

고장 극복 (Fault Tolerant) Ethernet 구현 방안의 비교 분석

김 세 목*, 고 윤 민*, 최 한 석*, 민 정 현*, Anh Pham Hoang*, 이 동 호**, 이 중 명***

A Comparison of Fault Tolerant Ethernet Implementation Approaches

Se Mog Kim*, Yun Min Ko*, Han-Seok Choi*, Jung Hyun Min*,
Anh Pham Hoang*, Dong Ho Lee**, Jong Myung Rhee***

요 약

최근 대부분의 첨단 임무 (mission critical) 시스템이 컴퓨터 노드 연결에 의한 Ethernet 네트워크 기반으로 구성되고 있어, 이에 따른 시스템 신뢰성 측면이 더욱 중요한 요소로 대두되고 있다. 본 논문은 첨단 임무 시스템의 신뢰도를 높이기 위하여 일반화되고 있는 고장 극복(Fault Tolerant) Ethernet의 개념을 이중화 관점에서 정립하고, 이를 위한 핵심 사항인 노드 이중화 구현 방식을 비교하여 제시한다. 고장 극복개념을 위한 기본 구성인 Ethernet 이중화를 독립형 이중화와 통합형 이중화로 구분하여 분석하였다. 또한 노드 이중화 구현 방식은 기존의 소프트웨어 접근 방식과 하드웨어 접근 방식은 물론, 최근 본 연구팀의 일부가 제안한 복합 (hybrid) 방식도 함께 비교 대상으로 하였다. 실시간이 요구되는 첨단 임무 시스템의 경우는 하드웨어 접근 방식 또는 복합 방식이 적합하고, Commercial-Off-The-Shelf (COTS) 사용이 요구되는 시스템의 경우는 소프트웨어 접근 방식이나 복합 방식의 선택으로 귀결된다.

ABSTRACT

Reliability is absolutely necessary in most recent mission critical systems which usually utilize the Ethernet based computer networks. A typical way to increase system reliability is to equip the fault tolerant Ethernet. In this paper we try to formulate the fault tolerant dual Ethernet concept and presents a comparison of its implementation approaches. Two types of dual Ethernet configuration are analyzed; the divided and the connected. Then the characteristics of three current implementation approaches which are the hardware based, the software based, and the recently proposed hybrid approach are compared. The results show that the hardware based or the hybrid approaches can be a better solution for the real time mission critical systems. Also for the systems which require the use of Commercial-Off-The-Shelf (COTS) hardware for fault tolerant Ethernet the possible choice is the software based or the hybrid approach.

Key-words : Fault Tolerant Ethernet, Implementation Approaches, Software based, Hardware based, Hybrid

1. 서 론

첨단 임무(mission critical)시스템에서는 대부분의 컴퓨터 시스템이 통상적으로 요구하는 고성능 및 저 전력 특성은 물론, 높은 신뢰도가

절대적으로 요구된다. 이러한 특성이 요구되는 시스템은 무인 비행체, 군용 무기체계, 공정 제어 네트워크, 그리고 항공장치 등이다. 따라서 이러한 시스템의 핵심 관점은 시스템 신뢰성 향상이다. 신뢰도는 고장률에 역 비례하므로 고장

* 명지대학교 통신공학과 Ubiquitous&Convergence 연구실 ((kimsemog, 80080146, koesnah, alswwjdgusl, anhph)@mju.ac.kr)

** 국방과학연구소 (hangboknara@naver.com)

*** 명지대학교 통신공학과 (jmr77@mju.ac.kr)

를 줄임으로써 신뢰도 증가가 가능하다. 현실적으로 고장이 없는 시스템은 불가능하므로 해결책은 고장 시에도 시스템이 동작을 유지하도록 하는 방안이 필요하며, 이는 고장극복(Fault Tolerant) 기능의 확보를 의미한다. 특히 다수의 컴퓨터, 스위치 및 관련 장치들이 연결된 최근의 네트워크화 첨단 임무 시스템에서의 고장 극복 기능 확보는 더욱 중요하게 된다.

본 논문에서는 현재 상용 표준으로 보편화된 Ethernet 장치의 이중화 방식 고장극복 방안의 비교분석을 제시하고자 한다. 그간 첨단 임무 수행을 위한 Ethernet 고장 극복의 다양한 방안이 연구 제시되었다. [1-5]. 기본적으로는 하드웨어 접근 방식과 소프트웨어 접근 방식(통상 layer 2 기반)이 주로 연구되었고, 최근 이러한 방식의 장점을 최대한 활용한 복합(hybrid) 방식이 당 연구팀에 의해 최근 제안된 바 있다 [6]. 그러나 이러한 방식들의 상세한 비교 분석에 대한 연구는 거의 제시되지 않고 있는 실정이다.

본 논문의 구성은 다음과 같다. 2장에서는 Ethernet 이중화 네트워크의 일반 개념을 제시하고, 3장에서는 이중화 고장 극복 방식의 비교 분석을 제시한다. 마지막으로 4장에서 결론을 제시하고자 한다.

II. Ethernet 이중화 개념

2.1 단일화 Ethernet 네트워크

Ethernet 프로토콜의 전송 지연에 따른 실시간 전송 어려움, 대역폭 제한, QoS(Quality of Service) 문제 등이 100Mbps 급 이상의 LAN switch 실용화로 해결되어, Ethernet은 bus 형태 대신에 통상적으로 그림 1과 같은 star 구조로 운용되고 있다. 즉 여러 노드가 개별 링크를 통하여 하나의 Ethernet switch에 연결된 구조로서, 개별 링크는 광 케이블 또는 twisted-pair 케이블로써 구성된다. 물론 twisted-pair cable의 경우 신호 대 잡음을 고려하여 노드와 스위치간의 거리가 100m 수준으로 제한되지만

repeater 사용 등으로 그 거리는 증대될 수 있다. 일반 네트워크와는 다르게, 첨단 임무 네트워크는 임무의 특수성 때문에 각 노드의 개수와 위치 등이 이미 사전에 정의 되어 고정적인 경우가 대부분이다. 이 경우 스위치의 mapping table이 고정적이며, 노드의 변화가 없으므로 switch의 소프트웨어 update 또한 불필요하다.

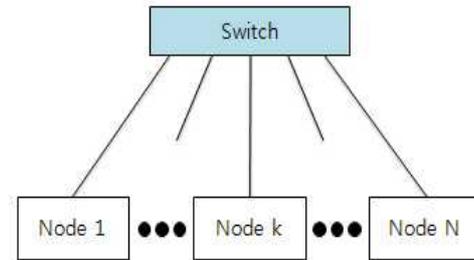


그림 1. 단일화 Ethernet 네트워크
Fig 1. Single Ethernet network

2.2 이중화 Ethernet 네트워크

2.2.1 독립형 이중화

그림 2는 독립형 이중화 Ethernet 네트워크를 예시하고 있다. 개념적으로는 2.1항의 단일화 Ethernet 네트워크를 두 개 복사한 구조이다. 각각 노드는 2개의 포트를 가지며, 각기 다른 스위치에 연결되어 있다. 스위치간의 연결이 없는 구조이므로 한 노드 안에서는 각각의 포트가 같은 IP를 사용하여도 네트워크상 충돌이 발생하지 않으므로 하나의 IP를 사용하는 것이 망 관리에 편리하다. 이와 같이 하나의 노드 당 한 개의 IP를 사용하는 경우, 두 개의 스위치의 configuration이 동일해 지는 편리성까지 더해진다.

일반적으로 네트워크 고장이란 스위치나 노드 고장의 장치 고장과 연결 케이블의 접속 불량 및 파손으로 인한 링크 고장으로 분류할 수 있다. 그림 2로부터 2개의 노드만으로 구성된 이중화 네트워크 시스템을 고려하면 그림 3과 같다.

이중화 네트워크 구조에서 두 개의 네트워크 구조는 서로 독립형이므로 표 1에 제시된 고장

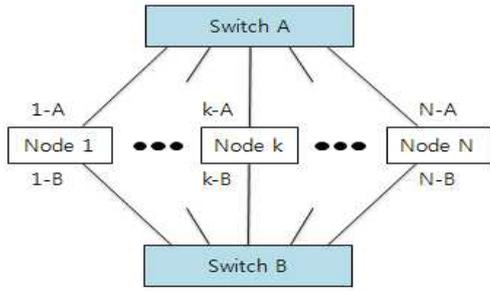


그림 2. 이중화 네트워크 구조
Fig 2. Dual Ethernet network configuration

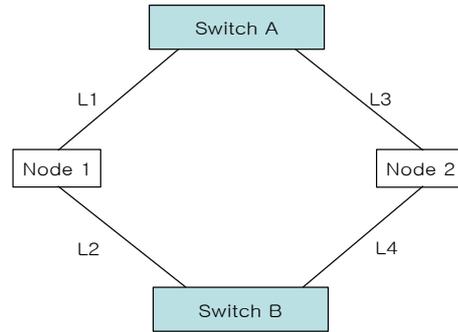


그림 3. 2개 노드의 독립형 이중화 네트워크 시스템
Fig 3. Separated dual Ethernet network system with 2 nodes

유형에 대하여 노드의 입장에서 두 개의 독립형 네트워크 중 어느 네트워크가 고장이며 어느 네트워크를 사용하여야 할지를 포트의 선택으로 결정하여야 한다.

일반적으로 첨단 임무시스템은 분산 시스템이며, 모든 노드끼리의 데이터 전송이 보장되어야 하므로 각각의 노드는 네트워크의 고장을 감지하고 두 포트 중 어느 포트를 사용하여야 할 것인가, 즉 어느 포트의 입력을 이용하고 어느 포트로부터 출력을 보내야할지의 판단 기준이 필요하다.

물론 이중화 네트워크에서도 고장복구의 한계가 존재하는데, 표 1에 제시된 바와 같이 (1) 이중화 링크 모두가 단선되어 노드가 단락된 경우와 (2) 노드는 단락되지 않았으나 두 개의 네트워크에서 한 개씩 단선이 된 경우이다. (2)의 경우는 이중화의 문제라기 보다는 독립형 이중화의 문제이다.

표 1. 링크 고장 유형
Table 1. The type of link failure

고장 유형	고장 링크(예)	독립형 이중화로 복구여부
단일 링크 단선	L1	가능
하나의 네트워크에서 이중 링크 단선	L1 and L3	가능
노드 단락	L1 and L2	불능
각각의 네트워크에서 단일 단선	L1 and L4	불능

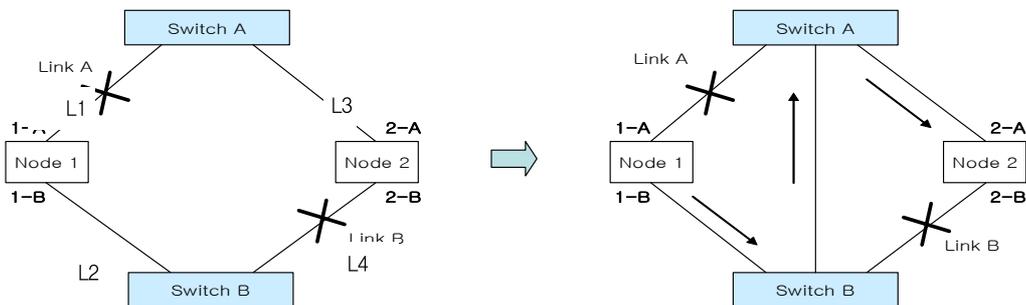


그림 4. 스위치간 연결로 두 개의 링크 고장 문제 해결

2.2.2 통합형 이중화

독립형 이중화 네트워크 구조에서는 스위치간의 연결이 없으므로 스위치가 분리되어 두 개의 네트워크가 서로 독립형인 구조를 가지게 된다. 이는 단순히 두 개의 네트워크를 중복 사용하는 데서 기인하는 것으로 이중화 네트워크의 유지보수가 단순화되는 효과가 있으나, 표 1에서와 같이 특정한 링크 여러 유형의 경우 복구가 불가능하므로 링크 자원을 충분히 활용 못하는 단점이 있다.

그러나 스위치를 연결하여 독립형인 두 개의 네트워크를 통합함으로써 이러한 특정 유형에 대한 여러 복구 불가능 문제를 극복하는 개념이 통합형 이중화이다 (그림 4 참조).

이는 그림 5와 같은 Dual IP 기반 구조로서 구성되는데 각 포트별로 다른 IP 어드레스를 부여하여 독립형인 두 쌍의 네트워크를 통합하는 개념이다. 이 경우 스위치와 스위치는 제한 없이 연결되어 있어 노드의 한 개의 포트는 상대 노드의 두 개의 포트가 보낸 패킷을 동시에 수신하게 된다. 네트워크 상태에 따라 송신포트의 IP를 선택하여 네트워크 오류 감시/복구를 하게 된다. 전체적으로 네트워크 부하는 독립형 이중화 네트워크와 비교하여 약 2배 증가하며, 노드에서의 dual IP 기반 감시/복구 알고리즘이 필요하다.

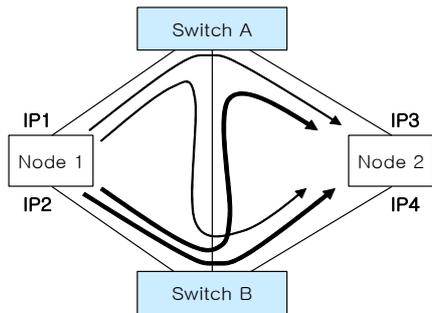


그림 5. Dual IP 기반 통합형 이중화 네트워크
Fig 5. Dual IP based a integrated dual Ethernet network

III. Ethernet 이중화 구현 방식 비교

3.1 소프트웨어 접근방식

3.1.1 개념

소프트웨어 접근방식은 각 노드가 자신을 제외한 모든 노드에 자신의 존재를 알리는 하트비트(heartbeat) 신호를 주기적으로 보내고, 각 노드에서 하트비트 신호를 받지 못하는 포트를 링크 고장으로 간주하여 하트비트 신호가 지속되는 포트를 채택 사용하여 네트워크 고장을 감지하고 복구하는 방식이다.

노드 개수를 N으로 가정하고, 스위치를 A, B 그리고 스위치 A에 접속된 노드 k의 포트를 k-A, 스위치 B에 접속된 노드 k의 포트 k-B로 가정한다. (그림 6 참조).

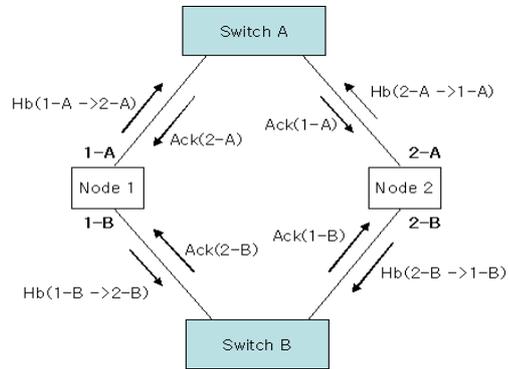


그림 6. 하트비트 동작원리
Fig 6. The heartbeat operation

k-A 포트는 자신을 제외한 N-1 포트 {m-A | m=1,...,k-1, k+1,..., N} 로 특정한 하트비트 신호 패킷을 주기적으로 보낸다. 주기를 T_h 라 할 때 각 노드는 T_h 초 마다 특정한 하트비트 패킷과 acknowledgement 패킷을 기대하게 된다. 이러한 방식은 스위치는 초당 $2*N(N-1)/T_h$ 개의 부가적 패킷을 처리하는 것이고, 개별 링크에는 $2*N/T_h$ 의 부가적 패킷이 전송 되므로 네트워크에 부담을 주게 된다. 예를 들어 50개 노드가 100msec 주기(T_h)로 하트비트를 보내는 경우 링크에는 초당 1×10^3 개의 패킷, 스위치에는 4.9×10^4 개의 패킷이 추가적으로 부여된다. 이 경우 ping 최소 패킷 형태인 ICMP 패킷을

64Bytes로 산정하면 링크 당 0.512Mbps의 추가 부담을 주게 된다.

3.1.2 고장감시 및 복구 방법

각 노드는 A포트와 B포트 두 개 포트 중에서 주 포트(primary port)와 보조 포트(secondary port)를 정한다. 정상시에는 주 포트 입력 시그널만 사용을 하고 보조 포트의 입력은 참조하지 않는다. 고장감시는 각 노드의 포트들이 하트비트 주기 T_h 초 마다 정의된 하트비트 패킷을 모니터링하는 것으로 이루어진다. 만약 특정 노드에서 주 포트가 하트비트 패킷을 수신하지 못한 경우, 해당 링크의 이상을 선언하고 B 포트의 하트비트 패킷 수신 여부를 확인한다. 만약 B 포트 마저 하트비트 패킷을 수신하지 못하였을 경우 해당 노드와의 네트워크 단절을 선언한다. 이 경우 물리적 복구방법 (off-line maintenance)이 강구되어야 한다.

3.1.3 특징

소프트웨어 접근방식은 특별한 하드웨어를 사용하지 않고 기존의 COTS (Commercial-Off-The-Shelf) 이더넷 카드를 이용하기 때문에 구성이 간단하고 가격이 저렴한 장점이 있다. 그러나 네트워크 부하가 노드 수 제곱($O(N^2)$)과 복구 최소시간인 하트비트 주기의 곱에 비례하기 때문에 다중 노드 시스템 경우 복구 최소시간을 최소화 할 수 없다는 것이 이 방식의 중대한 단점이다. 또한, 각 노드끼리 하트비트를 송수신하며 네트워크 고장을 감시하고 복구하는 방법이므로 노드 CPU에 필요치 않은 부하를 주며, 링크 고장이 여러 노드에서 나타나기 때문에 고장 부분의 위치추적을 위해서는 모든 링크 상태를 취합하는 소프트웨어적 계산을 필요로 한다.

3.2 하드웨어 접근방식

3.2.1 개념

하드웨어 접근방식은 각 노드에서 두 개의 포트신호를 비교하여 물리적 신호 존재여부를 이용하는 방식이다. 즉 각 노드는 물리적 신호가 없는 포트를 링크 고장으로 간주하고 물리적 신

호가 지속되는 포트를 사용하여 네트워크 고장을 감지/복구하는 방식이다 (그림 7 참조).

3.2.2 고장 감시 및 복구 방법

하드웨어 접근방식의 어려운 점은 물리적 신호 부재를 정의하는 방법이다. Ethernet은 physical layer 관점에서 보면, 패킷을 보낼 때만 전기적 시그널이 발생하는 통신망이므로 일

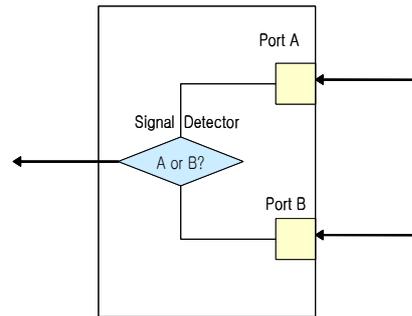


그림 7. 포트 시그널 감지
Fig 7. Signal detection at each port

반적으로 물리적 신호 부재 시간은 일정한 규칙이 없는 랜덤한 특성을 지닌다. 그러므로 물리적 신호의 부재를 선언하기 위하여 다음의 두 가지 방법을 고려할 수 있다.

첫째, 네트워크 특성에 따라 정의되는 확률적인 최소 시간단위 즉 임의의 한 노드가 한 패킷을 받은 후 다음 패킷을 받을 확률적 최소 시간단위 이상을 기다린 후 물리적 시그널 부재라고 선언하는 방법과,

둘째, 송신노드에서 두 개의 포트가 같은 양의 패킷을 시간차가 적게 보낸다는 가정 하에 수신 노드의 포트사이의 부재시간 차이가 허용 시간을 넘을 경우 신호 부재로 선언하는 방법이 있다.

3.2.3 특징

하드웨어 접근 방식은 소프트웨어 접근방식과는 다르게, 네트워크에 추가적 부하를 주지 않고 어느 링크에 고장이 있는지를 판단하기위해 상위 레벨로의 접근이 필요 없다는 장점이 있다. 반면 구현 시 제작사 고유의 하드웨어를 필요로 하여 현장 임무 특성에 맞는 패러미터

(parameter) 최적화가 제작 업체에 의존되는 단점이 있다.

3.3 복합 (hybrid) 방식

3.3.1 개념

복합방식이란 [6]에서 처음 정의된 방식으로 하드웨어 접근방식을 소프트웨어적으로 해결하는 방법이다. 이 경우 하드웨어 접근방식으로 구현되는 포트 선택 알고리즘을 소프트웨어적 차원에서 NIC카드 선택 알고리즘으로 구현하는 것에서 고려되어야 할 점은 어느 레벨에서 알고리즘을 구현할 것인가 하는 점이다. 하드웨어에 밀접한 driver 레벨로 내려갈수록 고장 감지와 복구 알고리즘의 반응시간을 단축할 수 있는 장점이 있으나 NIC 하드웨어에 종속적이 되어 범용성을 갖추기 어려운 문제가 있다.

현행 이더넷 Network Interface Card (NIC) 하드웨어는 Physical Layer와 Media Independent Interface (MII) 그리고 Media Access Control (MAC)과 같은 계층구조를 이루고 있다. 가장 이상적인 경우는 MII 레벨에서의 알고리즘 구현이다. 즉 Ethernet 송/수신기의 PHY 레벨의 블록에서 일반적으로 표시되는 RX_DV 신호는 유효한 신호가 수신되었을 경우 나오므로 두 카드의 RX_DV 신호를 비교하면 패킷수신의 성패를 판단할 수 있다. 하지만 일반적으로 MII 레벨의 시그널은 직접 접근하기 용이하지 않기에 MAC layer나 그 위 레벨에서 알고리즘이 구현될 수 있다.

3.3.2 고장 감시 및 복구방법

[6]에서 사용된 혼합 방식은 WINPCAP Library와 MS사의 SDK를 활용하여 구현한 방법으로 포트 전환 timer에 기반 하여 주 포트와 보조 포트를 선택하게 된다.

그림 8은 고장시의 포트 전환 개념으로서, 정해진 시간 내에 패킷 수신 확인되면 해당 timer가 reset되고, 그 시간 내에 패킷 수신 없는 경우에는 timeout이 선언되어 포트 전환 (switching)이 수행됨을 보인다.

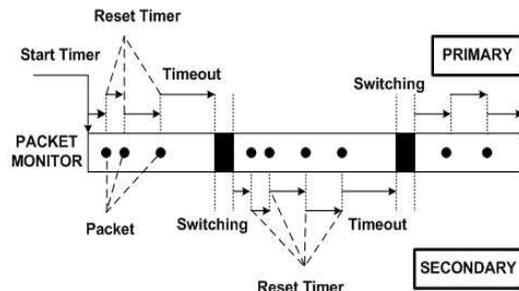


그림 8. 포트 전환 메카니즘
Fig 8. Port Switch-over mechanism

3.3.3 특징

혼합방식은 하드웨어 접근방식의 소프트웨어 구현 방식이므로 고장 시 전환 소요시간은 소프트웨어 접근 방식보다 매우 단축된다. 또한 소프트웨어 접근 방식과 마찬가지로 특정 하드웨어가 필요하지 않으므로 COTS 기반의 네트워크 카드 사용이 가능하다.

3.4 방식별 비교

세 가지 이중화 방식을 요약하면 표 2 와 같다.

표 2. 고장 극복 Ethernet방식 비교
Table 2. A comparison for FTE implementations

특성 \ 방식	하드웨어 접근방식	소프트웨어 접근방식	혼합 방식
고장 시 전환 (실시간성)	매우 빠름	느림	빠름
COTS 사용	불가	가능	가능

IV. 결 론

본 논문에서는 고장 극복 (Fault Tolerant) Ethernet 이중화 네트워크의 개념을 독립형 그리고 통합형으로 구분 정립하고, 독립형 방식의 고장 유형별 분석을 표1에 제시하였다. 아울러 네트워크화 첨단 임무 (mission critical) 시스템에 적합한 Fault-Tolerant Ethernet 구현 기법

을 비교 분석하였다. 즉 기존의 하트비트 신호에 근거한 소프트웨어 접근방식과 하드웨어 접근방식 외에 최근 본 연구진의 일부가 [6]에서 제안한 복합방식도 함께 정성적으로 비교하여 표2에 제시하였다.

향후 실시간적 침단 임무에 절대적으로 사용이 유리한 하드웨어 접근방식과 혼합방식의 실제 적용을 위해서는 전송 데이터 유형에 따른 최적 고장 전환 시간이 확률적 기반 하에서 도출되어야 한다.

아울러 소프트웨어 접근방식의 각 노드에 의한 분산적 하트비트에 의한 시간 지연을 최소화하기 위하여, 하트비트서버를 이용한 중앙 통제적 모델도 고려되어야 하며, 이 경우 하트비트 서버의 이중화는 함께 고려되어야 한다.

또한 이중화외에 다중화로의 발전도 해결해야 할 요소이다. 독립형 이중화의 경우 단순한 일반화로 독립형 다중화에 도달할 수 있다. 하지만 통합형 다중화는 간단치 않은 문제이다. 기본적으로 스위치간의 mesh 구조 하에서의 링크 오류를 효과적으로 빠르게 판단하고, 최적의 경로를 판단하여 접속하는 알고리즘 연구가 필요하다.

참고문헌

[1] J. Huang, S. Song, L. Li, P. Kappler, R. Freimark, J. Gustin, and T. Kozlik, "An Open Solution to Fault - Tolerant Ethernet : Design , Prototyping and Evaluation, Invited Paper", *Proceeding of the 18th IEEE International Performance, Computing, and Communications Conference, Phoenix*, February 1999.

[2] S. Song, J. Huang, P. Kappler, R. Freimark, J. Gustin, and T. Kozlik, "Fault - Tolerant Ethernet for IP - Based Process Control Networks", LCN, p. 116, *25th Annual IEEE International Conference on Local Computer Networks (LCN'00)*, 2000.

[3] C. Edmonds, "Programmable Controller

Networking - Dual Cable, Redundancy, Multiple Networks and Application", *ISA 92 Advances in Instrumentation and Control*, Vol.47, Part2, October 1992.

[4] Automation Research Corporation, "Ethernet - Based Control Network Strategies, Automation Strategies Report", *Automation ARC, Three Allied Drive, Dedham, MA 02026*, October 1997.

[5] Jan - Erik Sarparanta, "Fault - Tolerance Requirements and Solutions", *Department of Computer Science Helsinki University of Technology*, May 1998.

[6] Anh Pham Hoang, Jong Myung Rhee, Se Mog Kim, and Dong Ho Lee, "A Novel Approach for Fault - Tolerant Ethernet Implementation", *Communication Engineering Department, Myongji University, Rep. of Korea, International conference, NCM 2008(IEEE CS proceeding)*, September 2008.

저자약력

김 세 목(Se Mog Kim)



2008년 부경대학교
전자공학과졸업
2007년 명지대학교 통신공학과
석박사과정
1999년~2001년 (주)네오디지털
2002년~2008년 (주)성원정보통신
2008년~현 재 (주)하나시스텍

고 윤 민(Yun Min Ko)



2008년 명지대학교
통신공학과 졸업
2008년~현재 명지대학교
통신공학과 석사과정

최 한 석(Han-Seok Choi)



2008년 명지대학교 통신공학과
졸업
2008년~현재 명지대학교
통신공학과 석사과정

민 정 현(Jung Hyun Min)



2008년 명지대학교 통신공학과
졸업
2008년~현재 명지대학교
통신공학과 석사과정

Anh Pham Hoang



2005년 HoChiMinh
University of Technology,
Computer Science and
Engineering
2008년~현재 명지대학교
통신공학과 석사과정

이 동 호(Dong Ho Lee)



1990년 경북대학교 전자공학과 졸업
1992년 경북대학교 전자공학과 석사
2005년~현재 경북대학교 전자공학과
박사과정
1992년~현재 국방과학연구소
선임연구원

이 종 명(Jong Myung Rhee)



1976년 서울대학교 전자공학과 졸업
1978년 서울대학교 전자공학과 석사
1987년 North carolin State University,
ECE Dept.(Ph. D.)
1978년~1997년 국방과학연구소
책임연구원
1997년~1999년 데이콤 연구소 부소장
1999년~2005년 하나로텔레콤 CTO
부사장
2006년~현재 명지대학교
통신공학과 교수

<관심분야> Military Communications, Fault
Tolerant System, Ad-hoc, Data
Link, Convergence