

FC을 기반으로한 기가급 패킷스위치의 설계 및 구현

正會員 이 현 태*, 이 근 우**, 손 장 우*, 오 영 렬*, 이 재 용*, 이 상 배*

Design and Implementation of the Gigabit Packet Switch based on Fibre Channel

Hyeun Tae Lee*, Keun Woo Lee**, Jang Woo Son*, Young Yul Oh*, Jai Yong Lee*, Sang Bae Lee* *Regular Members*

I. 서 론

최근 Client-Server 구조에 의한 분산 처리 시스템이 보편화되고 컴퓨터 기술의 발달로 인해 수백 MIPS급의 고속 컴퓨터가 일반화되고 있으며, 복수 서버를 사용하는 병렬 컴퓨터 기술이 발달하고 있다. 이에 따라 상호 고속 연결망에 대한 요구가 높아지고 있으며 이러한 고속 상호 연결망은 수 Gbps이상의 높은 대역폭과 적은 전달 지연의 성능이 요구된다[3][6]. 공중망에서는 ATM 기술이 사용될 것이 확실하나 LAN에서는 ATM뿐만 아니라 다양한 기술이 사용될 수 있을 것이고, 특히 수 Gbps급의 고속 통신을 필요로 하는 응용에서는 기가급 이상의 대역폭과 적은 지연을 제공하는 통신망을 필요로 한다. 최근 이러한 요구에 따라 미국 ANSI의 X3T11 기술위원회에서는 기존의 다양한 망 프로토콜과 채널 프로토콜을 함께 지원하고 기가급의 전달 성능을 제공하는 FC(Fiber Channel) 프로토콜을 연구하며 표준화하였다[1][2][4][5].

FC 통신 방식은 ATM-LAN과는 달리 가변 길이의 비교적 큰 패킷 단위로 전송하여 오버헤드로 인한 망

부담률이 1.7%로 매우 작다. ATM은 기존 패킷 망의 문제인 Serialization delay를 해결하기 위해 짧은 길이의 셀을 사용하므로, 망 부담률이 9.5%로 크다. 또한 기존의 공유 매체형 LAN은 여러 사용자가 매체를 공유하므로 개개의 사용자에게 그 LAN이 제공할 수 있는 용량을 독점적으로 제공해 줄 수 없다. 따라서 기가급의 대역폭과 낮은 지연을 요구하는 고속 통신에는 적합하지 않다. 이에 반해, FC은 스위치 구조를 기반으로 하므로 기가급의 전용회선 서비스의 제공이 가능하다. 사용자에게 연결이 설정된 동안 전체 대역폭을 제공해주는 회선 교환 서비스가 FC이 제공하는 기본 서비스이다. FC은 이와 같은 회선 교환 방식의 연결형 서비스 이외에 패킷 교환 방식의 비연결형 서비스를 함께 제공할 수 있어 다양한 유형의 서비스를 제공할 수 있다. FC 통신 방식은 현재의 가능한 기술을 이용해 기가급의 데이터 전송을 제공할 수 있어, Gigabit LAN에서 유력한 기술로 인식되고 있다[1][4][5].

본 연구에서는 가변 길이이고 비교적 긴 길이의 패킷 단위로 교환을 수행하는 Giga bps급의 고속 패킷 스위치 구조에 적합한 구조로서 크로스바를 기반으로 한 스위치 구조를 고려하였다. 크로스바 스위치는 입력 포트와 출력 포트간의 내부 경로는 무 충돌 특성을 갖지만 출력 포트에 대한 충돌을 해결해야 한

*연세대학교 전자공학과
 **해동정보통신(주) 전무이사
 論文番號:96250-0817
 接受日字:1996年 8月 17日

다. 본 논문에서는 버퍼를 갖지 않는 크로스바 스위치 구조에서 출력 포트 중재 기능의 구성 방식에 대하여 기술한다. 출력 포트 충돌을 해결하기 위한 버퍼링 방식은 입력 버퍼링 방식을 고려하고 출력 포트의 중재 기능은 출력 포트 단위로 분산되어 수행하는 구조를 고려하였다. NXN 크기의 크로스바 스위치 구조에서 입력 포트와 출력 포트 중재기 간에는 논리적으로 N^2 개의 요구 신호와 N^2 개의 응답 신호가 필요하다[26]. 본 논문에서는 입력 포트와 출력 포트 중재기 간의 제어선의 복잡도를 줄이고 구현이 용이한 특성을 갖는 경로 제어 방식을 제안하였다. 제안된 경로 제어 방식은 입력 포트를 시분할 방식으로 폴링하는 구조를 갖는 경로 제어 버스를 사용한다. 일반적인 입력 큐 방식 스위치 구조의 성능을 개선하고 FC의 연결형의 회선 기반 서비스 제공이 용이하며 회선 기반 서비스 중에 비연결형 서비스의 제공이 용이한 스위치 구조로서 이중 크로스바 평면 구조를 갖는 스위치 구조를 제안하였다. 제안된 경로 제어 방식과 제안된 전체 스위치 구조에 대한 성능을 시뮬레이션을 통하여 분석하였다. 그리고 제안된 스위치 구조를 기반으로 스위치 시스템 구현을 위한 시스템 구성, 크로스바 채널의 구현 방식, 그리고 경로 제어 매커니즘에서의 그룹 라우팅의 구현 방식에 대하여 기술하였다.

본 논문의 구성은 다음과 같다. 제2장에서는 FC 프로토콜 표준에 대해 간단하게 서술하고, 제3장에서는 제안된 FC 스위치 구조 및 스위치 경로 제어 방식을 설명한다. 제4장에서는 제안된 경로 제어 방식 및 스위치 구조에 대하여 시뮬레이션을 통하여 그 성능을 살펴본다. 제5장에서는 FC 기반의 스위치 시스템인 GIGAWAYS 시스템의 구성과 크로스바 채널 구현 방식 및 그룹 라우팅 구현에 대하여 기술하고 제 6장에서 결론을 맺는다.

II. Fibre Channel Standards

2.1. 개요

FC 프로토콜은 기존의 컴퓨터 및 주변장치간의 연결 채널인 SCSI(Small Computer System Interface)나 HIPPI(High Performance Parallel Interface) 등이 갖고 있는 거리, 지연, 어드레싱 및 다중화의 제한을 극

복하기 위하여 미국 ANSI(American National Standards Institute)에서 새롭게 표준화한 프로토콜이다[1][4]. 전송속도를 1Gbps 까지 제공하고 전송 거리를 최대 10Km까지 확장 가능하며 기존의 다양한 채널(SCSI, HIPPI, IPI(Intelligent Peripheral Interface) 등) 및 LAN 프로토콜과 Internet 프로토콜을 다양하게 지원할 수 있고[13], STP(Shielded Twist Pair) 및 UTP(Unshielded Twist Pair), 광섬유 등의 다양한 매체를 접속할 수 있다[9]. FC은 회선 기반 서비스를 제공하여 매우 낮은 지연을 얻을 수 있어 기가급의 전달 속도를 제공하기에 매우 유리하다. 또한 연결형 및 비연결형 서비스, 혹은 이들의 복합적인 서비스를 제공하여 데이터 응용 서비스뿐만 아니라 멀티미디어와 같은 다양한 형태의 실시간 트래픽 처리를 위한 서비스도 제공한다 [1][10].

2.2. FC 서비스 클래스

다양한 통신 요구 사항을 만족하기 위하여 FC에서는 서로 다른 서비스 등급을 정의하고 있다. 클래스 1 서비스는 두 통신 노드 포트(N_Ports)간에 전용 데이터 경로를 제공하는 서비스 등급으로서 한번 설정된 전용 연결은 N_Port간의 프레임(FC에서의 데이터 전송 단위) 교환 기능을 수행하는 FC 패브릭(Fabric)에 의해 유지 및 관리되며, 두 N_Port 사이에 설정된 연결 내에서는 가용한 최대 대역폭을 보장해준다. 클래스 2 서비스는 클래스 1과 같은 전용 연결 설정 없이 프레임 단위로 다중화된 프레임 전달을 제공하는 비연결형 서비스(Connectionless service)이다. 패브릭은 발신 노드 포트로부터 보내온 송신 프레임에 대하여 착신 노드 포트로서 순서 보장 전달(In order delivery)을 반드시 보장하지는 않는다. 착신 N_Port에서는 전달에 대한 결과로 전달 응답(Ack)이나 비전달에 대한 통지(BSY, RJT)를 제공해야 한다. 클래스 3 서비스는 클래스 2와 마찬가지로 전용 연결 설정 없이 프레임 단위로 다중화된 프레임 전달을 제공하는 비연결형 서비스로 서비스 특성은 클래스 2 서비스와 동일하나, 클래스 2와 차이점은 착신지 노드 포트가 프레임 수신 후 전달 통지(ACK, BSY, RJT)를 보내지 않는다는 점이다. 클래스 4 서비스는 새로운 실시간 서비스 요구 사항을 지원하기 위하여 정의한 서비스 클래스이며, 클래스 1 서비스와 같이 연결 중심(Connection

oriented) 양방향 통신을 제공한다. 혼합(Intermix) 서비스는 두 노드 포트 사이에서 클래스 1 연결이 설정되어 있는 동안 비연결형의 클래스 2, 3 프레임을 삽입하여 전달하는 서비스이다. 혼합 서비스는 클래스 1 서비스 제공을 하기 위해 잠재하는 대역폭을 항상 사용하지 않는 점을 이용하여 클래스 2, 3 서비스를 제공하는 방법이다[1][9].

2.3. FC 패브릭

FC은 네트워크의 토폴로지를 제한하지 않는 프로토콜로서, 가능한 토폴로지로는 점대점(point-to-point) 방식, 중재 루프 (Arbitration Loop)[14] 또는 스위치 형태[11][12]가 있다[1][4][9]. FC을 응용하는 시스템에서의 토폴로지의 선택은 시스템 성능이나 확장성, 비

용 등을 고려하여 결정되는데, 본 연구에서는 스위치 토폴로지를 기반으로 하는 FC 패브릭을 구현하였다. 이러한 스위치 패브릭 토폴로지(이하 패브릭)는 노드 포트 사이에 FC 프레임을 전달하기 위한 전송 토폴로지이며, FC 표준이 지원해야 하는 모든 서비스를 제공할 수 있다.

그림 1은 FC 패브릭의 구조를 나타낸다. 패브릭은 한개 이상의 패브릭 단위 요소(Fabric Element)인 FC 스위치로 구성되며, FC 스위치는 패브릭의 기능을 완전하게 수행할 수 있는 최소한의 단위 시스템이다. 여기서 N_Port는 노드를 패브릭에 접속시키기 위한 기능을 가지며, 패브릭 포트(F_Port)는 N_Port를 패브릭에 접속할 수 있도록 하며, E_Port는 스위치간 링크(ISL: Inter-Switch Link)를 통하여 스위치간 접속 기능을 수행한다. 스위치는 시작 포트에서 목적 포트 로 프레임을 전달하기 위한 물리적인 경로를 제공하며, 스위치 제어기는 주소 할당, Login/Logout과 같은 링크 서비스 처리, 스위치 초기화, 스위치 제어 기능 등을 수행한다[7][11][12].

III. 초고속 패킷 스위치 구조

3.1. 크로스바 기반 스위치 구조

최근 분산 컴퓨터 응용과 같이 Giga bps이상 대역폭과 적은 latency를 요구하는 응용에서는 기존의 ATM과 같이 고정이고 작은 길이의 패킷(셀) 단위로 교환하는 방식보다는 큰 크기의 가변 길이 패킷 단위로 교환하거나 필요한 경우 회선 교환과 같이 회선을 설정하고 시작점과 목적점간에 전체 대역폭을 이용하여 전송하는 latency가 적은 전달 서비스를 요구한다. 대표적인 예로, FC에서는 가변 길이의 패킷 단위로 교환을 수행하고 회선 교환과 같은 서비스(클래스 1 서비스)를 제공한다. 본 논문에서는 이러한 특성의 서비스에 적합한 구조로서 크로스바를 기반으로 한 패킷 스위치 구조에 대해 기술한다.

크로스바 구조는 단일 단계를 갖고 입력과 출력간에 하나의 경로만 존재하고 무 충돌 특성을 갖는 스위치 구조로 설명할 수 있다. 크로스바 구조는 근본적으로 회선 교환 구조에서부터 근본을 찾을 수 있으며 최근에는 멀티 프로세서 구성에서 프로세서들과 메모리 모듈간의 상호 연결 망 구조로서 많이 연구되

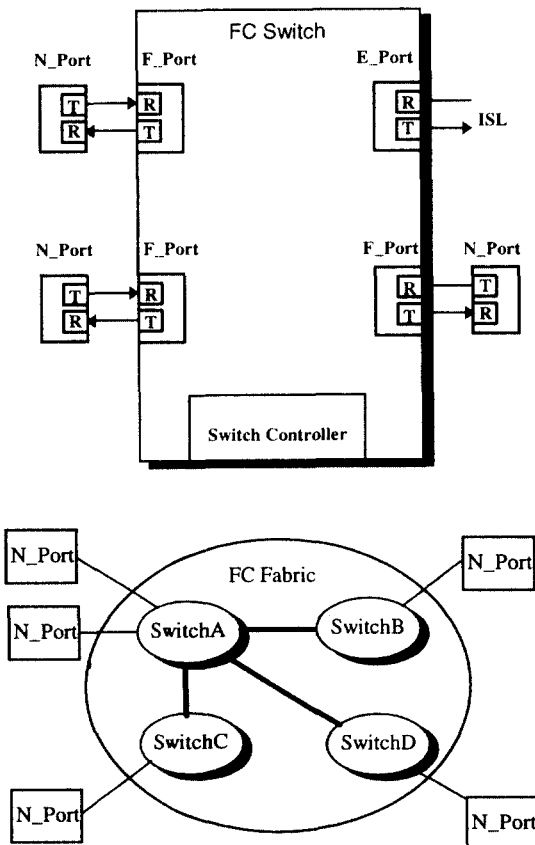


그림 1. FC 패브릭 구조
Fig 1. FC Fabric architecture

어 왔다[25][26]. 그리고 ATM 스위치에서도 비교적 소규모의 스위치에서나 기본 단위 요소 스위치의 구현으로 많이 사용되고 있다. 이는 크로스바가 갖는 무 충돌 특성과 적은 규모에서 구조가 간단한 점 그리고 모듈러한 크로스바 엘리먼트의 특성 등의 장점 때문이다. 그러나 크로스바는 규모가 커질 수록 기하급수적($O(N^2)$)으로 복잡도가 증가하는 단점을 갖고 있어 큰 규모로 확장하는 데 어려움이 있으며, 크로스바 구조를 갖는 self-routing 구성에서는 입력과 출력 짝간에 서로 다른 지연으로 인하여 공정성을 위한 인위적인 지연을 필요로 하는 등 단점도 갖고 있다[27].

크로스바 스위치의 내부 경로는 무 충돌 특성을 갖지만 출력 포트 충돌은 여전히 존재한다. 따라서 이를 해결하기 위하여 버퍼를 설치한다. 일반적으로 크로스바를 기반으로 한 스위치 구조를 출력 포트 충돌 해결을 위한 버퍼의 위치에 따라 입력 버퍼 구조, 크로스포인트 내부에 버퍼를 갖는 구조, 입력과 크로스포인트에 버퍼를 갖는 형태, 그리고 출력 포트에 버퍼를 갖는 구조로 분류하고 있다[15][27]. 출력 포트에 버퍼를 갖는 방식은 제어가 간단하고 높은 성능을 얻을 수 있으나 크로스바 경로의 전달 능력이 입력 포트의 속도보다 최대 N 배 고속으로 처리해야 하며 출력 버퍼의 속도도 최대 N 배 빠른 속도로 접근이 가능해야 하므로 Giga bps급의 스위치에서는 구현이 어렵다. 크로스바에 버퍼를 갖는 구성은 크로스포인트에 FIFO와 주소 필터를 갖고 있는 경우이다. 이러한 구조는 크로스바 내부에 많은 수의 버퍼를 요구하므로 FC과 같이 긴 길이의 패킷을 사용하는 구조에서는 버퍼 메모리의 구현 부담이 크다. 반면, 입력 버퍼 방식은 성능이 떨어지고 출력 포트 중재 기능이 복잡하지만 구현이 용이하고 고속 교환 구조에 적합하다. 본 논문에서는 입력 버퍼를 사용하고 내부적으로 버퍼를 갖지 않는 크로스바 구조를 사용한다. 크로스바 구조를 갖는 스위치 구현 방식을 제어 위한 경로와 데이터 경로가 동일한 구조와 제어 경로와 데이터 경로가 분리된 구조로 나누어 볼 수 있다. 전자는 패킷의 데이터와 라우팅 정보가 동일한 경로를 통하여 전달되는 구성으로 셀프 라우팅 특성을 갖는 크로스바 구성이나 크로스포인트 내부에 버퍼를 가진 구성과 같이 라우팅을 위한 정보 경로와 데이터 경로가 별도로 구성되지 않는 구조이다. 후자는 경로 제어를 위

한 별도의 채널을 입력 포트와 출력 포트 중재 기능 사이에 갖고 있어 데이터 경로와 경로 제어 채널은 전혀 다른 방식에 의해 구성될 수 있는 구조로서, 이 경우 데이터 전달은 두단계 즉, 경로를 설정하는 단계와 데이터 전달 단계로 나누어진다. 후자와 같이 데이터 경로와 제어 경로가 분리된 구조는 데이터 경로 방식이나 전달 매체에 제약을 받지 않고 고속화할 수 있다는 장점을 갖고 있다. 즉, optical link와 같은 고속 링크를 통하여 직렬로 전송된 데이터를 경로가 설정된 상태에서는 직렬 상태를 병렬 데이터로 변환하지 않고 그대로 출력 포트로 전달 할 수 있으며 optical 경로를 이용하면 신호를 optical 신호 상태로 교환할 수 있다. 본 논문에서는 큰 대역폭과 적은 전달 지연을 요구하는 응용에 적합한 비교적 긴 길이의 가변 길이 패킷 교환 구조에서 데이터와 경로 제어 경로가 분리된 크로스바를 기반으로 한 스위치 구조에 대하여 크로스바 출력 포트 중재를 위한 경로 제어 방식, 성능 개선 방안 및 구현 사항에 대하여 기술한다.

3.2. 크로스바 출력 포트 중재 방식

입력 포트에서 출력 포트로 경로를 설정하기 위해서는 각 입력 포트 제어기에서 해당 출력 포트 중재기로 경로 설정을 요구해야 하고 출력 포트 중재기에서는 설정 응답 조건을 만족하면 해당 입력 포트 제어기로 응답 신호를 보내야 한다. 그림 2와 같이 $N \times N$ 크기의 크로스바와 출력 포트 중재기로 구성된 스위치를 고려할 때 입력 포트와 중재기 간에 N^2 개의 요구선과 N^2 개의 응답선, N^2 개의 크로스바 엘리먼트 제어선, 출력 포트로부터 중재기로 N 개의 출력 상태 표시선을 포함해야 한다[26]. 이와 같이 입력 포트와 출력 포트 중재기 간에 공간적으로 완전히 독립적인 제어 신호선을 제공하는 완전 상호 연결 중재 구성 방식 (FIA: Fully Interconnected Arbitration)은 다음과 같은 점에서 제약점을 갖고 있다. 1) FIA 구성 방식은 입력 입력 포트와 중재기 간에 논리적으로 N^2 개의 요구선과 N^2 개의 응답선을 확보해야하며 목적점 포트에 대한 논리 주소로 부터 물리 주소를 찾기 위한 주소 룩업(lookup) 기능이 입력 포트 마다 필요하다. 이러한 룩업 기능은 구현 부담이 크므로 다수의 입력 포트에서 공유하는 집중형 룩업 기능으로 구현

할 수 있으나 이 방식 또한 또 다른 신호선이 요구되며 경로 제어에서의 성능 병목을 야기할 수 있다. 2) 입력 포트에서의 록업 기능의 부담을 줄이기 위하여 논리적 주소를 그대로 사용하고 각 출력 포트 중재기에서 자신의 논리 주소를 감지하여 경로 제어를 할 수 있는데 이 경우 각 출력 포트 중재기는 각 입력 포트로부터의 논리 주소 정보와 제어를 위한 신호선이 필요하게 된다. FC의 경우 24비트의 논리 주소를 사용하므로 32X32 크기의 스위치인 경우 적어도 768개 이상의 신호선을 필요로 하게 된다. 따라서 하나의 출력 포트 중재기를 칩으로 구성하는 경우 전원편과 부가적인 제어선을 고려한다면 하나의 칩으로 중재기를 구현하는 데 큰 제약 점이 된다.

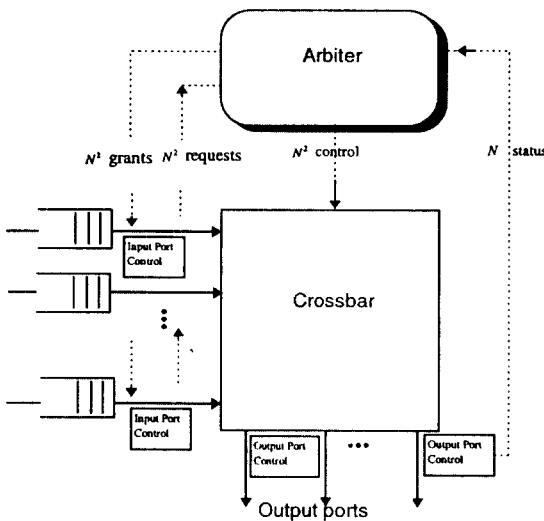


그림 2. 크로스바를 사용하는 스위치 구조에서 일반적인 스위치 중재기 구성

Fig 2. Switch arbiter configuration in crossbar based switch architecture

실제 상용 크로스바 소자를 사용하여 구현할 경우에 있을 수 있는 제약 사항으로 일반적인 크로스바의 크로스포인트 접점을 연결 혹은 제거하고자 할 때 크로스바 제어를 위한 구성 메모리를 액세스하는 동작으로 이루어진다. 따라서 하나의 크로스바 접점 제어를 위해서는 한번의 구성 메모리 액세스 동작을 필요로 한다. 따라서 아무리 출력 포트 중재를 분산하여

수행하여도 실제 크로스바 구성 제어에 병목이 생기게 되어 크로스바 제어 요구에 대한 별도의 중재 기능이 필요하게 된다.

본 논문에서는 입력 포트와 출력 포트 중재기 간의 신호선과 제어 로직이 간단하면서 중재 기능의 고속 병렬 처리가 용이한 입력 포트와 출력 포트 중재기 간의 경로 제어 구조를 제안한다.

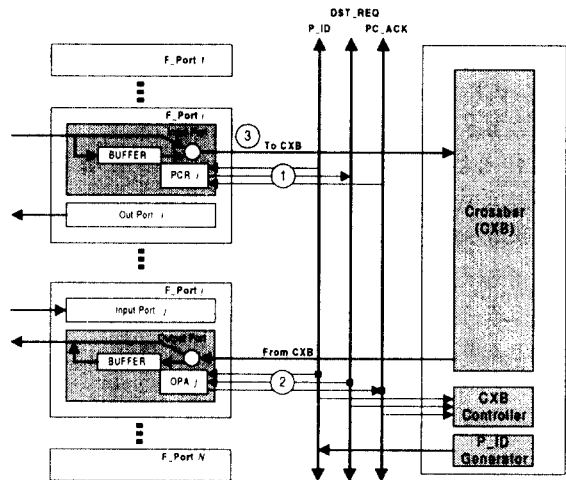


그림 3. 제안된 IPA 방식의 경로 제어 구조

Fig 3. Path control architecture using IPA method

제안된 출력 포트 중재를 위한 경로 제어 구조는 그림 3과 같이 입력 포트의 경로 제어 요구 기능(PCR: Path Control Requester)과 출력 포트 제어 기능(OPA: Output Port Arbiter)간 시분할 방식으로 입력 포트의 경로 설정 요구를 처리하는 경로 제어 버스(PCbus: Path Control bus)를 사용하고 각 출력 포트마다 독립적으로 분산된 OPA를 갖는다. 경로 제어 버스는 경로 제어 요구를 위한 주소 정보 및 제어 정보 비트를 포함하는 DST_REQ 버스 신호선, 경로 제어 요구에 대한 출력 포트 중재기의 응답을 나타내는 PC_ACK 버스 신호선, 그리고 현재 타임 슬롯에 해당하는 입력 포트를 나타내는 P_ID(Polling Identifier) 버스 신호로 구성된다. 제안된 입력 폴링 방식을 사용하는 경로 제어 방식(IPA: Input Polling Arbitration)을 통한 경로 설정 과정은 다음과 같다.

- ① P_ID 발생기는 단위 타임 슬롯 간격으로 각 입력 단을 라운드 로빈 방식으로 폴링한다. 각 입력 단의 PCR는 자신의 P_ID가 되고 전송할 프레임이 있으면 DST_REQ에 목적지 노드의 주소(D_ID)를 실는다.
- ② 목적지 출력 포트의 OPA는 DST_REQ 버스 신호를 매 타임 슬롯 마다 검사하여 자신의 출력 포트에 해당 하는 주소가 감지되면, 현재 출력 포트로 전송 중인 프레임이 없고 출력 버퍼가 꽉 찬 상태가 아니면 경로 설정 완료를 나타내는 응답을 PC_ACK 신호에 실는다.
- ③ 크로스바 제어 기능(CXB Controller)에서는 매 타임 슬롯 마다 PC_ACK를 검사하여 경로 설정 완료 응답이면 해당 P_ID와 DST_ID 주소 정보를 가지고 해당 크로스바의 접점을 연결한다.
- ④ 경로 설정을 요구한 PCR에서는 해당 PC_ACK 신호를 검사하여 경로 설정 완료 응답이면 설정된 크로스바 경로를 통하여 프레임을 전송한다.

이 과정이 그림 3에 ①, ②, ③으로 표현되어 있다. 설정된 경로를 해제하는 과정은 DST_REQ의 제어 신호를 통하여 해제 요구를 나타내고 해제 요구에 대한 동작이 완료되면 크로스바 제어기에서 해당 크로스바의 접점을 절단한다.

제안된 IPA 방식에서 다수의 입력 포트로부터의 경로 제어 요청에 대하여 하나의 입력 포트가 선택되어 경로 쟁어 완료 응답을 보내는 출력 포트 중재 동작은 PSR, Pcbus, OPA의 상호 동작에 의해 이루어지고 입력 포트 선택은 입력 프레임 도착 시각, P_ID 발생 순서, 출력 포트 상태에 의해 결정된다. 제안된 IPA 방식은 앞에서 설명한 바와 같이 경로 설정을 위하여 PCR과 OPA간에 입력 폴링 방식의 공유 버스를 사용하므로 경로 제어를 위한 신호선 수를 줄일 수 있고, 각 포트에 대한 단위 폴링 시간인 타임 슬롯 단위로 크로스바 접점 제어가 일어나므로 크로스바 접점 제어를 위한 별도의 중재를 필요로 하지 않는다. 본 연구에서는 DST_REQ(28 bit), PC_ACK(11 bit)와 P_ID(8 bit)를 포함하여 47 비트의 경로 제어 버스 선을 사용하여 구현하였다.

3.3. 이중 스위치 평면 구조

제안된 입력 폴링을 사용하는 공유 버스형 경로 제어 구조는 실제 구현 측면에서 신호선의 복잡도, 구현 환경 등을 고려하여 제안된 방식으로 입력단에 FIFO 입력 버퍼를 갖는 입력 큐 방식 스위치가 갖는 HOL(Head Of Line) 블러킹과 제안된 경로 제어 방식에서 입력 단 폴링으로 인한 대역폭의 낭비로 성능 수율이 떨어지게 된다. 이에 대한 성능 향상 방안으로 다음과 같은 접근을 고려할 수 있다.

먼저 입력 큐 방식의 스위치가 갖는 HOL 블러킹 효과를 줄이기 위하여 다음과 같은 방법을 사용할 수 있다. 첫 번째로는 입력 버퍼는 그대로 FIFO 형태의 버퍼를 사용하고, 스위치의 채널 처리 능력을 크게 하거나 스위치 평면을 다중 구조로 하여 스피드 업 효과를 얻는 방법이다[16][18]. 두 번째 방법은 스위칭은 단일 평면 구조를 사용하고 입력 버퍼를 다중 구조로 관리하는 방법이다[26][27]. 본 연구에서는 스위치의 성능을 향상시키고 FC의 Intermix 서비스를 효과적으로 처리하기 위하여 그림 4와 같이 두개의 입력 버퍼를 사용하고 이중 스위치 평면과 출력 포트에 두개의 출력 버퍼를 사용하였다. 제안된 구조는 선두 프레임에 출력 충돌이 있을 경우나 선두 프레임을 전송 중에 있을 때 다른 입력 버퍼에 있는 프레임의 경로 설정을 요구할 수 있으므로 스위치의 성능을 향상시킬 수 있다.

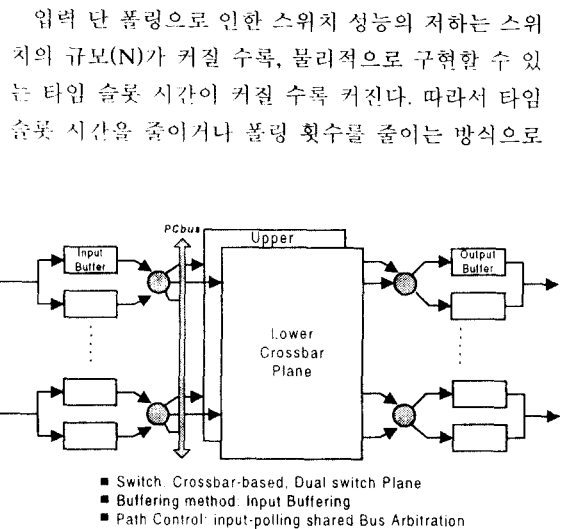


그림 4. 제안된 FC 스위치 구조

Fig 4. Proposed FC switch architecture

성능을 개선할 수 있다. 앞서 3.2절에서 제시한 IPA 경로 제어 방식에서는 P_ID 발생기에서 모든 입력 포트의 P_ID를 순차적으로 발생한다. 따라서 보낼 프레임이 없거나 현재 프레임을 전송중인 포트에 대해서도 P_ID를 발생한다. 이러한 점을 개선하여 입력 포트에서 경로 제어가 없는 경우에 폴링 순서에서 제외하는 방식(IPA+)을 사용할 수 있다. 이는 각 입력 포트에서 P_ID 발생기로 경로 제어 요구 상태를 나타내는 신호를 사용하므로 구현할 수 있다.

IV. 성능 분석

본 4장에서는 시뮬레이션을 통하여 제안된 IPA 방식의 경로 제어 구조와 이중 평면 구조를 갖는 스위치 구조의 성능을 분석한다. 제안된 경로 제어 구조와 스위치 구조는 그림3 및 그림4와 같으며 스위치 규모는 실제 구현한 스위치 규모인 N=32로 설정하였다. 시뮬레이션 내용은 크게 두 가지로 수행한다. 첫번째로, IPA 방식의 성능을 평균 입력 큐 길이(mean input queue length), 평균 큐잉 지연 시간(mean queueing delay)과 수율(throughput)에 대해 분석하고, 입력 포트와 출력 포트 중재 기간에 완전 상호 연결 구조를 가지므로 경로 제어 구조에 따른 별도의 처리 부담을 갖지 않고 입력 프레임이 동기 되어 출력 포트 단위로 중재되는 FIA 방식과 그 성능을 비교 분석한다. 두 번째는 제안 스위치 구조에 대한 실험으로서 수율을 성능 평가 인자로 하여 분석하며 단일 입력 큐/단일 평면 구조, 단일 입력 큐/이중 평면 구조, 그리고 3.3절에서 성능 향상 방안으로 제안된 이중 입력 큐/이중 평면 구조 등에 대해 성능을 비교한다.

각 입력 단으로 유입되는 traffic intensity는 동일하고 그 트래픽의 목적지가 각 출력 단으로 균등하게 분포되어 있는 homogeneous system을 가정한다. 시간 축은 입력 단을 폴링할 때 입력 단 간의 스위치 오버 시간과 같은 크기의 타임 슬롯 시간(SOT: Switch Over Time)으로 나누어져 있다. 여기서 타임 슬롯 시간은 입력 단 폴링 시 이웃하는 입력 단 간에 폴링 시각의 차로 정의한다. 한 타임 슬롯의 실제 길이는 경로 제어 기능의 처리 시간, 크로스바 점점의 제어 시간 등에 의해 결정된다. 본 시뮬레이션에서는 한 타임 슬롯 시간을 실제 구현 값에 가까운 100 nsec로 가

정하였다. 프레임은 각 입력 포트에 비동기적으로 도착하고 도착한 프레임들에 대해서 별도의 동기화 과정을 수행하지 않는다. 프레임의 길이는 FC의 프레임 길이(PL: Packet Length)가 2 KBytes이고 링크 속도가 1 Gbps이므로 160 타임 슬롯 시간이 되고, 따라서 프레임의 서비스 시간 D(=프레임 전송 시간/SOT)는 160으로 일정하다. 이때, 각 입력 큐의 선두 프레임들의 중재 방식은 앞장에서 언급한 JPA 방식으로 수행된다.

4.1. IPA 중재 방식에 대한 분석

IPA 중재 방식의 성능에 대한 시뮬레이션 결과는 그림 5와 같다. IPA 중재 방식과 FIA 방식과 그 성능을 비교 분석하였다. 입력 부하에 대한 입력 큐에 존재하는 평균 프레임 수와 큐잉 지연 시간이 그림 5(a), (b)에 도시되어 있다. 입력 부하가 약 50%를 넘으면서 입력 큐에 backlog가 급격히 상승하여 이후 큐가 완전히 차게 되는 것을 볼 수 있다. 또한 그림 5(c)에서 수율(throughput)이 53.7 %에서 포화가 됨을 볼 수 있다. 이는 FIA 방식의 수율이 59.3%인데 비해 제안된 경로 제어 방식을 적용한 경우 수율이 약간 떨어짐을 알 수 있다. 이는 제안된 경로 제어 방식에서는, 각 입력 큐의 선두 프레임이 출력 링크가 사용중이 아닌 경우에도 자신의 P_ID가 되어야만 서비스가 되므로 발생하는 현상이다. 즉, IPA 방식이 구현 지향적으로 설계되었으므로 구현 비용은 FIA 방식보다 간단하지만, 성능 면에서는 입력 폴링 버스 방식으로 인한 수율의 감소를 확인할 수 있다.

그림 6은 여러 가지 프레임 길이 값에 대하여 스위치의 크기에 따른 IPA 방식의 saturated throughput을 비교하고 FIA 경우와 비교하였다. 주어진 링크 속도(1Gbps)와 SOT(100nsec)에 대하여 프레임 길이(PL)가 크고 스위치 크기(N)가 작을 수록 IPA 방식의 성능이 FIA 방식에 가까워짐을 알 수 있다. 이것은 프레임 길이가 길고 스위치 크기가 작을 수록 입력 포트의 폴링 주기가 짧아지므로 입력 폴링에 의한 영향이 작아 지기 때문이다. 즉, 스위치 크기가 고정되면 프레임 서비스 시간 D가 스위치 성능 차이의 주요 원인이 된다. 이때 프레임 서비스 시간을 결정하는 SOT는 실제 구현 기술에 의해 제약을 받는다.

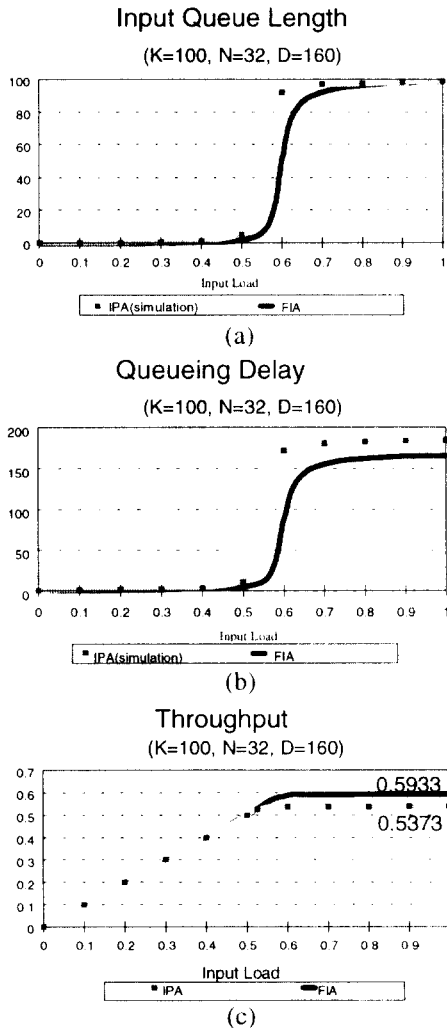


그림 5. IPA 방식에 대한 성능 분석(1B+1P)
Fig 5. Performance of IPA method(1B+1P)

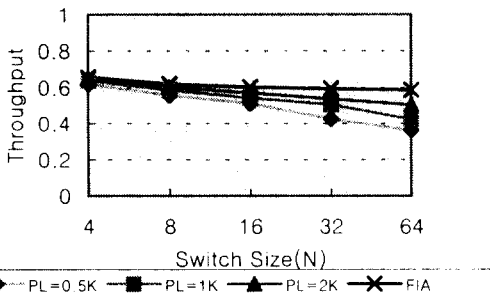


그림 6. 스위치 크기(N)와 프레임 길이(PL)에 따른 IPA 방식의 saturated throughput

Fig 6. Throughput of the IPA method for various PL and N

4.2. 제안 스위치 구조에 대한 분석

제안된 이중 입력 버퍼와 이중 스위칭 평면을 갖는 스위치 구조에 대하여 시뮬레이션을 통하여 성능을 분석하였다.

시뮬레이션은 다음의 4 가지 모델에 대해 수행하였다.

- 실험 모델1(1B+1P): 각 입력 단에 한 개의 입력 큐를 갖고 단일 스위칭 평면을 통해 프레임을 전송하는 구조로 앞 절의 실험 모델과 같다.
- 실험 모델2(1B+2P): 각 입력 단에 한 개의 입력 큐를 갖고 이중 스위칭 평면을 통해 프레임 전송하는 구조
- 실험 모델3(1B+4P): 각 입력 단에 한 개의 입력 큐를 갖고 4개의 스위칭 평면을 통해 프레임 전송하는 구조
- 실험 모델4(2B+2P): 각 입력 단에 두 개의 입력 큐를 갖고 이중 스위칭 평면을 통해 프레임 전송하는 구조로 입력 버퍼링은 랜덤 방식으로 저장하는 것으로 가정

표 1에 위의 4가지 모델에 대하여 여러 가지 프레임 길이에 대하여 IPA 방식과 FIA 방식의 saturated throughput을 구하였다. 여기서, Speed up factor (L)는 스위칭 평면의 개수를 나타낸다. 따라서 실험 모델 1, 2, 3은 각각 L=1, 2, 4인 경우이다. 실험 모델 4는 실험 편 이상, FC 경우인 PL=2Kbyte 경우에 대해서만 값을 구하였다. L이 1이고 PL=2K인 1B+1P의 경우, HOL 블록킹 현상과 입력 단 폴링 시 링크 용량의 낭비로 인해 수율이 53.7%에서 포화됨을 볼 수 있다. L=2이고 PL=2K인 1B+2P 모델의 경우 이중 평면 구조로 인해 선두 프레임이 서비스될 확률이 커지므로 수율이 향상되어 약 76%에 이른다. 2B+2P의 경우, 이중 입력 큐 구조로 인해 선두 프레임에 출력 충돌이 생길 경우 다른 큐에 있는 프레임이 경로 설정을 요구할 수 있을 뿐만 아니라, 한 입력 큐의 선두 프레임이 전송중인 동안 두 번째 큐의 선두 프레임이 경로 설정을 요구할 수 있어 효율적인 프레임 처리를 할 수 있으므로 수율이 82.5%를 나타낸다. L=4인 경우는 83.1%로 2B+2P 경우와 1% 이내의 차이로 비슷한 성능을 나타낸다. SOT를 줄이거나 폴링 횟수를 줄이는 방식을 사용하면 입력 폴링에 의한

표 1. 제안된 스위치 구조의 수율

Table 1. Saturated throughput for the three arbitration schemes

Arbitration Scheme	PL = 0.5KByte			PL = 1KByte			PL = 2KByte			
	L = 1	L = 2	L = 4	L = 1	L = 2	L = 4	L = 1	L = 2	L = 4	2B + 2P
IPA	0.4245	0.5844	0.6240	0.5049	0.7494	0.8308	0.5373	0.7637	0.8313	0.8253
IPA +	0.4931	0.8036	0.9696	0.5433	0.8543	0.9830	0.5684	0.8730	0.9898	0.9778
FIA	0.5933	0.8918	0.9965	0.5933	0.8918	0.9965	0.5933	0.8918	0.9965	0.9881

효과를 줄일 수 있으므로 보다 성능을 개선할 수 있다. 본 실험에서는 P_ID 발생기가 모든 입력 포트를 round-robin 방식으로 폴링하는 것을 개선하여 입력 포트에서 경로 제어 요구가 없는 경우에는 폴링 순서에서 제외하는 방식(IPA+)에 대하여 성능 분석하였다. IPA+ 경우 IPA 방식 보다 성능이 개선됨을 볼 수 있다. 32X32 스위치 규모에서 2B + 2P 모델인 경우 IPA+ 방식을 사용하면 거의 출력 버퍼 구조에 가까운 성능을 나타냄을 볼 수 있다. 즉, 3232의 스위치 규모에서, IPA+ 방식으로 경로를 설정하고 각 입력 포트에 두 개의 버퍼를 두고 이중 스위치 평면 구조를 가지면, 스위치 전달 능력을 N배로 하지 않아도 출력 버퍼 구조에 가까운 성능을 얻을 수 있음을 알 수 있다[16][20].

V. FC 기반 패킷 스위치 구현

5.1. 스위치 시스템 구성

기존 IEEE MAC 프로토콜이 비연결형 전달 모드로 동작하고 ATM이 연결형 전달 모드를 제공하는 것과 달리, FC에서는 Class 1과 같이 발신 포트와 착신 포트간에 연결이 설정되면 해제될 때 까지 경로 자원이 할당되는 형태와 Class 2, 3, 4와 같이 매 프레임마다 다중화하여 교환을 수행하는 서비스를 제공한다. 따라서 FC 서비스를 이용하는 상위 프로토콜은 서비스 요구에 따라 적절한 서비스 클래스를 선택할 수 있으며, FC 스위치에서는 각 서비스 클래스를 전달하기 위한 적절한 교환 기능을 가져야 한다[8][11][12][21][22][23][24].

그림 7은 본 연구에서 설계한 FC 스위치 시스템의 구성을 나타낸다. FC 스위치 시스템은 다수의 통신 단말의 FC 통신 기능을 수행하는 노드 포트(N_Port)

기능과 직렬 전송 링크를 통하여 접속되어 FC프레임을 수신하고 송신하는 기능을 수행하는 입출력 포트 제어 기능(IOC: Input/Output port Controller), 각 IOC간의 데이터 경로를 제공하는 크로스바 스위치(CXB: Cross-bar), 경로 제어 응답에 따라 CXB의 해당 접점(cross point)을 연결 혹은 절단하는 기능을 수행하는 크로스바 제어 기능(CXBC: Crossbar Controller), IOC간의 경로 제어를 위한 신호를 교환하는 경로 제어 버스(PCbus: Path Control bus), 스위치 시스템을 초기화하고 Login/Logout을 포함하는 링크 서비스 처리와 시스템 제어 및 관리 기능을 수행하는 스위치 제어 기능(SWC: Switch Controller), 그리고 FC에서 정의하고 있는 Directory, Alias 및 Network Management 서비스 기능을 수행하는 FC 서버 기능으로 구

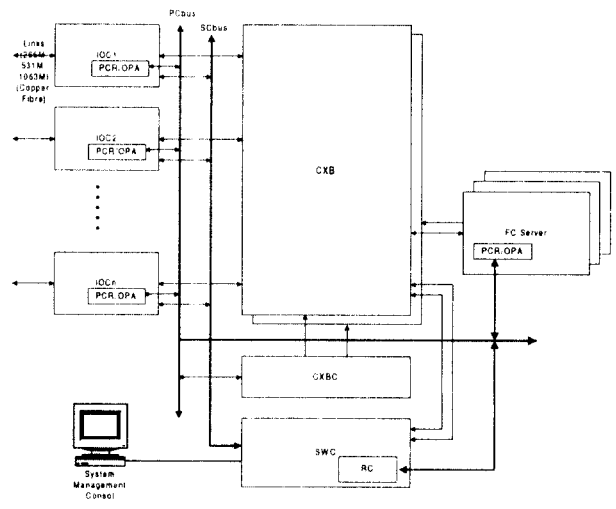


그림 7. FC 스위치 시스템 구성도
Fig 7. FC switching system configuration

성된다. FC 서버 기능은 크로스바의 채널에 접속되어 입출력 포트에서 수신된 서버 기능 요구 프레임의 목적점 주소에 따라 적절한 서버로 프레임이 전달되어 처리된다. 스위치 제어 기능도 크로스바 채널로 접속되어 관련 링크 서비스 프레임을 송수신하여 처리한다.

5.2. 크로스바 채널 구현 방식

크로스바를 통하여 데이터 경로를 구성하는 방법으로 입력 버퍼에 저장된 데이터를 단위 병렬 데이터 단위로 전달하는 병렬 방식과 각 채널 단위로 직렬 데이터 스트림 상태로 크로스바를 통해 전달하는 직렬 방식이 가능하다. 병렬 방식은 비교적 저속의 병렬 신호선을 통하여 전달하므로 일반적인 저속 신호선을 통하여 고속의 데이터 채널을 구성할 수 있다. 그러나 크로스바의 크기가 클 수록 신호선과 크로스바 핀 수가 커지므로 물리적인 크기가 크고 많은 핀을 구동하기 위한 전력 소모가 커지게 된다. 또한 모듈간의 거리가 클 경우에는 속도의 제약을 받게 되고, 신호 타이밍을 위해서는 별도의 타이밍 신호선을 각 채널마다 사용하거나 중앙의 클럭 발생 회로로부터 동기된 클럭 신호를 모듈마다 분배하여 사용해야 한다. 직렬 방식은 입력 버퍼에 저장된 데이터를 직렬화하여 크로스바를 통하여 출력 포트에 전달한다. 직렬로 데이터를 전송하므로 크로스바의 물리적 구성이 간단하여 비교적 큰 용량의 크로스바를 구현할 수 있다. 그러나 직렬로 데이터를 전달하기 위해서는 Giga bps급의 고속 신호를 전달할 수 있는 크로스바가 필요하고 고속 신호 처리를 위한 임피던스 제어를 위한 PCB, 컨넥터, 및 종단 회로 등에 대한 고려가 필요하다. 따라서 전기적으로 직렬화 하는 데는 이러한 요소의 기술적 제약으로 전달할 수 있는 속도의 제약을 받게 된다. 특히 본 논문에서 제안한 구조와 같이 출력 버퍼를 갖는 구성에서는 입력 포트와 출력 포트에서의 직렬화기(serializer)와 병렬화기(deserializer) 이외에 별도로 직렬 크로스바로 전달하기 위한 직렬화기를 필요로 하고 클럭 복원을 위한 코딩 기능도 필요하게 되어 부담이 커지게 된다. 직렬 크로스바를 사용하는 구조는 입력 버퍼에 저장된 데이터를 출력으로 전달할 때 8B/10B 인코딩하고 직렬화기를 통하여 직렬 신호 스트림으로 변환하여 크로스바를 통하

여 출력 포트에 전달되어 전기적 신호를 광 신호로 변환하여 광 링크를 통하여 목적점 노드 포트에 전달한다. 직렬 방식에서의 특성으로 신호가 직렬 신호 상태로 크로스바에 의해 교환되며 각 채널간에는 서로 동기화 되어 있지 않으므로 교환된 신호의 송신기와 수신기간에 신호의 동기가 이루어지기 위한 시간이 필요하다. 본 연구에서는 이런 점을 고려하여 병렬 방식으로 컨넥터를 포함하는 보드 간 신호선 처리, 크로스바의 속도, 메모리(FIFO)와 액세스 속도 그리고 물리적 설계의 제약을 고려하여 패리티를 포함하는 20 비트의 병렬 데이터로 처리하도록 크로스바를 구성하였다.

FC에서의 전송 단위인 프레임의 경계를 프레임의 처음을 나타내는 SOF(Start Of Frame) Ordered Set과 프레임의 끝을 나타내는 EOF(End Of Frame) Ordered Set으로 구분한다. Ordered Set은 8B/10B 코드에서 특별히 정의된 1 바이트(Comma character)와 4 바이트의 데이터로 구성되어 프레임의 경계 및 종류, 링크 명령 등으로 사용한다. 전송된 프레임을 입력 버퍼에 저장할 때 일반 데이터가 아닌 Ordered Set 정보를 저장하여 출력 포트에 전달하기 위하여 IFT(Internal Frame Tag)를 발생하여 프레임 데이터의 앞과 뒤에 붙여서 입력 버퍼에 저장하고 이를 출력 포트에 전달한다. 출력 포트에서는 링크를 통하여 전송할 때 IFT 정보를 사용하여 프레임의 SOF와 EOF를 만들어 보낼 수 있다. 그리고 입력 버퍼에 저장하거나 크로스바를 통하여 전달할 때 프레임의 처음과 끝을 나타낼 수 있는 비트 혹은 별도의 메커니즘을 포함하여야 한다.

5.3. 그룹 라우팅

단위 FC 스위치 시스템은 ISL(Inter Switch Link)를 통하여 상호 연결되어 보다 큰 규모의 FC 패브릭을 구성할 수 있다. 이 때 FC 스위치 시스템간에 복수의 ISL을 사용할 수 있으며 이 때 복수 개의 ISL 링크에서 사용하지 않는 적절한 링크를 선택할 수 있도록 경로 제어 메커니즘에서 지원하여야 한다[11].

Hunt Group은 FC Fabric에서 제공하는 기능으로서 다수의 N_Port로 구성된 Hunt Group을 지칭하는 Hunt Group Identifier를 착신지 주소로 사용하는 프레임에 대하여 패브릭에서 Hunt Group을 구성하는 다수의 N Port 중에 사용중이 아닌 N Port를 선택하

여 연결하는 기능이다[10]. 이는 클래스 1 연결의 경우는 전화에서의 같이 통화율을 높일 수 있고, 클래스 2, 3, 4인 경우는 총 전달 능력을 높임과 동시에 Latency를 줄일 수 있다. 또한 Hunt Group은 패브릭에서 제공하는 서비스로서 높은 신뢰도를 요구하는 시스템에 응용할 수 있다.

이러한 복수 ISL 라우팅과 Hunt Group 라우팅을 지원하기 위하여 경로 제어방식에서 실시간으로 라우팅을 제공하여야 한다. 앞에서 경로 제어 버스를 통하여 시작점 입력 포트와 목적점 출력 포트간의 크로스바 경로를 제어하는 방식을 설명하였다. 설명된 경로 제어 메커니즘에서는 시작점 IOC에서 자신의 P_ID의 다음 타임 슬롯의 DST_REQ에 목적점 포트의 논리적 주소, 즉 24비트의 FC Address를 실는다. 이 때 DST_REQ에 실린 논리적 주소에 해당하는 출력 포트가 응답하는 방식으로 다음과 같은 두 가지 방식을 고려할 수 있다. 첫째는, 제 3장에서 예시적으로 설명한 방식으로 각 출력 포트에서 자신의 논리적 주소 정보를 갖고 있어 DST_REQ와 검사하여 자신의 주소에 대한 경로 설정 요구에 응답하는 방법과 둘째는 논리적 주소를 물리적 주소로 변환하는 별도의 기능을 두는 방법을 고려할 수 있다. 전자는 각 출력 포트마다 자신의 주소를 가지므로 CAM과 같은

복잡한 룩업 로직을 필요로 하지 않고 간단한 비교 로직 만으로 구현이 가능하다. 그러나 그룹 주소(Group Address)나 복수의 스위치간 링크(ISL: Inter-Switch Link) 처리와 같이 여러 개의 포트 중에 선택해야 하는 경우에는 적절한 방법이 아니다. 따라서 본 연구에서는 후자와 같이 별도의 논리적 주소로 물리적 주소로 변환하는 기능(RC: Routing Control)을 두어 그룹 라우팅과 복수 ISL의 라우팅 기능을 갖도록 구현 하였다.

그룹 라우팅을 위하여 3장에서 경로 제어 절차를 다음과 같이 수정하였다. 수정된 절차는 다음과 같다.

- ① 각 입력단의 PCR는 자신의 P_ID가 되고 전송할 프레임이 있으면 DST_REQ에 목적지 노드의 주소(D_ID)를 실는다.
- ② SWC의 RC 기능에서는 매 타임 슬롯 마다 유효한 DST_REQ의 목적점 주소(혹은 그룹주소)를 물리적인 포트 주소(PTN:Port Number)로 변환하여 PTN 버스에 실는다.
- ③ 각 출력 포트에서는 매 타임 슬롯마다 PTN을 검사하여 자신의 포트이면 적절한 응답 신호(PC_ACK)를 발생한다.
- ④ CXBC에서 해당 크로스바의 접점을 연결한다.

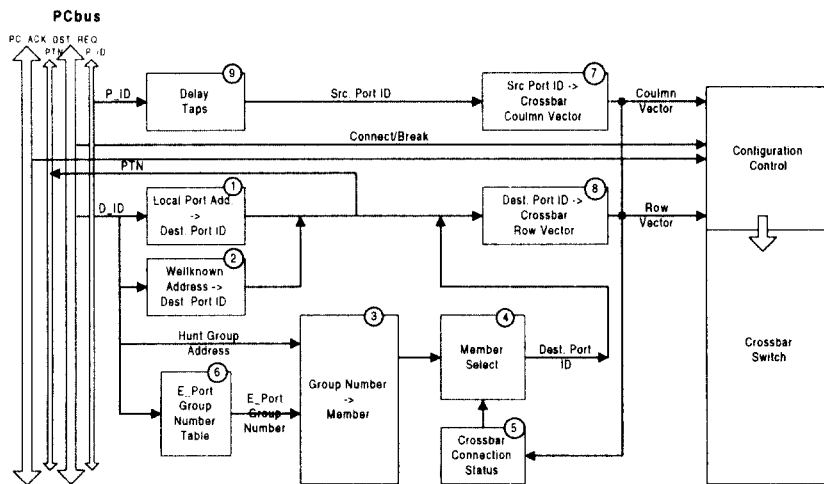


그림 8. 그룹 라우팅을 지원하는 경로제어기 구성

Fig 8. Oeganization of routing controller supporting group routing

⑤ 경로 설정을 요구한 PCR에서 설정된 크로스바 경로를 통하여 프레임 전송한다.

다수의 스위치와 연결된 구성에서 어느 스위치로 라우팅할 것인가에 대한 선택은 Spanning Tree와 같은 경로 선택 방식을 통하여 구현할 수 있으나 본 연구에서는 FC에서 주소를 Domain(8비트), Area(8비트), Port(8비트)로 구분하여 정의하고 있고[11] 스위치의 지리적 물리적 구성을 고려하여 주소를 할당하여 사용하는 방식이 가능한 데 프로토타입 시스템에서는 각 스위치 시스템이 서로 다른 Domain 주소를 갖도록 하여 스위치간 라우팅 기능을 간단하게 구현할 수 있도록 하였다.

앞에서 설명한 바와 같이 IOC에서의 경로 설정 및 크로스바를 통한 데이터 전송 기능, 출력 포트 제어 기능, RC에서의 논리적 주소를 물리적 주소로 변환하는 기능, 그룹 주소 및 복수 ISL 라우팅 기능 등 각 기능이 매 타임 슬롯 단위로 파이프라인 형태로 병렬로 수행되어야 한다. 따라서 각 기능 단위의 구현에서 가장 긴 시간을 소요하는 로직에 따라 타임 슬롯의 시간 간격을 결정하게 된다. 제4장의 성능 분석에서와 같이 단위 타임 슬롯 길이는 시스템의 성능에 영향을 주게 되며 특히 스위치의 규모가 커질수록 심각해지는 것을 확인하였다. 따라서 파이프라인으로 처리되는 단위 기능 단계가 최소화될 수 있도록 파이프라인 단계를 설계하여야 한다. 구현된 파이프라인 로직 설계에서는 메모리 접근 시간, FPGA의 입력과 출력간의 지연 그리고 칩간 혹은 보드간의 신호의 전달 지연으로 결정된다.

그룹 라우팅을 지원하는 RC기능과 CXBC 기능의 블럭 다이어그램은 그림 8과 같다. 만약 DST_REQ의 D_ID가 스위치의 로컬 포트 주소이면 논리적 D_ID 주소가 물리적 포트 주소(PTN)으로 변환된다①. 특별히 정의된 주소도 이를 지원하는 해당 포트 주소로 변환된다②. 만약 Hunt Group 주소가 감지되면 그룹의 멤버를 검사하고③, 현재 크로스바의 상태를 검사하여④ 가용한 하나의 포트를 선택한다⑤. 만약 D_ID가 자신의 스위치 주소가 아니면 해당 E Port 그룹 번호를 찾아서⑥ Hunt Group 동작과 같이 해당 E Port 그룹 번호에 속하는 멤버를 찾아서⑦ 가용한 E Port 주소를 PTN에 실는다. 만약 해당 목적점 포트에서

연결 완료 응답을 보내면 크로스바 구성 제어를 위한 행 벡터 주소와 열 벡터 주소를 발생 시킨다⑧. Delay Tap⑨은 파이프라인 동작에서 필요한 적절한 지연을 위해 사용한다. 이러한 동작은 매 타임 슬롯 단위로 파이프라인 동작에 의해 수행된다.

VI. 결 론

본 논문에서는 출력 포트 제어를 위한 신호선의 구현이 간단하고 높은 성능과 신뢰도를 제공할 수 있는 스위치 구조로서, 스위치 경로는 이중 평면의 크로스바로 구성되고 입력 단에 이중 버퍼를 갖고 있고, 출력 단에는 각 크로스바 스위치 평면별로 독립적인 출력 버퍼를 가지는 스위치 구조를 제안하였다. 출력 포트 중재를 위하여 시분할 방식으로 입력 포트의 경로 제어 요구를 폴링하여 입력 포트와 출력 포트 중재기 간의 신호선의 복잡도를 줄일 수 있는 IPA 방식을 제안하였다. 이중 입력 버퍼와 이중 스위치 평면을 사용하는 제안된 스위치 구조는 FC에서의 Intermix 서비스의 지원이 용이하고 프레임 처리를 위한 부하 분담에 의하여 전체 스위치 처리능력을 증가시킬 수 있는 장점을 가진다. 또한 이중의 입력 큐를 사용함으로써 출력 충돌을 감소시키고 스위치 경로 설정 지연에 따른 유효 대역폭 감소를 줄일 수 있는 특성을 가진다. 이러한 구조에서 프레임 경로 선택과 연결 설정 등을 위하여 보다 효율적인 스위치 분산 제어 구조 방식을 제안하였다.

제안된 IPA 방식의 경로 제어 기법에 대하여 시뮬레이션을 통하여 성능 분석한 결과 경로 제어 부담을 갖지 않는 FIA 방식에 비해 성능이 다소 떨어지는 것을 확인하였다. 이러한 입력 포트 폴링에 의한 성능 저하는 입력 포트 폴링을 개선한 IPA+ 방식에 의해 줄일 수 있어 FIA 방식에 가까운 성능을 얻을 수 있었다. 제안 스위치 구조에 대한 시뮬레이션을 통하여 IPA+ 출력 포트 제어 방식을 사용하고 이중 입력 버퍼와 이중 스위치 평면을 갖는 스위치 구조가 N배의 스피드 업을 하지 않고 거의 이상적인 출력 버퍼에 가까운 성능을 얻을 수 있고 FC의 Intermix 서비스를 효과적으로 수행할 수 있는 효율적인 구조임을 보였다.

그리고 FC 패킷 스위치 시스템의 구현에서 크로스바를 경로를 구현하는 방식으로 병렬 크로스바 방식

을 사용하였으며, 복수 ISL과 Hunt Group과 같은 그룹 라우팅을 실시간으로 처리하기 위한 경로 제어 알고리즘과 경로 제어를 실현하였다.

참 고 문 헌

1. Fibre Channel Association , Fibre Channel: Connection to the Future,1994.
2. H. T. Kung, "Gigabit Local Area Networks:A System Perspective," IEEE Communication Magazine, p79-89, April 1992.
3. Don E. Tolmie, "Gigabit Networking," IEEE LTS, p28-36, May 1992.
4. D. Getchell and P. Rupert, "Fibre Channel in the Local Area Network," IEEE LTS, pp. 38-42, May 1992.
5. K. B. Ocheltree, T. C. Tsai, R. Montalvo, "A Comparison of Fibre Channel and 802 MAC Services," Supercomputing93, 1993.
6. 이영희, "기가비트 통신기술," 전자통신동향분석, p123-134, 1995년 4월.
7. KilHwa Kim, KiDong Nam, HyeunTae Lee, YoungHee Lee, "Gigabit LAN based on Fibre Channel," ICOIN-9, p159-164, Dec. 1994.
8. HyeunTae Lee, YeongWeon Hwang, SeogKi Lee, "YoungHee Lee, A High Performance Switch Architecture for Fibre Channel," APCC95, p298-302, June 1995.
9. ANSI X3. 230-1994, Fibre Channel-Physical and Signalling Interface (FC-PH), Rev. 4.3, 1994.
10. ANSI X311/Project 901D/Rev 2.0, Fibre Channel-Enhanced Physical and Signalling Interface (FC-EP), Rev. 2.0. 1994.
11. ANSI X311/Project 959D/Rev 1.40, Fibre Channel-Switch Topology (FC-SW), Rev 1.40, 1994.
12. ANSI X311/Project 958D/Rev 3.0, Fibre Channel-Fabric Generic Requirements(FC-FG), Rev 0, 1994.
13. ANSI X311/Project 955/Rev 0.91, Fibre Channel-Link Encapsulation(FC-LE), Rev 0.91, 1995.
14. ANSI X311/Project 960D/Rev 4.5, Fibre Channel-Arbitrated Loop(FC-AL), Rev 4.5, 1995.
15. H. Ahmadi and W. E. Denzel, "A survey of modern high-performance switching techniques," IEEE JSAC, Vol. 7, No. 7, pp. 1091-1103, Sep. 1989.
16. Peter Newman, "A fast packet switch for the integrated services backbonenetwork," IEEE JSAC, Vol. 6, No. 9, pp. 1468-1479, 1988.
17. M. J. Karol, M. G. Hluchyj, and S. P. Morgan, "Input versus output queuing on a space-division packet switch," IEEE Trans. Commun., vol. COM-35 pp. 1347-1356, Dec. 1987.
18. T. H. Szymanski and S. Z. Shaikh, "Markov chain analysis of fast packet-switched crossbars and banyans with input, output and combined input-output queuing," The Columbia University Tech. Rep. CU/CTR/TR 214-90-44, Oct. 1990.
19. J.S.-C.Chen and T. E. Stern, "Throughput analysis, optimal buffer allocation, and traffic imbalance study of a generic nonblocking packet switch," IEEE J. Selected Area Commun., vol. 9, no. 3, pp439-449, April 1991.
20. Y. C. Jung, C. K. Un, "Performance analysis of packet switches with input and output buffers," Computer Networks and ISDN Systems 26, pp1559-1580, 1994.
21. A. Varma, S. Murthy, and R. Bryant, "Using camp-on to improve the performance of a Fibre Channel Switch," in Proc. IEEE Supercomputing '93, pp. 247-255, 1993.
22. K. Malavalli and B. Stovhase, "Distributed computing with Fibre Channel fabric," IEEE COMPCON, pp. 269-274, Feb. 1992.
23. A. Anzaloni et al., "Fibre Channel(GCS)/ATM interworking:A Design solution," IEEE GLOBECOM '93, pp. 1127-1133.
24. T. M. Anderson, R. S. Cornelius, "High-Performance Switching with Fibre Channel," IEEE COMPCON, 92. 2., pp. 261-264.
25. R. Hoebelheinrich, R. Thomsen, "Multiple Crossbar Network:A Switched High Speed Local Network," 14th Local Computer Network Confer-

ence, p 285-292, 1989.

- 26. Y. Tamir, H. C. Chi, "Symmetric Crossbar Arbiters for VLSI Communication Switches," IEEE Tran. on Parallel and Distributed Systems, Vol. 4, No. 1, 1993.
- 27. Ra'ed Y. Awdeh, H. T. Mouftah, "Survey of ATM Switch Architecture," Computer Networks and ISDN Systems 27, pp1567-1613, 1995.



이 현 태(Hyeun Tae Lee) 정회원

1983년 2월: 경북대학교 전자공학과 학사 졸업

1986년 2월: 연세대학교 전자공학과 석사 졸업

1986년 2월~현재: 한국전자통신연구소 통신시스템연구단 선임연구원

1993년 2월~현재: 연세대학교 전자공학과 박사 과정 재학중

※주관심분야: B-ISDN/ATM, Switching, Gigabit networks, Multimedia application

이 근 우(Keun Woo Lee) 정회원

1989년 2월: 한남대학교 전자기공학과 학사 졸업

1983년 4월~1996년10월: 한국전자통신연구소 선임연구원

1996년10월~현재: 해동정보통신(주) 전무이사

※주관심분야: B-ISDN/ATM, ATM Signalling, Multimedia application

손 장 우(Jang Woo Son) 정회원

1987년~1992년: 연세대학교 전자공학과 학사 졸업

1992년~1994년: 연세대학교 전자공학과 석사 졸업

1994년~현재: 연세대학교 전자공학과 박사 과정 재학중

※주관심분야: B-ISDN/ATM, Switching, Gigabit networks

오 영 렬(Young Yul Oh) 정회원

1989년~1993년: 연세대학교 전자공학과 학사 졸업

1993년~1995: 연세대학교 전자공학과 석사 졸업

1995년~현재: 연세대학교 전자공학과 박사 과정 재학중

※주관심분야: ATM traffic management, Gigabit networks, WDM

이 재 용(Jai Yong Lee) 정회원

1977년~1982년: 국방과학연구소 연구원

1983년~1986년: Iowa state university 연구조원

1987년: Iowa state university 조교수

1987년~1992년: 포항공대 전자계산학과 조교수

1987년~1994년: 산업과학기술연구소 겸직 연구원

1992년~1994년: 포항공대 전자계산학과 부교수

1994년~현재: 연세대학교 전자공학과 부교수

※주관심분야: Protocol engineering, Multimedia application, Network management.

이 상 배(Sang Bae Lee) 정회원

1954년~1958년: 공군 사관학교 학사 졸업

1951년~1961년: 서울대학교 전자공학과 학사 졸업

1962년~1964년: Stanford university 석사 졸업

1961년~1967년: 공군사관학교 조교수

1967년~1969년: 원자력 연구소 연구관

1969년~1979년: 서울대학교 전자공학과 조교수

1972년~1975년: 영국 Newcastle university 전자공학과 박사 졸업

1978년: 한국통신기술연구소 위촉연구원

1979년~현재: 연세대학교 전자공학과 교수

1982년: 영국 Newcastle university 교환 교수

※주관심분야: Computer networks, B-ISDN/ATM, High speed protocol.