

主題

개방형 하드웨어 인터페이스 기술

한국전자통신연구원 김 봉 태

차례

- I. 서 론
- II. 공통 스위치 인터페이스
- III. 네트워크 프로세서
- IV. 스위치 패브릭
- V. 결 론

I. 서 론

통신망이 멀티서비스를 지원하는 패킷 기반망으로 급속하게 변화되어 감에 따라 서비스별 상이한 QoS를 효과적으로 지원하면서 새로운 서비스의 도입 및 폐기를 신속하게 처리할 수 있도록 하드웨어 종속 계층과 미들웨어 계층, 응용서비스 계층으로 구분하여 서로 독립된 수평 구조를 갖는 개방형 통신 시스템 관련 기술연구 및 표준화가 MSF(Multi-service Switching Forum)를 중심으로 활발하게 추진되고 있다.

개방형 시스템의 하드웨어 기술은 반도체 기술의 급속한 발전과 더불어 시스템 구성 요소 및 기술의 수명주기가 점점 짧아지고 있으며, 새로운 기술 추세에 신속하게 대응하기 위하여 시스템 구성하는 주요 기능 모듈을 표준화 하고, 이들간 상호 정합 규칙을 정하며, 각 기능 모듈을 프로그램 가능하게 하는 일련의 작업들이 네트워크 시스템의 주요 구성요소 기술을 개발 하는 벤더들을 중심으로 추진되어 상당한

진전을 이루었다.

이러한 작업들의 대표적인 예가 CSIX(Common Switch Interface) 포럼 및 CPIX(Common Programming Interface) 포럼의 활동이라 하겠다. 본 고에서는 CSIX 및 CPIX 포럼의 표준 인터페이스 규격 내용을 요약하여 소개하고, 개방형 시스템의 주요 하드웨어 기술로 부상하고 있는 네트워크 프로세서 기술 및 초고속 스위치 패브릭 기술 동향을 요약 소개한다.

II. 공통 스위치 인터페이스(CSIX) 규격

CSIX규격은 패스트 이더넷, 기가비트 이더넷 등의 LAN 스위치, 라우터, ATM 스위치, TDM 스위치 등의 네트워크 시스템의 데이터 전달 경로를 그림 1과 같이 이들 네트워크 시스템의 외부와 접속하는 PHY/MAC 기능 블록과 데이터의 어드레스 변환,

통신 프로토콜 프로세싱, 큐잉, 스케줄링 등 트래픽 처리기능을 담당하는 트래픽 매니저(Traffic Manager), 그리고 이들을 상호 접속하여 교환기능을 수행하는 공통 스위칭 패브릭(Switching Fabric)으로 구분하고, 트래픽 매니저와 스위치 패브릭 간에 공통으로 적용할 수 있는 인터페이스 규격을 정의한 것이다.

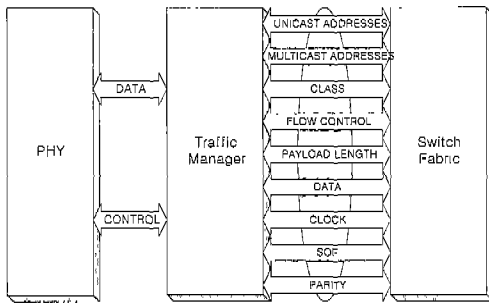


그림 1. CSIX 표준규격 개념

특기할 것은 CSIX규격은 백 플레인(Backplane) 인터페이스 규격이 아니라 동일 보드상에서 적용 가능한 병렬 데이터 버스 인터페이스 규격이라는 것이다. 따라서, CSIX에 기반한 대용량의 네트워크 시스템을 구성하기 위해서는 논리적으로는 스위칭 패브릭과 트래픽 매니저가 분리되어 있으나, 물리적으로는 스위치 패브릭과 트래픽 매니저가 동일 보드상에서 구현되어야 함을 뜻한다. 이를 위해, 스위칭 패브릭 기능 블록을 코어 스위치 패브릭과 확장 스위치 패브릭으로 분리 구성하고, 이들 간을 상호 접속하는 고속의 직렬 정합 링크가 적용되기도 한다.

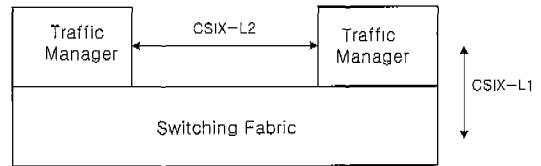


그림 2. CSIX의 계층적 구조

이러한 시도는 1998년 영국에 기반을 둔 초고속 스위치 패브릭 전문 회사인 Power-X사와 미국의 트래픽 처리 기술 전문 회사인 XaQti사가 주축이 되어 전기적, 논리적 인터페이스 규격 초안을 제시하였으며, 이후 전세계의 주요 통신 칩 제조사들이 대부분 참여하는 산업체 컨소시엄으로 발전하여 스위치 패브릭과 고성능 네트워크 프로세서로 구성된 트래픽 처리 모듈간의 인터페이스 표준화를 시도하였다. CSIX 규격의 주요 특성은

- 트래픽 매니저(TM)와 스위치 패브릭(SF) 간의 표준 인터페이스 규정
- 동일 보드상의 15~20cm의 거리에서 포트당 최대 32Gbps 데이터 전달 속도 달성
- 셀 기반 및 패킷 기반 모드를 동시에 지원
- Unicast, broadcast, multicast 동작 지원
- 최대 4,096개의 CSIX port 수용
- TM-SF, TM-TM간 in-band signaling
- 8-bit traffic class variable

CSIX 규격은 논리적으로는 그림 2에서의 같이 스위치 패브릭과 트래픽 매니저 간의 L1(Level 1) 상호 작용, 그리고 스위치 패브릭을 경유하여 연결되는 트래픽 매니저들 간의 L2(Level 2) 상호 작용 등 2계층의 계층적 구조를 갖는다.

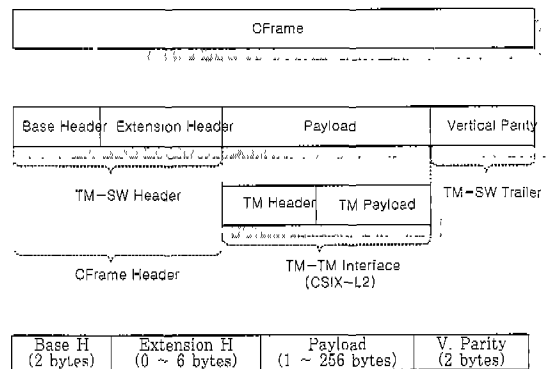


그림 3. C-Frame 포맷

트래픽 매니저와 스위치 패브릭 간의 정보 전달을

위해서 그림 3과 같이 정보교환 단위로서의 C-Frame을 정의하고 있다. C-Frame은 헤더와 페이로드, 그리고 병렬 데이터 버스의 버티칼 패리티 정보를 담고 있는 프레임으로 구성된다. C-Frame 헤더는 다시 베이스 헤더와 확장 헤더로 나뉘어 진다. 페이로드는 1~256바이트 까지 가변할 수 있으며, C-Frame의 최대 길이는 266 바이트 이다. TM과 TM간의 인터페이스 메시지 길이가 최대 C-Frame 페이로드 길이를 초과하지 않을 때는 하나의 C-Frame이 사용된다. TM과 TM간의 메시지 길이가 최대 C-Frame 페이로드 길이를 초과하면 여러 개의 C-Frame이 사용된다.

그림 4와 그림 5는 CSIX기반 스위칭 시스템 구성 예를 보인 것으로서, 그림 4은 비교적 소규모의 시스템 구성을, 그림 5는 대규모의 시스템 구성을 위해서 스위치 패브릭 내에 커넥터를 두고 스위치 패브릭을 모듈화 하여 구성한 예를 보여주고 있다.

데이터는 점대점 병렬 데이터 버스에 32비트의 배수, 즉 $n \times 32$ bits($n=1,2,3,4$)로 전달되며, 32비트 데이터 버스 100MHz로 동작시키면 2.5Gbps (OC-48), 64비트 버스에 200MHz 속도로 동작시키면 10Gbps (OC-192)가 된다. 그림 6은 32비트 인터페이스의 물리적 구현을 나타내고, 그림 7은 64비트 인터페이스의 물리적 구현을 나타낸 것이다.

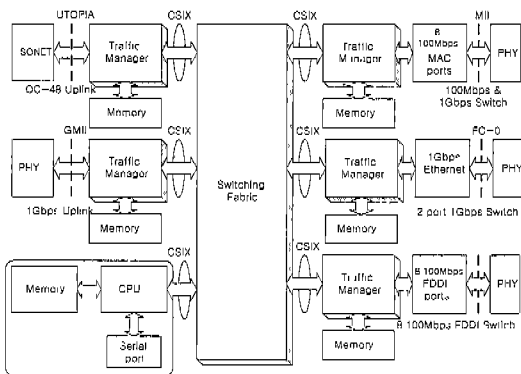


그림 4. CSIX기반 통합 스위치 구성 예

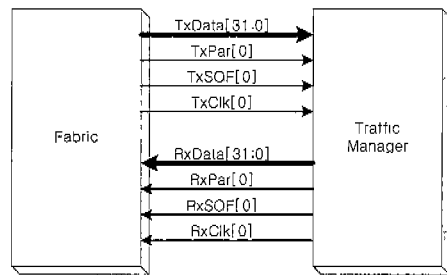


그림 6. 32-bit 인터페이스

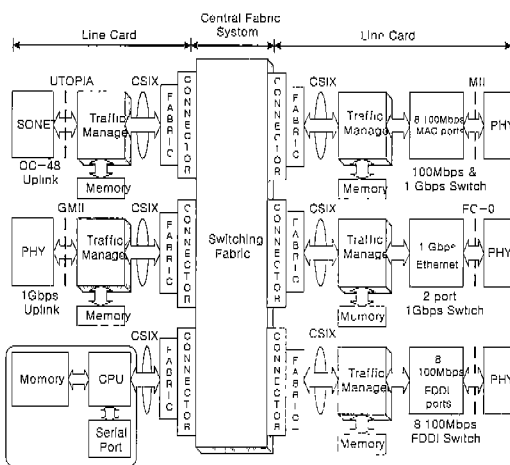


그림 5. CSIX기반 모듈화 스위치 구성 예

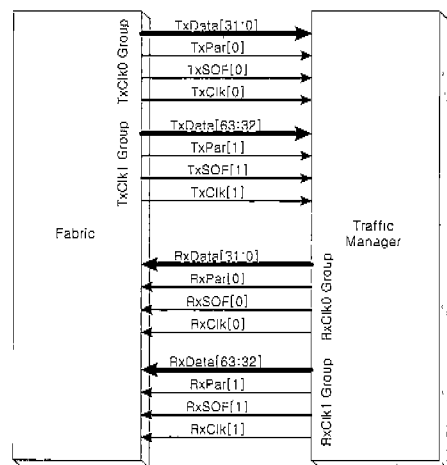


그림 7. 64-bit 인터페이스

Ⅲ. 네트워크 프로세서

CSIX 표준이 네트워크 시스템 설계에 있어서 모듈화 개념을 적용하여 시스템에 융통성을 부여하고, Time-to-market을 단축하는 등 시스템의 가격 경쟁력을 향상시키는 계기를 제공하였다면, 장시간의 개발 기간과 고비용의 ASIC 기반 네트워크 시스템의 정합 모듈(트래픽 매니저)의 설계에 있어서 큰 전환점을 맞이하게 된 것은 그림 8에 보인 것과 같은 고성능 네트워크 프로세서(Network Processor) 기술의 출현이라 할 수 있다.

네트워크 프로세서는 정합기능 수행에 자주 활용되는 기능 블록들과 8~32개의 고성능 RISC 또는 CSIC 구조의 프로세서 Core를 하나의 반도체 칩에 집적하여, 통신 프로토콜 프로세싱, 패킷 스케줄링, 큐잉, 등의 작업을 고속으로 수행할 수 있으며, 필요에 따라 기능을 업그레이드 하거나 새로운 기능을 손쉽게 추가 할 수 있는 프로그램이 가능한 디바이스 기술로 정의 할 수 있다.

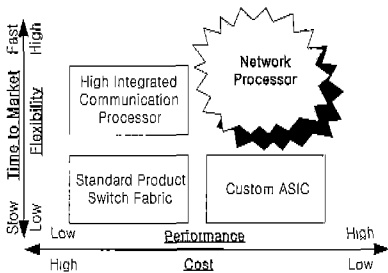


그림 8. 네트워크 시스템 기술대안 비교

네트워크 프로세서는 제조 회사마다 자기 다른 설계 방법을 취하고 있고, 네트워크 프로세서를 적용한 네트워크 시스템이 전체 네트워크에서 어떤 용도로 사용되는가에 따라라도 차이가 있다. 네트워크 프로세서가 갖는 가장 큰 장점은 프로그래밍 가능성이지만, 네트워크 프로세서 제조업체들이 프로그래밍 가능성을 제공하는 방식에는 크게 두 가지가 있다.

하나는 네트워크 시스템 생산업체에서 필요한 기능을 요구하면 네트워크 프로세서 제조업체에서 그 기능을 위한 코드를 작성해 주는 파라미터를 이용한 프로그래밍 방식이 있다. 이 경우에 네트워크 시스템 개발자는 자신이 개발하고 있는 시스템에 해당하는 특정한 값만 대입을 하면 된다. 대부분의 제조업체들이 취하고 있는 또 다른 방식은 소프트웨어 라이브러리를 제공하고 구매자가 그 라이브러리를 바탕으로 시스템을 구성하는 방식이다.

파라미터를 이용한 프로그래밍 방식을 제공하는 회사로는 Switchcore사와 Maker/ Conexant 를 들 수 있다. 구매자가 프로세서에 바로 프로그래밍을 할 수는 없지만 파라미터를 통해 변화를 줄 수 있도록 하는 방식이다. Agere/Lucent가 생산하는 FPP(Fast Pattern Processor) 같은 경우에는 FPL(Functional Programming Language)이라는 자체 프로그래밍 언어를 사용하게 하는데, 특정한 기능을 부여하는 데 있어서 최적의 방식으로 작업을 수행할 수 있게 하는 장점이 있는 반면에 프로그래밍이 FPL에서 제공하는 기능들에 한정되기 때문에 다양한 기능 구현이 힘들다.

데이터 분류(Classification) 작업은 가능한 경우의 수가 많고, 분류 작업 자체가 갖는 복잡성 때문에 고속의 일반적인 방법으로는 고속의 처리가 어렵다. Agere의 FPL은 이러한 통신 시스템에서의 분류의 문제를 해결하기 위한 분류 수행 언어이다. 분류 수행 언어(Classification Language)는 프로그래머가 분류 작업을 쉽고 효과적으로 표현할 수 있도록 하는 언어를 말한다. C언어와 비교해서 같은 작업을 수행하는데 더 짧은 코드로 가능하고 따라서 프로그래밍 오류를 줄일 수 있으며, 출시에 걸리는 시간이 짧고 여러 상황에 대해서 약간의 수정으로 재사용도 가능하다.

C-Port/Motorola사는 C/C++을 사용하여 프로그래밍을 가능하게 함으로써 서로 다른 플랫폼에서도 동일한 코드를 사용할 수 있게 한다. 네트워크 프

로세서의 구매자에게 친숙한 C나 C++같은 언어를 사용하게 하는 것은 동일한 네트워크 프로세서를 사용하더라도 개발자에 따라 차별된 제품을 구현할 수 있게 한다는 장점을 갖지만 최적의 코드를 제공할 수는 없다.

네트워크 프로세서는 초기에는 코어 스위치나 캐리어급의 라우터와 같은 대형 어플리케이션에 적합할 것이고 중형이나 소형 기업 근거리망 시장에서의 네트워크 프로세서의 사용은 계속 증가할 것으로 전망되고 있다. 이러한 시장에서의 성공을 계속 이어가기 위해서는 네트워크 프로세서 기술에서도 변화가 있을 것으로 보인다.

내장형 메모리(Embedded Memory) 기반의 네트워크 프로세서가 차세대 네트워크 프로세서의 주요 기술로서 부각되고 있다. 내장형 메모리를 사용하면 외부의 SRAM을 사용하는 방법보다 효과적인 패킷 프로세싱이 가능하다.

다른 한 방향은 네트워크 프로세서와 함께 사용할 수 있는 협동 프로세서(Co-processor)를 사용하는 것이다. 예를 들어 패킷 분류와 같은 특정 작업의 수행을 협동 프로세서를 통해 구현하는 것으로 네트워크 프로세서에 가중되는 부하를 크게 줄일 수 있다.

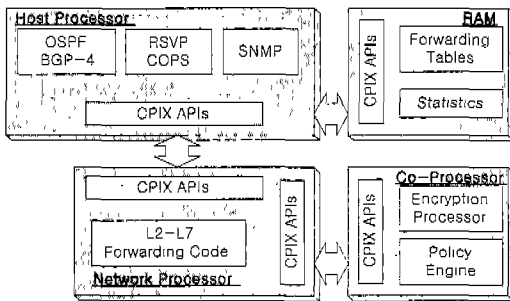


그림 9. CPIX 표준화 개념

CPIX포럼을 중심으로 이런 방법들에 관한 논의가 진행되고 있고, CPIX 포럼은 그림 9와 같이 네트워크 프로세서와 다른 데이터나 통신 구성요소 사이의

인터페이스를 표준화 하는 것을 목표로 하고 있으며 대부분의 네트워크 프로세서 제조 회사들이 참여하고 있다.

IV. 스위치 패브릭

인터넷 트래픽의 증가에 따라 단위 스위치의 포트 속도가 622Mbps, 2.5Gbps, 10Gbps, 40Gbps 등으로 고속화 추세에 있으며, 고속 대용량 패킷 스위치의 성능은 중재 제어 알고리즘의 성능에 의해서 좌우된다. 스위치의 중재 제어 알고리즘은 스위치의 포트 수가 많아질수록 성능이 떨어지게 되고, 스위치 포트의 속도가 높아질수록 구현이 어려워진다. 스위치 중재 제어 알고리즘 중 널리 알려진 것은 PIM, RRM, iSLIP, 2DRR, WFA, MUCS, RRGs 및 CORP들이 있으며, 이들 중재 제어 알고리즘을 사용하는 스위치는 입력단 패킷이 목적지별로 저장 및 관리되므로 HOL(Head-Of-Line) 블럭킹 현상이 제거되어 높은 처리율을 얻을 수 있다.

한편 테라 비트급 스위치 구성을 위해서는 스위치 포트 속도와 총 스위치 용량이 높아짐에 따라, 내부 속도 증가가 요구되는 공유 버퍼형이나 출력 버퍼형 보다는 입력력 버퍼형 스위치를 채택하고, 입력력 버퍼형 스위치의 성능이 낮은 단점을 보완하기 위하여 VOQ(Virtual Output Queuing)구조, 내부 speed-up 등을 도입하고 있다.

그러나 이들 입력 버퍼 방식의 VOQ 스위치는 많은 출력 큐가 필요하고, 입력 부하가 클 때는 스위칭 지연이 증가하며, 높은 처리율 얻기 위해서는 다수의 반복 경합 제어가 필요할 뿐만 아니라, 경합 제어 장치가 고속으로 동작하여야 한다. 이러한 기술적인 문제로 인하여 현재 이용 가능한 0.18um급 CMOS 반도체 기술을 사용하는 경우 스위칭 동작이 수행되는 단위 크로스바 스위치의 현실적인 구현 가능 최대 용량이 약 100Gbps로 단위 공유 버퍼 방식의 구현

가능 최대 용량인 40Gbps에 비교하여 그다지 높지 않다.

저속의 단위 스위치 소자들을 활용하면서 다단 스위치 구조를 채택하는 경우 이론적으로는 Tbps급의 대용량 스위치 구성도 가능하나, 경제성이 떨어져 제품화되지는 않고 있다. 현재 상기 임출력 버퍼형 스위칭 방식에 따른 스위치 소자들을 사용하여 320Gbps급 스위칭 시스템이 상용화되었으며, 스위치 패브릭 내부에 WDM 기술을 적용한 640Gbps급 스위치가 개발 중에 있어, 2003년경에는 수 Tbps급 상용 스위치가 출시될 수 있을 것으로 예측되고 있다.

V. 결 론

본 고에서는 개방형 시스템 하드웨어 기술을 설명하기 위하여 공통스위치 인터페이스 규격 및 개방형 네트워크 시스템의 주요 구성 요소인 트래픽 매니저와 스위치 패브릭, 그리고 트래픽 매니저를 프로그램 가능하게 구현하는 방식의 하나로 관심의 대상이 되고 있는 네트워크 프로세서의 기술 동향을 요약 기술하였다.

기술을 주도하던 중소 규모의 스위치 패브릭 기술 전문 기업들과 네트워크 프로세서 기술 전문 기업들이 대규모 네트워크 시스템 부품 제조사 들에 인수 합병 된 이후, 개방 시스템을 향한 표준 포럼 들의 활동이 다소 주춤거리고 있으나, 최근 CSIX 포럼과 CPIX 포럼이 통합하여 NP 포럼을 결성하고 새롭게 활동을 시작할 계획이다.

IP 기반 서비스의 확산으로 멀티서비스를 지원할 수 있는 패킷 기반망으로의 통신망 통합이 가속화 되고 있으며, IP 기반 서비스의 확산에 따른 서비스 저가화 압력에 따라 네트워크 시스템의 개방 시스템화 가 가속화될 것으로 전망된다.

※참고문헌

- [1] 유상경, 안운영, 김봉태, "네트워크 프로세서 기술동향", 전자통신동향분석, 2000.11.
- [2] G. Fiche, W. Lorcher, "Study of multiplexing for ATM traffic sources", in *ITC14*, Eseevier Science, pp.441-452, 1994.
- [3] B. Hubbs, "A Survey of Highly Integrated Ethernet DataComm Devices", IEEE Aerospace Conference, Vol.4 PP. 489-498, 1998
- [4] N. Cravolla, "The sky's the limit", EDN Magazine, pp.108-118, Nov. 24 1999
- [5] M. Levy, "Development tolls unleash network processors' Power", EDN Magazine, pp.115-120, Feb. 3, 2000
- [6] Agere, <http://www.agere.com>
- [7] Conexant, <http://www.conexant.com>
- [8] C-pot, <http://www.cportcorp.com>
- [9] Vitesse, <http://www.vitesse.com>
- [10] Network Processors Conference, <http://www.networkprocessors.com>
- [11] Network+Interop 2000, <http://www.key3media.com/interop>
- [12] "CSIX: Common Switch Interface Specification," Draft3.4, Apr. 10, 2000.
- [13] "Power X and XaQti Co-founders of the CSIX Consortium," http://www.px.uk.com/px_csix.htm.
- [14] "The Need for a New Common Switch/Fabric Interface to Support Next Generation Networks," <http://www.csix.org/csixpp.htm>.



김 봉 태

1983년 서울대학교 공과대학 학사(전자공학)

1991년 미국 NCSU 석사 (컴퓨터공학)

1995년 미국 NCSU 박사 (컴퓨터공학)

1983년~현재 한국전자통신연구원, 네트워크기술연
구소, 책임연구원, 부장

1984년~1986년 미국 ITT Telecom 방문연구원

관심분야 : 네트워크 시스템, 모뎀기술, 정보가전