《主 題》

고속 LAN

金 炯 秀* · 李 光 哲* · 申 炳 喆* · 尹 鍾 浩**
(*한국과학기술원(KAIST) **한국항공대학)

■차 례■

Ⅰ. 서론
Ⅱ. Fast Ethernet
Ⅲ. HIPPI(High-Performance Parallel Interface)
Ⅳ. 고속 Ring Network
Ⅴ. 기타 고속 LAN
Ⅵ. 결론

Ⅰ. 서 론

네트워크 및 디스크 시스템과 같이 컴퓨터에 부착되는 것들과 컴퓨터들의 처리 속도는 언제나 더 빨라지고 가격은 더 싸지고 있다. 이러한 부분시스템들의 성능 향상은 산업에 큰 변화를 초래한다. 컴퓨터 기술이 발전하여 여러 사람이 그것을 동시에 사용할 수 있을 만큼 강력하게 되었을 때 timesharing 기술이 발전했고 각 개인들이 그것들을 구입할 수 있을 만큼 컴퓨터 구성 요소들의 가격이 내려갔을 때 퍼스널 컴퓨터의 출현을 볼 수 있었으며, Ethernet와 같은 값싼 LAN이 발전하게 되자 분산 시스템과 분산 파일 시스템이 나타나게 되었다.

또다른 큰 변화들도 나타나고 있다. 몇년전에는 슈퍼 컴퓨터들이 할 수 있었던 것들을 이제는 저가의 중앙 처리 칩(CPU)들이 할 수 있을 만큼 급격히 처리 속도가 빨라지고 있다. DEC(Digital Equipment Corporation)의 Alpha[1]와 같은 칩들은 단지 몇 nanoseconds의 instruction cycle 시간을 가지며, 앞으로 일초에 수십억의 instruction을 수행할 수 있도록 하는 CPU도 곧 가능하게 될 것이다. 계산량이 매우 방대하고 슈퍼 컴퓨터나 고성능의 그래픽스 워크스테이션에서나 가능하다고 여겨졌던 voice recognition, high-quality real-time computer graphics, handwriting recognition, 그리고 real-time animation 및 video와 같은 멀티미디어 서비스들이 보통의 퍼스널 컴퓨터에서 가능할 만큼 이러한 프로세서들은 충분한 계산 능력을 가지고 있다. 이러한 새로운 멀티미디어 응용 서비스들은 그 데이타를 전송하기 위하여 강력한 네트워크의 출현을 재촉하고 있다.

본 연구에서는 급속히 변화하는 여러가지 고속 LAN 시스템에 관해서 살펴보고자 한다. 고속 LAN 기술중 이미 표준화도 잘 되어 있고 우리에게 친숙한 것으로는 FDDI-I과 DQDB등이 있는데 이는 여러 논문에서 이미 소개가 되어 있으므로 본 연구에서는 다루지 않았다. 본 연구에서는 우선 표준화 및 상용화가 급속히 진전되고 있는 Fast Ethernet을 검토하였다. 또한 고속의 point-to-point 접속 방식으로 활용되어 LAN이라고 보기가 다소 어렵지만 슈퍼 컴퓨터 사이의 고속 접속 표준으로 사용되는 HIPPI에 관해서 살펴보고자 한다. 광대역 장거리 통신망에서 ATM방식의 표준으로 채택되면서, LAN에서도 ATM의 cell format을 따르는 버스 방식과 링(ring)방식의 표준안이 거론되고 있는바, 버스 방식으로는 이미 DQDB가 ATM cell format을 따라 표준화되었다. 링 방식에 있어서는, 전통적으로 링 방식에 강한 IBM등이 제안한 포로토콜

을 놓고 미국내 ANSI(American National Standards Istitute)에서 표준화 작업을 진행중이다[27]. 그러나 국제 표준을 제정하는 ISO/JTC1에서는 ATM과 유사한 format을 채택한 링 방식의 LAN 표준안으로 일본과 한국에서 제안한 ATMR(ATM Ring) 및 HMR (High-Speed Multimedia Ring)을 놓고 의견 조정중이다. 본 연구에서는 링 방식의 고속 LAN으로서 그동안 많이 연구되어온 Orwell Ring, FCR(Fast Cambridge Ring)과 Meta Ring, 그리고 ATMR 및 HMR등에 관하여 소개하고자 한다. 이외에도 널리 알려지지 않은 고속 LAN으로서 HANGMAN, Manhattan Street Network, AN2 Switch, Atomic LAN, CSMA/RN ring등을 대략적으로 살펴보고자 한다.

Ⅱ. Fast Ethernet[24][25][26]

2.1 서 언

Fast Ethernet은 이미 참고문헌[24]에서 소개되었으므로 여기서는 간단히 요약만 하고자 한다. 광섬유를 이용한 고속LAN이 다수 등장하고 있는 가운데 건물내에 설치된 기존의 케이블을 그대로 사용하면서 현재 보다 높은 전송 속도를 얻을 수 있는 fast ethernet에 대한 연구가 활발히 이루어지고 있다. Fast ethernet은 기존의 UTP(Unshielded Twisted Pair) 케이블을 그대로 사용하면서 반경 100 m 이내에 있는 PC 나 워크스테이션들을 100 Mbps의 고속으로 연결 시켜줄 수 있기 때문에 저렴한 고속 LAN 시스템을 원하는 많은 사용자들의 관심을 끌고 있다.

Fast Ethernet 프로토콜로 현재 IEEE802 위원회에서 표준안으로 검토중인 것은 100 Mbits/s CSMA/CD 프로토콜과 100 BASE-VG(Voice Grade) 프로토콜 두가지인데 모두 그림 2-1과 같이 허브를 중심으로한 성형 구조를 갖는다. 이 구조는 기존에 사용중인 10 BASE-T CSMA/CD와 동일하다. 이 두가지 프로토콜의 차잇점은 100Mbits/s CSMA/CD가 10BASE-T CSMA/CD와 같은 CSMA/CD 방식의 MAC 프로토콜을 사용하는데 반하여 100BASE-VG에서는 충돌이 발생하지 않는 round-robin 방식의 새로운 MAC 프로토콜을 사용하는 것이다.

두가지 프로토콜에 대하여 살펴보기에 앞서서 Electronics Industries Association/Telecommunications Industries Association(EIA/TIA)-568에서 규정한 케이블 종류에 대하여 잠깐 살펴본다. EIA/TIA-568 규정은 표 2-1과 같은 4 가지 종류의 케이블을 규정한다.

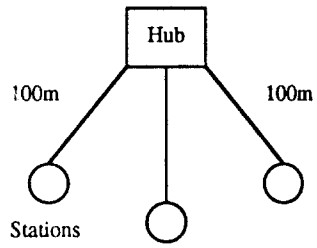

그림 2-1. Network Topology

케이블 종류	최대전송속도	특 징
Category 3 UTP	10Mbits/s	100-ohm, 4-pair
Category 4 UTP	16Mbits/s	100-ohm, 4-pair
Category 5 UTP	100Mbits/s	100-ohm, 4-pair
STP	155Mbits/s	150-ohm, 2-pair
Coaxial	100Mbits/s 이상	50-ohm
Optical fiber	고속	62.5 또는 125마이크론

표 2-1. EIA/TIA-568 케이블의 종류 및 특징

Category 3 UTP 케이블은 ethernet 및 4 Mbits/s 토큰링에 사용된다. Category 4 UTP 케이블은 16 Mbits/s 토큰링에 사용된다. Category 5 UTP 케이블은 CDDI (FDDI over copper)용으로 적합하다. STP는 ATM 용으로 적합하다.

2.2 100 Mbits/s CSMA/CD 프로토콜

이 프로토콜은 802. 3 CSMA/CD의 부가 표준으로 제안 되었으며 MAC 부분은 기존의 CSMA/CD 방식을 따르지만 라인 코딩은 100 Mbps에 부적합한 기존의 맨체스터 코드 대신에 새로운 라인 코딩 방식인 4B/5B, 5B/6B, 8B/6T, 7B/5T등을 사용하였다. 이러한 코딩 방식들은 아직까지 표준화가 이루어지지 않고 회사 마다 다른 코딩 방식을 사용하고 있다. 이 프로토콜의 장점은 10 Mbits/s CSMA/CD와 100 Mbits/s CSMA/CD를 같이 사용할 수 있다는 것이다. 즉 100 Mbits/s용 케이블을 다시 설치할 필요 없이 기존에 설치된 케이블에 100 Mbit/s CSAM/CD 카드와 허브를 설치하여 100 Mbps 속도를 얻을 수 있다.

2.3 100BASE-VG 프로토콜

이 프로토콜은 CSMA/CD 방식과는 전혀 다른 MAC 방식을 사용하며 802. 12라는 새로운 WG에서 연구할 예정이다. 이 프로토콜의 특징은 충돌이 발생

하지 않는 round-robin 방식의 demand assignment MAC 프로토콜을 사용하며 기존의 802. 3 MAC 프레임을 그대로 사용하고 category 3 UTP 케이블을 사용할 수 있다. 각 스테이션은 2 종류의 우선순위 중 하나를 가질 수 있으며 전송하는 패킷은 CSMA/CD 에서 처럼 broadcast 되는 것이 아니라 지능을 가진 허브가 수신 주소를 식별하여 목적지로만 전달 된다. 100BASE-VG MAC 프로토콜의 동작은 그림 2-2와 같이 이루어 진다.

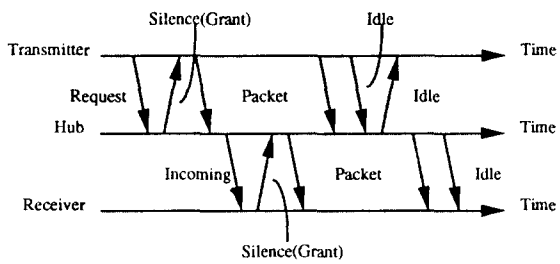

그림 2-2. 100BASE-VG 프로토콜의 동작

Ⅲ. HIPPI(High-Performance Parallel Interface)

HIPPI는 25m 길이까지의 동축 케이블에서 800Mb/s 또는 1600Mb/s의 데이타율로 전송할 수 있는 단방향 일대일 접속이다. 그림 3-1은 HIPPI 스위치를 통해 2개의 접속이 형성된 HIPPI 스위치의 한 예이다. 논리적인 쌍방향 전송은 두개의 HIPPI 접속으로 가능하다. 이것은 네트워크 프로토콜은 아니며 단지 두 장치들을 연결하는 데이타 채널이다. 하지만 HIPPI 스위치는 LAN으로서 많은 컴퓨터 시스템을 서로 연결할 수 있도록 지원한다[9]. 이것은 1987년 1월에 ANSI(American National Standards Institute) X3T9. 3의 WG(Working Group)에 의하여 제안되었으며 1988년 말쯤에 처음으로 구현되었다. 개발에 있어 몇가지 목표들은 데이타를 전송하기 위한 "fire hose"를 제공하는 것, KISS(Keep It Sweet and Simple), 빠른 시일내에 완성하는 것, 선택 사항을 적게 사용하는 것, 새로운 silicon을 요구하지 않는 것 등이었다. 그러므로 물리 계층에서의 800 또는 1600 Mb/s의 선택 사항과 다른 계층에서의 몇가지 선택 사항이 있으나 가능한 한 최소로 하였다.

HIPPI는 그림 3-2과 같이 HIPPI-PH, HIPPI-SC,

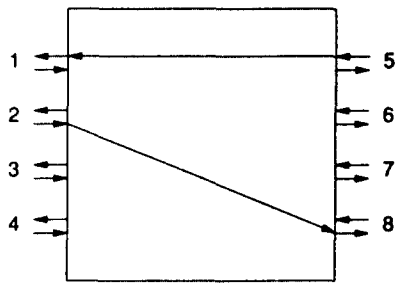

그림 3-1. A HIPPI Switch with Two Connections

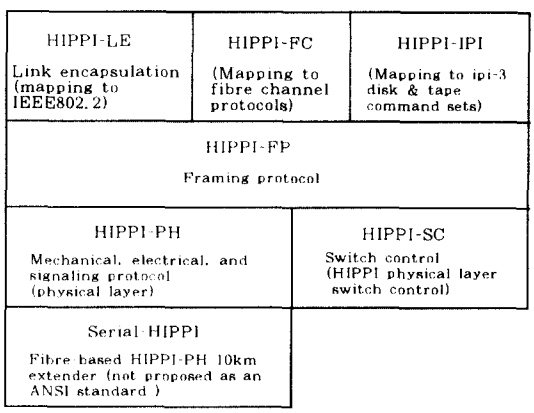

그림 3-2. HIPPI documents

HIPPI-FP, HIPPI-LE, HIPPI-FC, HIPPI-IPI, 그리고 Serial HIPPI 등으로 구성되어 있다.

HIPPI-PH(Physical Layer)는 25m 정도 떨어져 있는 두 지점사이에 50 쌍의 twist -pair 케이블을 규정하며 800 Mb/s의 경우 한 케이블을 사용하고 1600 Mb/s의 경우 두개의 케이블을 사용한다. 제어 신호와 데이타 정보들은 차등(differential) emitter-coupled logic(ECL) drivers를 사용하며 연속적인 25-MHz clock 신호에 타이밍을 맞춘다. 여기에서 사용되는 신호들은 다음과 같다.

· REQUEST : 송신단이 접속 확립을 요구한다.

· CONNECT : 수신단이 REQUEST에 대해 접속을 인정한다.

· READY : 송신단이 burst를 보내도록 한다.

· PACKET : 하나 또는 그 이상의 bursts를 하나의 패킷으로 묶는다.

· BURST : 256 words를 한 burst로 묶는다.

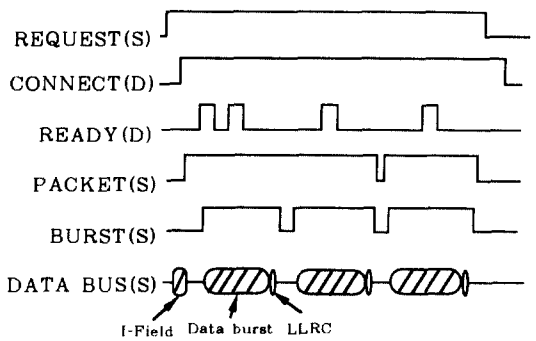

그림 3-3. Typical HIPPI waveforms

· DATA BUS : 32 또는 64비트

· PARITY BUS : DATA BUS의 odd byte parity 체크를 위한 4 또는 8비트

· CLOCK : 연속적인 25MHz, 40ns의 주기

· INTERCONNECT : 케이블이 연결되었고 power가 on되었음을 알린다.

그림 3-3은 HIPPI의 신호 파형의 예로 접속을 확립하고 2 bursts로 된 패킷과 1 burst의 패킷을 보낸 후 접속을 단절하는 과정을 보여 준다. 여기서 (S)는 송신단이 보내는 신호를, (D)는 수신단이 보내는 신호를 가리킨다. HIPPI는 전화 접속과 비슷하게 송신단에서 REQUEST신호를 보낸 후 수신단으로부터 CONNECT 신호를 받으면 접속이 확립되어 데이타 전송을 하게 된다. 접속 확립을 위해 송신단에서 REQUEST 신호와 함께 DATA BUS에 I-Field를 같이 보내는데 스위치에서는 I-Field를 토대로 접속할 출력 포트를 결정한다. 만약 접속하고자 하는 수신단이 다른 곳과 연결되어 있다면 스위치는 송신단에게 접속이 불가능하다는 것을 알리고 송신단은 다시 접속을 요구하든지 아니면 다른곳으로 접속을 요구할 수 있다. 또는 수신단과 다른 곳과의 접속이 끝날 때까지 기다리게 한 후 접속이 확립된다. 이것은 I-Field의 한 비트를 두어 위의 두 방법중 하나를 선택할 수 있게 한다. 접속이 확립되면 데이타를 그림 3-4에서와 같이 패킷의 형태로 전송하는데 패킷은 일련의 burst들로 구성되어 있다. Burst들은 다시 256 words로 구성되며, 1 word는 800Mb/s의 경우 32비트이고 1600Mb/s의 경우 64비트이다. 패킷에 있는 burst는 처음 것과 마지막 것일 때만 256 word 보다 작을 수 있다. 그리고 패킷에 있는 burst 수는 가변적일 수 있다. 데이타 전송이 끝난 후 송신단이나 수신단에서 REQUEST나 CONNECT신호를 단절함으로써 접속을 끊을 수 있다.

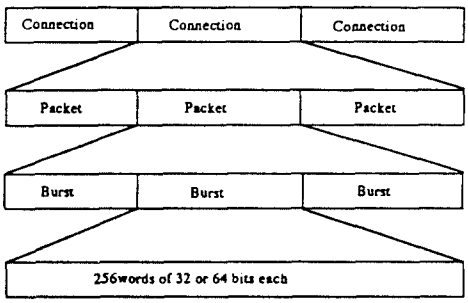

그림 3-4. HIPPI data frame

그림 3-3에서 LLRC는 각 burst에 대한 1 word의 even-parity checksum이다. HIPPI 수신단은 송신단이 256 words 까지의 burst를 전송할 수 있음을 알리기 위하여 그림 3-3에서처럼 READY 신호를 보낸다. 수신단은 현재 자신이 버퍼에 저장할 수 있는 능력에 따라 여러개의 READY 신호들을 송신단에 보낼 수 있다. 상위 계층에서의 Window와 유사하게 송신단에서 이 여러개의 READY 신호들은 큐에 저장되고 데이타 전송이 필요할 때마다 round trip handshake 지연시간 없이 이미 저장된 READY 신호를 이용하여 데이타 burst들을 보낼 수 있다.

HIPPI는 실시간 주변 장치에 대해 지원할 수 있도록 하는 제안이 있었으나 그러기 위해서는 새로운 신호들을 더 첨가해야 하기 때문에 그 제안은 거절되었다. 이것은 앞에서도 언급했듯이 KISS 원칙에 따라 되도록 구현을 간단하게 하기 위한 것이다.

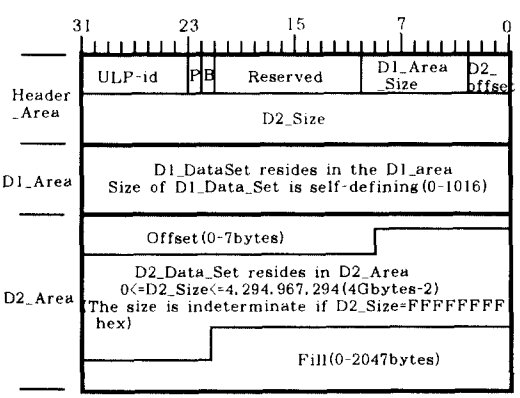

그림 3-5. HIPPI packet format

HIPPI-PH는 간단히 BYTE에 관한 parity check와 burst(수평, 수직적으로)에 관한 parity check를 근거로 한 에러 검출을 규정한다. 이것은 한 burst안에 있는 모든 1-비트, 2-비트, 3-비트, 그리고 홀수 비트 에러를 검출할 수 있으며 4-비트의 경우 99.99 퍼센트까지 가능하다. 64 비트 CRC(Cyclic Redundancy Check)를 사용하자는 제안이 개발 끝무렵에 있었으나 HIPPI가 이미 낮은 에러율을 가지고 있다고 가정할 수 있었고 그 제안을 수용하려면 새로운 칩이 필요했고 하드웨이를 다시 설계해야 하며 HIPPI를 구현하는 사람들에게 HIPPI 규정이 결점을 가지고 있으며 안정하지 않다는 것을 암시할 수 있다는 등의 이유로 기설되었다.

HIPPI-FP(Framing Protocol)은 데이타를 전송하기 위한 패킷의 형식을 그림 3-5과 같이 규정한다. 패킷은 D1과 D2의 두 데이타 부분과 헤더 부분으로 나누어진다. D1은 1016 bytes로 구성된 제어 정보들을 가지며 패킷의 첫 burst에서만 P = 1로 한 후 포함된다. D2는 사용자 데이타 부분으로 4 gigabytes-2bytes 까지 규정할 수 있으며 길이가 무한한 데이타의 흐름이나 크기를 알 수 없는 것에 대해서는 정의되어 있지 않다. 데이타 부분을 D1과 D2로 나눈 이유는 패킷이 전송되어 오는 동안에 D1의 데이타 정보를 상위계층으로 먼저 보내어, D2의 사용자 데이타 부분이 계속 전송되어 오는 동안에 상위 계층에서 D1의 제어 정보를 이용할 수 있게 하기 위한 것이다. D2의 offset부분은 D1과 D2를 word단위로 나눌 수 있도록 0에서 7bytes까지 둘 수 있다. 헤더 부분의 ULP-id는 패킷이 속하는 상위 계층 프로토콜 확인자이고 B 비트는 D2 데이타 부분이 첫번째 burst 안에서 시작되는지 두번째 burst의 첫부분에서 시작되는지를 나타낸다. D2가 burst의 경계에서 시작하도록 하는 것은 제어 정보와 데이타를 분리하여 각각 다른 버퍼에 저장하기가 쉽기 때문이다.

HIPPI-LE(Link Encapsulation)는 HIPPI 상에서의 전송을 위하여 IEEE 802.2 LLC 데이타에 대한 헤더 부분을 규정한다. 이것은 TCP/IP와 같은 많은 통신 프로토콜이 HIPPI를 이용할 수 있도록 한다. 송신단과 수신단에서 각각 48비트의 IEEE Universal LAN 주소를 가지고 있는 HIPPI-LE 헤더는 HIPPI-FP D1 데이타 부분에 있고 IEEE 802.2와 다른 정보들은 D2 데이타 부분에 존재한다. HIPPI-FC(Mapping to Fibre Channel)은 HIPPI를 근거로 한 장치와 광 채널 장치 사이를 서로 연결시키기 위한 방법을 규정하며 HIPPI의 상위 계층 프로토콜이 광 채널로 구성된 물리 계층상에서 전송을 할 수 있도록 mapping을 제공한다.

HIPPI-IPI(Mapping to IPI-3)은 HIPPI를 통하여 IPI(Inteligent Peripheral Interface)제어 방식에 의해 IPI-3 Disk와 IPI-3 Tape 명령들을 전송하는 것을 규정한다. 이것은 현재 HIPPI를 통해 연결되어 있는 고성능의 저장 시스템을 지원하기 위하여 사용되고 있다.

HIPPI-SC(Switch Control)은 HIPPI의 단방향 일대일 접속에 대해 그림 3-1과 같이 LAN에서 많은 장치들을 서로 연결시키기 위해 사용되는 크로스 바 스위치의 사용과 제어를 규정한다. Serial-HIPPI(Fibre Extender)는 10Km까지의 거리에 대하여 투명한(transparent) 광 링크를 정의하며, 이는 구현하는 사람들 사이의 약속이지 표준은 아니다.

HIPPI는 Multicasting을 위한 기능이 부족하다. Multicasting을 위해 multicasting 서비스를 제공하기로 되어 있는 서버 포트에 직접 특별한 multicast 포트를 지원할 수는 있다. 하지만 비디오나 오디오 회의를 위해 만족할 만한 multicast기능을 제공하지는 못한다.

Ⅳ. 고속 Ring Network

4.1 CFR(Cambridge Fast Ring)[14]

4.1.1 서언

CFR은 slotted ring 프로토콜을 사용하며 송신노드 소거 방식을 사용한다. 즉 링을 순환중인 패킷은 링을 1회 순환한 후 그 패킷의 송신노드에서 제거된다. 시스템의 최대 처리율은 링크 용량을 초과할 수 없으며 데이타를 전송하고 나면 그 데이타가 다시 자기에게 돌아올 때까지는 빈 슬롯이 도착되더라도 다음 데이타를 전송할 수 없기 때문에 한 노드에서 전송할 수 있는 최대 처리율은 링 내의 슬롯 수에 반비례하여 작아진다.

CFR에서는 송신노드로 되돌아온 슬롯을 송신노드에서 다시 사용할 수 있느냐 혹은 없느냐에 따라서 두 가지 종류의 슬롯을 사용한다. 하나는 normal slot이고 다른 하나는 channel slot 이다. Normal slot은 되돌아온 슬롯을 송신노드에서 사용할 수 없으며 empty slot으로 만들어서 downstream node로 보내주어야 한다. Channel slot은 되돌아온 슬롯을 송신노드에서 다시 사용할 수 있다. 그러나 되돌아온 슬롯을 송신노드에서 사용할 수 있기 위해서는 다음과 같은 문제가 해결되어야 한다.

CFR에서는 송신노드에서 슬롯의 마지막 비트까지 수신한 후에야 패킷이 제대로 수신측에 전달되었는지 알 수 있고 다음 패킷을 전송할 수 있다. 이것은 토큰링에서 처럼 수신노드에서 패킷을 잘 받았다는 표시를 슬롯의 끝에다 표시하기 때문이다. 만일에 송신노드에서 마지막 비트까지 확인하지 않고 패킷을 전송한 후에 이전 패킷에 오류가 있음을 알고 다시 패킷을 전송한다면 수신노드에서는 순서가 바뀐 패킷을 수신하게 된다. 이 문제를 해결하기 위해서는 송신노드에서 마지막 비트를 받을 때까지 슬롯 전송을 중지하는 방법, 수신노드의 상위 계층에서 해결하는 방법, 그리고 슬롯에 있는 response field를 사용하여 취소 시키는 방법등이 있다. 이중 첫번째 방법은 노드에서의 지연이 너무 많이 커질 염려가 있고 두번째 방법은 상위 계층을 복잡하게 할 우려가 있어서 CFR에서는 세번째 방법을 택하였다.

슬롯의 크기를 결정하는 것도 중요한 문제인데 너무 크게 잡으면 짧은 데이타들을 전송하는 경우에 남은 부분들을 padding해서 전송해야 하기 때문에 대역폭의 낭비가 심하고 음성과 같은 경우에는 지연시간이 길어지고 한 패킷을 잃었을 때 미치는 영향이 커지는 단점이 있다. 반면에 슬롯의 크기를 너무 작게하면 슬롯 헤더로 인한 대역폭의 낭비가 많아지고 링내의 슬롯 수가 많아지기 때문에 한 노드에서 전송할 수 있는 최대 처리율이 작아지는 단점이 있다. 이상의 문제점을 고려하여 CFR에서는 한 슬롯에 들어가는 데이타의 크기를 32 byte로 정하였다.

4.1.2 슬롯 구조

CFR의 슬롯 구조는 그림 4-1과 같다. 슬롯의 헤더는 6 bytes를 사용하고 데이타 필드는 32 bytes를 사용하였다. 물리 링크는 100 MHz로 동작되는데 두개의 슬롯을 사용하고 슬롯 사이에 2 bytes 간격을 두면 최대 80 Mbps의 대역폭을 얻을 수 있다.

CFR에서는 channel mode에서도 데이타들 간에 순서를 보장해주기 위하여 다음과 같은 독특한 에러 제어 방법을 사용한다. 수신노드에서 데이타를 정확히 받으면 bad CRC로 만들어서 송신노드로 전송한다. Bad CRC를 만드는 방법은 CRC 필드의 마지막 4 비트를 inverting 함으로써 만든다. 수신노드에서 데이타를 정확히 받지 못하면 good CRC를 만들어서 송신노드로 전송한다. 송신노드에서는 bad CRC를 받으면 수신노드에서 데이타를 정확히 받은 것으로 알고 재전송을 하지 않는다. 송신노드에서 good CRC를 받으면 오류가 발생한 것으로 간주하여 보내고 있는 데이타를 취소시킨다. 취소시키는 방법은 bad CRC를 전송함으로써 한다. 이를 다시 요약하면 수신노드에서는 정확한 데이타가 수신되면 이를 받아들이고 bad CRC를 전송하며 잘못된 데이타가 수신되면 이를 받지 않고 CRC를 고쳐서 good CRC를 전송한다. 송신노드에서는 bad CRC를 받으면 수신노드에서 데이타를 잘 받은 것으로 간주하고 good CRC를 받으면 에러가 발생한 것으로 간주하여 전송중인 데이타를 취소시키고 에러가 발생한 데이타를 재전송한다.

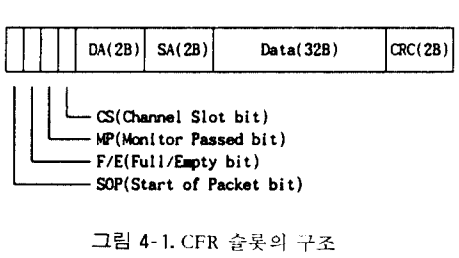

그림 4-1. CFR 슬롯의 구조

4.2 Orwell Ring[15]

4.2.1 서언

Orwell ring은 Cambridge slotted ring의 시뮬레이션 모델로부터 만들어졌다. Orwell ring의 특징을 살펴보면 slotted ring 프로토콜이며 수신노드소거 방식에 의한 공간 재사용(spatial reuse) 개념을 사용하여 링크 용량보다 높은 처리율을 얻을 수 있도록 하였다. 시뮬레이션으로 수행해본 결과 140 Mbits/s Orwell ring으로 최대 지연을 200 s 이내로 제한하는 조건에서 170 Mbits/s 까지 처리율을 얻을 수 있었다. 음성과 같이 지연에 민감한 서비스는 호설정 전에 미리 최대 액세스 지연을 계산하여 QoS(Quality of Service)를 만족시켜 줄 수 있는지 판단한다. 만약 만족시켜 줄 수 있다고 판단되면 호를 연결시키고 대역폭을 할당해준다. 데이타와 같이 어느 정도의 지연을 허용하는 서비스는 지연에 민감한 트래픽을 모두 전송한 후에 남은 대역폭을 사용한다. 이와 같은 방법으로 Orwell ring에서는 음성과 데이타를 모두 서비스해 줄 수 있다. 노드간에 발생할 수 있는 공정성 문제를 해결하기 위하여 리셋 사이클 알고리즘을 사용하였다. 이것은 각 노드마다 독자적으로 망의 부하를 판단하여 한 리셋 사이클 동안에 전송할 수 있는 슬롯의 갯수를 계산하고 계산된 만큼만 전송함으로써 다른 노드에 미치는 영향을 최소화하는 방법이다.

4.2.2 Orwell 액세스 프로토콜

각각의 노드는 독립적으로 동작된다. 노드가 ACTIVE(ready to tx)상태인 경우에는 빈 슬롯이 도착될 때마다 데이타를 실어 전송하고 카운터값을 하나씩 증가시킨다. 카운터 값이 할당치(Di)에 이르면 PAUSE 상태로 되고 더 전송할 데이타가 있더라도 전송을 중지한다. 따라서 downstream node에도 전송 기회를 주게된다. IDLE/PAUSE 노드들은 지나가는 빈 슬롯들의 Trial bit(T)를 1로 세트하여 trial slot을 발생시킨다. 이때 수신주소는 자신의 주소로 하여 슬롯이 링을 한바퀴 회전하고 돌아왔을때 자기가 보낸 것인지 확인할 수 있도록 한다. Trial slot 이라 하더라도 Full/Empty(F/E) bit는 0으로 리셋되어 있기 때문에 어느 노드라도 전송할 데이타가 있으면 이 슬롯을 사용할 수 있다. 이 때에는 이 슬롯의 F/E bit는 1로 세트되고 수신주소도 바뀌게 된다. 그러나 링 내의 모든 노드가 IDLE/PAUSE 상태라서 자신이 보낸 trial slot이 그대로 자신에게 돌아오게 되면 trial 시도가 성공한 것으로 알고 리셋 슬롯(F/E = 1, T = 1)을 발생시킨다. 리셋 슬롯은 full slot 이므로 다른 노드에서 인터럽트를 걸수 없으며 모든 노드로 broadcast 된다. 리셋 슬롯을 수신한 노드는 카운터 값을 다시 0으로 리셋하고 ACTIVE 상태로 바뀌어 새로운 리셋 사이클을 시작한다.

리셋 주기를 T ms 이내로 제한하면 Di/T 이상의 대역폭이 보장된다. 그러나 리셋 주기는 부하에 따라서 가변적이기 때문에 각 노드에서는 대역폭을 할당하기 전에 리셋 주기를 계산하여야 한다. 예를 들어 노드에 새로운 호시도가 발생하게 되면 보장된 대역폭을 할당해주어야 하는데 먼저 리셋 주기를 조사하고 새로운 호를 수락하였을때 변하게될 리셋 주기를 계산하여 이값이 T ms 보다 작으면 호를 수락하고 Di를 새로 할당된 대역폭만큼 증가시킨다. T ms 보다 클 때에는 호를 거절한다.

비동기 트래픽인 데이타 서비스를 위해서는 처음부터 일정한 대역폭을 미리 할당하여 처리한다. 동기 트래픽이 많은 경우에는 리셋 사이클이 천천히 발생하기 때문에 비동기 트래픽에 돌아가는 대역폭이 적게 되고 동기 트래픽이 적은 경우에는 리셋 사이클이 빨리 발생하기 때문에 비동기 트래픽에 많은 대역폭이 할당되게 된다.

4.2.3 슬롯의 구조

Orwell ring에서는 시간 지연에 민감한 실시간 트래픽과 덜 민감한 데이타 트래픽을 모두 서비스해주기 위해서 정보를 고정된 길이를 갖는 작은 슬롯으로 나누어 전송한다. 음성과 같은 실시간 트래픽은 패킷 길이가 길 경우에 패킷화하는데 시간이 많이 걸릴 뿐 아니라 망 내에서의 지연도 길어지기 때문에 되도록이면 짧게 하는 것이 좋다. 그러나 너무 짧게 하면 패킷 헤더로 인한 오버헤드가 커지기 때문에 처리율이 떨어지게 된다. 이와 같은 사정을 고려하여 Orwell ring에서는 슬롯 길이를 20 byte로 정하였다. 슬롯의 구조는 그림 4-2와 같다.

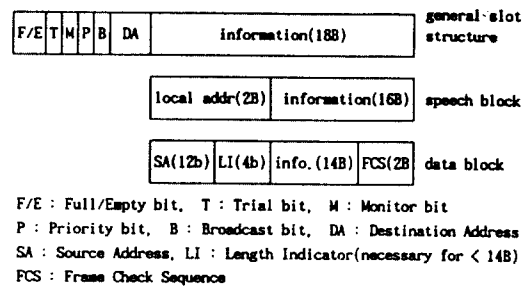

그림 4-2. Orwell ring 슬롯의 구조

4.3 MetaRing[16]

4.3.1 특징

MetaRing은 1989년 IBM Watson 연구소에서 처음 시제품이 개발되었으며 이중 링 구조를 가지고 있고 buffer insertion mode와 slotted ring mode로 동작된다. 수신노드에서 패킷을 소거하는 방법을 통해 공간 재사용이 가능하기 때문에 링크 용량 보다 더 큰 처리율(throughput)을 얻을 수 있다. Buffer insertion mode 에서나 slotted ring mode 에서 모두 랜덤 액세스 방식을 사용하기 때문에 낮은 부하에서도 액세스 지연 시간을 작게 할 수 있다. 그러나 일반적으로 공간 재사용 방식을 사용하면 링크 액세스 권한이 upstream node에게 유리하게 되기 때문에 노드들간에 공정성 문제가 발생하게 된다. 이를 해결하기 위한 방안으로 Orwell 이나 ATMR에서는 리셋 사이클이라는 방법을 사용하였는데 이경우 사이클 사이에 생기는 유휴시간(idle time)으로 인해서 링의 전달 시간이 길어짐에 따라 처리율이 민감하게 영향을 받는 단점이 있었다. MetaRing에서는 그림 4-3과 같이 데이타 트래픽과 반대 방향으로 회전하는 별도의 제어 신호를 사용하여

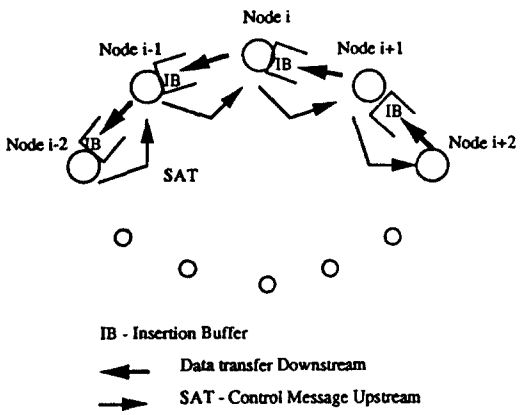

그림 4-3. MetaRing의 구조(한방향)

공정성 문제를 해결하였다. 제어 신호는 codeword들 중에서 데이타로 사용하고 남은 여분의 codeword들을 할당하여 사용하는데 이들은 데이타 트래픽 보다 높은 preemptive 우선 순위를 가지고 있어서 데이타 트래픽을 전송하던 중이라도 언제든지 이를 중단하고 제어 신호를 먼저 전송할 수 있다. 이외에도 MetaRing에는 우선 순위가 다른 여러 트래픽들을 처리해주기 위한 알고리즘들이 제안되어 있으며 동기 트래픽과 비동기 트래픽을 통합하여 서비스해줄 수 있도록 구현되었다.

1989년에 제작된 시제품은 100 Mbits/s 이중 링을 사용하여 700 Mbits/s 까지 지원하며 현재는 기가비트 버전이 개발중이다.

4.3.2 액세스 제어 모드

MetaRing은 두가지 모드로 동작할 수 있는데 한 가지는 buffer insertion mode이고 다른 한 가지는 slotted ring mode 이다. Buffer insertion mode에서는 각 노드마다 세개의 버퍼(입력 버퍼, 출력 버퍼, insertion buffer)가 필요하다. 입력 버퍼는 착신 주소가 자신으로 되어 있는 패킷이 들어오면 이를 저장하는 역할을 하며 출력 버퍼는 외부로 부터 들어오는 데이타를 링으로 전송하기 전에 일시적으로 저장하는 역할을 한다.

Insertion buffer는 링으로 부터 유입되는 데이타를 일시적으로 저장하기 위해 사용된다. 그러나 링 트래픽은 insertion buffer 내에서 cut-through 방식으로 전달되기 때문에 패킷의 마지막 비트가 도착될 때까지 insertion buffer에 머물 필요는 없으며 착신측 주소 필드를 조사하여 자신으로 입력되는 트래픽이 아니면 즉시 다음 노드로 전송할 수 있다.

Buffer insertion mode에서는 외부로 부터 입력되는 데이타와 링에서 순환되고 있는 데이타 사이에 충돌이 발생할 수 있는데 이때에는 링에서 순환되고 있는 데이타에 non-preemptive 우선 순위를 두어 해결하였다. 즉, 출력 버퍼에 전송할 데이타가 있더라도 insertion buffer에 전송할 데이타가 있으면 이것을 먼저 전송한다. 출력 버퍼에서 이미 데이타를 전송하고 있으면 전송이 끝날때까지 링 트래픽은 insertion buffer 내에서 대기하여야 한다. 따라서 이 모드로 동작하는 경우 낮은 부하에서는 전송 지연이 작으나 부하가 많아짐에 따라서 전송지연이 커지는 단점이 있다. 이를 보완하기 위하여 MetaRing에서는 slotted ring mode와 병행하여 사용된다.

Slotted ring mode에서는 buffer insertion mode에서와 동일한 하드웨어 구성을 가지고 있다. 다만 buffer insertion mode에서 insertion buffer로 사용되던 것을 slotted ring mode에서는 슬롯 동기를 위한 elastic buffer로 사용하는 것이 다르며 buffer insertion mode에서는 가변길이 패킷을 사용하지만 slotted ring mode에서는 고정길이 패킷을 사용한다는 차이가 있다.

4. 3. 3 공정성 제어 알고리즘

노드들 간에 균등한 액세스 권한을 주기 위하여 SAT(SATisfied) 제어 신호에 의한 공정성 제어 알고리즘이 사용되었다. SAT는 데이타 트래픽과 반대 방향으로 전송된다. SAT를 수신한 노드는 이전에 SAT를 upstream node로 전송한 이후로 이미 n 개 이상의 패킷을 전송하였거나 출력 버퍼에 전송할 데이타가 없으면 즉시 SAT를 upstream node로 forwarding 해주고 그렇지 않은 경우에는 n 개의 패킷을 전송하고 난 후에 SAT를 upstream node로 전송해주고 이 노드는 다시 k 개(k n)까지 패킷을 전송할 수 있게된다. 즉 이 알고리즘에 따르면 모든 노드들은 SAT가 링을 한 바퀴 회전하는 동안에 최소한 n 개에서 최대 k 개까지 패킷을 전송할 수 있다. 전송 도중에 SAT 신호가 유실될 경우에 대비하여 각 노드들은 타이머를 가지고 있으며 이것이 타임아웃 될때까지 SAT 신호가 돌아오지 않으면 유실된 것으로 간주하고 다시 SAT 신호를 발생시킨다. 또한 SAT 신호를 가지고 있는데 또 다시 SAT가 전송되어 오면 중복된 것으로 간주하고 하나를 폐기시킨다. SAT 제어 신호를 사용한 공정성 제어 알고리즘을 그림 4-4에 나타내었다.

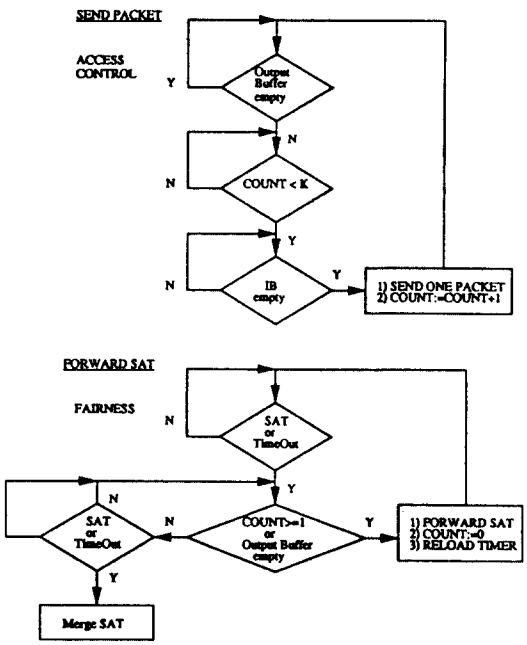

그림 4-4. MetaRing의 공정성 제어 알고리즘

4.3.4 동기 트래픽과 비동기 트래픽의 통합전송

MetaRing에서 동기 트래픽과 비동기 트래픽의 통합 전송 방식은 FDDI의 Timed Token 프로토콜과 유사하다. 즉 동기 트래픽에 대해서는 필요한 만큼의 대역폭을 미리 할당해 두고 나머지 대역폭을 비동기 트래픽에 할당해주는 방식을 사용한다. 동기 트래픽이 발생한 경우에 필요한 대역폭을 보장해주기 위하여 각 노드들은 다음과 같은 방법으로 대역폭을 예약한다.

i) 요청된 QoS로 부터 얼마만한 데이타를 전송할 것인지 계산한다(n 데이타 유니트).

ii) 전송 경로를 결정한다.

iii) 전송 경로 사이에 있는 모든 노드들에게 n 데이타 유니트 만큼의 대역폭을 예약해 줄것을 요청한다.

iv) 전송 경로 사이에 있는 모든 노드로 부터 긍정 수신이 오면 예약이 성공한 것으로 보고 동기 트래픽을 전송한다. 긍정 수신이 오지 않은 경우에는 예약이 실패한 것으로 보고 다시 전송 경로 사이에 있는 모든 노드들에게 앞서 예약한 n 데이타 유니트 만큼의 대역폭을 취소할 것을 요청한다.

비동기 트래픽을 전송하기 위해서 ASYNC-EN Rotation Protocol이 사용되었다. 이 프로토콜에서는 ASYNC-EN 제어 신호가 링을 회전하면서 각 노드마다 ASYNC-EN의 회전 주기를 측정하여 그 값에 따라서 비동기 트래픽을 전송할 것인지 결정한다. 회전 주기를 측정하기 위하여 각 노드마다 ASYNC-EN-TIMER(t)라는 타이머를 사용한다.

노드가 비동기 트래픽을 전송하기 위해서는 다음과 같은 조건을 모두 만족하여야 한다.

i) SAT 알고리즘에 의해 satisfied 되지 않았을것.

ii) 빈 슬롯이 있을 것.

iii) ASYNC-EN-TIMER(t) r

'r'은 링의 슬롯 수를 나타낸다. 노드에서 동기 트래픽이 큐에 많이 쌓이게 되면 ASYNC-EN 신호를 받더라도 즉시 전송하지 않고 한 슬롯 동안 가지고 있다가 전송하기 때문에 upstream node 에서의 ASYNC-EN-TIMER(t) 값이 r 보다 커지게 되고 따라서 비동기 트래픽의 전송을 제한하게 된다.

4.4 ATMR[17][18][19]

4.4.1 특징

ATMR은 Orwell ring의 리셋 사이클 알고리즘을 개선한 프로토콜이다. 리셋 사이클이란 global fairness mechanism의 한 방법이다. Orwell ring에서는 리셋 사이클을 위해 2 개의 사이클(trial + reset)이 필요하였으나 ATMR 에서는 이것을 1 사이클(reset)로 줄임으로써 성능을 개선하였다. 또한 우선순위 트래픽 제어를 위해서 리셋 인터럽션 메카니즘을 사용하였다. ATMR이 갖고 있는 대표적인 특징을 살펴보면 다음과 같다.

분산 액세스 방식의 slotted ring network

서로 반대 방향으로 회전하는 이중 링 구조

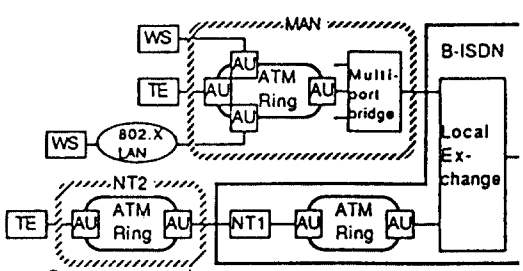

그림 4-5. B-ISDN 구성과 ATMR

- ATM과 유사한 셀 포맷 및 구조로 B-ISDN과 호환성 용이
- 수신노드소거 방식에 의한 공간 재사용
- "윈도우 사이즈"와 "리셋 메카니즘"을 사용한 분산 공정성 제어
- 다중 서비스 계층 지원을 위한 우선순위 제어

ATMR을 B-ISDN에 여러가지 형태로 적용한 모습을 그림 4-5에 나타내었다.

4.4.2 ATMR 네트워크와 노드 구성

1) 이중 링 구성

ATMR은 슬롯이 서로 반대 방향으로 돌아가는 두 개의 링으로 구성된다. 물론 한개의 링으로도 구성할 수 있다. 두 링은 정상 상태에서는 독립적으로 동작하지만 이상 상태에서는 하나의 링으로 동작한다.

2) ATMR 구조

ATMR의 구조는 ATM의 구조와 유사하게 물리 계층, ATMR 계층 그리고 ATM 적응 계층으로 이루어져 있다.

가) 물리 계층

ATM의 물리 계층과 동일한 기능을 갖으며 STM-n (Synchronous Transfer Mode - n)을 지원한다.

나) ATMR 계층

ATMR 계층에서는 다음과 같은 기능을 수행한다.

- ATMR 셀 생성
- 액세스 제어 : 공정성과 우선순위 제어
- 잘못된 셀의 제거

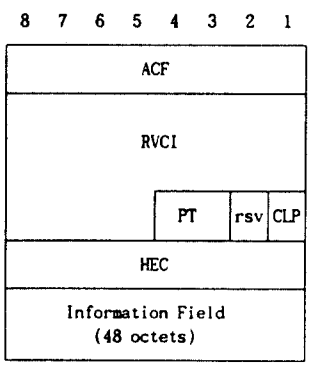

그림 4-6. ATMR 셀 포맷

- 윈도우 사이즈 할당
- RVCI(Remote Virtual Channel Identification) 변환
- 셀 다중화/역다중화

다) ATM 적응 계층

CCITT 광대역 통신 프로토콜에서 제정한 ATM의 적응 계층과 동일한 기능을 한다.

3) ATMR 셀 포맷

ATMR 셀 포맷은 그림 4-6에서 보는 바와 같이 ATM 셀과 유사하다. 차이점은 ACF(Access Control Field)와 RVCI(Remote VCI) field 이며 나머지 필드들의 기능은 ATM과 동일하다. ACF 필드의 기능은 다음과 같다.

- 리셋 셀과 데이타 셀의 구분
- Busy Address (BA) 구분
- Monitor 상태 여부

ACF 필드에 지정된 값은 표 4-1과 같다.

표 4-1. ATMR의 Access Control Field

Value	Definition
X 0 0 0 0 0 0 0 X 0 0 0 0 0 0 1 . . X 1 1 1 1 0 0 0	Busy address
X 1 1 1 1 1 0 0 X 1 1 1 1 1 0 1 X 1 1 1 1 1 1 0 X 1 1 1 1 1 1 1	Reset_4 Reset_3 Reset_2 Reset_1
1 X X X X X X X	In the monitored state
0 X X X X X X X	Not in the monitored state

RVCI는 ATMR 셀의 수신 주소를 나타낸다. 일반적으로 RVCI 값을 지정할 때 계층적 주소 지정 방법 (RVCI = AUA + LCN ; Access Unit Address, Logical Channel Number)을 사용하면 수신 주소 확인을 빨리 할 수 있기 때문에 노드에서의 지연을 줄일 수 있다. RVCI 값이 0 이면 이것은 empty cell 임을 나타낸다.

4.4.3 ATMR 동작

1) MAC 프로토콜에 사용되는 파라미터들

ATMR 프로토콜에서 사용되는 파라미터들은 다음과 같다.

- NCLASS (우선순위 클래스의 수)

NCELL (링에 있는 셀의 갯수)

IV-TMR-Q_i (큐 타이머의 초기값)

우선순위 i 클래스의 타임아웃 값을 나타낸다.

AUID (Access Unit Identifier)

IV-CTR-W_i (윈도우 카운터의 초기값)

AU에서 한 리셋 사이클 동안에 최대로 전송할 수 있는 우선순위 i 클래스의 셀 수를 나타낸다.

TLEVEL (전송 레벨)

value(i) : AU는 제일 우선순위가 높은 셀로 부터 i 클래스 우선순위 셀까지 전송할 수 있다.

QUE_i (송신큐 안에 있는 우선순위 i 클래스 셀의 갯수)

TMR-Q_i (큐 타이머)

리셋 사이클이 시작되고 나서 경과 시간을 나타낸다. 리셋 사이클이 시작되면 TMR-Q_i 값은 IV-TMR-Q_i 값으로 초기화 되고 시간이 경과됨에 따라서 감소된다. 이 값이 0 이 되면 타임아웃 된 것으로 간주하여 Reset_i 셀을 보내서 링의 전송 레벨을 조정한다.

CTR-W_i (윈도우 카운터)

리셋 사이클이 시작될때 IV-CTR-W_i값으로 초기화 되고 우선순위 i 클래스 셀을 전송할 때마다 값이 하나씩 감소한다.

2) 분산 공정성 제어

Slotted ring protocol에서 공정성 제어를 하지 않으면 그림 4-7에서 보는 바와 같이 한 노드에서 대역폭을 모두 사용하기 때문에 다른 노드에서 보낼 데이타가 있더라도 보낼 수 없게 되는 hogging 현상이 발생하게 된다. ATMR에서는 이러한 hogging 현상을 방지하고 각 AU 간에 대역폭을 공정하게 할당하기 위하여 윈도우 사이즈와 리셋 메카니즘을 사용한다. 각 AU는 리셋 사이클이 시작 될 때 CTR-W의 초기값이 IV-CTR-W(Window Size)로 주어지며 셀이 전송될 때 마다 CTR-W 값은 하나씩 감소한다. CTR-W 값이 0 이되면 그 리셋 사이클 내에서 이미 할당된 대역폭을 다 사용한 것이 되므로 전송할 데이타가 있더라도 더이상 전송하지 않고 다음 리셋 사이클이 올때까지 기다린다. 이 상태를 INACTIVE 상태라 한다. 노드가 ACTIVE 상태에서 INACTIVE 상태로 될 때에는 다른 AU들의 상태를 체크하여 모두 INACTIVE 상태일 경우에 리셋 셀을 전송한다. 리셋 셀을 전송하기 전에 먼저 다른 AU들의 상태를 살펴보아야 하는데 이러한 목적으로 ATMR 셀 헤더의 ACF 필드 내에 있는 busy address를 사용한다. ACTIVE 상태에 있는 AU들은

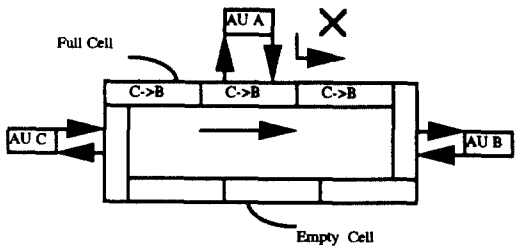

그림 4-7. Hogging in Slotted Ring

busy address 필드에 무조건 자신의 AUID를 기입하기 때문에 링을 한바퀴 돌고온 셀 헤더에 자신의 AUID가 있으면 다른 AU들이 모두 ACTIVE 상태인 것을 알 수 있다.

3) 우선순위 제어

서로 다른 QoS 요구조건을 갖는 여러 서비스들을 만족시키 주기 위하여 우선순위 제어 기능이 있다. ATMR에서는 지연에 민감한 트래픽을 우선 전송하는 우선순위 제어 방법을 사용한다. 이를 위해서 전송 레벨이라는 개념을 사용하였다. 전송 레벨이란 어떤 우선순위 그룹이 링을 액세스 할 수 있는지 결정하는 파라미터이다. 예를 들어 링의 전송 레벨이 2라고 하면 우선순위 그룹 1 과 2 만이 링을 액세스 할 수 있음을 나타낸다(낮은 수가 우선순위가 높은 그룹을 나타낸다).

전송 레벨은 리셋 셀에 의해서 결정된다. 즉 Reset_1 셀을 전송하면 링의 전송 레벨은 1이되어 우선순위 그룹 1 만이 링 액세스가 허용된다.

Reset_i 가 발생되면 링 내에서는 다음과 같은 과정이 이루어 진다.

(a) i가 1과 K(가장 낮은 우선순위 그룹)사이이면 링의 전송 레벨이 i로 바뀌고, i가 K+1 이면 K로 바뀐다.

(b) i 보다 높은 우선순위 그룹의 윈도우 카운터를 리셋한다.

(c) i 보다 높은 우선순위 그룹의 큐 타이머를 재시동한다.

4) 타이머 인터럽션에 의한 제어

전송 레벨이라는 개념만을 가지고는 우선순위 제어를 충분히 할 수 없다. 즉 다음과 같은 두 가지 경우에는 높은 우선순위 정보가 낮은 우선순위 트래픽으

로 인해 지연 요구사항을 지킬 수 없게 된다.

경우 1 : 한 노드에서 우선순위가 낮은 트래픽을 계속적으로 발생 시킴으로써 downstream node에서 우선순위가 높은 트래픽을 전송할 수 없는 경우.

경우 2 : 높은 우선순위 트래픽 전송을 완료하였지만 낮은 우선순위 트래픽 전송이 완료되지 않아 리셋 사이클이 끝나지 않은 경우. 이 때 이 리셋 사이클이 길어짐으로써 다음 리셋 사이클 동안 전송해야될 높은 우선순위 트래픽이 지연 요구조건을 만족시키지 못하게 된다.

위 문제를 해결하기 위해 리셋 인터럽션이 사용되었다. 먼저 경우 1을 해결하기 위해서는 각 노드마다 우선순위 그룹 별로 제한된 시간안에 할당된 트래픽을 전송할 수 있는지 예측 가능하여야 한다. 경우 2를 해결하기 위해서는 큐 타이머를 사용하였다. 즉 TMR-Q_i가 종료되면 인터럽션 리셋(Reset_i+1)을 발생시키고 따라서 우선순위가 i 보다 높은 셀들은 다시 쿼타를 할당받아 새로운 리셋 사이클을 시작한다.

우선순위 그룹이 2개인 경우에 앞에서 설명한 우선순위 제어 알고리즘은 그림 4-8과 같다. 그림 (a)는 두 그룹이 모두 QoS를 만족하도록 처리된 것으로 이 때에는 정상적인 리셋(Normal Reset)이 발생하여 새로운 리셋 사이클이 시작된다. 두 우선순위 그룹에 모두 새로운 쿼타가 할당된다. 그림 (b)는 경우 1에 해당되며 타이머 T2 안에 높은 우선순위 트래픽을 다 전송하지 못하면 전송레벨을 올려서 우선적으로 전송한다. 높은 우선순위 트래픽의 전송이 끝나면 다시 전송레벨을 낮춘다. 그림 (c)는 경우 2에 해당되며 T1이 타임아웃 되면 낮은 우선순위 트래픽의 쿼타가 남아있더라도 리셋 셀을 발생시켜서 높은 우선순위 트래픽이 다시 쿼타를 할당받도록 한다.

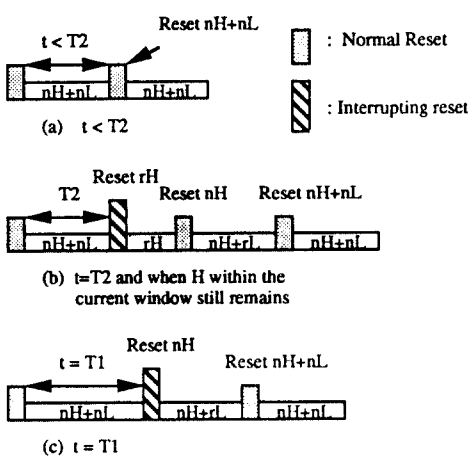

그림 4-8. ATMR의 우선순위 제어 알고리즘

5) 오류 셀의 제거

전송 도중의 오류나 예기치 않은 AU의 제시로 인하여 링에 존재하지 않는 AU 로 향하는 셀이 발생할 우려가 있다. 이러한 오류 셀을 방치하면 링 내에서 무한 루프를 돌기 때문에 처리율을 급격히 떨어뜨리게 된다. 이와 같은 현상을 방지하기 위하여 셀 헤더 중 ACF 필드내에 모니터 비트를 두고 오류 셀인지 여부를 판단하고 오류 셀인 경우에는 폐기한다.

4.5 HMR(High-speed Multimedia Ring)

4.5.1 서언

HMR은 gigabit 이상의 속도에서도 대역폭을 효과적으로 사용하고 멀티미디어 트래픽을 제공해주는 것을 목표로 제안되었다. HMR은 서로 반대 방향으로 회전하는 두개의 링으로 구성된다. Slotted ring protocol을 사용하며 ATMR이나 MetaRing에서와 같이 수신노드소거에 의한 공간 재사용 방식을 사용하여 링 고 용량 이상의 높은 처리율을 얻을 수 있다. 이 프로토콜에서는 공정성 제어와 우선순위 제어를 위해서 다중 윈도우 토큰을 사용하였다. 이 방법은 MetaRing에서 SAT 제어신호를 사용하여 노드들간의 공정성을 제어하는 방법과 유사한데, MetaRing에서는 우선순위 제어를 위해 SAT 제어신호에 파라미터를 사용하여 각 노드들에게 다음에 전송할 수 있는 우선순위 그룹을 알려주는데 반하여 HMR에서는 각 우선순위마다 하나의 윈도우 토큰을 할당하였다. 이 토큰들은 셀이 전송되는 방향과 반대 방향으로 전송되면서 만나는 노드마다 새로운 쿼타를 할당해준다. 다시 말해서 각 노드는 토큰이 한바퀴 회전하는 동안 그에 해당되는 우선순위 트래픽을 미리 할당된 쿼타 만큼만 전송할 수 있기 때문에 노드들간의 공정성이 유지된다.

4.5.2 HMR 네트워크와 노드 구성

1) 네트워크 토폴로지

HMR의 기본 구성은 슬롯이 서로 반대방향으로 돌

아가는 이중 링 구조이다. 이중 링 구조는 신뢰성이 있고 특히 수신노드소거 방식과 최단경로라우팅 방식을 사용할 경우 처리율이 링크 용량의 4배 까지 얻을 수 있는 잇점이 있다. 두 링은 정상 상태에서는 독립적으로 동작하지만 고장이 발생하면 단일 링으로 동작하며 물론 처음부터 단일 링으로 동작할 수도 있다. 또한 그림 4-9와 같이 이중 버스 구조로 구성할 수도 있다.

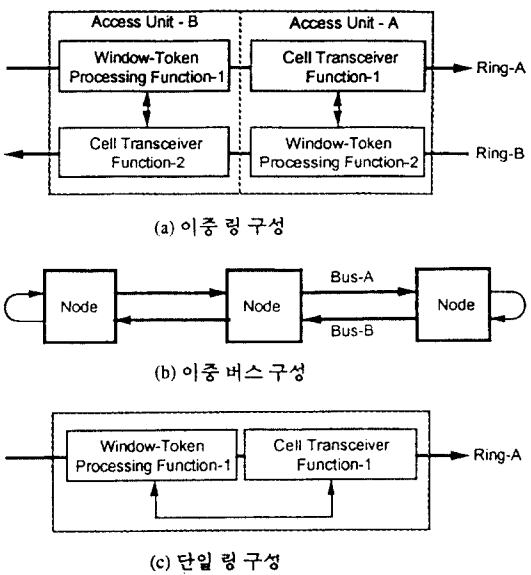

그림 4-9. HMR 네트워크 토폴로지

2) HMR 노드의 구조

HMR 노드의 구조는 그림 4-10과 같이 세 계층으로 이루어져 있다. HMR 노드의 구조는 ATM 노드의 구조와 유사하게 만듦으로써 B-ISDN과의 호환성을 높였다. 물리계층은 Transmission system과 Transmission Convergence Function으로 구성되는데 다양한 전송속도를 갖는 전송시스템들이 접속될 수 있고 이것들은 Transmission Convergence Function의 기능으로 HMR 계층에 영향을 미치지 않고 확장될 수 있다. HMR 계층은 셀을 만들고 전송하며 매체 접속 기능을 수행한다. 적응계층에서 하는 일은 ATM의 적응계층의 기능과 유사하다. 단지 HMR에서는 connectionless data를 위해서 MAC Convergence Function 이라는 기능을 추가하였다.

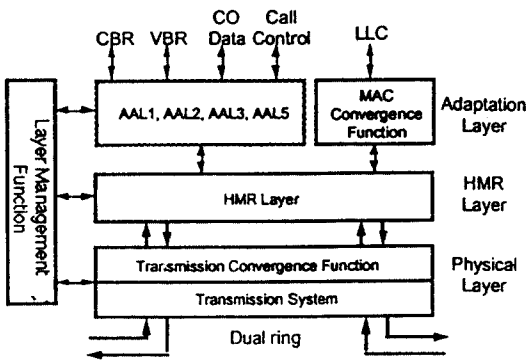

그림 4-10. HMR 노드의 구조

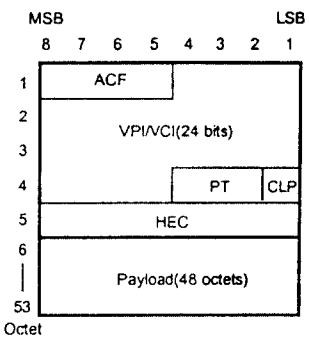

그림 4-11. HMR 셀 포맷

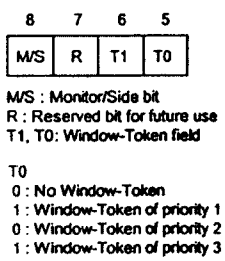

그림 4-12. HMR의 ACF 필드

3) HMR 셀 포맷

HMR의 셀 포맷은 그림 4-11에서 나타내었듯이 ATM의 셀 포맷과 유사하다. 차이점은 ATM 에서 GFC(Generic Flow Control) 필드로 사용하는 4 bit를 HMR에서는 ACF 필드로 사용한 점이다.

가) ACF(Access Control Field)

ACF 필드는 그림 4-12와 같다. Window-Token field는 셀이 윈도우 토큰을 포함하고 있는지 아닌지, 포함하고 있으면 어떤 우선순위 그룹 토큰인지를 나타낸다. 2 bit를 사용하기 때문에 HMR에서는 우선순위 그룹을 세개까지 가질 수 있다.

M/S(Monitor/Side) bit는 링 구성에서는 모니터 비트로 사용되고 버스 구성에서는 side bit로 사용된다. 모니터 비트는 garbage cell을 발견해서 폐기시키기 위해 사용된다. 송신노드에서는 셀을 전송할 때 모니터 비트를 0으로 리셋해서 전송한다. 링 초기화시에 노드들 중 하나를 모니터 노드로 선정하게 되는데 모니터 노드에서는 들어오는 셀들의 모니터 비트를 조사해서 0 이면 1로 세트시켜서 전송하고 1 이면 garbage cell로 판단하여 폐기시킨다. 이중 버스 구조에서는 이 필드가 side bit로 사용된다. 이중 버스 구조에서는 링 A를 위한 윈도우 토큰과 링 B를 위한 윈도우 토큰이 동일한 링을 회전하기 때문에 이들을 구분해주어야 하는데 이를 위해 side bit가 사용된다.

나) VPI/VCI

이 필드는 B-ISDN addressing scheme을 따르거나 Multimedia Ring Network(MRN) addressing scheme을 따른다.

(1) B-ISDN addressing scheme

B-ISDN과의 호환성이 강조되는 곳, 예를들어 B-NT 2에는 B-ISDN의 UNI와 동일한 포맷을 사용한다. 즉 VPI에 8 bit, VCI에 16 bit가 할당되어 사용된다. 이 방법에서는 모든 노드들이 VPI/VCI space를 공유하기 때문에 수신노드를 확인하기 위해서는 24 bit를 모두 분석하여야 한다. 따라서 노드에서의 지연이 커지는 단점이 있다.

(2) MRN addressing scheme

VPI, VCI가 각각 12 bit 씩 할당된다. VPI는 링에 접속된 노드의 주소를 나타내고 VCI는 한 노드에서 여러 채널을 동시에 열 경우에 각각의 채널을 구분하기 위해 사용된다. 이 방법을 사용하면 수신노드를 확인하기 위하여 12 bit만 조사하면 되기 때문에 앞의 방법에 비하여 노드에서의 지연이 적어지는 장점이 있다. HMR에서 제안한 VPI/VCI 구조는 그림 4-13과 같다. 셀은 그림에서 보듯이 네 가지 종류로 나누어지는데 각각 management cell, CO cell, CL cell 그리고

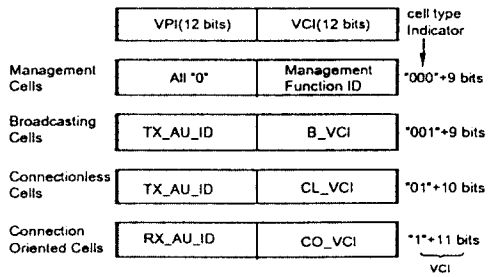

그림 4-13. HMR의 VPI/VCI 구조

broadcasting cell 이라고 한다. CO cell은 수신노드소거 방식을 사용하고 CL cell과 broadcasting cell은 송신노드소거 방식을 사용한다. CL cell은 모든 노드에서 일단 수신한 후에 MAC Convergence Entity로 전달해 주고 이 곳에서 셀의 수신주소를 조사하여 자신에게 오는 것이면 받아들이고 그렇지 않으면 폐기시킨다.

4.5.3 HMR 동작

1) MAC 프로토콜에 사용되는 파라미터들

HMR 프로토콜에서 사용되는 파라미터들은 다음과 같다.

- TOKEN_TMR_i(윈도우 토큰 타이머)

우선순위 그룹 i의 윈도우 토큰의 회전시간을 측정하기 위한 타이머이다.

- WC_i(윈도우 카운터)

Token_i가 한바퀴 회전하는 동안에 노드에서 전송할 수 있는 우선순위 i 셀의 최대 갯수를 나타낸다.

- RING_LENGTH(Ring Length)

링의 전달지연시간을 cell time으로 나타낸 값이다.

- NUM_NODE(Number of Node)

링에 접속된 노드의 수를 나타낸다. 이 파라미터 값은 수신노드로 가는 가장 짧은 경로를 찾는데 사용된다.

- CYCLE_MIN_i(Window-Token Rotation Time Lower Threshold), CYCLE_MAX_i(Window-Token Rotation Time Upper Threshold)

낮은 우선순위 셀 송신을 제한하기 위해 사용된다.

- QUOTA_i(Quota of Priority i)

QUOTA_i는 Token_i가 노드를 통과할 때마다 WC_i에 할당되는 값이다. 이 값은 Token_i가 한 바퀴 회전하는 동안에 그 노드에서 최대로 전송할 수 있는 우선순위 i 그룹 셀 수를 나타낸다.

- TX_Q_i(Transmission Queue Indicator) : 우선순

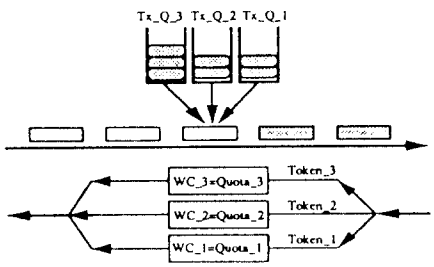

그림 4-14. HMR의 액세스 제어 모델

위 그룹 i의 송신큐가 비어있는지 여부를 가리킨다.

- TRT_i(Window-Token Rotation Time Indicator) : 각 노드마다 우선순위 그룹 i의 바로 이전 토큰 회전시간을 저장하기 위해 사용된다.

2) 액세스 제어

HMR의 기본적인 액세스 제어 모델은 그림 4-14와 같다. 각 노드에는 우선순위 별로 송신큐(TX_Q_i : i는 우선순위)와 윈도우 계수기(WC_i)가 있다. TX_Q_i는 송신큐가 비어있는지 여부를 나타낸다. WC_i는 Token i가 노드에 들어와서 다음 노드로 전송될 때 미리 주어진 값(쿼타)으로 세트되며 i 우선순위 셀을 전송할 때마다 하나씩 감소한다. 송신큐에 있는 셀은 WC_i가 0 보다 큰 경우에만 송신될 수 있다. 우선순위 별로 독립적인 토큰(Token i)이 링을 회전하는데 노드는 Token i 수신시 i 보다 같거나 높은 우선순위의 윈도우 계수기가 0 이거나 혹은 송신큐에 전송할 셀이 없는 경우에만 Token i를 다음 노드로 통과시킨다. 이렇게 함으로써 우선순위가 높은 토큰일수록 노드에서의 지연이 적으므로 링을 빠르게 회전할 수 있다. Token i가 링을 한바퀴 회전하는 동안에 모든 노드들은 미리 할당된 우선순위 i 그룹 트래픽을 전송할 수 있기 때문에 공정성 문제가 해결된다.

3) 최대 액세스 지연

셀의 최대 액세스 지연은 토큰의 회전시간(TRT i)에 의하여 결정된다. 좀더 정확히 표현하면 셀의 최대 액세스 지연은 다음 식으로 나타낼 수 있다.

Max. Access_Delay_i = 2 x TRT_i

이 식은 윈도우 토큰의 회전주기를 특정값 이내로 제한 시키면 셀의 최대 액세스 지연도 역시 특정값 이내로 제한시킬 수 있음을 의미한다. 윈도우 토큰의 회전주기를 특정값 이내로 제한 시키기 위해서는 회전주기 동안에 전송가능한 셀수(쿼타)를 제한하여야 한다. 우선순위 그룹이 하나 뿐일 경우에는 모든 노드에 할당된 총 쿼타를 특정값 이내로 제한하면 되기 때문에 간단하지만 우선순위 그룹이 여러개일 경우에는 각 우선순위 그룹 별로 최대 지연 시간에 대한 요구가 다르고 낮은 우선순위 트래픽의 영향으로 우선순위가 높은 트래픽이 QoS를 만족시켜 주지 못하는 경우가 발생할 수 있다. 이와 같은 문제를 해결해 주기 위해서는 우선순위 별로 QoS 만족 여부를 항상 체크하여 높은 우선순위 트래픽이 QoS 요구조건을 벗어나는 일이 발생하기 전에 낮은 우선순위 트래픽의 송신을 제한하는 우선순위 제어 알고리즘을 적용하여야 한다.

HMR에서 우선순위 제어 알고리즘은 Token_i가 통과할 때마다 재시동하는 타이머 Token_TMR_i와 낮은 우선순위 트래픽을 제한하기 위해 사용하는 변수 Cycle_MIN_i와 Cycle_MAX_i를 사용하여 다음과 같은 두가지 방법으로 우선순위가 높은 셀의 최대 액세스 지연을 QoS 이내로 만족시킨다.

i) Cycle_MIN_i에 의한 송신제한

모든 노드에서 Token_TMR_i < Cycle_MIN_i 이면 i 보다 낮은 우선순위의 셀 송신을 중지한다.

ii) Cycle_MAX_i에 의한 송신제한

Token i가 오면 Token_TMR_i의 값을 TRTR_i(Token Rotation Time Register i)에 저장하고 TRTR_i ≥ Cycle_MAX_i 이면 i 보다 낮은 우선순위의 셀송신을 중지한다.

HMR의 각 노드에서 셀을 전송하기위한 조건과 윈도우 토큰을 통과시키기 위한 조건을 요약하면 각각 표 4-2, 표 4-3과 같다.

Send Condition_1	TX_Q_1=not EMPTY and WC_1 > 0
Send Condition_2	(TX_Q_1=EMPTY or WC_1=0) and TOKEN_TMR_1 ≤ CYCLE_MIN_1 and TRT_1 ≤ CYCLE_MAX_1 and TX_Q_2=not EMPTY and WC_2 > 0
Send Condition_3	(TX_Q_1=EMPTY or WC_1=0) and (TX_Q_2=EMPTY or WC_2=0) and TOKEN_TMR_1 ≤ CYCLE_MIN_1 and TRT_1 ≤CYCLE_MAX_1 and TOKEN_TMR_2 ≤ CYCLE_MIN_2 and TRT_2 ≤CYCLE_MAX_2 and TX_Q_3=not EMPTY and WC_3 > 0

표 4-2. 노드에서의 셀 송신 조건

표 4-3. 노드에서의 윈도우 토큰 통과 조건

Window_Token Priority	Passing Conditions
Token_1 (Highest priority)	WC1_=0 or TX_Q_1=EMPTY
Token_2	(WC_1=0 or Tx_Q_1=EMPTY) and (WC_2=0 or Tx_Q_2=EMPTY) and TOKEN_TMR_1 ≤CYCLE_MIN_1 and TRT_1 ≤CYCLE_MAX_1
Token_3 (Lowest priority)	(WC_1=0 or Tx_Q_1=EMPTY) and (WC_2=0 or Tx_Q_2=EMPTY) and (WC_3=0 or Tx_Q_3=EMPTY) and TOKEN_TMR_1 ≤CYCLE_MIN_1 and TRT_1 ≤CYCLE_MAX_1 and TOKEN_TMR_2 ≤CYCLE_MIN_2 and TRT_2 ≤CYCLE_MAX_2

Token_i가 도착했을 때 쿼타를 할당하는 방법에도 정적 쿼타 할당 방법과 동적 쿼타 할당 방법이 있다. 정적 할당 방법은 연결 설정시 쿼타를 추가하고 연결 해제시 쿼타를 회수하는 방법으로 CBR(Constant Bit Rate) 트래픽을 전송하는데 유용하다. 그러나 VBR (Variable Bit Rate) 트래픽은 대역폭이 일정하지 않으므로 동적 할당 방법을 사용하는 것이 유리하다. HMR에서는 윈도우 토큰의 회전시간과 송신큐에 쌓여 있는 셀의 수에 따라 쿼타를 계산하여 동적으로 할당하는 방법을 사용하고 있다.

V. 기타 고속 LAN

5.1 HANGMAN

HANGMAN은 1987년, 영국의 Bristol에 있는 Hewlett-Packard 연구소에 의해 개발된 전형적인 gigabit cell LAN이다. 이것은 3가지 면에서 DQDB나 CBR(Cambridge-Backbone-Ring)과 다른 특징을 가지고 있다.

첫째 HANGMAN은 그림 5-1과 같이 노드라는 장치를 통해 여러개의 시스템들을 광 케이블에 연결시킨다 [10]. 시스템과 네트워크 사이에 노드는 실제로 데이타를 전송하는 물리 계층과 MAC계층, 그리고 네트워크에 독립적인 서비스를 제공하고 사용자 접속을 관리하는 프로토콜로 구성되어 있다.

둘째 HANGMAN은 256바이트의 비교적 긴 길이의 셀 크기를 가진다. HANGMAN의 설계자들은 B-ISDN과 같은 작은 길이의 셀 크기를 가지는 외부 네트워크와의 인터페이스보다 LAN에서의 데이타 트래픽에 대한 성능을 최적화하는데 더 관심을 두어 256바이트의 셀 크기를 선택하였다.

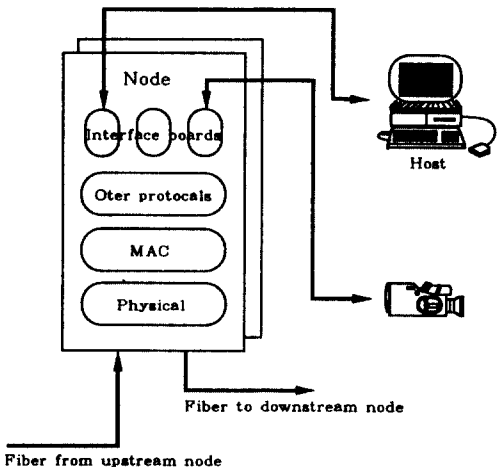

그림 5-1. Overview of the HANGMAN node

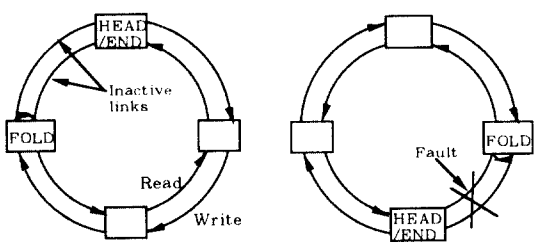

그림 5-2. The physical dual ring of HANGMAN

세째 CBR이나 DQDB의 dual bus와는 달리 HANGMAN의 구조는 그림 5-2과 같이 논리적으로 겹쳐진 bus이다. 각 노드는 bus에 두 번 연결되는데 한번은 쓰기 전용이고 다른 한번은 읽기 전용으로 연결되어 있다. Bus의 맨 첫부분인 HEAD/END노드는 slot generator로 쓰기 전용 버스에 빈 슬롯을 만들고 각 노드들은 거기에 데이타를 실어 보낸다. 그 슬롯은 FOLD노드를 거쳐 읽기 전용 버스로 보내지고 수신단 노드에서 그 슬롯의 데이타를 읽어 들인다. 만약 두 노드 사이에서 하나의 링크나 한 노드에 문제점이 생기면 HEAD/END노드와 FOLD노드 사이의 논리적으로 단절된 부분이 연결되고 고장난 부분의 양 노드중 하나는 HEAD/END노드가 되고 다른 하나는 FOLD노드가 된다.

HANGMAN은 DQDB와 마찬가지로 공평성에 문제점이 있을 수 있다. HEAD/END 노드에 가까이 있는 노

드들은 다른 노드들과 다른 확률로 접속 가능성이 있을 수 있다. 이 문제점을 해결하기 위하여 S(Simple)라는 액세스 프로토콜을 사용한다[5]. 여기서 각 노드들은 하나의 카운터를 가지는데 슬롯에 데이타를 보낼 때마다 1씩 감소하여 카운터의 값이 0이 되면 그 노드는 데이타 전송을 멈추게 된다. 그 후에 입력 전용 버스에 빈 슬롯이 지나가면 그것은 다른 노드들이 전송할 데이타를 가지고 있지 않다는 것을 의미하므로 그 카운터는 다시 채워져서 P라는 카운터의 값을 가지게 된다. 이 카운터의 값, P는 각 노드들에 대해 그 값들을 달리하는 방법으로 각 노드에 적절한 밴드폭을 할당해 줄 수 있다. S 프로토콜에서는 노드의 위치와 데이타 트래픽에 따라 할당되는 밴드폭과 다음 빈 슬롯을 기다리는 지연 시간이 약간씩 달라질 수 있다. 하지만 모든 노드들이 항상 데이타를 보내려고 하는 최악의 경우에는 각각의 노드는 완전히 공평하게 데이타를 보낼 기회를 갖게 된다.

HANGMAN은 노드들이 동기적 전송을 할 수 있도록 하는 수단을 제공한다. 그러기 위하여 슬롯들을 2048개 단위로 사이클 그룹을 만들고 노드가 각 사이클의 일부분을 예약할 수 있도록 한다. 슬롯을 표시하기 위하여, slot generator는 예약된 슬롯의 송신단 주소 부분에 예약한 노드의 주소를 넣고 그 슬롯이 동기식 트래픽 전송을 위해 예약 되어 있음을 표시해 준다. 그러면 예약을 한 노드는 자기가 예약한 슬롯을 찾아 내기 위하여 동기식 전송을 위해 예약되었다고 표시된 슬롯의 송신단 주소 부분을 확인하면 된다. 데이타 율이 1 Gb/s일 때 예약된 한 슬롯은 500Kb/s의 밴드폭에 해당하므로 동기식 서비스는 500Kb/s 또는 그 이상의 밴드폭을 요구하는 응용에 사용될 수 있다. 그림 5-3는 동기식 슬롯과 비동기식 슬롯의 구성과 한 슬롯의 형태를 나타낸다. 한 슬롯의 첫부분에 있는 MFC(MAC frame control header)비트는 동기식 슬롯과 비동기식 슬롯을 구분짓기 위하여 사용된다. 만약 한 동기식 슬롯이 예약을 한 노드에 의하여 사용되지 않으면 그 슬롯은 그 이하의 노드에서 비동기식으로 사용될 수 있다. 물론 동기식 슬롯으로 예약되지 않은 슬롯은 비동기적으로 사용된다.

HANGMAN은 하나의 노드에 여러 호스트들을 연결할 수 있게 되어 있고, 이를 이용하여 원하는 실험을 할 수 있도록 구성되어 있다. 각 노드는 내부 프로세서를 가지고 있다. 이 프로세서는 현재 각 노드에서 트랜스포트 계층 프로토콜의 기능까지 수행할 수 있도록 구현되어 있다. 이러한 특징은 카메라와 비디오 스크린 장치를 연결할 때 그 사이에 컴퓨터등의 도움없이 노드에 직접 연결할 수 있게 해 준다. HANGMAN에서 각 노드는 16비트 노드 주소를 가지며 broadcasting도 가능하다. 하지만 앞 부분에서도 언급했듯이 HANGMAN은 ATM과 호환이 안된다는 단점을 가지고 있다.

5.2 MSN(Manhattan Street Network)

MSN은 1985년 BELL연구소의 Maxemchuk에 의해 개발된 것으로 망 네트워크의 한 형태이다[6][7]. 그림 5-4은 6x6 MSN의 형태를 보여 주며 Manhattan의 거리 모습에 비유하여 행을 street, 열을 avenue라 부르기도 한다. MSN은 구조의 특징상 짝수개의 행과 열로 되어 있으며 각 노드들은 두개의 입력 링크들과 두개의 출력 링크들을 가지고 있어 전송된 패킷에 대해 routing을 결정하

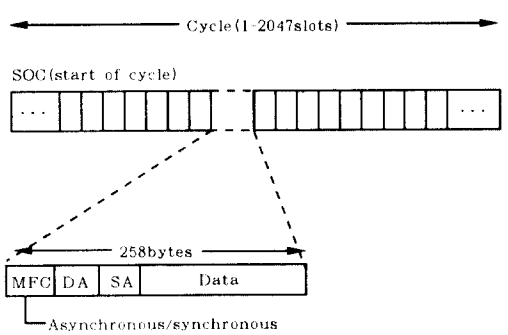

그림 5-3. The slot template

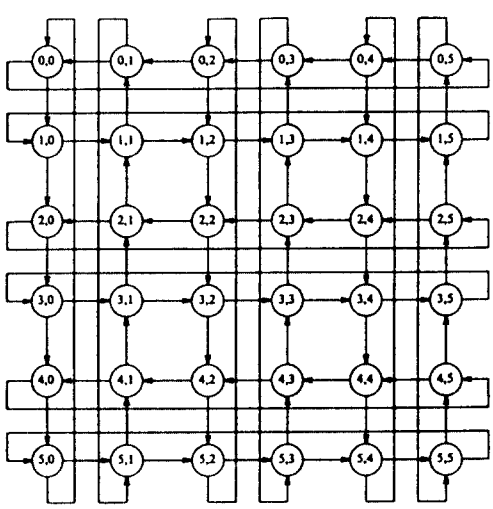

그림 5-4. 6x6 node MSN

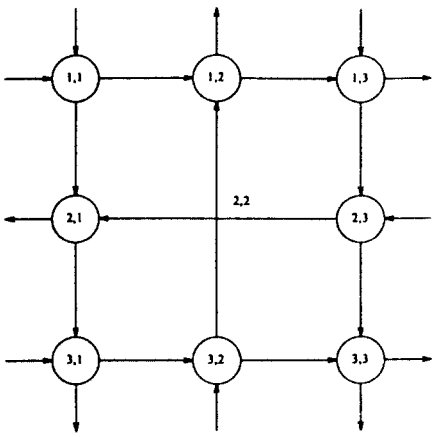

그림 5-5. Operation of the MSN when nodes fail

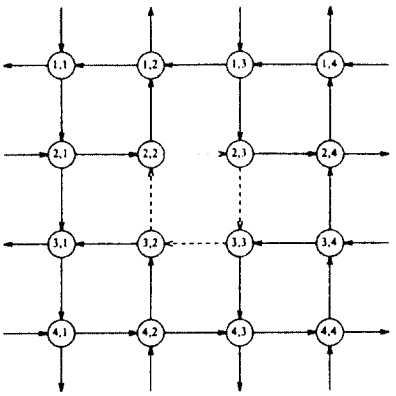

그림 5-6. Operaton of the MSN when links fail

여야 한다. 하지만 routing 규칙은 단순해야 할 필요가 있다. 행이나 열의 링크가 한 방향으로 어어지면 그 인접 링크는 반대 방향으로 이어져 있고 끝부분이 서로 연결되어 있어 cyclic 구조를 이루고 있다.

각 노드는 망에서의 위치에 따라 주소를 가지는데 행과 열의 쌍으로 되어 있다. 예로 (0,0)는 첫째행, 첫째열 노드에 대한 주소이고 (2,3)은 세째행, 네째열 노드의 주소이다. MSN은 cyclic 구조를 가지고 있기 때문에 routing 규칙은 수신단 노드와 송신단 노드 사이의 상대적 거리에 의해 이루어 진다. 만약 두 셀들이 같은 방향으로 진행하고자하면 충돌이 발생한다. 이 때 한 셀을 다른 셀과 다른 링크로 routing시킨다. 그러면 한 셀이 좀 더 먼 경로를 통해 가겠지만 결국 수신단 노드에 도착할 수 있다.

만약 한 노드가 고장이 나면 그림 5-5에서와 같이 고장난 노드에서 두 링크들을 서로 연결시킨다. 링크에 문제가 생길 경우 다음의 규칙에 따라 해결한다.

- 행 링크에서 신호가 오지 않으면 열 링크로의 신호 전송을 멈춘다.
- 열 링크에서 신호가 오지 않으면 행 링크로의 신호 전송을 멈춘다.

노드 사이에 전송할 데이타가 없을 지라도 그 사이에는 비트들이 연속적으로 전송된다. 이것은 패킷간의 동기를 맞추기 위한 수단으로 사용되며 이 신호의 전송이 멈추면 노드나 링크가 고장났음을 감지할 수 있다. 그림 5-6은 노드(2,2)와 노드(2,3)사이에 링크가 고장났을 경우를 나타낸다.

MSN은 다수의 링크 고장이나 높은 부하에서 잘 동작할 수 있다는 장점은 있으나 multicasting이나 broadcasting의 기능이 없고 트래픽의 양에 따라 셀이 매우 먼 경로로 진행할 수 있다. 이것은 보장된 성능을 요구하는 응용에서는 나쁜 결과를 가져올 수 있다.

5.3 AN2 Switch

AN2는 Digital Equipment Corporation의 Systems Research Center에서 개발한 연구용 ATM switch이다. Switch의 크기는 16x16으로 LAN switch로 특별히 설계된 것이다. 이 switch는 gigabit로 데이타를 빠르게 전송할 수 있으며 트래픽의 분리 기능과 input buffering, 그리고 loop-by-loop flow control등을 사용한다[2].

AN2 switch를 통한 ATM 트래픽은 continuous bit-rate와 variable bit-rate로 나누어 진다. Continuous bit-rate를 지원하기 위하여 1000셀들을 가상적인 프레임으로 묶고 그 프레임안에 셀들을 할당해 주게 된다. Continuous bit-rate용도로 할당되지 않은 셀들은 variable bit-rate용도로 사용될 수 있다. Variable bit-rate로 사용될 경우 서로 경쟁이 발생할 수 있는데 경쟁에서의 손실을 줄이기 위하여 그림5-7에서와 같이 hop-by-hop flow control과 함께 input buffer를 둔다. 각 스위치에서 virtual circuit는 하나의 window 카운터를 가지고 있어 셀을 다음 스위치로 전송할 때마다 카운터의 값을 1씩 감소시킨다. 카운터의 값이 0이 되면 셀의 전송을 멈추고 buffer에 들어 오는 셀들을 저장하여야 한다. 다음 스위치에서 한 셀을 내보낼 때마다 ATM 헤더의 GFC비트로 acknowledgement를 전 스위치에 보낸다. 스위치가 acknowledgement를 받을

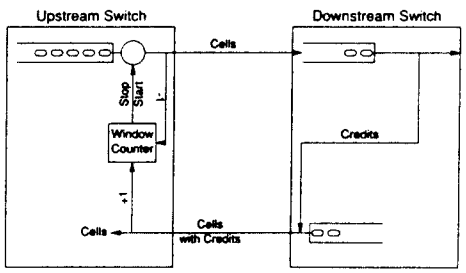

그림 5-7. AN2 Flow Control

때마다 카운터는 1씩 증가한다. 여기서 입력의 버퍼 공간이 홉의 round-trip시간과 같을 필요가 있다. 만약 버퍼 공간이 round-trip시간보다 작고 virtual circuit가 최대율로 셀들을 전송하고 있다면 acknowledgement를 받기 전에 셀들을 다 보내고 기다려야 한다. 만약 버퍼 공간이 round-trip시간보다 크면 virtual circuit는 셀들을 계속 보내게 되고 congestion이 발생하게 된다.

AN2 스위치는 서로의 존재를 자동적으로 알고 있으며 스위치들과 호스트들 사이의 route를 생성한다. 이 기능은 Autonet packet network를 위해 개발된 것으로 AN2 스위치 설계에 영향을 주었다[8]. 이러한 구성은 flooding 알고리듬을 이용하여 이루어지며 각 스위치들은 특별한 프로토콜을 이용하여 서로의 위상 정보들을 교환한다.

5.4 ATOMIC LAN

ATOMIC은 1987년 Caltech에서 만들어진 Mosaic chip을 이용하여 USC(University of Southern California)의 Information Science Institute에 의해 개발된 gigabit LAN이다[3]. 그림 5-8에서 나타나듯이 Mosaic chip은 둘레에 8개의 0.5 gigabit 통신 채널을 가지며 chip의 버젼에 따라 각 채널은 500에서 800 Mb/s까지의 범위에서 동작한다. ATOMIC은 MSN(Manhattan Street Network)구조와 유사하지만 끝부분이 서로 연결되어 있지 않다.

Mosaic은 각 패킷에 x 카운터와 y 카운터를 두어 송신단과 수신단의 상대적인 거리를 송신단에서 두 카운터에 적어 둔다. 이 패킷은 각 mosaic chip을 routing해 가면서 카운터의 값을 1씩 증가 또는 감소시킨다. 두 카운터의 값이 0이 될 때 이 패킷은 수신단에 도달한 것이다. x 카운터의 경우 패킷이 오른쪽으로 진행하면 1씩 감소하고 왼쪽으로 진행할 때 1씩 증가한다.

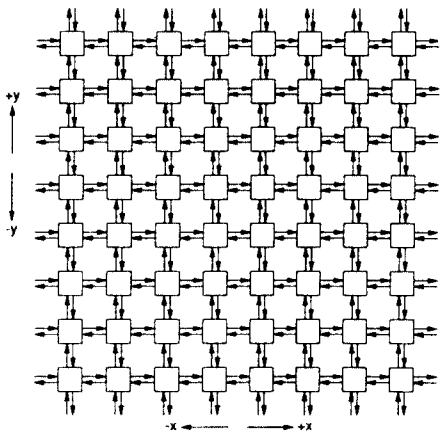

그림 5-8. Mosaic Chips in an 8x8 Mesh

y 카운터의 경우도 x 카운터와 유사하게 위로 진행하면 1씩 감소하고 아래로 진행할 때 1씩 증가한다. Routing의 방법은 먼저 x 카운터의 값이 0이 될때까지 x 방향으로 routing을 한 후 y 카운터의 값이 0이 될 때까지 routing하게 된다. ATOMIC은 끝부분이 서로 연결되어 있지 않기 때문에 끝부분에 호스트가 연결되어 있을 경우 routing이 불가능한 경우가 생길 수 있다. 이러한 문제점을 해결하기 위하여 한 패킷에 하나 이상의 x, y 카운터의 쌍을 가질 수 있게 하는 것이다. 그러면 처음 카운터의 쌍이 0이 될 때까지 routing을 한 후 다른 카운터 쌍이 있는지를 확인하여 아직 남아 있으면 그 패킷을 네트워크에 다시 내 보낸다. 그것에 대한 한 예가 그림 5-9에 나타나 있다. 만약 호스트 A에서 호스트 D로 routing한다면 x, y 카운터가 (2, 2)의 값으로 되어 있으면 올바르게 routing하여 갈 수 있다. 하지만 호스트 D에서 호스트 A로 패킷을 전송한다면 x 방향으로 routing하기 전에 y 방향으로 먼저 routing하여야 하므로 패킷안에 카운터의 쌍이 하나라면 패킷 전송이 불가능하게 된다. 하지만 패킷안의 카운터 쌍의 값을 (0,2), (-2,0)로 한다면 먼저 y 방향으로 routing할 수 있으므로 패킷 전송이 가능하게 된다. 이렇게 개선된 routing 방법은 Mosaic mesh에서 한 패킷을 여러 경로를 통하여 routing할 수 있도록 한다. 또한 Mosdaic mesh를 하나의 집중기로 사용하여 집중기들 사이에 routing할 수 있는 방법이 주어진다면 여러 Mosaic mesh를 연결하여 사용할 수 있다. 그림 5-9에서 호스트 B로부터 호스트 C로 routing하기 위하여 패킷의 카운터 쌍들에 [(1,0),(0,4),

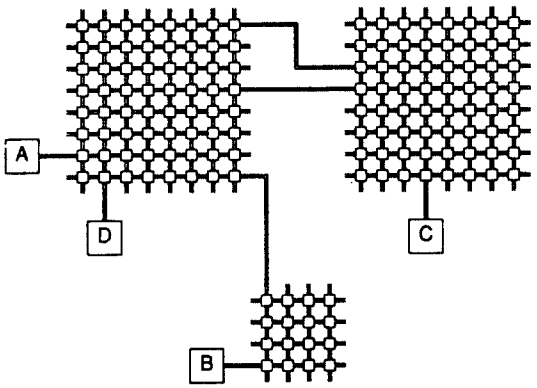

그림 5-9. ATOMIC LAN with 3 Mosaic Meshs

(0,4),(4,0),(0,-5)]을 지정해 주면 된다.

Mosaic은 가변적인 길이의 패킷을 전송할 수 있는 기능을 제공하는데 이 때 두 개의 패킷이 동시에 한 경로를 통과하여 진행하려 한다면 한 패킷은 다른 패킷이 그 Mosaic chip을 다 지나갈 때까지 기다려야 한다. Mosaic chip은 byte-to-byte 흐름 제어를 하므로 두 chip 사이의 길이가 0.6m를 초과할 수 없다. 이것은 네트워크를 구성하기에는 너무 짧은 거리이다. 이 문제를 해결하기 위하여 Mosaic chip과 함께 Slack chip을 함께 이용하여 window에 의한 flow control를 수행함으로써 제어 단위를 1byte에서 수 bytes까지로 늘릴 수 있다.

ATOMIC에는 주소를 관리하는 address consultant라는 것이 있다. Address consultant는 네트워크에 flooding을 하여 연결된 모든 노드들의 위치를 알아 낸다. 여기서 얻은 정보를 이용하여 임의의 두 호스트들 사이의 route를 지정하는 routing table을 작성한다. 한 호스트가 어떤 다른 호스트에게 데이타를 보내려고 할 때 그 호스트는 address consultant에게 routing 경로에 대해 문의한다. 집중기들 사이의 연결 상태에 따라 두 호스트들 사이에 여러 경로들이 있을 수 있으므로 address consultant는 집중기들 사이의 링크상에서의 트래픽을 조정하기 위하여 또는 링크의 고장난 부분의 복구를 위하여 같은 요구에 대해 다른 routing경로를 지정해 줄 수 있다. Address consultant는 특별한 multicast 주소에 대한 패킷들을 어디로 보내야 할 것인가를 mosaic mesh내의 프로세서들에게 알려 준다. 또한 address consultant는 호스트들에게 multicast 주소를 어떤 프로세서에게 보내야 하는가를 알려 주어 multicasting 기능을 수행한다.

ATOMIC은 트래픽을 정교하게 조정할 수 있고 multicasting 기능을 제공한다는 장점은 있으나 한정된 지연 시간내에 전송을 완료하지 못하는 단점을 가지고 있다.

5.5 CSMA/RN(Carrier Sense Multiple Access/ Ring Network)

CSMA/RN은 1991년 Old Dominion University의 한 연구팀에 의해 개발된 gigabit packet ring의 한 종류이다[4]. 이것은 그림 5-10에서와 같이 packet ring과는 좀 특이한 access 프로토콜을 가지고 있다. 링으로부터 전달되어 들어 오는 패킷은 일단 delay buffer에 저장되고 같은 패킷이 제어부로 직접 전달된다. 제어부는 전송된 패킷에 있는 정보를 통하여 delay buffer안의 데이타 흐름을 제어한다. 패킷의 수신단 주소가 제어부의 호스트에 해당되면 delay buffer안의 데이타를 호스트로 전송하고 그렇지 않으면 링으로 데이타를 그대로 내보낸다. 링이 ACTIVE일 때(링으로부터 패킷이 delay buffer로 전송될 때)는 호스트는 데이타를 전송할 수 없다. 링이 INACTIVE일 때 호스트는 제어부를 거쳐 패킷을 전송하는데 그 동안 링으로부터 새로운 패킷이 전송되면 제어부는 호스트의 패킷 전송을 멈추어야 한다. 링으로부터의 패킷은 일단 delay buffer에 저장되므로 제어부는 그 동안 호스트의 데이타 끝에 truncation mark를 한 후 중단할 수 있다.

만약 CSMA/RN이 비교적 짧은 길이의 셀을 사용한다면, delay buffer의 길이를 셀과 같게 하여 호스트가 셀을 전송하는 동안 링으로부터 새로운 셀이 도착하더라도 그것이 delay buffer에 저장되는 동안 호스트는 셀의 전송을 끝낼 수가 있다. 하지만 CSMA/RN은 가변적인 길이의 패킷을 사용하므로 충돌에 대한 대비 조치를 하여야 한다.

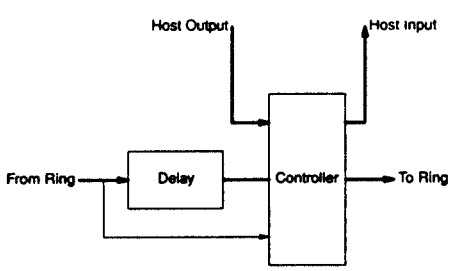

그림 5-10. CSMA/RN Interface

한가지 방법은 전송중인 호스트의 패킷 뒤에 abort token을 두고 수신단에서 그 패킷을 버리는 방법이 있지만 버려지는 패킷들에 의해 밴드폭의 낭비를 초래할 수 있다. 다른 한 방법은 delay buffer를 호스트의 transmission buffer에 연결하여 transmission buffer안의 호스트의 패킷 뒤에 링의 새로운 패킷을 두는 방법이다. 하지만 이 방법은 결과적으로 링 크기를 증가시키게 되므로 전송 시간이 가변적으로 증가하게 된다. 증가하는 가변적인 전송 시간은 적당한 최대 패킷 크기와 링에 연결할 수 있는 최대 갯수를 정하여 한정된 값으로 제한할 수 있다.

CSMA/RN은 수신단에서 패킷을 받은 후 제거하기 때문에 송신단에서 없애는 것보다 링의 밴드폭을 평균적으로 두 배로 사용할 수 있고 패킷들이 링에 있는 시간이 작아지므로 충돌이 발생할 가능성이 줄어든다.

VI. 결 론

지금까지 고속 LAN에 관한 세계적인 연구 동향 및 표준화 동향과 그동안 비교적 소개가 안된 것들을 중심으로 살펴보았다. 특히 사용가능성이 큰 고속망으로서 Fast Ethernet 및, LAN 과는 성격이 약간 다른 HIPPI protocol 을 소개하였다. 이어서 고속 Ring Network 중 영향력이 크다고 생각되는 것으로 CFR, Orwell Ring, Meta Ring 을 소개하고 ISO/IEC JTC1 SC6에서 표준안으로 검토중인 일본의 ATMR 및 한국의 HMR에 관하여 중점적으로 살펴보았다. 이외에도 소개된 LAN 기술로는 MSN, crossbar 방식의 AN2 Switch, Mosaic 방식의 Atomic LAN 및 링 타입의 HANGMAN, CSMA/RN ring, CFR 등을 살펴보았다.

본 연구에서 지면 관계로 다루지는 못하였으나 중요한 기술로는 미국의 ATM Forum을 중심으로 표준화가 진행중인 Star Switch방식의 ATM LAN과 IEEE 802. 9에서 연구중인 IS(Integrated Service) LAN 등이 있는데, 이미 상업화가 많이 진행중이므로 국내에서도 각별한 주의를 기울일 필요가 있다고 생각된다.

고속 LAN은 광대역 ATM 시스템에 비하여, 규모가 작은 대신 발전 속도가 빠르고 실지가 용이하다. 실제로 광대역 고속 전송 응용 시스템은 고속 LAN에서 우선 사용될 가능성이 크고 광대역 ATM 시스템에 파급 효과가 큰 만큼 많은 연구가 필요하다.

참 고 문 헌

1. R. L. Sites, Alpha Architecture Reference Manual, Digital Press, Bedford, Mass.
2. C. Partridge, Gigabit Networking, Addison-Wesley Publishing Company, 1994.
3. W. J. Dally, and C. L. Seitz, "Deadlock-Free Message Routing in Multi processor Interconnection Networks," IEEE Trans. on Computers, Vol. 36, No. 5, May, 1987.
4. E. C. Foudriat, K. Maly, C. M. Overstreet, S. Khanna, and F. Paterra, "A Carrier Sensed Mutiple Access Protocol for High Data Rate Ring Networks," ACM Computer Communication Review, Vol. 21, No. 2, Apr. 1991, pp. 59-70.
5. J. Limb, "A Simple Multiple Access Protocol for Metropolitan Area Networks," Proc. ACM SIGCOMM'90, Philadelphia, Sep. 24-27, 1990, pp. 69-78.
6. N. Maxemchuk, "Regular Mesh Topologies in Local and Metropolitan Area Networks," AT&T Technical Jour. , Vol. 64, No. 7, Sep. 1985, pp. 1659-1685.
7. N. F. Maxemchuk, "Routing in the Manhattan Street Network," IEEE Trans. on Communications, Vol. 35, No. 5, May, 1987, pp. 503-512.
8. M. D. Schroeder, A. D. Birrell, M. Burrows, H. Muarry, R. M. Needham, T. L. Rodeheffer, E. H. Satterthwaite, and C. P. Thacker, "Autonet : A High Speed, Self-Configuring Local Area Network Using Point-to-Point Links," IEEE J. Sel. Areas in Commun. , Vol. 9, No. 8, Oct. 1991, pp. 1318-1335.
9. D. Tolmie,J. Renwick, "HIPPI : Simplicity Yields Success," IEEE Network Mag. , Vol. 7, No. 1, Jan. 1993, pp. 28-33.
10. G. Watson,S. Ooi,D. Skellern,and D. Cunningham, "HANGMAN Gb/s Network," IEEE Network Mag. , Vol. 6, No. 4, Jul. , 1992, pp. 10-18.
11. G. Watson, and S. Tohme, "S^{++} : A MAC Protocol for Gbit/s Networks," Proc. 2nd Telecom Paris/Universitat Stuttgart Joint Wksp. on High-Speed Networks, Stuttgart, Germany, Sep. 1991.
12. D. J. Greaves, D. Lioupis, and A. Hopper, "The Cambridge Backbone Ring," IEEE INFOCOM'90, San Francisco, Jun. 1990, pp. 8-14.

13. A. Hopper, "Data Ring at Computer Laboratory, University of Cambridge," Computer Science and Technolodge;NBS Special-pub 500-31, US National Bureau of Standards, Washinton D. C. , 23-23, Aug. 1977.
14. Andrew Hopper, "The Cambridge Fast Ring Networking System," IEEE Trans. on Computers, Vol. 37, No. 10, Oct. 1988, pp. 1214-1223.
15. R. M. Falconer et al, "Orwell : a protocol for an integrated services local network," BT Technology J. , Vol. 3, No. 4, Oct. 1985.
16. I srael Cidon, and Yoram Ofek, "MetaRing-A Full-Duplex Ring with Fairness and Spatial Reuse," IEEE. Trans. on Communications, Vol. 41, No1, Jan. 1993, pp. 110-120.
17. H. Ohnishi, "ATM RING PROTOCOL AND PERFORMANCE," ICC. '89, pp. 394-398. 18. Kasuo Imai, "ATMR : Ring Architecture for Broadband Networks," IEEE. Globecom '90, 1990, pp. 1734-1738.
19. "Asynchronous Transfer Mode(ATM) Based Medium Access Method and Physical Layer Specifications," ISO/IEC JTC1/SC6 N 7662, Temporary No : ISD-43r, 1992. 7.
20. 전병천, "멀티미디어 통신을 위한 Giga bits/s급 공유매체 접근제어 프로토콜 설계," 박사학위논문, 충남대학교, 1992. 8.
21. 전병천, 김대영, 한치문, "다중 윈도우토큰을 사용한 ATM-LAN 프로토콜," JCCI-'93, pp. 79-82.
22. 전병천, 김대영, "멀티미디어 통신을 위한 기가비트 LAN/MAN 프로토콜," 한국통신학회 추계종합학술발표회 논문집, 1992, pp. 69-72.
23. 두소영, 김태균, 김대영, "HMR 프로토콜을 사용한 GFC 프로토콜의 성능분석," 한국정보과학회 '93 가을 학술발표논문집, 1993, pp. 603-606.
23. Stephen Saunders, "Ethernet Gears Up for 100 Mbit/s," Data Communications, Jan. 1993, pp. 35-38.
24. 윤종호, 신병철, 최준균, "FAST ETHERNET 프로토콜 : 100-Mbits/s CSMA/CD and 100BASE-VG," 한국통신학회지, 제 10권 제 12호, 1993년 12월, pp. 899-910.
25. Johnson, "802. 3 higher speed CSMA/CD," IEEE802. 3 WG Document, Denver Jul. 1993.
26. Albrecht and D. Cunningham, "100BASE-VG IEEE Tutorial," IEEE802. 3 WG Document, Denver, Jul. 1993.
27. Raj Jain, "FDDI : Current issues and future plans," IEEE Commun. Mag., Sep. 1993, pp. 98-105.

金 炯 秀

• 1963년 2월 17일생
• 1985년 2월 : 연세대학교 전자공학과 학사
• 1987년 2월 : 한국과학기술원 전기 및 전자공학과 석사
• 1987년 3월 ~ 1992년 2월 : 한국통신 통신망연구소 전임연구원
• 1992년 3월 ~ 현재 : 한국과학기술원 전기 및 전자공학과 박사과정
※ 주관심분야 : High Speed Network, Multimedia 통신, B-ISDN

李 光 哲

• 1989년 2월 : 과학기술대학교
전기 및 전자공학과 졸업
• 1993년 3월 ~ 현재 : 한국과학기술원
전기 및 전자공학과 석사과정

申 炳 喆

• 1952년 11월 2일생
• 1975년 2월 : 서울대학교 전기공학과 졸업
• 1977년 2월 : 한국과학기술원
전기 및 전자공학과 석사 졸업
• 1984년 8월 : 한국과학기술원
전기 및 전자공학과 박사 졸업
• 1977년 2월~1980년 2월 : 한국전자통신연구소 연구원
• 1987년 3월~1988년 2월 : SRI International에서 Post Doc. 연수
• 1984년 9월~현재 : 한국과학기술원 전기 및 전자공학과 부교수(현재)
※ 주관심분야 : Multimedia 통신, High Speed Network, Wireless Communication, ATM Network 분야

尹 鍾 浩

• 1957년 12월 18일생
• 1984년 2월 : 한양대학교 전자공학과 학사
• 1986년 2월 : 한국과학기술원
전기 및 전자공학과 석사
• 1990년 8월 : 한국과학기술원
전기 및 전자공학과 박사
• 1991년 9월 ~ 현재 : 한국항공대학교
항공통신정보공학과 조교수
※ 주관심분야 : 고속패킷망, LAN, 변복방식 및 장치 개발, 성능분석