10GbE 스위치간 링크 집합을 위한 프레임 분배방식

이호영*○ 이숭희** 김연수** 조현성**
*인제대학교 광대역정보통신공학과, **인제대학교 전자정보통신공학과
2002b803@gurum.inje.ac.kr○, shlee@ijnc.inje.ac.kr, {2002b610, 2002b612}@gurum.inje.ac.kr

Frame Distribution Methods for Link Aggregation between 10GbE Switches

Hoyoung Lee*○ Soonghee Lee** Younsu Kim** Hyunsung Cho**
Dept. of Broadband Information & Communication Engineering, Inje Univ*
Dept. of Electronic Information & Communication Engineering, Inje Univ**

요 약

본 논문은 10GbE로의 망 전개에 따라 구현될 10Gb 스위치에 필요한 계층2 부가서비스들 중에서 여러 개의 물리적인 링크를 하나의 논리적인 링크로 집합하여 대역폭의 증가를 가져오는 링크 집합(Link Aggregation) 기술에 필요한 프레임 분배방식에 대한 것이다. 링크 집합 기술은 분배 알고리즘의 설계에 따라 스위치에서 성능의 차이를 가져오므로 링크 집합의 장점을 살리려면 좋은 성능을 가지는 분배 알고리즘이 필요하다. 우선 링크 집합의 간단한 구현 방법을 살펴보고 기존에 제시된 스위치와 스위치 사이에서의 프레임 분배방식인 정적/동적 프레임 분배 방식과 우리가 새로이 제안한 패딩을 이용한 분배방식을 비교한 후, 정적/동적 프레임 분배 방식과 패딩을 이용한 분배방식을 컴퓨터 시뮬레이션을 통해 성능을 비교하였다.

I 서 론 *

인터넷 서비스의 급속한 확산으로 인해 초고속 통신망의 구축에 대한 사용자의 요구가 증대되고 있고 대도시의 주요 빌딩군으로 데이터 트래픽이 집중됨에 따라 MAN 구간의 병목 현상이 가중되고 있다. 또한, 초고속 인터넷 서비스 이용자들이 발생시키는 트래픽 양은 일반 인터넷 서비스 이용자들이 발생시키는 트래픽 양을 훨씬 초과하고 있어서 보다 큰 대역폭과 고속의 링크 제공이 요구되었다.

이를 위해 현재 망에서 기존의 링크를 사용하여 사용자가 원하는 대역폭을 얻기 위한 방법으로 802.3ad로 표준화된 링크 집합 기술을 적용할 수 있다. 이 기술은 여러 물리적인 링크들을 집합하여 하나의 논리적인 링크로 사용하는 기술이다. 아래의 그림 1은 망을 구성하는 각각의 장치 사이에 적용된 집합된 링크를 보여주고 있다.

그림 1에서 볼 수 있듯이 링크 집합은 스위치와 스위치, 스위치와 서버, 서버와 서버 사이에서 하드웨어 업그레이드 없이 적용될 수 있다.[3]

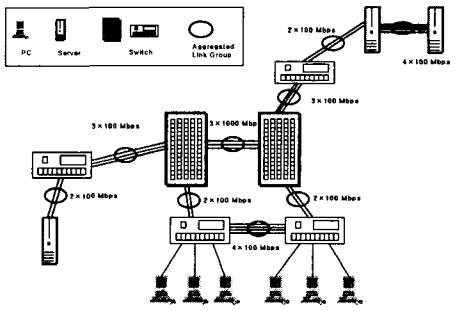

그림 1. 망 상에서의 링크 집합

이러한 링크 집합에서는 장치들 간에 여러 개의 물리적인 링크가 하나의 논리적인 링크로 연결되어 있기 때문에 프레임들을 적절히 분배하여 전송하여야 한다. 분배를 위한 알고리즘은 표준안에 규정되어 있지 않다. 단,

- 해당 포트로 프레임이 전송되어질 때 프레임의 순서가 바뀌면 안 된다.
- 프레임의 중복이 발생하면 안 된다.

위의 두 사항을 고려하여 분배 알고리즘을 설계하여야 한다[1].

* 본 연구는 한국전자통신연구원 위탁과제로 수행되었습니다.

분배 알고리즘의 설계에 따라서 링크 집합 기능을 제공하는 스위치의 성능이 많이 달라질 수 있다. 이러한 분배 알고리즘으로는 기존에 제시된 스위치와 스위치 사이에서의 정적/동적 프레임 분배 방식이 있으나, 수신측에서 집합된 링크를 모두 사용하지 못하는 단점이 있다[2]. 따라서 이 방식을 보완한 패딩을 이용한 분배방식을 제안하였다.

본 논문에서는 II장에서 기존의 프레임 분배방식과 패딩을 이용한 분배방식을 비교하고 III장에서는 컴퓨터 시뮬레이션을 통해 기존의 분배방식과 패딩을 이용한 분배방식을 비교하여 성능 분석 후 IV장에서 결론을 맺는다.

II 정적/동적 분배 방식과 패딩을 이용한 방식의 비교

기존의 정적/동적 프레임 분배 방식은 수신측의 MAC 주소에 따라 링크를 고정하여 프레임의 순서 불일치를 해결한다. 그러나 이 분배 방식은 수신측 단말 입장에서는 링크가 고정되어 있기 때문에 링크 집합에 적용한 링크를 모두 사용하지 못해 링크기술의 장점을 살리지 못한다[2]. 이를 해결하기 위해 링크 집합 기술의 장점을 살리면서 프레임의 순서 불일치도 일어나지 않도록 하기 위한 방법으로 패딩을 이용한 분배 방식을 제안한다.

이 방식은 프레임을 여러 링크로 보내기 전에 프레임을 최고 길이로 패딩을 하여 순차적으로 전송을 하게 되어 여러 개의 링크 중 어느 링크로 전송하더라도 프레임의 순서 불일치는 생기지 않으며 수신측에서 보더라도 링크를 모두 사용할 수가 있다. 또한, 구현 시 기존의 분배 방식보다 더 간단하다. 원래 이 방식은 스위치와 스위치 사이에 적용되던 기술로서 스위치와 서버 또는 서버와 서버사이에 소프트웨어에 의해서 구현해야 하는 경우에는 적합하지 않은 것으로 여겨졌다. 그러나 10GbE 스위치 등에서 Network Processor를 이용한 경우를 가정하면 이 방식의 적용이 가능해진다.

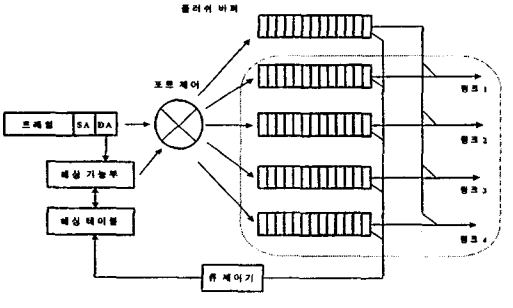

그림 2 정적/동적 프레임 분배방식의 예

그림 2와 그림3은 정적/동적 프레임 분배방식과 패딩을 이용한 분배방식을 나타낸 것이다[2]. 패딩을 이용한 방식은 최고 길이로 패딩을 하고 있기 때문에 해싱 기능부, 해싱테이블 및 큐 제어기가 필요가 없게 된다. 또한, 4개의 링크로 순차적으로 프레임을 분배하므로 프레임의 순서 불일치는 나타나지 않으며 기존 방식보다 훨씬 간단하다. 하지만 스위치에서 프레임을 최대 프레임 크기로 패딩해야 하기 때문에 프레임의 크기가 작을 경우 링크의 효율이 떨어지는 단점이 있다.

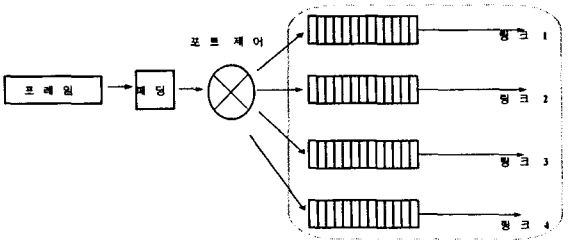

그림 3. 패딩을 이용한 분배방식의 예

III 성능 분석

기존의 정적/동적 프레임 분배방식과 제시한 패딩을 이용한 분배 방식을 컴퓨터 시뮬레이션을 통해 성능 분석을 하여 비교하였다.

시스템의 모델링은 송신측과 수신측 사이에 4개의 링크로 연결되어 있으며 각 링크는 프레임이 전송되는 동안 오류 없이 이상적으로 동작한다고 가정하였다. 각 프레임은 지수분포를 따르는 길이를 가지도록 발생시킨 다음 LAN 환경에서와 같이 프레임 형태를 얻기 위해 프레임의 최대 길이를 1,500 바이트, 프레임의 최소 길이를 64 바이트로 제한하였다. 따라서, 발생된 프레임의 서비스 시간 분포는 이론적인 지수 분포를 따르지 않는다. 그리고 입력 트래픽은 MMPP(Markov Modulated Poisson Process) 분포를 따르며 시뮬레이션 시 전송되는 프레임 수는 800,000만개(2초에 해당)로 가정한다. 스위치에서는 프레임들을 최대 프레임 길이로 패딩을 하고, 패딩된 프레임들은 4개의 링크에 1에서 4번 순으로 차례대로 분배된다고 가정하였다. 이때, 송신측 스위치의 각 링크별 출력 버퍼 크기는 4K 바이트와 8K 바이트 등 두가지의 시스템을 가정하고, 시뮬레이션에서는 수신측이 각각의 링크에 무한버퍼를 갖는 시스템으로 가정한다.

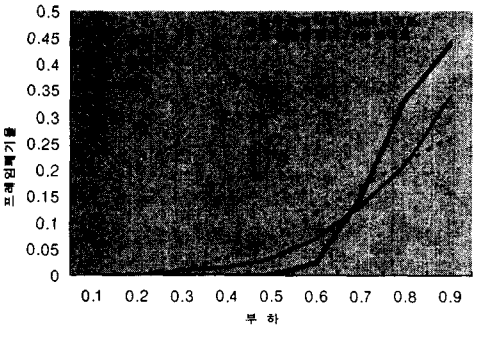

그림 4. 정적 / 동적 프레임 분배 방식과 패딩 방식의 프레임 폐기율
(버퍼크기 4 K바이트, 평균 프레임 길이 929 바이트)

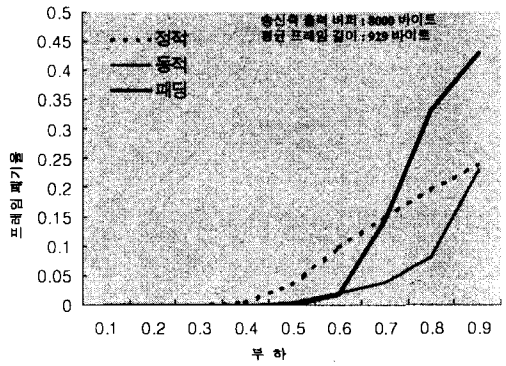

그림 5. 정적 / 동적 프레임 분배 방식과 패딩 방식의 프레임 폐기율
(버퍼크기 8 K바이트, 평균 프레임 길이 929 바이트)

그림 5와 그림 6은 정적/동적 프레임 분배 방식과 패딩을 이용한 분배 방식의 프레임 폐기율을 나타내고 있다. 패딩을 이용한 분배 방식은 버퍼 크기와 상관없이 프레임 폐기율이 거의 같은 것을 알 수 있다. 부하가 0.66일 때까지는 정적/동적 분배 방식보다 프레임 폐기율이 작은 것을 볼 수 있으나 부하가 0.66이상부터는 정적/동적 분배 방식보다 프레임 폐기율일 커짐을 볼 수 있다. 이는 부하가 증가할수록 출력 버퍼내의 프레임들이 패딩에 의한 오버헤드가 급속도로 증가하기 때문에 폐기율이 정적 및 동적 분배 방식보다 커지게 된다.

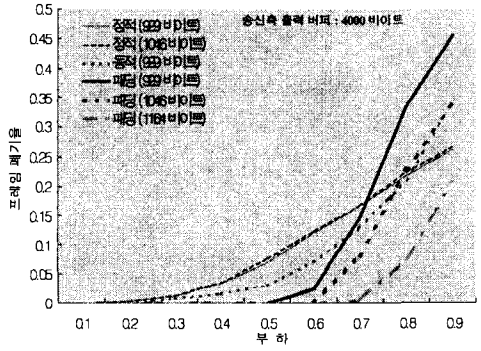

그림 6. 평균 프레임 크기에 따른 프레임 폐기율
(버퍼크기 4 K바이트)

그림 6은 정적 및 동적 분배 방식과 패딩 분배 방식을 평균 프레임 크기를 929, 1,046, 1,164 바이트로 달리 하여 프레임 폐기율을 알아보았다. 정적 분배 방식은 프레임의 크기에 따른 프레임 폐기율의 변화가 작고 동적 분배 방식은 부하가 0.8까지는 정적 분배 방식보다 폐기율이 작지만 그 이상에서는 많아진다. 패딩을 이용한 분배 방식은 평균 프레임 크기가 929 바이트는 부하가 0.66까지는 프레임 폐기율이 작지만 그 이상일 때는 정적/동적 분배 방식보다 많아지는 것을 볼 수 있고 1,046 바이트에서는 부하가 0.75이상일 때 정적/동적 분배 방식보다 프레임 폐기율이 많아진다. 또한 평균 프레임 길이가 1,164 바이트에서는 정적 및 동적 분배 방식보다 프레임 폐기율이 적은 것을 알 수 있었다. 이는 같은 전송시간에서 프레임 길이가 크면 출력 버퍼에서 출력되는 데이터 크기가 커지므로 출력 버퍼에서 받아 들일 수 있는 데이터 양이 많아지기 때문이다. 이를 볼 때 평균 프레임 크기가 커질수록 정적/동적 프레임 분배 방식보다는 패딩을 이용한 분배 방식의 성능이 더 효율적이다.

III 결론

링크 집합의 구현에 중요한 역할을 하는 프레임 분배 방식 중 기존의 정적/동적 프레임 분배 방식과 새로이 제시한 패딩을 이용한 분배 방식을 비교하여 성능 분석을 하였다. 기존의 분배 방식보다 패딩을 이용한 분배 방식은 링크 집합의 이점을 살릴 수 있고 구현 또한 간단해진다. 컴퓨터 시뮬레이션을 통해 버퍼 크기에 따른 프레임 폐기율을 확인하였고 평균 프레임 크기에 따른 프레임 폐기율을 비교 분석하였다.

패딩을 이용한 분배 방식은 프레임의 크기가 작은 경우에는 기존의 방식에 비해 전송 효율이 낮은 문제점이 있지만 프레임의 크기가 클 경우에는 전송 효율이 훨씬 좋아지는 것을 알 수 있었다. 이는 길이가 큰 파일일 경우에는 충분히 성능을 달성 할 수 있음을 말해준다.

앞으로 좀 더 나은 성능을 가지는 분배 알고리즘을 위해서는 링크의 추가 및 삭제 등 여러 경우에 대한 구체적 방안 등에 대한 연구가 추가로 진행되어야 한다.

참고 문헌

[1] IEEE, IEEE Standard 802.3, 2000 Edition
[2] 전우정, 윤종호, "통합링크기능을 가진 매체접근제어 기용 프레임 분배 방식의 성능분석", 한국통신 학회지, 2000. 3.
[3] Link Aggregation according to IEEE 802.3ad White Paper, 2002
[4] IEEE P802.3ad Link Aggregation Task Force Public archive area, http://grouper.ieee.org/groups/802/3/ad/public/
[5] Solving Server Bottlenecks Link Aggregation/FEC*/GEC* http://www.pentium.co.kr/network/connectivity /solutions/server_bottlenecks/bot_sol2.htm
[6] Link Aggregation 802.3ad, http://www.riverstonenet.com/support/configdb/0032.html