

초고속 기가비트급 광 IP 패킷의 포워딩을 위한 새로운 Lookup 장치의 구현

이정준, 홍준혁, *강승민, *송재원
경북대학교 정보통신학과, *경북대학교 전자공학과
전화 : 053-940-8632 / 핸드폰 : 011-815-8952

Implementation of A Multigigabit Lookup Scheme for Optical IP Packet Forwarding

Jung Jun Lee, Jun Hyuk Hong, *Sung Min Kang, *Jae Won Song
Dept. of Information & Communication, Kyungpook National University
*Dept. of Electronic Engineering, Kyungpook National University
E-Mail : lee2j@hanmail.net

Abstract

This paper reports a very fast lookup scheme for Optical IP Packet forwarding. A LD by driven Pattern Generator generate a optical IP Packet encapsulated by any header of level1 and level2. A high speed Lookup scheme for a forwarding has been implemented by EEPLD with tiny SRAMs for optical internetworking. With SRAM of a 10ns access time and ~400kB, the Lookup scheme has achieved very high speed lookup time about 100ns for 2 memory accesses.

I. 서론

데이터 트래픽의 수요는 끝없이 폭발적으로 증가하고 있다. 최근 몇 년간 웹 중심으로 성장하는 경향에 의하여 IP/ATM/Frame Relay를 통한

음성 정보와 디지털 가입자 회선 및 케이블을 통한 비디오, 고속 가입자 액세스에 대한 수요가 계속 늘어날 전망이다. 또한 전자 상거래에 인터넷과 엑스트라넷의 채택은 대역폭 수요와 기능성의 요구라는 두가지 측면에서 IP 서비스 인프라에 커다란 변화를 가져올 것이다. 이러한 사회적 요구에 부합하기 위해서는 광인터넷워킹 기술이 필수적이다.

광 인터넷워킹에서 고성능 인터넷워킹 장치(스위치와 라우터)는 광 네트워킹 기술을 통해서도 연결된다. 이러한 장치들은 광섬유와 직접 연결될 수도 있고, 인터넷워킹 장치와 SONET/SDH 네트워크 요소 등의 다양한 클라이언트에게 파장 라우팅을 제공하는 광 네트워크 레이어에 연결될 수도 있다. 어떤 경우이든 스위치나 라우터는 통계적으로 다중화된 패킷이나 셀에 의해 섬유나 파장에서 제공되는 최대용량을 사용하여 데이터 용량을 효율적으로 사용할 수

있다.

이러한 연구는 최근 WDM 장비와 직접 인터페이스하기 위한 고속 스위치 및 라우터 개발을 가속화시키고 있으며, 기존 SONET/SDH 백본 시스템의 문제점을 해결할 수 있을 것으로 예상된다. 그러나 실제로 WDM에 맞는 광 인터페이스를 위한 고속 스위치 및 라우터 구현을 위해서는 고속 포워딩 엔진 개발이 필수적이다. 따라서 본 연구는 메모리 용량을 최소화하고, WDM 시스템 기반 테라비트 IP 라우터용 Lookup 알고리즘을 개발하여, 고속 포워딩 엔진을 하드웨어로 구현하고 이를 이용하여 광 IP 패킷을 라우팅하는 장치를 구현 하였다.

II. 제안된 Lookup 알고리즘

“직접방식”과 “간접방식”은 메모리의 크기가 너무 크거나 고속의 메모리인 SRAM에 적용할 수 없다는 것을 알 수 있다. 이러한 문제점은 본 논문이 제안한 새로운 lookup 알고리즘에 의하여 해결될 수 있으며 그림 1은 위에 제안된 Lookup 기법의 개념도이다.

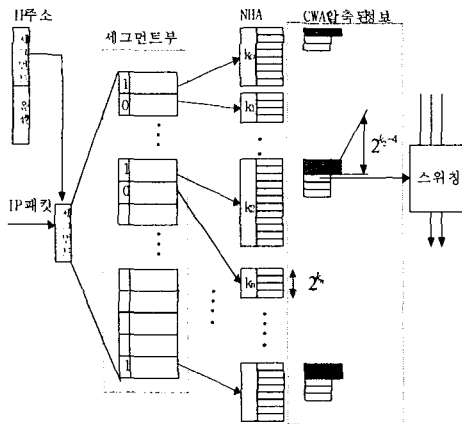


그림 1. 제안된 Lookup 기법의 개념도

첫째, 세그먼트 표에 가변 유효 필드를 두고, 이를 근거로 세그먼트별로 NHA(Next Hop Array)의 크기를 가변하여 메모리 크기를 줄이고, 둘째,

프리픽스 길이가 긴 경우 압축 알고리즘을 수용하여 메모리 크기를 최소화하고자 하였다. 셋째, 압축 기법을 적용함으로써 새로 가공된 정보가 저장되는 CWA(Code Word Array)/CNHA(Compressed Next Hop Array)에서, CWA내의 Map(NH의 8bits 정보를 1bit로 나타내어 CW당 16bits씩 배열)이 Hx(0000)인 경우 Base(이전 Map까지의 “1”의 개수를 저장)자리에 다음홉의 출력 포트값을 담으므로 2회의 메모리 접근비용을 극대화하였다. 이는 10ns의 SRAM과 Altera사의 EEPLD 기반의 EPM7192시리즈에 50MHz 클럭을 적용할 경우를 대비하여 VHDL로 모델링하고 하드웨어로 구현하였다.

III. Lookup장치의 구현 및 실험

그림 2는 구현된 Lookup 장치의 외형 사진이다. Altera사의 EEPLD EPM7192 5개와 주변 회로로 구현되었다. 디버깅을 위하여 DIP 타입 IC를 사용하였다. 메모리는 고속의 SRAM 400KB를 사용하였다.

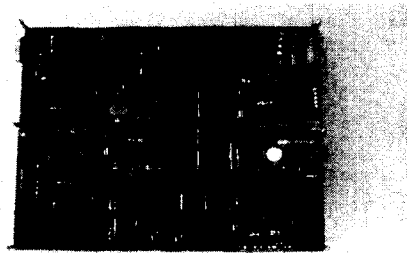


그림 2. Lookup 장치

IPMA 사이트에서 구한 라우팅 정보를 이용하여 Lookup 실험을 하였다. 패킷 발생기에서 구현된 IP 패킷은 두 개의 파장에 실려 커플러에 전달된다. 두 개의 파장에 독립적으로 변조된 신호는 광섬유를 통해 전달되고 이는 WDM 소자에서 분리된다. 각각 분리된 광 IP 패킷은 각각의 Lookup 장치에 전달되어 포워딩 과정(그림 3)이

수행된다.

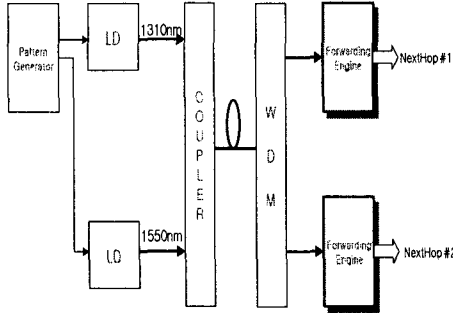


그림 3. 광패킷 포워딩 장치

각 포워딩 장치는 독립적으로 3개의 포트가 있으며 IP의 원격지 주소에 의해 해당 포트에 라우팅된다. 해당 Lookup 장치에서의 포워딩 과정을 살펴보면 다음과 같다.

프레임형식으로 전송되어 온 데이터 중에서 광 IP packet를 걸러 낸 후 IP 헤더부의 5번째 Word(32bit)를 목적지 주소로 인식을 하여야 한다. 그리고 상위 16bit는 세그먼트표를 참조하기 위한 인덱스로 사용하여야 한다. 이는 IP 헤더 분석부에서 수행한다.

이에 병행해서 세그먼트표의 최상위 1bit의 값을 읽고 동작하는 엔진 Mode를 결정하고, 최하위 4bit(k-1의 값)와 중간 27bits의 값(NH 또는 포인터)을 읽고 추가의 메모리 접근을 결정하게 된다. 여기서 세그먼트표의 27bit 값이 255이하이면 이 값이 직접적인 NH(Next Hop)의 출력포트 번호가 된다. 실제 출력 포트는 3개를 두고 NH로 가는 출력 포트 번호가 0~51까지는 포트 1로, 52~127까지는 포트 2, 128 ~ 255까지는 포트 3으로 스위칭되도록 하였다. 세그먼트표의 어떤 엔트리의 27bit 값이 256이상이고 압축여부부(d)의 값이 "0"이면 두 번째의 메모리 접근(NHA)이 요구되며, 세그먼트표의 어떤 엔트리의 최초 1bit의 값이 "1"이고 중간 27bits의 값이 256이상이면 역시 두 번째의 메모리 접근(CWA)이 요구된다. 또한 세그먼트표의 어떤 엔트리의 최초 1bit의 값이 "1"이고 중간 27bits의 값이 256이상이면 두 번째

의 메모리 접근(CWA)이 요구되고 해당 CWA의 CW(Code Word)의 Map의 값이 "0"이 아니면 3 번째의 메모리 접근이 필요하다. 예를 들어 IP 주소가 128.11.208.0인 패킷이 들어온다고 할 때, 주소 상위 16bits인 128.11로 세그먼트표의 해당 엔트리를 찾고 엔트리의 최초 1bit의 값 "1"과 중간 27bits의 값 "256"을 가지고 해당 CWA를 결정한다. 그리고 나서 윗셀길이(k값) "8"를 가지고 IP 주소의 하위 16bits 중 최초 4bits의 값을 가지고 CWA에서 해당 엔트리를 결정한다. CWA의 해당 엔트리의 Map(상위 16bits)값이 "H_x(6000)"이고 Base(하위 16bits)의 값 "H_x(0008)"이므로 이 두 값을 가지고 CNHA의 해당 엔트리 출력 포트가 된다.

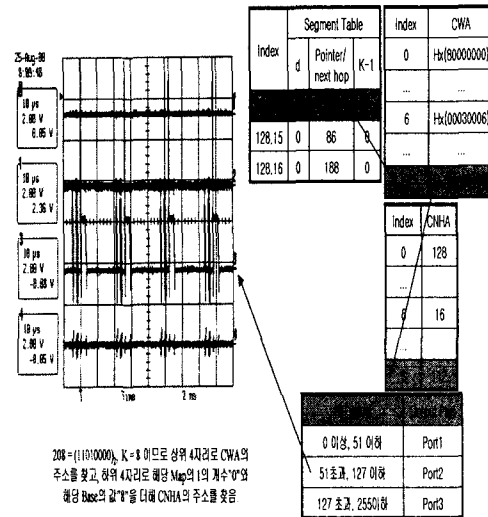
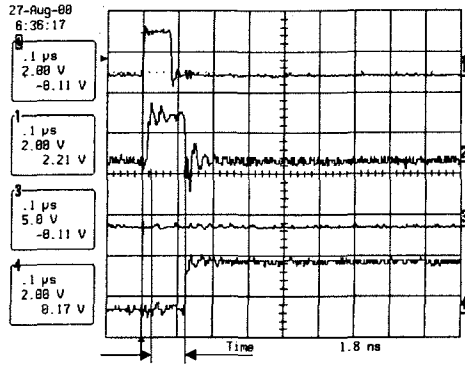


그림 4 IP 주소가 128.11.208.0인 경우

실제 출력은 포트 2로 결정된다. 3번의 메모리 접근으로 CNHA를 이용하여 출력 포트를 결정하는 측정 결과를 그림 4에 나타냈다. 측정 데이터를 보면 NH의 값이 112로서 포트 2번으로 스위칭된다는 것을 알 수 있다.

표 1에서 보는 바와 같이 IPMA 웹사이트에서 얻은 실시간 라우팅표를 본 기법에 적용한 결과 ~316KB의 메모리가 소요됨을 확인하였다. 또한 이들의 메모리 접근 회수는 2회 이내가 95%이었던

으며, 이 때 최대 100ns의 Lookup 시간이 소요 (그림5)되었다.



Port 1 → Port 3 Switching Time : 100ns

그림 5. 스위칭 시간 측정-메모리 2회 접근(NHA data = NH Port인 경우)

이는 10Mpps의 처리 속도로서 10Gbps의 신호 처리 능력을 의미하는 것이다.

표 1 여러 Lookup 기법들의 비교.

Lookup 기법	메모리 접근수	비율	스위칭 속도	포워딩 표 크기	SRAM 사용
본 기법	1	15%	50ns	~316KB	Yes
	2	80%	100ns		
	3	5%	250ns		
DIR-24-8	1/2	50%		33MB	No
DIR-21-3-8	1/3	-		9MB	No
SFT	2/9	-		150 ~ 160KB	Yes

IV 결론

본 논문은 하드웨어로 구현가능하고 고속의 Lookup이 가능한 새로운 알고리즘을 제안하고, 이를 VHDL로 모델링하여 성능을 미리 검증한 뒤 하드웨어로 실제 구현하고, 이를 원시적인 형태의 광 IP패킷을 발생시켜 주소에 따라 해당 출력 포트로 포워딩됨과 고속의 스위칭이 가능함을 확인하였다.

향 후 더욱 빠른 ASIC과 5ns이하의 초고속

SRAM을 사용할 경우 테라비트급 스위치, 라우터 구현은 물론, 원격지 주소를 파장과 직접적으로 연계한 광라우팅 시스템의 기초를 마련 할 것으로 기대된다.

참고문헌

- [1] W. Doeringer, G. Karjoth, and M. Nassehi, "Routing on longest-matching prefixes," IEEE/ACM Trans. Networking, Vol. 4, pp. 86-97
- [2] M. Doegermark, A. Brodnik, "Small forwarding tables for fast routing lookups," in Proc. ACM SIGCOMM'97, France.
- [3] Nen-Fu Huang, Shi-Ming Zhao, "A Novel IP-Routing Lookup Scheme and Hardware Architecture for Multigigabit Switching Routers," IEEE Journal of Selected Areas in Communications, Vol. 17, No. 6, June 1999, 1093-1104.
- [4] Michigan University and Merit Network, Internet Performance Measurement and Analysis(IPMA) Project [Online], Available <http://nic.merit.edu/~ipma/>.