

2×1 광버퍼 스위치

정희원 서창진*

2×1 Optical Buffering Switch

Chang-jin Suh* *Regular Members*

요 약

비동기식 네트워크에서 일어나는 광패킷들의 일시적인 쇄도를 데이터의 광/전기 신호변환 없이 해결하기 위하여 지연선(DL : Delay Line) 스위치가 전광스위치로 사용되고 있다. 이 방식에서는 여러 개의 지연선과 2×2 광스위치를 준비하여 패킷이 통과하는 지연선의 길이를 적당히 조절하여서 원하는 만큼 시간을 지연시키는 방법으로 버퍼링 기능을 수행한다.

이 논문은 FIFO 2×1 광버퍼를 소개하고 그 스위치의 구조와 제어방식에 대하여 논하고 있다. 기존의 광버퍼 기술기반 상에서 제안된 이 스위치는 블로킹이 발생하지 않으며 매우 간단한 방법으로 자체 라우팅(self-routing)을 하는 유용한 특성뿐 아니라 최소 지연시간 및 binomial 분포에서의 최소 패킷손실율과 같이 최적의 성능을 발휘한다. 제안된 구조는 크기, 잡음, 파워, 발열 등의 문제를 가지고 있는 가장 적은 수의 광소자만을 사용하여서 구현 비용 및 운영상의 문제를 최소화하여 실용화를 가로막고 있던 많은 제한을 없애준 매우 혁신적인 네트워크 구성요소이다.

ABSTRACT

Switched Delay Lines(SDL) as all-optical switches are used in asynchronous network to handle bursty optical cells without optical/electrical nor electrical/optical conversion. SDL has many DLs(Delay Lines) and controls the total delay by providing different routing paths of a cell delay with 2×2 optical switches. The total delay is determined by the total length of DL's that a cell visits and this performs buffering capability of cells.

This paper introduces an innovational FIFO 2×1 optical buffering switch based on current all-optical technologies, and describes its structure and control strategy. This approach not only has many useful properties like nonblocking, self-routing, but also shows the best performance such as minimum cell loss rate under binomial distribution and minimum delay. The proposed structure is a very practical network component with minimum implementation cost and minimum maintenance problem by using minimum number of optical elements that currently impose many physical constraints of size, noise, power or heat.

I. 서론

1970년대 이후 눈부신 발전을 이룩한 광통신의 발전 결과, tera(10^{12})급 신호전송을 이론적으로 가능하게 되었으며, 실제적으로 giga bps 단위의 많은 전송설비가 가동되고 있다. 이 엄청난 신호전송능력은 그 제어의 어려움 때문에 한 개의 대역폭을 고정된 수의 중간속도의 대역폭을 가진 채널들로 분

배시키는 방법으로 이용해 오고 있다. 이러한 동기 방식은 구조가 간단하며 관리하기 편하여 전화와 같은 CBR(Constant Bit Rate) 정보통신이 주류를 이루는 지금까지의 네트워크에서 뛰어난 투자비용대비 효과를 볼 수 있었다.

반면 앞으로의 초고속네트워크에서는 수백 Gbps의 대용량 고속통신뿐만 아니라 많은 중간 속도의 서비스와 원격검침과 같은 kbps대의 저속에 이르기까지

* 송실대학교 컴퓨터학부(cjsuh@computing.soongsil.ac.kr)

** 논문번호 : 98392-0901, 접수일자 : 1998년 9월 1일

지 다양한 범위의 서비스를 제공해야 하며, 또 한편으로 실시간에서 전송량이 계속 변하는 VBR (Variable Bit Rate) 서비스도 제공할 수 있어야 한다. 만일 동기식 전송 방식만으로 이 모든 서비스를 수용하고자 한다면 다양한 서비스들을 일정한 대역폭의 채널들로 할당하는 과정에서 대역폭의 큰 낭비를 초래하게 된다. 이러한 손실을 줄이기 위하여 사용자마다 할당되는 동기식 채널 대신 ATM에서 사용되는 것과 같은 가상채널을 사용하여 여러 서비스들이 주어진 대역폭을 함께 사용하는 비동기식 전송이 등장하였다.

비동기 방식은 대역폭 분할에 따른 낭비가 발생하지 않는 대신 스위치 노드에서 여러 가상채널이 특정 링크를 점유하는 과정에서 스케줄링이 필요하다. 만일 두 개의 가상채널에서 동시에 특정 전송 링크의 할당을 요구한다면 한 개의 가상채널을 선택하여 그 곳의 패킷을 즉시 전송하고, 다른 가상채널내의 패킷은 전송 링크가 사용이 가능할 때까지 지연시키는 식의 버퍼링 기능을 갖추어야 한다. 패킷들을 버퍼링하기 위해서는 기억소자가 필요하지만 불행히도 현재는 광기억소자가 존재치 않으며 가까운 미래에도 광기억소자가 상용화될 가능성은 희박하다.

이러한 구현상의 한계를 극복하기 위하여 광소자에 의한 버퍼링 대신, 광/전기 및 전기/광 신호변환 장치를 이용하여 전기 영역에서 버퍼링을 한 후 광 전송을 행하는 방법이 보편적으로 사용되고 있다. 이 방식은 보낼 수 있는 광전송 속도가 그보다 현격히 낮은 전기소자의 처리속도 때문에 실제 가능한 광전송보다 훨씬 낮은 속도의 광선로를 이용하고 있다.

전송속도를 극대화할 수 있는 다른 시도는 광데이터 자체는 그대로 유지시킨 채 약간의 제어에 필요한 정보만을 전기신호로 변환하여 이를 통하여 광패킷을 제어하는 전광(all-optical) 스위치방식이다. 이 방식은 보편적으로 통용되는 전기적 제어방식의 광기들을 사용할 수 있지만 광데이터의 버퍼링에 필요한 광기억소자의 역할을 대체할 광지연소자가 필요하다. 본 논문은 광지연소자로서 DL을 이용한 SDL(Switched Delay Line) [1-4]방식을 근간으로 한다.

SDL은 DL을 이용한 전광스위치 방식이다. 여기에서는 여러 개의 지연선들을 준비하여 패킷이 통과하는 지연선의 길이를 2×2 광스위치를 적절히 제어하여서 원하는 만큼 시간을 지연시키는 방법으

로 일반메모리와 같은 버퍼링 기능을 수행한다. 이때 패킷의 총지연시간은 패킷이 경유한 모든 광지연소자에서 지연된 시간의 합이다.

우리가 다룰 전광스위치는 매 채널마다 하나의 주파수를 할당하는 방식을 사용하는 일반 전광스위치방식과는 달리 하나의 주파수대에 여러 채널이 비동기적으로 뒤섞이는 것을 허용하고 있으며, 이 논문은 어떤 방법으로 패킷들을 손실하지 않고 공통의 채널로 삽입할 수 있는가를 연구한다. 이 방식은 전통적인 전광스위치의 기본구조를 그대로 답습하였고, 사용되는 소자에 대하여 아무런 제한을 가하지 않는다. 따라서 광스위치 소자 특성 및 제작에 관한 부분은 생략한다. (이 부분에 관해서는 전광스위치에 관한 특집을 다루고 있는 [8]을 참고하기 바람.) 그 대신 DL과 2×2 스위치를 어떻게 구성하여야 손쉬운 제어만으로도 많은 버퍼용량을 얻을 수 있고, 패킷손실율을 줄일 수 있는가하는 네트워크 성능의 관점에서만 논하고자 한다.

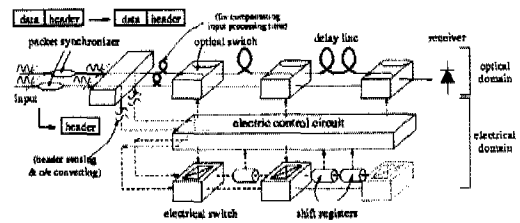


그림 1. SDH의 하부구조

앞으로 논할 전광스위치는 다음과 같은 방법으로 동작한다(그림 1 참조). 송신노드에서는 패킷의 payload는 높은 주파수대역에서 고속으로 전송되며, 헤더는 이보다 훨씬 낮은 주파수대역에서 저속으로 전송한다. 또한 헤더는 payload보다 조금 일찍 전송된다. 이렇게 코딩된 패킷이 스위치노드에 다다르면 synchronizer의 도움으로 패킷의 위상을 맞추어 재배열된다. 그리고 헤더부분만을 디코딩하여 전기적인 신호로 바꾼 후, 패킷의 입력여부를 전기적 제어 회로로 결정하여 제어에 필요한 전기신호가 생성된다. 조금 후 payload가 입력되면 payload는 광회로를 빠르게 통과하며, 이보다 낮은 속도로 전기신호는 전기회로를 지난다. 이 두 회로는 동일한 토폴로지를 가지고 있고 서로 일대일 대응되는 광스위치와 전기스위치가 동일한 상태(\times 와 $=$)를 항상 유지하기 때문에 payload와 이와 관련된 헤더로부터 만들어진 전기신호는 항상 토폴로지 내에서 동일한 위치에 놓인다. 전기신호와 광신호의 속도의 차는

payload의 길이에 비해서 전기신호의 길이가 짧은 것을 이용하고 대부분의 전기신호를 병렬화하여 극복된다. 광스위치와 전기스위치의 상태는 전기제어 영역에서 내린 결정에 따라 정해진다.

광지연소자는 필요할 때까지 정보를 보관하는 일반메모리와 달리 단지 일정시간만을 지연시킬 수 있다. 이처럼 제한된 광지연소자의 기능을 통신에 이용하기 위하여 초기에는 케환방식(feedback)이 제안되었다. 이 방식은 한 패킷이 케환루프를 통하여 여러 번 동일한 지연소자를 통과하게 설계되었는데 케환횟수가 거듭될수록 신호 레벨은 저하하는 반면 신호잡음이 증가하는 물리적 한계를 안고 있다. 이러한 단점을 줄이기 위하여 가능한 한 지연소자를 통과하는 횟수를 최소화 할 수 있는 무케환(feedforward)방식이 등장하였다.

전광스위치 구조로서 일반적인 스위치 형태를 가진 구조가 제안되었으나^[1] 한 두 개의 기억소자를 사용하였기 때문에 deflection network에서 패킷손실율을 줄이는 목적으로 사용됐을 뿐 본 논문에서 처럼 단독소자로서 사용할 수는 없는 수준이다. 좀 더 발전된 형태가 Manhattan Street 네트워크에서 모든 노드가 사용하는 송수신 장치로 제안되었다^[2]. 이곳에서는 DL을 여러 단에 걸쳐 직렬로 배치하여 각 DL에 한 개씩의 패킷이 보관될 수 있게 구성한 결과 패킷의 우회확률(deflection probability)을 현저히 감소시켰다. 한편 DL의 길이를 기본길이에서 여러 길이로 늘여 동일한 단에서 병렬로 배치한 Staggering switch가 $N \times N$ 스위치로서 선보였다.^[3]

이와는 별도로 [4]에서는 2단으로 고정된 2×2 스위치를 소개하여 단 2개의 서로 다른 길이의 DL만을 이용하여 패킷손실율을 얼마까지 낮출 수 있는가를 알아보았다. 그 결과 3개의 단위길이의 DL을 사용한 경우보다 더 우수한 결과를 얻을 수 있었다. SDL의 가격은 2×2 스위치 개수에 주로 좌우되는데, DL의 개수를 줄이면 2×2 스위치 소자의 개수도 줄게되어 SDL의 전체 제작비용을 줄일 수 있다. 그러나 대신 패킷이 입력되는 순서대로 출력이 보장되어야 하는 FIFO (First-In First-Out) 규칙이 더 이상 성립되지 않게 되고, 지연시간이 증가하는 부작용이 따랐다. 이 단점들은 다중길이의 DL로 입력된 패킷은 그 DL에서 정해진 지연시간이 완료되는 시간슬롯에 비로소 접근할 수 있는 성질 때문이다. 만일 모든 패킷이 DL의 중간부분을 통과하고 있을 경우 그 시간슬롯동안은 출력링크는 사용되지 않고 낭비된다.

본 논문에서는 2×1 다단 광수신기를 소개한다. 이 방식은 FIFO가 보장된 상태에서 사용된 2×2 스위치 소자 수(n)가 증가함에 따라 버퍼용량이 기하학적으로 증가하는(2ⁿ) 구조를 제안하고, 제안된 구조의 성능 및 제반 성질들을 다루고 있다. 이 방식은 다중길이의 DL을 사용하여 실제적으로¹⁾ FIFO 성질과 무블로킹 성질을 동시에 만족하는 최초의 시도임과 동시에 여러 최적 성질이 규명된 완벽한 솔루션이다.

2×1 버퍼링 스위치는 다음과 같은 경우에 이용된다. 첫째 한 쌍의 FIFO 2×1 스위치로 FIFO 2×2 스위치를 만든다. 이는 버퍼용량이 사용되는 스위치의 수에 지수적으로 비례하는 무블로킹 FIFO 2×2 버퍼링 스위치는 이론적으로 존재할 수 없기 때문에 큰 버퍼용량이 필요할 경우 가장 경제적인 방법이다. 2×2 스위치는 FDDI 링, 두 개의 단방향 초고속 링을 잇는 브리지나 Manhattan Street 네트워크에서의 송수신소자로 사용할 수 있다. 두번째로는 2×1 버퍼링 스위치는 스위치 내부의 특정 기능을 구현한다. [5]에서는 N 개의 입력이 각각 N/2 개의 두 출력단으로 나뉘는 다중채널분할(multi-channel splitter)을 여러 단계에 걸쳐 나누는 방법으로 $N \times N$ 스위칭 및 버퍼링을 행하는 ATM 스위치를 제안하였다. 이 방식은 전기적으로 작동되지만 각 다중채널분할기의 후단에 2×1 광버퍼링 스위치를 사용하면 $N \times N$ 전광(all-optical) 스위치로 사용이 가능하다.

우리가 논할 방식은 CORD 프로젝트에서 CRO라는 이름으로 노드가 두 개인 2.488Gbit/s 네트워크에서 80Mbit/s의 헤더채널을 사용하고 패킷의 형태를 ATM 셀의 규격으로 가지게 한 채 구현되었다^[7]. 이 실험에서는 그림 1에서처럼 DL의 수를 두 개로 한정된 상태에서 구동시킨 결과 여러 부하 상황에서도 훌륭히 동작될 수 있었음을 확인하였다. 이 논문에서는 CORD에서 DL의 수를 2로 고정된 제한을 없애고 임의의 자연수 d로 확장하였으며, 제어방식도 이론을 체계화하여 복잡도를 최소화하였다.

본 논문은 다음과 같이 구성된다. II장에서는 제안된 2×1 광버퍼 스위치의 구조를 설명하고 사용

1) [3]에서 제안된 방법을 이용하여 길이가 1,2,3 .. L인 L개의 DL을 병렬로 연결하여 FIFO를 유지시킬 수도 있겠지만 이는 병렬로 연결하는 과정에서 많은 (4L) 스위치 소자를 사용해야 하기 때문에 $N \times 1$ 스위치로서는 실효성이 없다.

될 제어 방법을 정의한다. III장에서는 2×1 버퍼링 스위치에서 블로킹의 문제를 다룬 후 블로킹이 원천적으로 발생하지 않는 제안된 광버퍼 스위치의 여러 최적성질들을 열거하고 있다. 이때 사용되는 증명은 VI장에 언급된다. IV장에서는 제안된 스위치의 성능을 분석하고 있다.

II. 2×1 버퍼링 스위치

2.1. 스위치 구조

두 개의 광선로로부터 패킷들이 입력되어 한곳의 광선로로 출력되는 멀티플렉서를 가정하자. 시간은 시간슬롯으로 나누어져 있고, 시간슬롯이 개시되는 시점에 맞추어 패킷이 입력된다. 시간슬롯 길이는 광소자와 전기제어 신호들의 상태변수가 변하기 시작한 후 안정된 새로운 상태를 유지할 때까지의 소요되는 동작지연시간과 패킷전송시간, 그리고 동기화의 최대 허용 오차 시간의 합으로 결정된다. 입력된 패킷들은 입력되는 시점에 따라 우선순위가 결정되며 동일하게 입력된 두 패킷의 경우는 임의로 우선순위가 주어진다. 이 우선순위에 따라 출력은 가장 우선순위가 높은 패킷, 즉 스위치 내의 패킷 중 가장 일찍 입력된 패킷이 출력되므로 FIFO는 보장된다. 만일 스위치가 우선순위를 어겨서 미리 출력되거나, 원하는 시간에 출력이 불가능할 경우 블로킹이 발생했다고 말한다. 블로킹이 없는 스위치는 버퍼용량내의 범위에서 매 시간슬롯마다 패킷을 출력링크로 하나씩 전달하므로 출력링크의 이용률을 극대화할 수 있다.

제안된 2×1 버퍼링 스위치는 (d+1)개의 광스위치소자를 이용하여 2^d개까지의 패킷을 블로킹 없이 처리한다. 2×2 스위치소자는 광패킷들이 순차적으로 통과할 수 있게 직렬로 배치되며 두 인접한 스위치소자들은 DL과 광링크로 연결된다. 마지막에 위치한 광스위치소자는 양쪽 입력 중 하나를 선택하여 이를 출력포트로 전달한다. 만일 두 개의 패킷이 동시에 발생한다면 그 중 하나는 정상적으로 출력될 수 있지만 나머지 하나는 더 이상 버퍼링이 불가능하여 그림 2의 아래쪽 포트를 통하여 폐기된다.

그림 2에서와 같이 k번째 (1≤k≤d) 광스위치소자 SW_k와 그 출력단에 위치한 k번째 DL과 k번째 광링크를 k번째 단계(stage)라 하자. k번 단계로 진입되는 패킷이 DL과 광링크 중 어느 쪽으로 출력하가는 k단의 2×2 광스위치소자의 상태에 따라 결

정된다. 광스위치의 상태는 단순통과상태(=)와 전환통과상태(×)로 나뉘어지는데 매 시간슬롯마다 한 개의 안정된 상태만을 가진다. 단순통과상태에서 SW_k는 상단의 입력선으로 입력되는 패킷을 상단의 출력선으로 출력하게 되고 동시에 하단의 입력선으로 입력하는 패킷을 하단의 출력선으로 출력한다. 전환통과상태의 경우는 동일한 입력조건에서 출력되는 장소가 상단과 하단이 서로 맞바뀌게 된다. 하단으로 출력된 패킷은 지연시간 없이 (k+1)단계의 광소자로 입력되며, 상단으로 출력된 패킷은 DL_k의 길이에 해당하는 시간슬롯 동안 DL_k에서 지연된 후 (k+1)단계 광스위치소자로 입력된다. 이처럼 k단계의 스위치소자의 상태는 패킷이 k단에서 얼마나 지연시킬 것인가를 결정한다. k단의 DL은 2^{k-1}의 길이를 가지고 있으므로 최대로 머무를 수 있는 시간슬롯은 모든 DL을 다 경유할 경우 2^{d-1}이 된다.

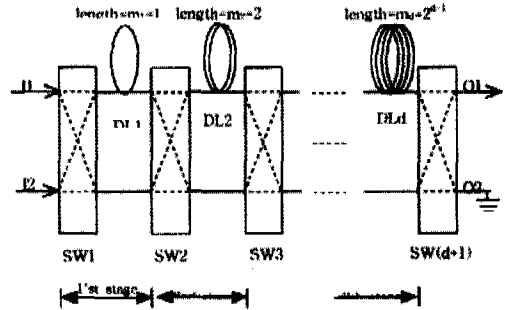


그림 2. d단계로 구성된 2×1 버퍼링 스위치의 구조

광버퍼 스위치 용량 C는 최대 처리 가능한 패킷의 수로 정의된다. 이 정의에 따르면 버퍼링 기능이 없이 한 개의 2×1 스위치소자만으로 이루어진 2×1 스위치는 한 시간슬롯 당 두 개의 입력선 중 한 곳에서만 패킷이 입력될 경우 이의 처리가 가능하므로 용량 C는 1이다. 제안된 d단계로 구성된 2×1 버퍼링 스위치 용량을 C_d로 표기할 때 스위치는 최대 2^{d-1} 개의 패킷을 DL에 저장할 수 있고 또한 입력단의 1개의 패킷을 추가적으로 조절할 수 있으므로 다음의 관계가 성립한다.

$$C_d \leq 2^d, \quad d \text{는 } 0 \text{ 혹은 자연수} \quad (1)$$

2.2. 제어방식

제어방식은 광기기보다 훨씬 동작속도가 느린 전기회로에 의하여 구현되기 때문에 극히 단순해야 한다. 여기서 제안된 제어방식 F는 단 두 개의 상

대변수 S, R로 간단히 이루어지는 만큼 이점에서 매우 뛰어난 방법이다. 시스템 변수 S는 현재 스위치내의 DL에 있는 패킷의 총수를 뜻하는데 매 시간슬롯마다 경신된다. 이를 기준으로 입력되는 패킷들에게 R이라는 버퍼링시간이 주어진다. R=3의 값이 입력되는 패킷에 주어지면 이 패킷은 스위치에서 세 시간슬롯을 지체한 후 출력된다. S와 R은 최소값 0을, 최대값 C-1(C=버퍼용량)을 가지며, 패킷이 하나도 입력되지 않은 초기상태에서 R은 0을 유지한다. 한 시간슬롯동안 입력되는 패킷의 수에 따라 제어방식 F는 시간슬롯변수 t를 이용하여 다음과 같이 정의된다. 이때 R₁, R₂는 각각 패킷 P₁, P₂에 주어지는 변수 R을 뜻한다. 한편 S_t는 시간 t에서의 변수 S로서 아래에서처럼 t시간에 입력된 패킷의 수에 따라 변화된다.

- ⊙ 패킷 0개 입력 : S_t = max (S_{t-1}-1, 0)
- ⊙ 패킷 1개(P₁) 입력 : S_t = S_{t-1}, R₁ = S_{t-1}
- ⊙ 패킷 2개(P₁, P₂) 입력 :
 0 ≤ S_t < C-1 일 때
 : S_t = S_{t-1}+1, R₁ = S_{t-1}, R₂ = S_{t-1}+1
 S_t = C-1 일 때
 : S_t = S_{t-1}, R₁ = S_{t-1}, R₂ = 0

마지막 줄은 용량을 초과하여 패킷이 입력되는 경우의 제어방식을 보여주고 있는데 두 개의 입력 패킷 중 P₂가 스위치의 다른 출력인 O₂를 통하여 제거되는 과정을 보여준다. 이때 초과된 패킷은 R=0의 값을 부여받아 입력하자마자 또다른 출력단을 통해 버려진다. 이처럼 폐기여부는 패킷이 입력될 때 결정되며 일단 입력이 허용된 패킷들은 모두 예외 없이 R 시간슬롯 후에 출력된다.

지연시간 R을 배정 받은 패킷은 R시간을 정확히 맞추기 위하여 통과하는 DL과 광링크를 선택해야 하는데 m_k(DL_k의 길이)가 2^{k-1}인 상황에서 R만큼 지체하기 위해서는 R을 2진수 r_{d+1}...r₁로(r₁이 제일 낮은 자리수임) 표기한 후 2진수 값이 1인 자리수를 모아 그 길이를 가지는 DL만을 경유해야 한다. 즉 k단계에서 광패킷은 자신의 r_k가 1인 경우는 길이가 2^{k-1}인 DL을, r_k가 0인 경우는 하단의 광링크를 선택하는 방법으로 스위치 내부에서의 라우팅은 R이 정해질 때 확정된다.

R은 스위치 내에서 패킷이 경유해야 할 경로를 담고 있으므로 패킷의 헤더에 보관되어 광패킷과 함께 스위치 내에서 움직인다. 이처럼 제안된 구조

는 라우팅 정보가 라우팅 경로를 따라서 움직이는 자체라우팅(self-routing)을 수행한다. 반면 k단의 스위치 SW_k의 상태제어는 k단을 진입하는 패킷들의 r_k값으로 결정된다. 만일 두 개의 패킷이 입력되었는데 두 r_k값이 동일하다면 스위치는 두 패킷 중 하나는 원하는 링크로 보낼 수 없다. 제안된 구조는 3절에서 언급하다시피 이러한 충돌은 발생하지 않는다 (정리 1참조). 이 성질을 이용하여 SW_k(k=1,2,..,d)는 자신을 통과하려는 패킷을 하나만 찾더라도 그 패킷의 헤더 부분에 기록된 r_k값을 읽고 다음 SW_k의 상태를 결정할 수 있다. 만일 상단(I₁ 혹은 DL_{k-1}, k=1, .., d-1)에서 입력되는 패킷의 r_k값이 1이거나 하단에서 입력되는 패킷의 r_k값이 0일 경우는 = 상태를, 그 외의 경우는 × 상태를 유지한다. 한편 DL_d에서부터 패킷이 SW_{d+1}로 입력될 경우는 = 상태를, 그 외의 경우는 × 상태를 유지시킨다. 한 SW_k의 입력단자에서는 그 곳에 패킷이 입력되는지의 여부(existence bit)와 입력될 경우 그 r_k bit만을 제어신호로 사용한다. 즉 양쪽 링크에서의 총 4 bit로서 SW_k (k=1,..,d+1)의 상태를 제어할 수 있다. 위에서 살펴본 바와 같이 개개의 스위치소자는 분산제어가 가능하고, 제어의 복잡도는 또한 버퍼용량과 관계없이 최소화되었다.

III. 스위치의 성질

스위치 용량 내에서 블로킹 발생불가 조건과 블로킹 여부를 미리 알아내는 작업은 2×1 수신기 동작에 매우 중요하다. 이는 현재 DL에 보관되고 있는 패킷 중 블로킹 당하여 제거되는 패킷의 이후의 패킷은 FIFO를 지키기 위하여 제거해야 하기 때문이다. 제안된 구조에서는 블로킹이 발생할 수 없음을 정리 1.의 결과에 따라서 이러한 작업이 근본적으로 필요하지 않는다. 아래에서는 여러 정리들이 언급되는데 증명 부분은 6장 부록에서 별도로 언급된다.

정리 1. 그림 2에서 정의된 2×1 버퍼링 스위치에서 제어방식 F가 사용될 경우 블로킹은 발생하지 않는다.

블로킹은 k단 스위치소자에 동시에 도착한 두 개의 패킷이 모두 상단으로 전달되어야 하거나 모두 하단으로 출력되어야 할 때 발생한다. 제안된 버퍼링 스위치는 블로킹을 회피하는 여타의 방법을 동

원하지 않고도 DL의 길이 시퀀스를 특정한 값으로 고정함으로써 간단하게 블로킹을 피하고 있다. 제안된 구조에서 d개의 DL 길이는 $\{1, 2, 4, \dots, 2^{d-1}\}$ 의 기하급수적인 값만을 가지고 있다. 이 순열조합은 0과 스위치 용량 C-1 사이의 모든 정수를 가장 적은 수의 정수로 표시하게 한다. 한편 제안한 스위치의 용량은 $C_d=2^d$ 인데 이는 식 (1)을 고려해 볼 때 블로킹 발생이 없는 상태에서 이를 수 있는 가장 큰 용량이다. 이러한 사실을 요약하여 정리2를 얻을 수 있다.

정리 2. 제어방식 F를 사용하고 d개의 DL 및 d+1개의 스위치소자를 가진 블로킹 발생이 없는 스위치를 가정하자. 이 중 그림 2에서 정의된 2×1 버퍼링 스위치는 가장 큰 스위치 용량 C를 가지고 있으며, 또한 사용된 d개의 DL들은 가장 짧은 길이를 가지고서 최대용량 C를 이루고 있다.

서로 다른 길이의 DL을 배치하는 방법은 총 d!가지가 있지만 블로킹을 발생하지 않는 것은 오직 그림 2 한가지 경우임을 정리 3은 알려준다.

정리 3. 제어방식 F의 사용을 가정하고 총 d개의 DL의 총길이가 2^d-1 을 가정할 때, 그림 2에서 정의된 2×1 버퍼링 스위치에서 블로킹을 발생하지 않는 유일한 조건은 m_k (DL_k 의 길이) = 2^{k-1} 이다. ($k=1, 2, \dots, d$)

2장에서 언급한 스위치소자 제어방식과 위에서 언급한 정리들을 종합하면 그림 2에서 제안된 구조가

- ⊙ 최소 개수의 DL과 최소 개수의 스위치소자와
- ⊙ 최소 길이의 DL만을 사용하여
- ⊙ 최소한의 간단한 제어로 분산제어가 가능한
- ⊙ 최대 용량을 이루는

유일한 블로킹 없는 스위치 구조라는 것을 알 수 있다.

또한 제어 방식은 입력되는 패킷에게 최단의 출력시기를 부가하므로 자연히 패킷지연시간은 최소화된다. 만일 DL_k 의 길이를 늘려 2^{k-1} 이상의 길이를 가지게 한다면 이 성질은 없어지지만 변수 S 값 배정방법을 달리하여 FIFO 규칙은 유지시킬 수 있다. 즉 m_1, m_2, m_3 가 1, 2, 5라면 용량은 $(1+2+5)+1=9$

로 제안된 구조의 용량인 8보다 1만큼 늘어날 수 있다. 현재 3개의 패킷이 스위치 내에 있고 2개의 패킷이 입력된 상황을 가정하자. 이 경우 한 패킷 P_1 은 R=3을 배정 받지만 다른 하나 P_2 는 4 대신 5의 값을 R로 배정 받는다. 이는 {1, 2, 5}로 만들 수 있는 $3(=1+2)$ 다음으로 큰 정수 합은 5이기 때문이다. 그 결과 P_1 이 출력된 다음 시간슬롯에는 아무 패킷도 출력되지 않고 그 다음 시간슬롯에 가서야 P_2 가 출력된다. 이번 경우와 그림 2에서 제안된 경우 둘을 비교해 보면 이번 경우가 지연시간이 길어지는 것은 당연하지만 패킷손실율에서 어느 쪽이 적은지는 예측하기 어렵다. 만일 양쪽 입력에서 이산분포로 패킷이 입력된다면 아래의 결과를 얻는다.

정리 4. 두 입력단에서 이산분포로 패킷이 입력될 경우 그림 2에서 정의된 2×1 버퍼링 스위치는 $m_k = 2^{k-1}$ ($k=1, 2, \dots, d$)일 때 최소의 패킷손실율을 나타낸다.

정리 4는 임의의 m_k 값 중 어떤 m_k 값을 가질 때 최소의 패킷손실율을 가질 수 있는가는 점접하였다. 패킷손실율을 줄이려면 용량을 늘려야 하므로 용량을 높이기 위해서 무작정 m_k 의 값을 증가해 보자. 이 경우 0과 C-1사이의 합성 불가능한 지연시간들 때문에 알맞는 만큼의 지연시간을 만들 수 없으므로 출력단에 패킷을 보낼 수 없는 경우가 발생한다. 즉 출력단의 이용도가 감소한다. 정리 4.는 본 논문에서 제안한 구조가 이산분포에서 가장 적은 패킷손실율을 보인다는 사실을 밝혀냈다. 이는 출력단의 이용도를 희생하고 스위치 용량을 늘이는 방법은 패킷손실율의 관점에서 볼 때 손해임을 알려준다.

위에서 살펴본 바와 같이 스위치 성능 면에서 제안된 2×1 버퍼링 스위치는

- ⊙ 최소 버퍼 지연시간과
- ⊙ 이산분포에서 최소 패킷손실율을 가진다.

IV. 성능분석

제안된 구조와 기존의 방식에 대해서 패킷손실율을 조사했다. 기존의 방법 중에는 다중길이의 DL로 FIFO를 유지시키는 방법은 제안된 바 없으므로 단일길이만을 사용하여 FIFO를 구현할 수 있는 Quadro 방식을 기존 방식으로 채택하였다. (d+1)개 - d 개의 2×2 스위치소자와 한 개의 2×1 스위치

소자 -의 스위치소자를 사용하여 d-단계로 스위치를 구성할 경우 제안된 구조는 2^d의 용량을 가지는 반면 Quadro의 경우는 (d+1)의 용량 C를 가진다.

두 방식들은 무블로킹 성질 때문에 부하에 따른 입력이 이산분포를 이룰 경우 Discrete Markov chain을 통하여 패킷손실율을 쉽게 계산할 수 있다. 즉 2.2 제어방식에서 언급된 St를 상태변수로 정의 하면 단순화를 위하여 어떠한 수학적 가정을 하지 않더라도 2x1 광스위치는 길이가 (C-1)인 M/M/1의 경우로 전환된다. 이산분포 부하 r은 한 입력포트에서 어느 특정 시간슬롯에 패킷의 입력확률이 1/2인 경우로 정의되며, 두 입력단으로부터 시간슬롯 당 평균 r개의 패킷이 입력된다. 또한 임의의 두 슬롯 사이와 두 입력단 사이의 패킷 발생여부는 서로 독립적인 관계를 가진다. r=0.9의 의미는 완벽한 버퍼링을 수행할 경우 출력단에는 평균 90%의 타임슬롯에는 패킷이 출력됨을 뜻한다.



state = number of packets in the DLs

그림 3. 용량이 C인 무블로킹 2x1 버퍼링 스위치의 Discrete Markov Chain (state transition diagram)

그림 3은 용량이 C인 무블로킹 2x1 광스위치의 상태 변이도를 나타내고 있다. 만일 용량이 C인 경우는 0부터 C-1개의 패킷을 가지므로 C개의 상태가 존재하며, 패킷의 손실은 C-1 상태에서 r2가 발생했을 경우이다. 상태도에서 r1은 한 시간슬롯동안 입력된 패킷의 개수가 i일 확률이다. 평균부하 r은 따라서 $r = 0 \times r_0 + 1 \times r_1 + 2 \times r_2 = r_1 + 2 \times r_2$ 의 관계를 가진다. 한편 패킷손실율은 C-1에서의 상태값과 r2값의 곱으로 나타낸다.

그림 4에서는 제안된 구조를 "propose"로 명기 하고 실폭선으로 표현하였으며 Quadro의 경우는 점선으로 처리하면서 일반적인 부하상태인 r=0.6과 극히 과부하인 r=0.9인 두 가지 경우에 대하여 패킷손실율을 나타내고 있다. 패킷의 손실율은 대략적으로 버퍼의 용량에 지수적으로 비례하기 때문에 스위치와 버퍼용량이 일치하는 Quadro에서는 세로축이 log로 표시된 그림2에서는 선형적으로 감소한다. 한편 버퍼용량이 소요되는 스위치소자 수에 대해 지수적으로 증가하는 제안된 광수신기의 경우는 이보다 패킷손실율이 훨씬 빠르게 감소하는 것을 볼 수 있다. 각각 3개, 4개의 스위치로 제작된 제안된 스

위치는 Quadro의 경우 4개, 8개의 스위치를 사용한 경우와 동일한 성능을 보였다. 상용화된 ATM 스위치에서 채널당 버퍼용량이 수십 개 정도인 점을 고려할 때 제안된 스위치는 7-8개의 스위치소자만으로 이러한 버퍼용량을 구현할 수 있었으며 이 때에 10⁻¹⁰보다 적은 - 실제적으로 완벽한 - 패킷손실율을 얻을 수 있었다.

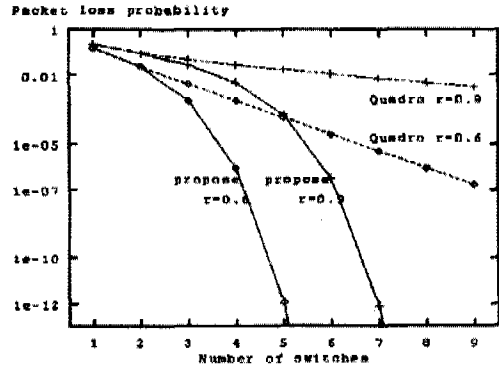


그림 4. 제안된 구조 및 Quadro에 부하(r=0.6, 0.9)의 이산분포가 가해졌을 때 소요된 스위치 수에 따른 패킷손실율 곡선

V. 결론

이 논문에서 제시된 2x1 무블로킹 버퍼링 광스위치는 최소비용으로 구현이 가능할 뿐 아니라 지연시간과 패킷손실을 면에서도 최적성능을 나타낸다. 이처럼 제안된 스위치는 현재 여러 관리측면에서 까다로운 기존의 광소자의 단점을 극소화시킬 수 있으며 뛰어난 버퍼링 능력을 보인다. 실제적으로 7, 8개의 2x2 스위치소자만으로 우수한 2x1 버퍼링 기능을 발휘하였다.

반면 동일한 성능을 기존의 방식을 사용한다면 64-128개의 광스위치소자가 소요되므로 구현비용이 엄청나게 상승할 뿐 아니라, 제작범위를 훨씬 넘어선 복잡한 제어방식, 대단위의 열 발생, 노이즈 등의 문제로 원천적으로 구현이 불가능하였다. 이처럼 제안된 스위치는 상용화의 길을 막고 있던 많은 문제점을 해결하였다. 이 장점으로 인하여 차세대 비동기식 초고속 광네트웍에서 버퍼링을 담당하는 전 광 스위치의 기본구조로 유용하게 쓰일 것이다.

VI. 부록

정리 1 증명

증명은 4단계로 이루어진다. 1-3단계에서는 용량 초과가 발생하지 않은 경우에 블로킹이 발생하지 않음을, 4단계에서는 용량초과시의 블로킹여부를 조사한다. 1-3 단계에서는 SW₁과 SW_k(1≤k≤d), 그리고 SW_{d+1}에서 블로킹이 일어나지 않음을 각각 증명한다. 증명은 블로킹의 정의를 각 단계에 적용하여 얻어진다. 즉 SW_j에 동시에 도착한 두 패킷이 SW_j의 동일한 출력포트로 출력하는 경우가 발생하는가를 점진한다.

(1단계) 두 개의 패킷이 도착하면 이들은 S_i과 S_{i-1}을 R의 값으로 배정 받는다. 두 연속된 정수는 하나는 짝수, 하나는 홀수이므로 홀수 R값을 배정 받은 패킷은 정확한 R값을 채우기 위해서 DL중 유일한 홀수길이를 가진 DL₁을 반드시 경유해야 하므로 SW₁의 상단으로 방출되어야 하며, 동일한 이유로 짝수 R값을 배정 받은 패킷은 DL₁은 절대로 경유해서는 안되므로 SW₁의 하단으로 전달되어야 한다. 이와 같은 동작으로 인해서 SW₁에서는 블로킹이 발생하지 않는다.

(2단계) 두 개의 패킷 P₁, P₂가 t₁, t₂ 시간에 입력되어 R₁, R₂의 값을 배정 받고서 t 시간에 SW₁에서 블로킹을 일으켰다고 가정하자. t₁과 t₂의 차이는 두 패킷이 SW₁에 도착하기까지 DL에서 지연된 시간의 차이인데 한 패킷은 모든 DL을 다 거치고 다른 한 패킷은 하나도 DL을 경유하지 않고 SW₁에 다다를 경우에 제일 큰 차를 보인다. 이를 정리하면 다음과 같다.

$$|t_1 - t_2| \leq \sum_{k=1}^i m_k = \sum_{k=1}^i 2^{k-1} = 2^{i-1} - 1 = m_i - 1 \quad (2)$$

또한 R은 모든 경우 S_i 혹은 S_{i-1}값을 가지며, S_i는 시간슬롯 당 최대 1이상 변화(증가하거나 감소할 수 없다)할 수 없다는 점을 이용하면

$$|R_1 - R_2| \leq |S_1 - S_2| + 1 \leq |t_1 - t_2| + 1 \leq m_i \quad (3)$$

을 얻을 수 있다. 출력단 O₁에서는 매 시간슬롯 당 한 패킷씩 출력되는 제어방식 F의 성질 때문에 i 단계에서 링크를 공유하고 있는 두 패킷 P₁, P₂는 i+1 단계 이후에서 라우팅 링크가 최소 한번 분리되어야 하고 이러한 라우팅 경로의 차이로 인하여 두 패킷의 출력 시점(t₁+R₁, t₂+R₂)이 달라진다. 출력시점의 차이는 경우한 DL의 길이의 차이이며, i 단계 이후 가장 짧은 DL이 DL_{i+1}이므로 최소한 m_{i+1}의 차이를 가진다.

$$\begin{aligned} m_{i+1} &\leq |t_1 + R_1 - (t_2 + R_2)| \\ &= |(t_1 - t_2) + (R_1 - R_2)| \\ &\leq |t_1 - t_2| + |R_1 - R_2| \end{aligned} \quad (4)$$

이를 정리하면

$$\begin{aligned} |R_1 - R_2| &\geq m_{i+1} - |t_1 - t_2| \\ &\geq 2 \cdot m_i - (m_i - 1) = m_i + 1 \end{aligned} \quad (5)$$

식 (3)과 식(5)을 비교해보면 이들은 서로 모순된다. 결국 가정한 블로킹의 발생은 일어나지 않는다.

(3단계) 제어방식 F는 블로킹 없이 한번에 한 패킷씩 출력하도록 되어 있으므로, 마지막 SW_{d+1}에 두 패킷이 입력되었다면 그 중 하나만 출력단 O₁으로 보낼 수 있고 나머지 하나는 별도의 출력단 O₂로 폐기해야 한다. 이는 2×1 버퍼링 스위치의 용량 이상의 패킷이 입력될 경우에만 발생한다. 반대로 스위치 용량이 초과되지 않는다면 SW_{d+1}에는 한 패킷씩만 입력되므로 블로킹이 발생하지 않는다.

(4단계) 용량을 초과한 패킷의 존재 비트는 0으로 세트된다. 일반 패킷과 용량초과 패킷이 동시에 특정 스위치소자 SW_i로 입력되면 스위치는 일반패킷을 기준으로 제어하면 SW_i는 항상 블로킹 없이 제어된다. ■

정리 3 증명

d개의 숫자로 1부터 2^d-1 사이의 모든 숫자를 만들 수 있는 숫자조합은 {1, 2, ..., 2^d-1} 뿐이다. 이 결과를 이용하여 d개의 DL로 블로킹이 발생하지 않으면서 가장 큰 용량 C를 갖게 하려면

$$\{m_1, m_2, \dots, m_d\} = \{1, 2, \dots, 2^{d-1}\}$$

의 관계를 만족시켜야 한다. 아래는 m_k (1≤k≤d)가 d개의 가능한 값 중 어떤 값을 가져야 하나를 블로킹의 발생여부와 관련하여 얻는다.

우선 m₁의 값을 정해보자. 두 패킷이 텅 빈 2×1 스위치로 입력될 때 하나는 R=0을 다른 하나는 R=1을 부여받아야 한다. R=0인 패킷은 아래쪽 링크를 통하여 출력되어야 하므로 위쪽 DL은 R=1인 패킷이 보내져야 하므로 DL₁의 길이는 1이어야 한다.

일반적인 m_k의 값을 수학적 귀납법을 이용하여 풀기 위해 m_i=2ⁱ⁻¹(1≤i≤k)의 값을 가지며 블로킹이 발생하지 않았다고 가정하자. 또한 (i) t=0의 시각에 패킷 P₁이 입력되어 R=2^{k-j}(1≤j≤2^{k-i})의 값을 받았고, (ii) t=2^{k-j}의 시각에 패킷 P₂가 입력되어 R=2^k

를 보유했다고 하자. 이는 2^k 개의 패킷이 2^{k-j} 시간 슬롯 중에 발생하면 가능하다. P_1 은 자신의 R 이 2^{k-1} 을 넘지 못하므로 k 번째 LSB가 0이므로 SW_k 에 도달하면 아래쪽 광링크를 선택하므로 블로킹을 막기 위해서 P_2 는 SW_k 의 상단 DL을 차지해야 한다. 반면 P_2 는 $R=2^k$ 의 값을 지니고 있으므로 $m_{k+1}=2^k$ 의 값을 가져야 한다. ■

정리 4 증명

정리 4에서는 모든 $\{m_i\}$ 에 대해서 블로킹이 발생하지 않는 구조를 찾을 수 있다고 가정한 채 m_i 값을 어떻게 결정해야지 Binomial 분포 하에서 가장 적은 패킷손실율을 가질 수 있는지를 점검한다. $m_k=2^{k-1}$ ($k=1, \dots, d$)의 경우를 경우1이라 칭하면 경우1은 정리 1에서 이미 블로킹이 발생하지 않음을 증명하였다. 한편 블로킹이 발생하지 않는 특수한 $m_i(1 \leq i \leq d)$ 조합이 존재한다고 가정하고 이를 경우 2라 한다면 우리는 경우2가 경우1보다 패킷손실율이 항상 더 커짐을 증명하고자 한다.

용량이 커질수록 버퍼링 할 수 있는 패킷의 수가 커지므로 용량 C 가 2^d 보다 작은 경우에 비해서 $C=2^d$ 일 경우 패킷손실율이 적어진다. 남은 부분은 용량 C 가 2^d 보다 커질 경우라도 패킷발생율이 커지는 현상을 증명하는 것이다. 이 경우에도 d 개의 숫자로 조합할 수 있는 숫자는 0과 $C-1$ 사이의 정수 중 총 2^d 개이다. 이들 숫자를 적은 숫자부터 차례로 나열하여 $q(i)$ ($i=0, 1, 2, \dots, 2^d-1$)라 하자. 우리는 양쪽 경우에 대해서 현재 스위치내의 패킷의 수로 상태변수를 정하여 Markov chain에 의한 확률식을 세울 수 있다. 이를 바탕으로 경우1에서 상태값이 j 가 될 확률을 $Pr(i, j)$ 로 나타내면

$$Pr(1, j) < Pr(2, q(j)) \tag{6}$$

의 관계를 얻을 수 있다. 경우1의 패킷손실율 대 경우2의 패킷손실율의 비는 $Pr(1, 2^d-1) : Pr(2, q(2^d-1))$ 로 나타낼 수 있는데 (6)의 관계에 따라서 경우1은 항상 경우2보다 적은 패킷손실율을 가진다. 자세한 증명은 [6]에 언급되어 있다.

참고 문헌

[1] F. Forghieri, A. Bononi, P. R. Prucnal, "Analysis and Comparison of Hot-Potato and Single-Buffer Deflection Routing in Very High

Bit Rate Optical Mesh Networks," IEEE Trans. on Comm., vol.43, no.1, pp.88-98, Jan. 1995.
 [2] I. Chlamtac, A. Fumagalli, "An Optical Switch Architecture for Manhattan Networks", IEEE, JSAC, vol.11, no.4, pp.550-559, May 1993.
 [3] Z. Haas, "The 'Staggering Switch': An Electronically Controlled Optical Packet Switch", Journal of Lightwave Technology, vol.11, no.5/6, pp.925-936, May/June 1993.
 [4] I. Chlamtac, A. Fumagalli, C.J. Suh, "Multi-Buffer Delay Line Architectures for Efficient Contention Resolution in Optical Switching Nodes", to appear in IEEE Trans. on Comm.
 [5] Hyong S. Kim, "Design and Performance of Multinet Switch : A Multistage STM switch Architecture with Partially Shared Buffers", IEEE/ACM Trans. on Networking, vol.2, no. 6, pp. 571-580, December 1994.
 [6] C. J. Suh, "Switched Delay Lines Using Multiple-length Delay Lines in Optical Network", dissertation, University of Massachusetts, Amherst, February 1996.
 [7] CRO I. Chlamtac, et al., "Contention Resolution by Delay Lines", IEEE JSAC, vol 14, No. 5, pp 1014-1029, June 1996.
 [8] Special Topic on "High Speed Point-to-point Optical Comm. System", IEEE JSAC vol 14, No. 5,

서 창 진(Chang-jin Suh)

정회원



1959년 5월 15일생
 1982년 : 서울대학교 제어계측 공학과(학사)
 1984년 : 서울대학교 제어계측 공학과(석사)
 1985년~1990년 : 한국전자통신 연구소

1996년 : Univ. of Massachusetts at Amherst, Dep. of Electric and Computer Eng. (Ph.D)
 1997년 이후 : 숭실대학교 컴퓨터공학과 교수
 <주관심분야> 스위치 이론(광스위치, ATM 스위치), WDM, 무선통신의 mac성능분석