

테라비트 라우터 기술 동향

전 종 암, 변 성 혁, 안 병 준, 이 형 호

한국전자통신연구원 네트워크연구소 라우터기술연구부

요 약

테라비트 라우터는 세계 유수의 라우터 개발 업체들이 경쟁적으로 연구 개발을 진행 중에 있으며 수 Tbps급의 라우터가 시장에 출시될 것으로 예측되고 있다. 본 고에서는 테라비트 라우터의 출현 배경과 요구 사항 및 핵심 기술에 대하여 분석하고, 상용 테라비트 라우터 시스템에 대한 개발동향 등에 대해서 살펴 보기로 한다.

I. 서 론

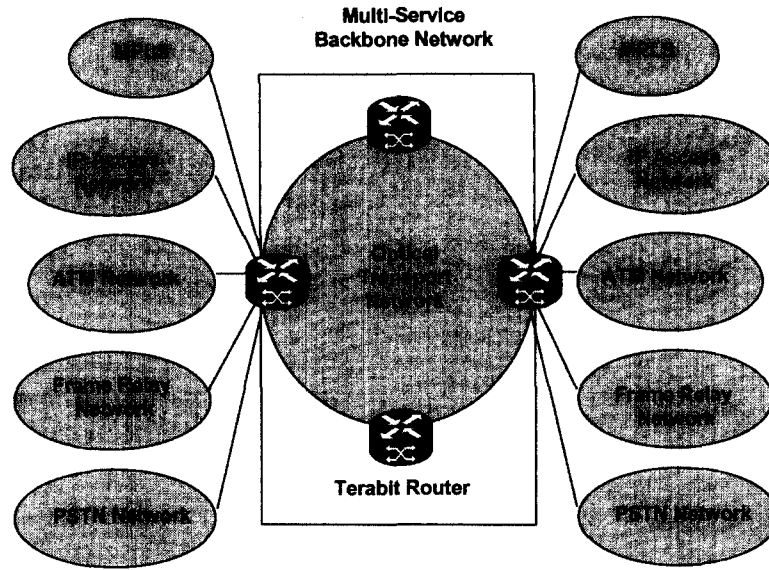
최근까지 이어온 인터넷 사용자의 폭발적인 성장으로 말미암아 인터넷 서비스 제공자는 기존 인터넷 백본이 가지고 있는 서비스 품질의 저하를 포함한 여러 문제점에 직면하게 되었다. 인터넷 서비스 제공자는 이러한 당면 문제의 해결 뿐 아니라 향후 예측되는 인터넷 서비스의 효율적인 수용을 위하여 새로운 방안을 모색하게 되었다. 최근 활발한 연구개발이 이루어지고 있는 테라비트 라우터는 이와 같은 요구 사항을 바탕으로 등장하게 되었지만 테라비트 라우터의 응용은 기존 인터넷 서비스를 포함한 기존 통신망 서비스 전반을 포함하는 멀티서비스 형태로서 그 응용영역을 확장하고 있는 추세이다.

현재 통신망 서비스의 제공은 응용 서비스의 종류에 따라서 서로 다른 네트워크를 통하여 제공되고 있으나, 이러한 방법은 통신망 사업자의

입장에서 중복적인 투자로 인한 경제적인 손실 뿐만 아니라 이를 운용, 관리하기 위한 운용 및 유지보수 비용 등 제반적인 비용의 증가를 초래하게 된다. 이러한 여러 가지 문제점을 해결하기 위한 방안으로 테라비트 라우터를 이용하여 멀티 서비스 백본 망을 구축한다면 기존 서비스의 수용 뿐만 아니라 향후 새로운 서비스를 효율적으로 수용할 수 있을 것으로 예측되고 있다.

이와 같은 구성의 네트워크는 여러 가지 장점들을 가지고 있다. 먼저 서비스 제공자 측면에서 살펴보면 노드 수 감소에 따른 장비투자비용 절감 및 네트워크 운용, 관리 비용의 절감효과를 기대할 수 있으며 불필요한 노드간 트렁크의 감소에 따른 장비의 효율적 운용 및 다양한 서비스 등급(서비스 제공 품질에 따라 서비스 등급을 구분하고 이에 맞게 요금체계를 구성하여 서비스를 제공하는 방법) 개발을 통한 매출 증대 효과를 기대할 수 있을 것이다. 사용자의 측면에서 살펴보면 보다 빠른 인터넷 접속이 가능하고 업무의 특성에 따라 다양한 서비스 등급을 적용 함으로써 비용을 절감하는 효과를 기대할 수 있다^[1]. <그림 1>은 테라비트 라우터를 이용한 멀티 서비스 백본 망의 구성 예를 나타낸 것이다.

테라비트 라우터는 현재 Cisco, Juniper, Avici, Lucent, Pluris 등의 라우터 개발 업체들이 경쟁적으로 연구개발을 이끌고 있으며 조만간 수 Tbps급의 라우터가 시장에 출시될 것으로 예측되고 있다. 이와 함께 유수의 테라비트 라우터용 칩 셋 회사들의 경쟁도 치열하게 이루어지고 있으며 테라비트 라우터용 주요 핵심 칩셋들의 출시를 앞 다투어 선전하고 있다. 이와 함께 학계에



〈그림 1〉 테라비트 라우터의 주요 응용분야

서도 테라비트 라우터 관련 주요 핵심 기술에 대한 연구가 활발히 이루어지고 있다.

본 고에서는 테라비트 라우터의 출현 배경과 요구 사항 및 핵심 기술에 대하여 분석하고, 상용 테라비트 라우터 시스템에 대한 개발동향 등에 대해서 살펴 보기로 한다. 이러한 내용을 다루기 위하여 본 고의 2장에서 테라비트 라우터가 멀티 서비스 백본 망의 장비로서 갖추어야 할 요구사항을 언급하고 3장에서는 이러한 요구사항을 만족하기 위하여 중점 개발되고 있는 핵심 기술 분야에 대하여 분석한 후 4장에서는 현재 출시되고 있는 상용 고속 라우터 시스템의 개발 동향에 대하여 언급한 후 5장에서 결론을 맺는다.

II. 테라비트 라우터 요구사항

테라비트 라우터가 백본 망의 장비로서 갖추어야 할 요구 사항을 살펴보면 다음과 같다.

1. 테라비트 확장성 기능

테라비트 라우터를 단일 새시로 구성하는 것보

다 여러 개의 새시로 구성하는 방법이 효율적이다. 이를 위하여 테라비트 라우터의 단일 새시 당 패킷 처리용량은 320 Gbps에서 640 Gbps 정도를 지원할 수 있어야 하며 이러한 새시가 다중으로 연결되어 테라비트 라우터를 구성할 수 있는 확장성을 제공할 수 있어야 한다. 테라비트 라우터의 스위치 포트 속도는 2.5 Gbps에서 40 Gbps 급까지 확장성 있게 지원함으로써 향후 개발되어 질 것으로 예측되는 40 Gbps급의 선로 속도를 수용할 수 있어야 한다.

2. 고속 패킷 처리기능

폭발적으로 늘어나는 인터넷 트래픽을 백본에서 효율적으로 수용하기 위해서는 2.5 Gbps에서 40 Gbps급의 광 선로 접속 기능과 고속 선로를 통하여 수신된 패킷을 wire-speed로 처리할 수 있는 고속 IP 검색 및 트래픽 관리 기능 등을 포함한 고속 패킷 처리 기능이 제공되어야 한다.

3. QoS 보장형 패킷 처리 기능

인터넷 검색 등과 같은 기존의 단순한 데이터 전달 서비스 이외에 인터넷 화상회의, 인터넷 방송 등 실시간 처리가 요구되는 응용에 대한 서비

스 요구가 급증하고 있으며, 기존 인터넷 서비스 및 공중망 서비스를 포함한 멀티서비스의 수용을 위해서는 서비스별 다양한 QoS를 보장할 수 있는 패킷 처리 기능을 제공할 수 있어야 한다.

4. 시스템 이중화 기능

인터넷 백본 장비로서 테라비트 라우터의 안정성에 대한 중요도는 기존의 공중망 장비인 교환기에 버금간다고 볼 수 있다. 이와 같은 시스템의 안정도를 높이기 위하여 테라 라우터 중요 기능에 대한 이중화는 완벽하게 지원되어야 한다. 기존 고속 라우터의 경우 백본에서 중요한 위치에 사용되는 장비는 시스템 단위의 이중화 즉 동일한 두 대의 장비를 설치하여 그 기능을 수행하지만, 고가 장비인 테라비트 라우터의 경우 그러한 방법은 매우 비 현실적이라고 볼 수 있다. 따라서 시스템 내의 주요 기능 요소인 스위치 패브릭, 전원, 팬에 대한 이중화를 포함하여 라우팅 프로토콜이 수행되는 라우팅 프로세서에 대한 이중화를 완벽하게 지원 하여야 한다.

5. DWDM 연동 기능

광 전송 망의 효율적인 활용을 위하여 DWDM (Dense Wavelength Division Multiplexing) 기술이 제안되었으며, 이러한 DWDM 기술은 광 전송 망의 핵심 기술중의 하나이다. 테라비트 라우터의 경우 광 전송 망에 접속되어 활용될 가능성이 매우 높으므로 DWDM 장비와의 연동기능을 제공할 수 있어야 한다.

6. 멀티서비스 제공 기능

최근 각광을 받고있는 MPLS (Multi-Protocol Label Switching), IPv6 (Internet Protocol Version 6) 등과 같이 현재의 통신시장은 하루가 다르게 변모해가고 있으며 이러한 요구 사항을 효율적으로 수용하기 위해서는 기존 통신 서비스를 효율적으로 수용할 수 있을 뿐만 아니라 신속하면서 저 비용으로 신규 서비스를 수용할 수 있는 기능을 제공할 수 있어야 한다.

III. 테라비트 라우터 핵심 기술 동향

테라비트 라우터의 요구 사항들을 만족시키기 위한 핵심 기술 분야로서 스위칭 능력의 고속화를 위한 기술들, IP 패킷 처리의 고속화를 위한 기술들, 인터넷 백본에서 QoS를 보장하기 위한 기술 등에 대하여 소개한다.

1. 스위칭 능력의 고속화

라우터 시스템 내부에서의 패킷 데이터 전송방식으로서, 초기 라우터에서는 시스템 버스를 사용하였으나 데이터 교환의 고속화를 위하여 스위치를 사용하게 되었다. 따라서 테라비트 라우터 시스템에서는 스위치 기술이 가장 주된 연구대상 기술중의 하나이다. 테라비트 라우터를 위한 스위치 설계 기술은 스위칭 단위 데이터 길이의 고정성 (fixed-length) / 가변성 (variable-length) 과 버퍼의 위치 등이 주된 고려 대상이다^[2].

고정 길이 스위칭 단위는 주로 ATM 셀 스위칭 방식에서 많이 연구되어 왔으며, 라우터의 고속화를 위해 가변길이를 갖는 패킷을 고정길이 단위로 나눔으로써 스위칭 기능의 고속화를 실현할 수 있다. 버퍼의 위치에 따른 스위치의 구조는 입력버퍼 스위치, 출력버퍼 스위치, 그리고 공유버퍼 스위치로 크게 구분된다. 테라비트 라우터에서는 링크 속도와 총 스위치 용량이 높아짐에 따라, 내부 속도 증가가 요구되는 공유 버퍼형이나 출력 버퍼형 보다는 입력 버퍼형 스위치의 확장성과 이점들을 유지하면서 HOL (Head of Line) 블로킹 문제를 극복한 VOQ (Virtual Output Queuing) 구조 등을 도입하고 있다. 최근 개발되고 있는 대부분의 고속 스위치 칩의 경우 이러한 형태로 개발되고 있다.

VOQ 구조가 스위치의 스루풋의 제한은 없었다고 하나, QoS 측면에서 출력 버퍼 스위치의 성능에는 미치지 못한다. 스위치에서 연결별 QoS 보장을 위해 주목 받고 있는 fair queuing 알고리즘은 출력 버퍼 스위치를 기본 가정으로 하고 있기 때문에, 입력 버퍼 스위치에서 fair queuing

〈표 1〉 고속 스위치 칩 셋 특징

회사 명	PMC-Sierra	IBM	AGERE(구 Lucent)	MindSpeed
칩셋 명	TTX	PRS-EQ	PI-40 (PI40X memory switch PI40C Crossbar switch)	CX27300 (CX27301 Queue Manager CX27302 Crossbar SW)
스위치 구조	VOQ with crossbar	Shared memory+ VOQ	3단(Memory-Space- Memory)	VOQ with crossbar
용 량 (full duplex)	40G~2.5T	40G~320G	40G~2.5T	40G~320G
PHY 인터페이스	LCS2	2.5G UniLink	Propriety Serial Link	SkyRail Link Protocol
스위칭 포트	Max 256 ports @ 10G	64 ports @ 1G, 10G ports with link bundling	Max 1024 ports @ 2.5G Port Grouping : 1, 2, 4, 8, 32 link grouping 지원	Max 256 ports @ 2.5G Port grouping : 1×OC192 or 4×OC48

알고리즘을 그대로 적용하기 곤란하기 때문이다. 따라서 입력 버퍼 스위치로서 출력 버퍼 스위치의 성능을 얻을 수 있는 방법에 대한 연구가 최근에 많이 이루어지고 있다^[3]. 일반적인 스위치의 구조는 speedup이 있는 CIOQ(Combined Input/Output Queueing) 구조에서 VOQ를 채택하는 것을 기본 가정으로 한다. CIOQ 스위치는 입력 버퍼와 출력 버퍼를 모두 가지는데, 출력 버퍼는 성능 향상을 위한 내부 speedup을 하기 때문에 장착되었다. Chuang 등은 CIOQ 스위치에서 2배의 speedup만으로 출력버퍼 스위치를 완전히 에뮬레이션 할 수 있음을 이론적으로 보였다^[4].

최근 스위치 칩 셋과 네트워크 프로세서 사이의 인터페이스와 관련하여 표준화가 진행되고 있다. CSIX-L1(Common Switch Interface Specification Level-1)라 불리는 이 표준 규격이 2000년 8월에 발표되었고, 현재 CSIX-L1를 지원하는 2.5 Gbps급 네트워크 프로세서가 출시되고 있다. 올해 안에 스위치 한 포트 OC-48(2.5 Gbps)을 수용하는 스위치가 상용화 될 것으로 예상된다. CSIX-L1의 경우 OC-48 급에 가장 적합한 인터페이스 규격을 제공하며 OC-192급의 인터페이스에도 적용이 가능하다. 그러나 CSIX-L1의 경우 최대 32 Gbps까지의

동작을 지원하므로 40 Gbps급 인터페이스인 OC-768의 경우 적용이 불가능하다. 따라서 이에 대한 새로운 CSIX 규격이 향후 정의되어야 할 것으로 예측된다.

향후 출시 예정인 테라 라우터 스위치 칩 셋의 경우 단위 스위칭 능력이 40G급 이상이며 최대 스위칭 능력은 320 Gbps에서 2.5 Tbps급 까지 확장할 수 있는 구조로 개발되고 있는 추세이다. 다음 〈표 1〉은 향후 출시 예정인 고속 스위치 칩 셋의 주요 특징을 나타낸 것이다.

2. IP 패킷 처리의 고속화

IP 패킷 처리의 고속화를 위하여 IP 패킷의 포워딩 엔진을 라인카드에 분산시키는 구조를 기본으로 한다. 라인 카드별로 포워딩 엔진을 분산함에 있어서 라인카드 링크 속도가 OC-48 내지 OC-192까지 올라가고 있기 때문에 고속의 IP 패킷 포워딩 엔진의 설계가 핵심기술로서 대두하게 되었다. 고속 IP 패킷 포워딩 엔진에서 활발한 연구가 이루어진 분야중의 하나인 IP 주소 룩업의 경우 IP 패킷 포워딩 엔진의 성능을 제한하는 요소가 아니라고 말할 수 있으며, 현재의 ASIC 기술로 10 Gb/s까지 라인 속도의 IP 주소 룩업을 할 수 있을 것으로 예측하고 있다^[5]. 일 예로서, Alliance사에서는 66 Mpps(Packet

Per Second)의 성능을 얻을 수 있으며 엔트리가 64K인 IP 포워딩 프로세서를 발표한 바 있다. MPLS 망 내부에서는 IP 주소 룩업이 필요 없지만 LER(Label Edge Router)에서는 수행해야 하며, LER의 입력 링크 속도는 망의 코어에 비해 낮을 것이므로 IP 주소 룩업은 더 수월한 문제이다.

IP 포워딩 엔진에서 구현이 어려운 것은 새롭게 요구되고 있는 패킷 구분(packet classification) 기능 혹은 패킷 필터링(packet filtering)이다. 패킷 구분은 원래 방화벽(firewall) 같은 보안을 위해서 필요했지만, 최근에는 IETF(Internet Engineering Task Force)의 Differentiated Service(DiffServ)를 위해서, 그리고 RSVP(Resource ReSerVation setup Protocol)나 MPLS 등과 같은 패킷 흐름별 서비스를 제공하기 위해 필요하게 되었다. DiffServ를 위해서는 패킷 흐름을 구분할 필요는 없지만 패킷이 누구로부터 오는 것이며 어떤 클래스에 속하는 것인지를 망의 에지(Edge)에서 구분하고 policing까지 해야 하며, 망의 코어에서는 TOS(Type Of Service) 값에 따라 패킷을 구분하여 차별화된 서비스를 제공해야 한다.

패킷 구분을 위해서는 목적지 IP 주소 뿐만 아니라, source IP 주소, source/destination port number, protocol field 등까지도 볼 필요가 있는데, 이를 미리 설정된 다양한 패킷 필터링 규칙들과 비교해야 하는 작업을 매 패킷마다 해야 되기 때문에 IP 주소 룩업보다 훨씬 어려운 작업이다. 패킷 구분을 고속화하기 위해서는 무엇이 최적의 패킷 구분자(packet classifier)인지에 대한 연구와 함께, 설정된 수백, 수천개의 패킷 구분 규칙에 따라 입력 패킷의 헤더에 있는 패킷 구분자 항목을 고속으로 검색하는 방법이 요구되며, 이러한 분야에 대한 연구는 아직 미진한 상태이다.

3. QoS 보장형 서비스의 지원

인터넷 전화나 방송, VPN(Virtual Private Network) 등과 같은 실시간 혹은 대역폭 요구

서비스들이 늘어남에 따라 best-effort 서비스에 의존하고 있는 인터넷에 QoS(Quality of Service) 기술이 주요 과제로 등장하게 되었다. 이와 같은 QoS 보장 문제는 최근 DWDM(Dense Wavelength Division Multiplexing)과 같은 광 전송망 기술에 의해 충분한 전송 대역을 제공함으로써 다소 완화될 수는 있지만 궁극적으로 QoS 문제는 여전히 남아있다는 것이 일반적인 견해이다.

인터넷에서 QoS 보장방법에는 섬세한(fine-grained) QoS를 보장할 수 있는 RSVP와 거친(coarse-grained) QoS를 보장하는 DiffServ가 있다. DiffServ는 흐름별로 트래픽을 관리하지 않고, 패킷을 클래스별로 분류해서 처리하기 때문에 RSVP와 같은 확장성의 문제가 없다. 순수 IP 라우터 기반의 백본망에서는 RSVP와 같은 자원예약(resource reservation) 프로토콜 없이는 섬세한 QoS를 보장할 수 없는데, RSVP는 확장성의 문제로 망의 코어에 사용되기 힘들기 때문에, DiffServ 기반의 거친 QoS를 보장해야 할 것이다. 음성 트래픽 같은 경우는 DiffServ에서 높은 우선순위의 클래스를 할당하는 것으로 충분히 QoS를 보장받을 수 있을 것으로 보이는 데, 이는 차세대 인터넷이 음성 트래픽을 수용한다고 하더라도 절대적인 양에서 데이터 트래픽보다는 훨씬 적을 것이라고 예상되기 때문이다.

이러한 QoS를 지원하기 위해서 패킷 구분, 버퍼 관리, 출력 링크에서의 패킷 스케줄링 등이 라우터에 포함되어야 할 기능이다. 패킷 구분은 서로 다른 요구조건을 갖는 패킷들을 클래스 별 혹은 흐름 별로 구분하는 기능으로서, MPLS나 DiffServ에서는 망의 입구에서 수행하여야 한다. 그리고 클래스별 혹은 흐름별 큐잉(per-flow queueing)을 통해 다른 클래스의 트래픽에 영향 받지 않고 CoS(Class of Service)나 QoS를 보장할 수 있어야 한다. 일반적으로 흐름별 큐잉이 섬세한 QoS 제어를 보장하므로, 흐름이 구분되는 MPLS 라우터에서는 흐름별 큐잉을 할 필요가 있으며, DiffServ 라우터는 흐름의 구분이

없으므로 클래스별 큐잉을 해야 한다.

버퍼 관리에서 폭주시 패킷 폐기 방법으로는 RED(Random Early Detection)가 주로 사용된다. 버퍼가 넘칠 때 패킷을 폐기하면 해당 라우터를 거치는 대부분의 TCP(Transport Control Protocol) 연결에서 폭주를 인식하고 TCP의 backoff 알고리즘에 의해 윈도우 크기를 동시에 급격히 줄이는 등 TCP 연결간의 동기화 문제가 발생한다. 이러한 TCP 동기화 문제를 해결하기 위해 코어 라우터는 RED로 폭주가 되기 전에 랜덤하게 패킷을 제거하여 TCP backoff 시점을 분산시키면 TCP 연결에 대해 큰 효과를 볼 수 있다. RED에 의한 패킷 폐기 시에는 큐 길이와 패킷의 계약 준수 여부, 우선 순위 등을 고려해야 한다.

저장된 패킷의 전송 시에는 각 큐별로 정해져 있는 대역과 지연 요구조건을 만족시킬 수 있도록 적절한 스케줄링이 필요하다. 스케줄링 방법으로는 WFQ(Weighted Fair Queuing)가 가장 좋다고 얘기되나 구현의 복잡도 등을 고려할 때, WRR(Weighted Round Robin)이나 DR(Deficit Round Robin) 등을 많이 사용한다.

테라비트 라우터는 주로 코어 백본 시스템에 적용될 것이다. 코어 백본 라우터는 단순한 트래픽 처리 능력의 증가만을 의미하는 것이 아니며, 여러 인터넷 서비스 제공자들이 백본을 통하여 서로 관계되어지기 때문에 효과적인 서비스를 위하여 에지 디바이스에서 트래픽을 분류하고 코어 라우터는 서비스 요구사항이 만족되도록 서로 전달해 줄 수 있는 역할을 하여야 한다. 따라서 코어 라우터도 에지에서 실행된 QoS 분류 정보를 알고 있어야 하고, 코어 라우터에서도 주어진 QoS 분류 정보를 기반으로 트래픽의 흐름이 제어 및 관리되어야 한다. VPN 서비스는 에지에서 결정된 QoS 체계가 코어에서도 여전히 보장되어야 하는 한 예가 된다⁶⁾.

4. 네트워크 프로세서 기술

상기 IP 패킷 처리의 고속화 및 QoS 보장형 서비스의 지원을 위하여 최근 개발이 활발히 이

루어지고 있는 분야가 네트워크 프로세서 기술이다. 테라비트 라우터의 구조는 OC-192(10Gbps)나 OC-768(40Gbps)를 수용하는 라인카드에 수백 기가~수 테라급의 스위치가 연결된 형태를 갖는다. 이러한 링크 속도에서의 패킷 포워딩은 각 라인카드 내에서 각각 독립적으로 처리되어야 하며, 이를 위해서는 고속의 프로세서가 장착되어야 한다. 또한, 위에서 살펴보았듯이, 테라비트 라우터의 구현에는 고속의 IP처리 및 QoS의 보장 등, 고속이면서도 유연한 패킷 처리가 가능해야 한다. 단순히 속도만이 중요하다면 전용 ASIC을 사용한 구현이 가장 효과적이다. 하지만 이 경우 유연성(flexibility)이 떨어져서 스케줄링 방식을 바꾼다든지 등의 새로운 QoS의 요구에 대처하기가 힘들다. 반면에 일반 프로세서를 통해 구현한다면, 유연하고 다양한 처리가 가능하나 속도의 한계를 갖는다.

이러한 단점들을 극복하는 새로운 대안이 네트워크 프로세서이다. 네트워크 프로세서는 하나의 범주에 넣기에는 너무 다양한 형태를 갖지만, 몇 가지 특징을 요약하면 다음과 같다. 첫째, 네트워크 프로토콜 처리에 적합한 축소된 명령어를 처리하는 코어가 있다. 통상 RISC 코어를 사용하지만, 컴팩트한 명령어 세트로 고속처리가 가능하다. 또한 병렬처리를 위해 4개 이상의 코어를 집적하는 추세이다. 둘째, 네트워크 데이터 처리에 적합한 입출력 구조를 갖는다. 물리층 디바이스와 직접 연결이 되거나, 네트워크 칩 벤더들이 사용하는 버스 구조의 입출력 인터페이스를 장착함으로써, 네트워크 데이터가 바로 네트워크 프로세서에서 처리되어 진행될 수 있다. 셋째로 넓은 대역폭을 갖는 내부 메모리를 들 수 있다⁷⁾.

네트워크 데이터의 처리에는 IP 룩업, QoS를 위한 정책 선택 등을 위해 빈번한 메모리 참조가 필요하다. 만일 7계층 룩업까지를 고려한다면, 물리적 네트워크 대역폭의 32배의 메모리 대역폭이 필요하다는 분석 결과도 있다. 즉 OC-192(10 Gbps) 인터페이스의 수용을 위해서는 320Gbps 대역폭의 메모리가 필요로 한다. 이러한 메모리 대역폭은 연결 핀 수의 한계 때문에 외부 메모리

〈표 2〉 네트워크 프로세서 특징

회사명	IBM	VITESSE	AGERE(구 Lucent)	AMCC(구 MMC)
처리속도	10G	10G	10G	Four 2.5G NP+10G TM
칩셋 구성	[Sanford] NPc10GS1, NPd10GS1, NPs10GS1	Denali-10	NP10, TM10	nP×7250(2.5G NP) nP×5710(10G TM) nP×5720(10G MM)
PHY 인터페이스	SPI-4 P1	1 SPI-4 or 4 SPI-3	SPI-4	UTOPIA L3, POS PHYL3 Flex Bus 3, RGGI, ViX-v3
Switch 인터페이스	SPI-4 P1	CSIX	2.5G SerDes	Vix-v3
트래픽 관리	16K policing 64K flow queues RED	16K egress queue WRR, WFQ, WRED, Shaping	4K queues RED, WRED, WFQ, Shaping	256K input flows, 512K output flows, Cell-based/packet-based scheduling WRR, WFQ, WRED

로는 구현이 어렵고, 내장(embedded) 메모리로만 구현이 가능하다. 다음 〈표 2〉는 향후 출시 예정인 대표적인 10G급 네트워크 프로세서의 주요 특징을 나타낸 것이다.

IV. 테라급 라우터 개발 동향

테라비트 라우터로 주장하는 시스템은 수년 전부터 발표되었지만 현재 시판되고 있는 실제 “테라비트 라우터”들은 80 Gbps에서 200 Gbps급의 시스템들이다. (라우터의 용량은 vendor마다의 계산법 차이 때문에 수배 이상 부풀려지곤 하는데, 본 논문에서는 실질적인 의미에서의 라우터 용량인 최대 실장 시 전체 라인카드의 full duplex 용량의 합으로 정의한다) 이러한 범주에 속하는 라우터로는 Cisco의 12416 GSR, Juniper의 M160, Avici의 TSR, Lucent NX64000, Pluris의 Teraplex 20 등이 있다. 이 중 12416 GSR, M160, NX64000은 단일 새시 구조여서 최대 용량이 제한적이지만, Avici TSR, Pluris Teraplex 등은 여러 개의 새시를 연결할 수 있는 확장 구조여서 최대 용량이 수 Tbps~십 수 Tbps까지 확장할 수 있다. 그러나 확장성을 강

조한 Avici와 Pluris의 경우 블록킹 구조의 스위치 네트워크를 사용하여 성능 저하는 피할 수 없을 것으로 보인다. 한편 각 라우터들은 고속 라우터용 상용 칩셋이 나오기 전에 개발된 것이라, 스위치와 라인카드의 대부분의 기능들을 custom ASIC으로 개발하였다. 각 라우터별 특징을 살펴 보면 다음과 같다.

1. Cisco 12416 GSR^[8]

라우터 시장의 대부분을 차지하고 있는 Cisco의 최고 용량의 라우터로서 16개의 slot을 갖는 단일새시 구조이다. Slot당 10G 용량의 라인카드를 장착할 수 있으며, RP(routing processor) module이 최소 하나의 slot을 점유하여 최대 150G의 라인카드를 설치할 수 있다. 스위치는 VOQ(virtual output queue) 구조를 이용한 crossbar 형태로서 스위치 패브릭은 5매로 구성되는데, 160G 용량을 위해서 4매가 필요하며 1매의 패브릭은 이중화를 위한 모듈로 4:1 이중화 방식을 지원한다. 테라급 이상의 용량을 구현하기 위해서 여러대의 12000 GSR 시리즈 라우터를 연결하는 방법을 제시하고 있다. 12416 GSR의 스위치 모듈을 교체하고 외부 스위치 rack을 통해 8대 내지 16대의 12416 GSR을 연결하여 1.2 Tbps(full duplex) 용량을 구현할 수 있을

것이라고 하나, 현재까지 구현되지는 않았다. OC-192c 라인카드까지 출시되었으며 라인카드 당 25Mpps의 wire-speed IP forwarding 성능을 갖는다고 한다.

2. Juniper M160^[9]

코어 라우터 시장에서 Cisco의 강력한 경쟁자로 부상한 Juniper의 최상위 모델로서 8개의 라인카드 slot을 가진 단일 새시 구조이다. 라인카드 Slot당 용량은 10G로서 최대 80G의 라인카드를 수용 가능하다. 전면에는 라인카드, 후면에는 스위치와 프로세서 카드를 장착한 midplane 구조를 가진다. 스위치 패브릭은 4매로 구성되며, 분산 공유메모리 구조의 스위치이다. 일반적인 고속 라우터가 라인카드별로 포워딩 엔진을 갖는 반면 Juniper는 IP 포워딩 기능이 스위치 패브릭 카드에 위치하는데, 스위치 카드당 40Mpps 급의 패킷포워딩 엔진인 IPP2(Internet Processor 2)가 장착되어 총 160Mpps의 포워딩 성능을 갖는 중앙집중형 IP 포워딩 방식을 채택하였다. Juniper M160은 OC-192c 라인카드를 가장 먼저 출시하는 등 기술적인 측면에서 앞선 면도 있으나, 최대 용량이 80G (full duplex)으로 비교 대상중 가장 작은 용량을 가진다.

3. Avici TSR^[10]

스위치 용량이 테라급으로 확장되기 위해서는 일반적인 중앙집중형 스위치 구조로는 곤란하다는 생각하에서 분산 스위칭 구조를 갖는 제품들이 개발되었는데, Avici의 TSR과 Pluris의 Teraplex가 이에 해당한다. TSR은 3-D toroidal mesh 형태의 구조로서, 정육면체들이 쌓여 있는 격자 구조에서 x, y, z 세 축의 각 방향의 양쪽 끝 노드들을 서로 연결해 세 방향에 대해 모두 링 구조를 갖도록 한 구조이다. 각각의 라인카드가 하나의 스위치 노드가 되어 상하, 좌우, 앞뒤의 6개 노드와 연결되어 있는 분산 스위치 구조를 갖는다. 하나의 랙은 40개의 라인카드를 수용하며, 총 14개의 랙을 연결하여 560개의 노드를 갖는 3-D toroidal mesh 구조로 확장할

수 있도록 하였다. 10G 라인카드를 개발할 경우 최대 5.6T까지 확장할 수 있는데, 현재 출시된 라인카드는 2.5G 라인카드 및 10G 라인카드(2매로 구성)이다. TSR의 경우 분산 스위치 구조여서 스위치 용량의 증설이 간단한 구조이나, 기본적으로 3-D toroidal mesh는 블로킹 네트워크이기 때문에 실제 스루풋은 꽤 낮을 것으로 판단되며, 따라서 실질적인 트래픽 처리 용량이 스위치 용량에 못 미칠 것으로 보인다.

TSR에서 주목할 만한 것은 link aggregation 기능인 Composite Link 기능의 지원이다. Composite link 기술은 물리적으로 여러 개의 링크를 하나의 논리적 링크로 묶는 것으로 최대 16개의 링크를 묶을 수 있으며 각 링크 멤버 간의 속도 차이가 4:1 이하일 경우도 가능하다. 이 기능은 TSR 간의 링크 속도를 물리적 링크의 한계 속도 이상이 가능토록 하는 것으로서, 구현상의 문제로 물리적 링크의 속도는 제한되면서 인터페이스 수를 늘일 수 밖에 없는 테라 라우터 간의 논리적 링크 용량을 인터넷 서비스 제공자가 원하는 만큼 높일 수 있어서 테라급 라우터에서 매우 유용한 기술이다.

4. Lucent NX64000^[11]

원래는 Nexabit의 제품으로 Lucent가 Nexabit을 인수함에 따라 Lucent의 제품군이 되었다. NX64000은 단일 새시 구조로서 16개의 라인카드, 4개의 스위치 카드를 갖는다. 스위치 용량은 Lucent측의 주장에 따르면 하나의 스위치 카드가 1.6T 용량이며, 4개의 스위치 카드로 6.4T 용량을 얻을 수 있다고 한다. 그러나 현재 출시된 라인카드는 2.5G OC-48c POS가 최고 속도여서 라인카드 용량으로는 40G급이라고 말할 수 있으며, 출시 예정인 10G 카드가 나올 경우 160G급 라우터가 될 수 있다.

5. Pluris Teraplex 20^[12]

Pluris는 하이퍼 큐브 방식의 interconnection network을 사용해서 19.2 Tbps까지 용량을 확장하는 방식을 채택한다. 하나의 새시에 16

개의 10G 라인카드, 16개의 스위치 카드를 갖는다. 라인카드와 스위치 카드간, 또는 새시와 새시간의 연결은 TeraConnect라는 10G 광 링크로 연결한다. 스위치 카드는 이중화되어서 논리적으로 한 새시에 8개의 스위치 카드가 있으며 4개는 새시 내부 연결용이고 나머지 4개가 새시간 연결용으로 사용된다. 하나의 스위치 카드는 9×9 스위치로 포트 속도는 10G이다. 하나의 새시가 노드가 되어 하이퍼큐브 네트워크를 구성하면 최대 128개의 새시까지 연결할 수 있다. 라인카드 용량을 기준으로 보면 하나의 새시는 150G(한 slot은 control processor용)이며, 최대 19.2 T까지 확장할 수 있다. 전체 새시의 절반을 스위치에 할당하면서 확장성을 강조한 구조인데, 하이퍼큐브 역시 블록킹 구조이기 때문에 실제 운용 가능한 트래픽 용량은 라인카드 용량보다 적을 수밖에 없을 것이다. Teraplex도 Avici의 composite link 기능과 유사한 IP bond라는 link aggregation 기능을 지원한다.

6. 국내 라우터 개발 동향⁷⁾

국내에서는 수백 Gbps급 이상의 고속 라우터는 개발된 사례가 없다. 현재 ETRI에서는 정보통신부 출연의 선도기반기술사업으로 80 Gbps급 라우터를 2001년 말 상용화를 목표로 개발 중에 있으며 2003년에 640 Gbps급 라우터, 2005년에 5.12 Tbps급의 라우터를 상용화하기 위한 연구 개발을 2001년 착수하였다. ETRI에서 개발할 테라비트 라우터는 2.5Gbps/10Gbps/40 Gbps POS 인터페이스, 10 기가비트 이더넷 인터페이스 및 WDM 인터페이스를 가지며, wire-speed IPv6 패킷 포워딩, MPLS, 멀티캐스트, QoS 라우팅, 지능형 형상 제어 및 트래픽 제어 기능 등이 제공될 예정이다.

V. 결 론

차세대 인터넷은 인터넷 트래픽의 급격한 대역

폭 증가와 QoS 보장 요구의 수용이 핵심 이슈로 대두되고 있으며 이를 위해 인터넷의 허부 구조는 변혁을 피할 수 없게 되었다. 테라비트 라우터는 이와 같은 배경을 바탕으로 등장하게 되었지만 테라비트 라우터의 응용은 기존 인터넷 서비스를 포함한 통신 서비스 전반을 포함하는 멀티 서비스 형태로서 그 응용영역을 확장하고 있는 추세이다. 본 고에서는 테라비트 라우터의 주요 요구사항으로서 테라비트까지의 확장성 기능, 고속 패킷 처리기능, QoS 보장형 패킷 처리 기능, 시스템 이중화 기능, DWDM 연동 기능 등에 대해서 언급하였다. 테라비트 라우터의 기술적인 문제에 있어서 최대 이슈는 고속 IP 포워딩 엔진과 테라비트까지 확장성을 제공할 수 있는 고속 스위칭 패브릭의 구현이며, QoS의 보장을 위한 버퍼 관리 및 스케줄링 등도 중요한 연구 과제중의 하나이다. 앞으로의 인터넷 백본 망은 광 스위치 및 전송기술이 갖는 고속성 등의 장점으로 궁극적으로는 광 인터넷으로 될 것으로 예상된다. 그러나 완전 광 기술에 의한 시스템 및 네트워크의 구성에는 기술적 한계가 여전히 존재하고 또 한 경제성 문제를 고려할 때 상당 기간 동안 테라비트 라우터와 같은 초고속 라우터 시스템과 일부 광 기술을 적용한 장치가 공존하게 될 것으로 보인다.

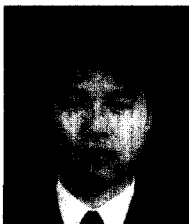
참 고 문 헌

- [1] 강원호, "Terabit Switch/Router (TSR)의 흐름 및 시장 분석", www.netmanias.com, 2000년1월
- [2] S. Keshav and R. Sharma, "Issues and trends in router design", IEEE Communications Magazine, vol.36, Issue.5, pp 144-151, May 1998
- [3] G. Nong and M. Hamdi, "On the provision of quality-of-service guarantees for input queued switches", IEEE Communications Magazine, vol.38,

Issue.12, pp 62-69, Dec. 2000

- (4) S.-T. Chuang, A. Goel, N. McKeown, and B. Prabhakar, "Matching output queueing with a combined input/output-queued switch," IEEE J. Sect. Commun., vol. 17, no. 6, pp. 1030-1039, Jun. 1999.
- (5) Ross Callon, "Technologies for the core of the Internet", Tutorial Presentation of NGN'99, 1999.
- (6) 변 성혁, 이 형호, "차세대 IP 스위치 및 라우터 기술", Telecommunications Review, 제10권 1호, pp23-35, 2000년 2월.
- (7) 이 형호, 김 봉완, 안 병준 "테라비트 라우터 기술", Telecommunications Review, 제11권 2호, pp237-247, 2001년 4월.
- (8) Cisco, <http://www.cisco.com/univercd/cc/td/doc/product/core/cisl2000/index.htm>
- (9) Juniper, <http://www.juniper.net/techpubs/>
- (10) Avici, <http://www.avici.com/documents.html>
- (11) Lucent, <http://lucentdoctrinary-svca.www.conxion.com/english/products.htm#ipsvc>
- (12) Pluris, <http://www.pluris.com>

저자 소개



金鍾岩

1987년 2월 경북대학교 전자공학과 (공학사), 1989년 2월 연세대학교 전자공학과 (공학석사), 1997년 2월~현재: 한국과학기술원 전기 및 전자공학과 박사과정, 1989년 2월~현재: ETRI 네트워크 연구소 라우터기술연구부 선임연구원, <주관심 분야: ATM 프로토콜 기술, 동기식 전송기술, 테라비트 라우터, QoS 기반 스위치 스케줄링 알고

리즘>



邊性赫

1991년 2월 한국과학기술원 전기 및 전자공학과 (공학사), 1993년 2월 한국과학기술원 전기 및 전자공학과 (공학석사), 1999년 2월 한국과학기술원 전기 및 전자공학과 (공학박사), 1999년 2월~현재: ETRI 네트워크 연구소 라우터기술연구부 선임연구원, <주관심 분야: ATM 장비, 테라비트 라우터, 인터넷 QoS, 스위치, 통신망 시뮬레이션>



安炳俊

1984년 2월 한양대학교 전자통신공학과 졸업 (공학사), 1986년 2월 한양대학교 전자통신공학과 졸업 (공학석사), 1999년 5월 Iowa State University 졸업 (Computer Engineering) (공학박사), 1986년 2월~현재: 한국전자통신연구원 라우터기술연구부 책임연구원, <주관심 분야: ATM, 트래픽 제어, QoS, 고속 라우터 기술>



李榮豪

1977년 2월 서울대학교 공업교육과 전자전공 (공학사), 1979년 2월 한국과학기술원 전기 및 전자공학과 (공학석사), 1983년 8월 한국과학기술원 전기 및 전자공학과 (공학박사), 1984년 12월~1986년 11월: 미국 AT&T Bell 연구소 방문 연구원, 1996년 9월~1998년 8월: 충남대학교 공과대학 전자공학과 겸임교수, 1995년 1월~1998년 12월: 대한전자공학회 회지편집위원장, 1996년 1월~1998년 12월: 대한전자공학회 전자교환연구회 전문위원장, 1996년 1월~1999년 12월: IEEE ComSoc APB MDC의장, 1998년 1월~현재: 대한전자공학회 이사, 상임이사, 1998년 5월~현재: 통신위원회 전문위원, 1999년 1월~현재: 한국통신학회 교환 및 라우팅 연구회 위원장, 1983년 8월~현재: 한국전자통신연구원 네트워크기술연구소 라우터기술연구부장, 책임연구원, <주관심 분야: BISDN망, ATM교환, 고속 LAN 및 라우터 기술, 인터넷, 신호처리, 패킷통신, 무선ATM, IMT2000, 지능망>