리에 대해서만 백업링크가 설정된 경우에 일반 데이터 프레임용으로 제공될 수 있는 대역 C_R 을 주 링크의 대역 C_P 에 따라 비교한 것이다. 또한 그림 8은 이러한 개선율을 비교한 것이다. C_P 값이 커짐에 따라 제안된 방법의 장점이 부각됨을 알 수 있다. 참고로 이러한 장점은 노드수 N 에는 무관하다.

나. QoS트래픽의 프레임 손실량

Constant Bit Rate를 가지는 QoS트래픽은 링크장애 시 손실되는 트래픽량은 장애감지소요시간 T_D 와 Hold-off Time T_H , 그리고 Connectivity Restoration Time T_C 에 영향을 받는다. 스페닝트리에 참여하지 않은 링크 단절시에는 망 단절에서 복구까지의 총 지연

시간 T 는 $T=T_D$ 뿐이지만, 스페닝트리에 참여한 링크의 단절시에는 이웃 브리지와의 새로운 스페닝트리 가지를 연결하는데 소요되는 시간이 추가되어 $T=T_D+T_H+T_C$ 로 주어진다. 이 과정에서 손실되는 프레임의 양은 $C_p T$ 이다.

다. 일반 프레임의 손실량

망복구시간 T 기간 동안 일반 데이터 프레임에 대한 프레임 손실양은 $M/M/1/K$ 큐잉모델에 의해 다음과 같은 식 (1)로부터 구할 수 있다. 여기서 l 은 데이터 프레임의 평균길이(비트)이다.

$$F_{Lost}(T, K) = \frac{\lambda(1-\rho)\rho^K}{1-\rho^{K+1}} \times T, \quad (1)$$

$$\rho = \lambda / (C_R / l)$$

브리지만 일반 프레임 전송용 대역인 C_R 은 모든 링크가 백업링크로 설정된 경우와 스페닝트리에 대해서만 백업링크가 설정된 경우에 따라 다르다.

그림 9는 CCM을 사용하는 경우 T 는 10msec마다 수신되어야 하는 3개의 CCM메시지 손실에 의한 장애감 지지연시간 및 추가적인 T_H 와 T_C 를 고려하여 $T=35msec$ 로 설정하고, $l = 1000$ 바이트, 스위치의 공유 버퍼메모리크기 31K바이트에 의한 $K=32$ 로 설정된 경우, 우회링크로 스페닝트리에 속하는 백업링크가 설정된 경우와 그렇지 않은 경우에 따른 C_R 과 입력량에 의한 프레임 손실개수를 비교한 것이다. 이 결과로 부터

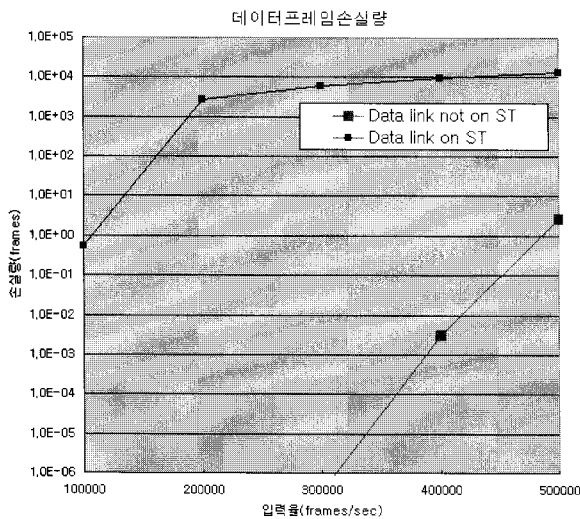


그림 9. 링크보호절체기간 동안의 데이터프레임 손실량 비교 ($T=35msec$ 의 경우)
 Fig. 9. Number of lost normal frames during link protection process ($T=35msec$).

링크 장애시 해당 데이터 프레임이 우회시킬 경우. 스페닝트리에 속하는 백업링크가 설정되지 않은 링크가 사용될 수 있도록 FDB를 설정하는 것이 유리함을 알 수 있다.

IV. 구현

1. 시스템 구성

제안된 방식이 포함된 PBB-TE시스템은 모두 Fedora 리눅스 운영체제(2.4.18.1)에서 개발되었으며, 그림 10과 같은 요소로 구성된다.

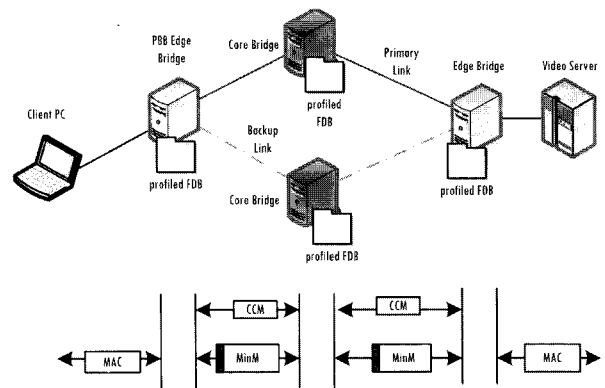
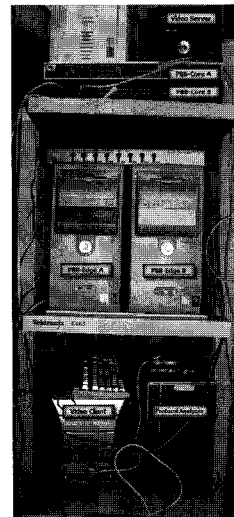


그림 10. 구현된 PBB_TE시스템과 망의 구조
 Fig. 10. Implemented PBB_TE network system.

- PBB 엣지브리지
- PBB 코어브리지
- Video 서버 및 클라이언트

가. PBB 엣지브리지
 엣지브리지는 리눅스커널용 브리지코드 기반에서

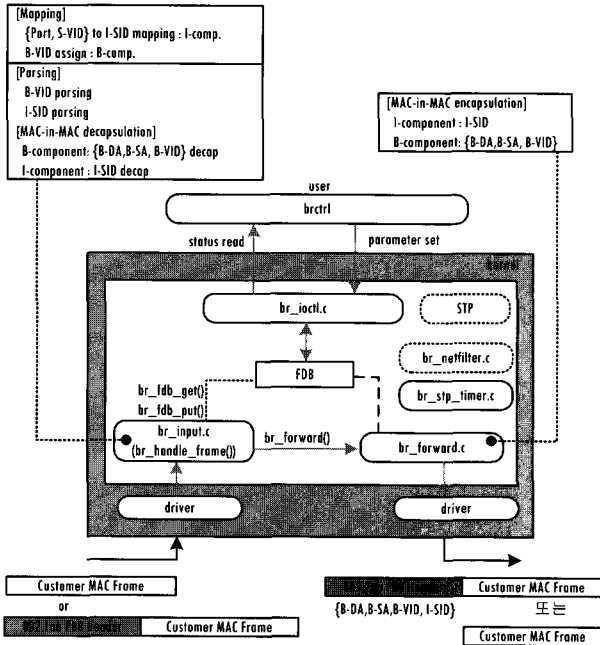


그림 11. PBB엣지브리지의 소프트웨어 구성
Fig. 11. Software architecture of the PBB edge bridge.

802.1ah MAC-in-MAC 수납과 복원기능을 수행하는 데이터평면 기능과 BRCTL을 이용한 고정 FDB 설정기능을 수행한다. 또한 CCM메시지 생성 및 수신기능, CCM송수신간격 설정기능을 구현하였다. 구현된 PBB 브리지모듈은 그림 11과 같으며 각 기능부의 동작은 다음과 같다.

- br_iput : 포트식별값에 따라 I-포트기능과 수신측 종단 PBB-TE 엣지브리지 포트인 B-포트 역할을 각각 수행한다. 즉 수신된 customer MAC프레임에 대하여 {수신된 port ID, S-VID}값에 대한 I-SID를 할당하는 I-Component기능을 수행하고 또한 B-VID도 할당하는 B-Component기능도 함께 수행한다. 또한 코어 망으로부터 수신된 B-MAC프레임에 대해서는 {B-VID, I-SID}를 파싱하면서 B-MAC헤더부분을 제거하여 customer장비로 전달한다.
- nbr_forward: MAC-in-MAC 수납과정을 수행하거나 B-MAC헤더가 제거된 customer MAC프레임을 전달한다.
- nbr_timer : CCM메시지 송신주기 및 수신만기시간용으로 사용한다.
- nbr_ioctl : 사용자 평면의 brctrl 유틸리티와의 API

를 제공하여 브리지내부 변수의 설정, FDB설정 및 동작상태 정보교환을 지원한다.

나. 코어 브리지

PBB-TE 코어 브리지는 B-MAC주소를 기반으로 단순한 프레임 전달기능을 수행한다. FDB는 수동으로 설정된다. 구현시 코어브리지는 일반 스위치를 사용하였다.

다. CFM기능과 링크장애 복구기능의 구현

802.1ag CFM 기능 중에서 CCM/RDI: Connectivity Check, Remote Defect Indication만 구현하였다. CCM 메시지의 주기적인 송신기능은 브리지 내부 타이머(기본값 10msec)를 추가하여 구현하였으며 그 주기는 관리자가 설정할 수 있도록 하였다. 상대방 브리지는 CCM메시지의 손실이 3개 이상 발생하면 미리 설정된 백업링크를 사용할 수 있도록 하였다.

링크고장방지 기능은 각 트렁크별 주 링크와 백업링크용 B-VID를 할당하고, 주 링크 고장시 즉 802.1ag에 규정한 continuity check messages(CCM) 가 수신되지 않으면 이것을 감지한 브리지는 해당 B-VID값을 protect용 B-VID로 대체하여 미리 설정된 백업링크가 설정된 포트로 중계한다.

그림 12의 예는 실험 망에서 루트브리지 B4를 중심

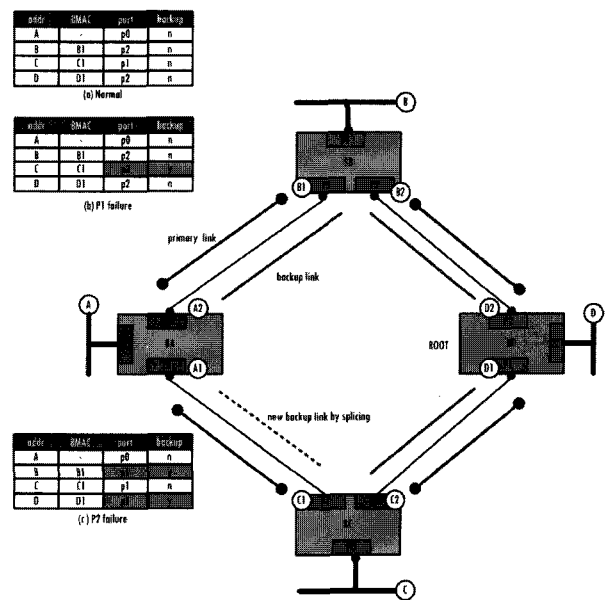


그림 12. 실험망에서의 링크보호절체의 예
Fig. 12. Example of link protection operation.

으로 백업링크용 스페닝트리를 형성한 상태에서 정상상태, BA의 P1포트가 연결된 링크가 단절된 경우, BA의 P2포트가 연결된 링크가 단절된 경우에 대한 BA의 FDB내용이다.

라. 프로토콜 분석도구 구현

PBB-TE 시험망을 위한 브리지구현에 필수적인 802.1ah PBB프레임과 802.1ag CCM메시지의 전송과정을 분석하기 위한 프로토콜 분석기용 윈도우 환경에서 개발하였다.

마. 시험

동작시험을 위하여 PBB-TE 엣지브리지 2대와 2대의 코어브리지를 2x2 그리드 방식으로 연결 하였으며, 자체 제작한 프로토콜 분석기도 프로그램 개발 시 사용하였다.

관리자는 그림 13과 같은 브리지의 설정화면을 통하여 FDB설정과 CCM메시지 생성주기를 설정 한다.

그림 14는 이후 MAC-in-MAC방식으로 수납된

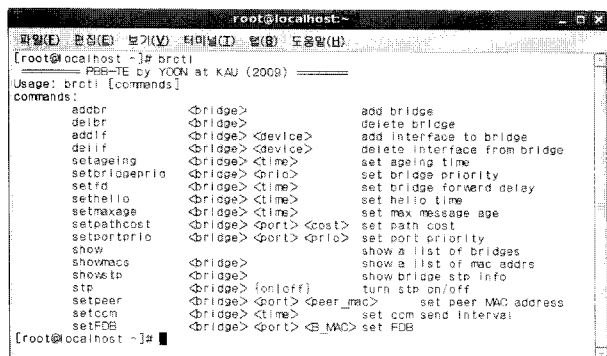


그림 13. PBB-TE 설정화면
Fig. 13. Console interface of a PBB-TE bridge.

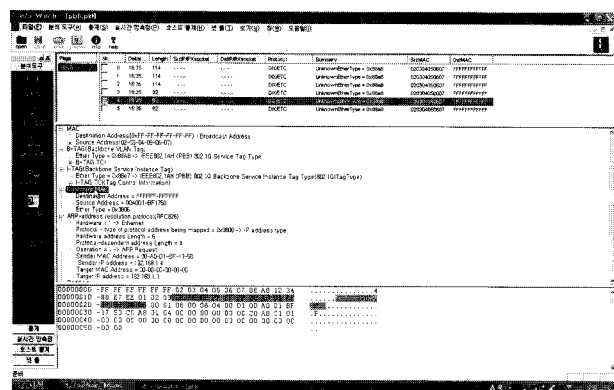


그림 14. IEEE802.1ah PBB 메시지의 예
Fig. 14. Decoded IEEE802.1ah PBB Frames.

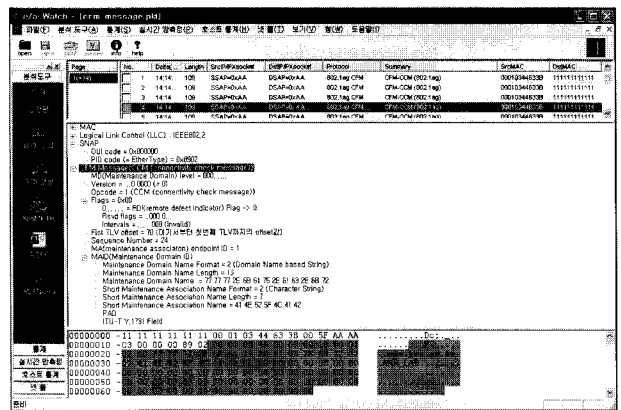


그림 15. 802.1ag CCM메시지의 예
Fig. 15. Decoded CCM messages.

PING메시지에 대한 것이며, 그림 15는 CCM 메시지의 전송 예이다. 시험과정에서 주 링크를 단절한 경우 35msec 이후 이를 감지하고 백업링크를 통한 비디오 스트림 전송이 지속되는 것을 확인하였다.

V. 결 론

본 논문에서는 브리지망에서 사용되는 스페닝트리를 활용하여 루트브리지에서 모든 노드로 연결된 스페닝트리 형태의 사전에 예약된 백업링크를 설정하고, 링크장에서 QoS가 제공되어야 할 트렁크에 해당되는 프레임에 대해서는 별도의 시그널링 절차없이 트리상의 백업링크에 프레임을 전달하여 이들이 목적지까지 루프없이 전달될 수 있는 새로운 방법을 제안하고 성능을 분석하였다. 제안된 방법이 적용된 PBB-TE 실험망 시스템을 리눅스로 구현하고 관련된 프로토콜 분석도구도 함께 구현하여 동작을 시험하였다. 앞으로 TCP/IP 클라이언트/서버 기반의 Path Computation Element(PCE) 기능과 각 브리지에서의 엄격한 트래픽 폴리싱 기능을 구현할 계획이다.

참 고 문 헌

- [1] A.Reid, P. Willis, I. Hawkins, and C. Bilton, "Carrier Ethernet," *IEEE Commun. Mag.*, pp. 96-103, Sept. 2008.
- [2] D. Allan et al., "Ethernet as Carrier Transport Infrastructure," *IEEE Commun. Mag.*, vol. 44, no. 2, pp. 134~140 Feb. 2006.
- [3] IEEE P802.1ah-2008, "IEEE Standard for Local and Metropolitan Area Networks- Virtual

Bridged Local Area Networks - Amendment 6: Provider Backbone Bridges.", 2008.

[4] IEEE P802.1Qay/D3.0, "IEEE Standard for Local and Metropolitan Area Networks -- Virtual Bridged Local Area Networks -- Amendment : Provider Backbone Bridge Traffic Engineering," 2008.

[5] IEEE802.1ag - Connectivity Fault Management, Dec, 2007.

[6] Ryoo, Song, Park, and Joo, "OAM and Its Performance Monitoring Mechanisms for Carrier Ethernet Transport Networks," *IEEE Commun. Mag.*, pp.97-103, March. 2008.

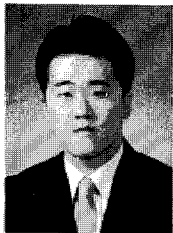
[7] Suh, Kim, and Shin, "ENDIST: Edge Node Divided Spanning Tree," *ICACT*, Volume 1, pp. 17-20, Feb. 2008.

[8] Aref Meddeb, "Smart Spanning Tree Bridging for Carrier Ethernets," *Globecom*, pp. 1 - 5, 2008.

[9] Mick Seaman, "IEEE 802.1aq Shortest Path Bridging," IEEE Draft Standard, work in progress.

[10] Aref Meddeb, "Multiple Spanning Tree Generation and Mapping Algorithms for Carrier Class Ethernets," *IEEE Globecom*, 2006.

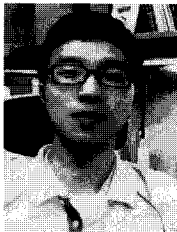
저 자 소 개



남 위 정(학생회원)
 2009년 한국항공대학교 정보통신공학과 학사 졸업
 2009년~현재 한국항공대학교 정보통신공학과 석사과정
 <주관심분야 : 컴퓨터통신망, 네트워크보안, NetFPGA, Future Internet >



윤 중 호(평생회원)
 1984년 한양대학교 전자공학과 학사 졸업
 1986년 한국과학기술원 전기 및 전자공학과 석사 졸업
 1990년 한국과학기술원 전기 및 전자공학과 박사 졸업
 1991년~현재 한국항공대학교 항공전자 및 정보통신공학부 교수
 <주관심분야 : 유무선통신망 설계 및 성능분석>



이 헌 주(학생회원)
 2010년 한국항공대학교 정보통신공학과 학사 졸업
 <주관심분야 : 컴퓨터통신망, 센서네트워크>



홍 원 택(정회원)
 1998년 성균관대학교 정보공학과 학사 졸업
 2000년 성균관대학교 컴퓨터공학 석사 졸업
 2000년~2002년 (주)콤텍시스템 기술 연구소 연구원
 2002년~현재 한국과학기술정보원 슈퍼컴퓨팅본부 기반기술개발실 선임연구원
 <주관심분야 : 네트워크 관리, 트래픽분석, 람다네트워킹>



문 정 훈(정회원)
 1999년 경북대학교 컴퓨터공학과 석사졸업
 2010년 현재 한국과학기술정보연구원 슈퍼컴퓨팅본부 기반기술개발실 선임연구원

<주관심분야 : 네트워크 가상화, 융합네트워크, 원격가상현실>