

지역망간의 트리전환을 이용하는 인터넷 멀티캐스트 라우팅 프로토콜

(An Internet Multicast Routing Protocol with Region-based Tree Switching)

김 원 태 * 박 용 진 **

(WonTae Kim) (YongJin Park)

요 약 본 논문에서는 인터넷 네트워크 구조 및 트래픽 특성을 고려하여 수정 계층적 네트워크 모델을 설계하고, 이를 기반으로 확장성, 데이터 손실율, 전송지연특성 그리고 자원효율성 등을 향상시킨 EDCBT(Enhanced Dispersed Core-Based Tree) 멀티캐스트 라우팅 프로토콜을 제시한다. EDCBT는 양방향성 분산공유트리방식으로, 기존의 방식들과는 달리 도메인 내부의 트리관리와 도메인간 트리관리 모두를 통합처리할 수 있는 구조로 설계되어 있다. 도메인 내부에서는 코어 라우터를 중심으로 CBT 형태의 멀티캐스트 트리를 구성하고, 이렇게 형성된 각 도메인간의 트리구성은 경계라우터들간의 상호작용을 통해 이루어지게 된다. EDCBT의 주요 특징으로서 지역망간 트리전환이라는 새로운 개념을 도입하였는데, 이를 통해 통신의 품질과 네트워크 이용율을 획기적으로 향상시킬 수 있다. 최종적으로 제안 프로토콜의 성능평가 및 코어 라우터의 위치에 따른 효과등을 종단간 전송지연 및 데이터 손실율, 망처리율과 같은 대표적인 평가기준에 따라 네트워크 모의실험용 소프트웨어인 MIL3사의 OPNet을 이용하여 검증한다.

Abstract We design a modified network architecture with considering current Internet network model and traffic characteristics, and propose EDCBT(Enhanced Dispersed Core-based Tree) multicast routing protocol, which enhances scalability, reliability, end-to-end delay and resource utilization. EDCBT adopts bidirectional dispersed shared trees and manages both of intradomain and interdomain multicast trees for a multicast group. Each regional multicast tree is established from its core router and they are interconnected by the operation between border routers on edges of each regional network. As a result, interdomain multicast tree can be easily established. We introduce a new concept named RBTS(Region-based Tree Switching), which dramatically enhances QoS and network utilization. Finally, protocol performance and the effect of core router location are evaluated with MIL3 OPNet network simulator, in terms of end-to-end delay, packet loss and throughput.

1. 서 론

멀티캐스트란 네트워크를 통해 다양하게 분포되어 있으나 같은 그룹에 가입된 수신자들에게 단일 송신자 혹은 다중 송신자로부터의 데이터를 전송하는 서비스이다 [1]. 많은 인터넷 응용서비스들이 기존의 최선노력(best

effort) 서비스에 의존하지 않고 보다 많은 대역폭과 차별화된 QoS를 요구하게 됨에 따라, 인터넷은 이러한 요구사항을 만족시키기 위해 점차 차세대 인터넷을 향해 진화하고 있다. 특히, 웹기반 다자간 화상회의 시스템, 인터넷 메스미디어, 디지털 도서관, 텔레-이머전(Tele-Immersion), 가상현실 게임 및 분산형 상호작용 시뮬레이션 시스템(DIS: Distributed Interactive Simulation) 등과 같은 인터넷 멀티캐스트 서비스들이 이러한 응용에 속한다. 이들 중 몇몇은 적어도 수십 ~ 수백Mbps의 대역폭, 엄격한 종단간 지연시간 및 많은 그룹 멤버의 동시 지원등을 요구하기도 한다 [2].

* 학생회원 : 한양대학교 전자공학과
wtkim@hyucc.hanyang.ac.kr

** 종신회원 : 한양대학교 전기전자공학부 교수
park@hyucc.hanyang.ac.kr

논문접수 : 1999년 5월 6일

심사완료 : 1999년 12월 13일

만일 이러한 서비스의 그룹구성원들이 점대점 방식으로 통신을 한다면 네트워크는 쉽게 포화상태에 도달할 것이며, 네트워크 처리율(throughput)은 결국 열악해 질 것이다. 비록 현재의 네트워크 환경이 과거의 환경에 비해 보다 향상되었다고는 하지만, 명백히 물리적 한계치를 가지고 있으므로 이는 피할 수 없는 문제이다. 더욱이 각 송신자는 모든 수신자에게 동일한 데이터를 반복적으로 송신해야 하므로 네트워크보다 먼저 송신자 측에서 병목이 발생하게 될 것이다. 그러므로, 이를 해결할 수 있는 유일한 해결책은 네트워크에서의 멀티캐스트 라우팅 서비스이다.

본 논문에서는 광역에 걸쳐 분포되어 있는 다수의 그룹 가입자들을 효율적으로 서비스할 수 있는 새로운 멀티캐스트 라우팅 프로토콜을 제시하고자 한다. 2절에서는 인터넷의 망 모델과 트래픽 특성을 제시하며, 3절에서는 기존 인터넷 도메인간 멀티캐스트 라우팅 프로토콜을 제시하며 광역 멀티미디어 통신 시 발생될 문제점을 분석한다. 4절에서는 제안하는 멀티캐스트 라우팅 프로토콜을 자세히 상술하며, 5절에서는 실제적 인터넷 모델에 기반하여, 제안하는 프로토콜의 OPNET 네트워크 모의실험을 수행한 결과를 제시한다. 제안 프로토콜의 평가기준은 종단간 전송 지연치, 데이터 손실을 및 망 처리율 향상치 등이다. 참고로, 본 논문의 EDCBT는 저자의 기존 연구결과인 DCBT를 향상시킨 모델이다 [3].

2. 인터넷 모델링

2.1 인터넷 망 모델링

본 절에서는 멀티캐스트 라우팅 프로토콜을 설계하기 위해 앞서 인터넷 네트워크 모델링에 대한 두 가지 가정을 세우고자 한다. 먼저, 지금까지 많은 인터넷 망 모델링 방법이 제시되었다 [4][5]. 그러나, 이들의 대부분이 그래프 이론에 기반하기 때문에 현재의 인터넷 구성에 비추어 볼 때 비현실적이거나 혹은 지역적으로만 적용할 수 있는 모델이다. 예를 들어, 임의의 두 노드간에 링크가 존재할 확률을 두 객체간 거리의 함수로서 표현하여 네트워크 모델을 구성하는 방법이 있다. 혹은 대표적인 네트워크 모델 즉, 링형, 버스형, 방사형등으로 표현하는 방법을 통해 인터넷을 모델링하기도 한다.

본 논문에서는 그림 1에서 보이는 바와 같이 인터넷 모델을 계층적 구조를 가지며, 각 계층은 전송망과 종단망으로 구성된 여러 망노드들이 임의로 연결된 모델을 채택한다. 전송망이란, 트래픽을 발생시키거나 소비하는 종단망이 아닌 주로 전송을 목적으로 구성된 망을 의미한다. 계층이 0 인 경우는 논리적 대륙단위의 망을 의미

하며, 하부조직망으로 내려갈수록 계층값이 증가한다. 각 계층은 여러 네트워크 노드들로 구성되며, 이들은 직접적인 경로외에도 여러 2차적 경로를 통해 물리적으로 연결되어 있다. 다시 말해, 다중 경로화(Multi-homed) 되어 있다. 하나의 망노드에서 다른 망노드로 데이터를 전송하거나 역방향으로 데이터를 수신하기 위해 하나의 입출력 인터페이스를 가지는데, 이를 경계 라우터(BR:Border Router)로 부르기로 한다. 예를 들어, 그림 1에서 망노드 A는 망노드 B, D 그리고 E로 각각 A.1, A.2 그리고 A.3의 경계 라우터들을 갖는다. 각 경계 라우터간에는 BGP-4와 같은 도메인간 라우팅 프로토콜에 의해 경로정보의 교환 및 결정이 이루어진다 [6]. 둘째, 각 링크와 노드는 다양한 특성과 성능을 가진다. 최근 들어, Tbps급의 전송율을 가지는 DWDM(Dense Wavelength Division Multiplexing)과 같은 고속 광전송기술의 개발로 전송망의 고속화가 촉진되고 있는 반면, 기존의 전용선과 같은 저속 전송서비스 역시 여전히 사용되기도 한다. 인터넷 전송노드로서 라우터에는 스위칭 기술이 접목되어 제 2 계층 스위칭이나 제 3 계층 스위칭등의 고속 IP 패킷 전송기술이 속속 출현하고 있다 [7][8]. 따라서, 인터넷에는 다양한 성능을 갖는 망노드와 링크들로 구성되며, 저성능의 시스템에 의해 통신 품질이 결정된다.

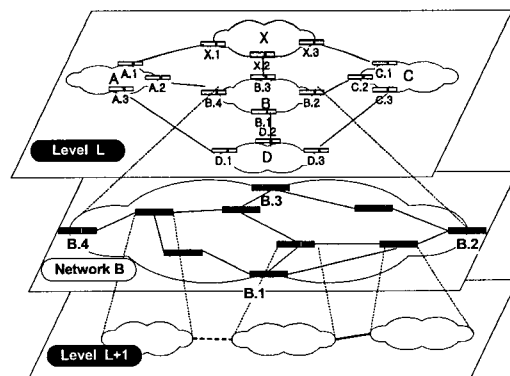


그림 1 수정 계층적 인터넷 구조

2.2 인터넷 트래픽 특성

인터넷 트래픽은 그 특징을 한마디로 특징짓기 어렵다. 이는 인터넷의 데이터 발생 자체가 매우 임의적이며, 응용프로그램에 따른 특성 역시 제각기 다르기 때문이다 [9]. 그러나, 인터넷 트래픽을 측정·분석한 결과 중 하나로서 트래픽 선호도(Traffic Favoritism)가 있다는 사실을 얻었다. 다시말해 인터넷을 통해 흘러 다니는

트래픽의 80% 이상이 20% 이하의 특정 도메인으로부터 집중적으로 발생되거나 소비된다는 사실이다. 또한, 인터넷의 링크 이용률이 시간에 따라 큰 폭의 차이가 생기는데, 트래픽이 몰리는 시간대와 한산한 시간대가 명백히 발생하며, 동시에 지역망간 트래픽의 비대칭성이 있음이 입증되어 있다 [10].

3. EDCBT 프로토콜

3.1 EDCBT 기본개념

3.1.1 기본 설계 방향

EDCBT의 기본 설계 방향은 다음 4가지로 정리할 수 있다.

- ① 광역 거대그룹(Wide Scale and Large group) 지원 측면: 광역에 걸쳐 넓게 분포되어 있으면서 동시에 그 객체 수가 대단히 많은 멀티캐스트 그룹을 지원할 수 있어야 한다.
- ② 멀티미디어 통신 지원 측면 : 망적체(Network Congestion)로 인한 멀티캐스트 통신 객체의 중단간 지연 시간 및 패킷 소실율을 최소화해야 한다.
- ③ 트래픽 관리(Traffic Engineering) 측면: 동적인 멀티캐스트 트래픽 관리를 통해 네트워크의 유한 자원을 최대한 활용할 수 있도록 한다.
- ④ 확장성(Scalability) 측면: 현재의 인터넷구조 및 가용 프로토콜을 적극 고려하여 확장성을 제공할 수 있도록 해야 한다.

3.1.2 멀티캐스트 트리 구성상 특징

EDCBT는 다중-코어기반의 멀티캐스트 라우팅 프로토콜이다. 그러나, 기존의 방식들과는 엄밀히 다른 특징들을 가지고 있다. 기존의 HPIM이나 OCBT등도 다중-코어를 정의하고 있으나 코어간 계층구조를 형성하고 있다 [11][12]. 따라서, 트리구성에 앞서 각 코어간 관계를 형성해야하므로 프로토콜 오버헤드와 확장성에 있어 문제가 있다. 또한, 각각은 근본적으로 PIM과 CBT를 기반하기 때문에 공유트리방식이 갖는 문제점들을 그대로 가지고 있다 [3] [13][14][15]. 반면, EDCBT에서는 각 코어 라우터가 개별적으로 지역망의 공유트리를 관리하고 나아가 지역망 트리를 연결시키는 지역망간 트리를 경계 라우터와의 연동을 통해 구성하며, 지역망 내부의 트리와 지역망간 트리는 독립적으로 구성·관리된다. 따라서, 기존의 다중 코어기반 멀티캐스트 라우팅 프로토콜에서 제공할 수 없는 확장성을 보장한다. EDCBT에 의해 구성되는 트리는 기본적으로 양방향성 공유트리이며, 지역망 내부 트리의 루트(root)은 코어 라우터이고, 지역망들간 트리의 루트는 루트망이

다. 지역망간 트리구성은 BGMP의 기본개념과 유사하나, BGMP는 단순히 서로 다른 멀티캐스트 라우팅 프로토콜을 가지는 이종망간의 연결에만 초점을 맞추고 있는 반면, EDCBT는 망내부와 망간 트리구성 모두를 통합처리할 수 있는 방식으로 설계되어 있다. 그림 2는 EDCBT에 의해 구성된 멀티캐스트 트리의 구조를 거시적 측면에서 도시한 것이다. 굵은 실선은 멀티캐스트 트리 링크를 의미하며 점선은 물리적 링크임을 의미한다.

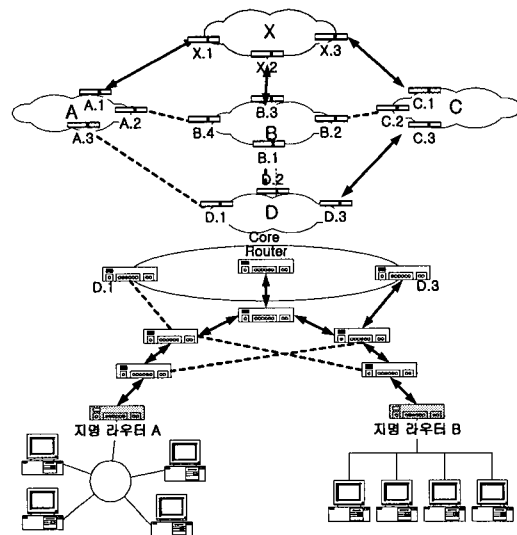


그림 2 EDCBT 멀티캐스트 트리 구성예

3.1.3 트래픽 관리 측면

기존의 멀티캐스트 라우팅 프로토콜과는 달리, EDCBT는 망의 상태에 따른 동적 트래픽 관리를 수행한다. 이를 지역망간 트리전환(RBTS: Region-Based Tree Switching)이라 부르기로 한다. RBTS는 병목이 발생하는 특정부분의 과부하를 줄여 줌으로써 통신품질(QoS: Quality Of Service)을 향상시키고 망의 전체적인 이용률을 높인다. 거대 멀티캐스트 그룹에 의한 망적체는 특정한 망으로부터의 트래픽에 의해 발생될 확률이 높고 특히 이러한 트래픽이 낮은 성능의 망노드와 링크에 흐르게 될 때 적체 확률은 더욱 높아진다. 이는 2.2절에서 기술한, 트래픽 선호도, 다중경로화 및 다양한 망성능에 근거를 두고 있다. 따라서, 이러한 트래픽을 다른 트래픽과 구별시켜 적체점을 우회하는 2차 경로를 통해 전송하면, 망적체를 어느 정도 해결할 수 있고 여분의 망자원을 충분히 활용할 수 있게 된다. 자세한 내용은 3.3절에서 다루기로 한다.

3.2 망 구성요소

EDCBT망의 구성요소는 다음과 같다.

① EDCBT 라우터: EDCBT 프로토콜을 구현한 모듈을 가진 라우터이다.

② 지명 라우터(DR: Designated Router): 사용자 단말과 직접 연결되어 있는 라우터로, IGMPv2(Internet Group Management Protocol version 2) 모듈과 EDCBT 라우팅 모듈을 모두 보유한다 [16].

③ 코어 라우터: EDCBT 라우팅 기능을 포함하며, 부가적으로 코어 라우터 기능을 가진다.

이들 외에 사용자 단말이 있는데, 사용자 단말은 IGMP 모듈과 멀티캐스트 그룹에 동적인 가입/탈퇴를 할 수 있는 멀티캐스트 응용서비스를 가진다.

3.3 EDCBT 트리 구성방법

3.3.1 지역망 내 트리 구성 및 관리

그림 3는 하나의 지역망 내에서 멀티캐스트 트리가 구성되는 과정을 보인 것이다. 초기에 DR Y와 지역 코어 라우터(RC: Region Core) 간에 이미 멀티캐스트 트리가 형성되어 있다고 가정한다. DR X는 IGMP에서 정의하고 있는대로 주기적으로 (1) IGMP Group Query 메시지를 자신의 서브넷에 브로드캐스팅한다. 서브넷에 속한 호스트 중 특정 그룹에 가입하고자하는 호스트는 (2) IGMP Group Report 메시지를 자신의 지명 라우터인 DR X에게 전송한다. 이때 IP TTL값은 1이므로 DR X를 넘어 전송될 수는 없다.

IGMP Group Report 메시지를 받은 DR X는 RC로의 최단경로에 있는 EDCBT_R A에게 EDCBT Group_Join 메시지를 전송하고 타이머를 구동시킨다. 대기시간이 초과하게 되면 가입 메시지를 재전송하게 된다. 정상적인 경우, EDCBT_R A는 EDCBT Group_Join 메시지를 받게되지만, 이미 멀티캐스트 트리가 코어 라우터와 형성되어 있기 때문에 더 이상 이 메시지를 코어 라우터 방향으로 전송하지 않고 바로 EDCBT Group_Join_Ack 메시지를 DR X에게 되돌려 준다. 결과적으로, EDCBT_R A는 DR X를 위한 정보를 자신의 멀티캐스트 라우팅 테이블에 다음과 같이 등록하게 된다.

(* , G ID, Parent interface: 1, Child interfaces: {2,3})

여기에서, *는 송신자에 관계없이 해당 멀티캐스트 채널을 통해 전송될 수 있음을 의미한다. G ID는 멀티캐스트 IP 주소이고, interface는 EDCBT_R A와 연결된 network interface를 의미한다. 특히, parent interface는 코어 라우터를 향한 인터페이스이며, child

interface는 하부 네트워크로 향하는 인터페이스이다. 한편, 멀티캐스트 트리 제거방식은 트리 형성과 유사하여, 그 역으로 과정을 수행하면 되므로 그룹탈퇴에 대한 내용은 본문에서 제외한다.

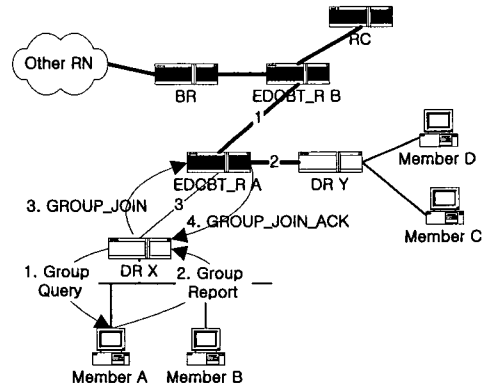


그림 3 멀티캐스트 트리 구성 예

3.3.2 지역망 간 트리 구성 및 관리

지역망 간 멀티캐스트 트리는 각 지역망 내에서 구성된 멀티캐스트 트리들을 해당 경계 라우터들 간의 상호 작용을 통해 연결한 공유트리이다. 즉, 각 지역망은 코어 라우터를 중심으로 공유트리를 구성하고, 이 공유트리는 다른 지역망의 공유트리와 멀티캐스트 연결을 맺음으로써 하나의 거대한 멀티캐스트 트리구조가 형성된다. 그러나, 이 과정에서 반드시 해결되어야 할 점은 첫째, 지역망 간 트리 구성 시 전체 트리의 루트 결정방법. 둘째, 경계 라우터 동작방식. 셋째, 중간 지역망이 그룹멤버가 전혀 없는 경우 즉, 중간 지역망이 전송망으로서만 동작할 경우의 해법 등이다. 다행히도 첫 번째 문제는 현재 IETF에서 제안되고 있는 MASC (Multicast Address Set Claim)을 적용함으로써 해결할 수 있다 [17]. MASC는 기존의 임의로 멀티캐스트 주소를 할당하던 방식에서 벗어나 지역망 별로 멀티캐스트 주소공간을 할당하여 주소체계를 조직화하기 위한 방안이다. MASC를 채용할 경우, BGP4+와 같은 도메인간 라우팅 프로토콜에 멀티캐스트 주소 정보를 실어 보낼 수 있기 때문에 모든 경계 라우터는 특정 멀티캐스트 주소에 대한 루트 지역망을 알 수 있게 된다. 이러한 방식은 현재 지역망간 멀티캐스트 트리를 구성하는 프로토콜인 BGMP에 적용되고 있다 [17].

중간 지역망이 멤버가 없는 전송망의 역할만을 담당하는 경우에 두가지 해결방안이 있을 수 있다. 첫째는 현재 BGMP에서 제안되고 있는 터널링 방식이다. 이

방식은 멀티캐스트 트래픽을 전송망의 경계 라우터간 에 터널 인터페이스를 통해 전송하는 방식으로, 간단하기는 하나 전송망 내에 그룹멤버가 발생할 경우 터널의 처리가 복잡해지고 IP 헤더를 한번 더 붙이기 때문에 오버헤드가 발생하는 단점이 있다. 둘째는 전송망 내에 3.2.2 절에서 설명한 지역망 내부의 멀티캐스트 트리를 구성하여 입출력 경계 라우터들을 연결해 주는 방식이다. EDCBT에서는 후자의 방식을 채택하였다.

3.4 지역망간 트리 전환 방안

도메인 간 트리 전환(RBTS)이란, 네트워크의 특정부분에 병목이 발생하는 경우에 특히, 멀티캐스트 트래픽에 의해 병목이 발생할 경우에 EDCBT에서는 원인이 되는 소스 지역망의 트래픽을 우회경로를 통해 전송하므로써 네트워크의 부하를 분산시키는 방안을 말한다.

3.4.1 기본방안

그림 4에서 RN A로부터 발생되는 멀티캐스트 트래픽의 전송율이 다른 지역망에 비해 매우 높고, X.3와 C.1간의 링크만이 지속이며 경계라우터로서 C.1만이 패킷처리 성능이 좋지 않은 경우를 예로 든다. 이 때, 전체 네트워크에 흐르는 트래픽은 하나의 멀티캐스트 그룹에 속한 트래픽임을 가정으로 한다. 3.4.2절에서는 보다 실제적인 경우를 위한 방안이 제시된다.

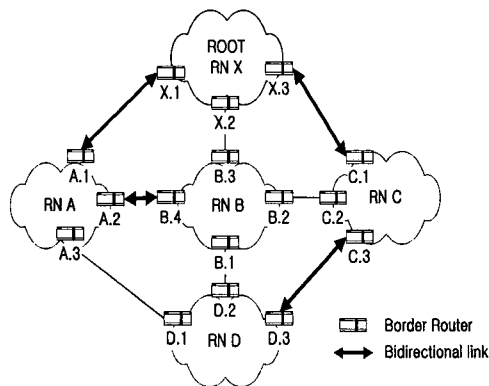


그림 4 지역망간 멀티캐스트 트리 구성 결과

C.1은 RN A로부터의 트래픽이 X.3-C.1 링크에 병목을 발생시키는 것을 트래픽감시(traffic monitoring)를 통해 알게된다. 현재 사용되고 있는 BGP는 CIDR를 채용하고 있기 때문에 소스주소를 통해 쉽게 지역망을 구별할 수 있다. 이 때, C.1은 RN A로 향하는 최단경로로서 C.2를 알고 있으므로 C.2로 EDCBT tree_switching(Req, C.1, C.2, ID(RN A)) 메시지를 전송한다. C.2는 tree_switching 메시지를 EDCBT Group_

Join(SBT, ID(RN A))로 변환하여 B.2에게 전송한다. 이 때, SBT는 Source Based Tree를 의미하며 ID(RN A)는 RN A의 AS(Autonomous System) 번호나 CIDR 주소체계등의 RN A를 의미하는 식별자를 말한다. 이 메시지를 받은 B.2는 자신의 지역망 코어 라우터와 공유트리를 형성한다. 따라서, B.2는 코어로부터 RN A를 포함한 다른 지역망으로부터의 모든 트래픽을 받게 된다. 그러나, C.2로부터의 EDCBT Group_Join메시지 내에 RN A로부터의 트래픽만을 보내라는 정보를 받았으므로 B.2는 모든 트래픽을 C.2에게 보내지 않고 단지 RN A로부터의 트래픽만을 걸러 보내게 된다. 이는 일종의 라우팅 정책(Policy Routing)이 될 수 있다. 한편, C.2는 B.2에게 Group_Join 메시지를 보냄과 동시에 RN C의 멀티캐스트 트리에 EDCBT tree_switch(SBT mode, Req, C.1, C.2)를 멀티캐스트한다. 이 메시지는 지명 라우터까지 전송되고, 이를 수신한 지명 라우터들은 C.2로 tree_switch(Ack, DR, C.2)를 보낸다. 이 메시지가 C.2까지 전송되는 경로 상의 EDCBT라우터들은 라우팅 테이블에 C.2를 루트로하는 SBT를 구성해 낸다. tree_switch(Ack, DR, C.2) 메시지가 C.2에 수신되면 C.2는 C.1에 tree_switch(Ack, C.2, C.1, ID(RN A))를 전송하여 지역망 RN C내에서 C.2를 루트로하는 SBT구성이 완료되었음을 알린다. C.1은 Group_Join(SBT mode, C.1, X.3, ID(RN(A)))를 X.3에게 전송하여 더 이상 X.3-C.1 링크를 통해 RN A로부터의 트래픽을 보내지 말라는 메시지를 전달하여 트리전환 프로세스를 완료한다. 결과로서 구성된 멀티캐스트 트리구조를 그림 5에 도시한다. 점선의 양방향성 트리는 RN

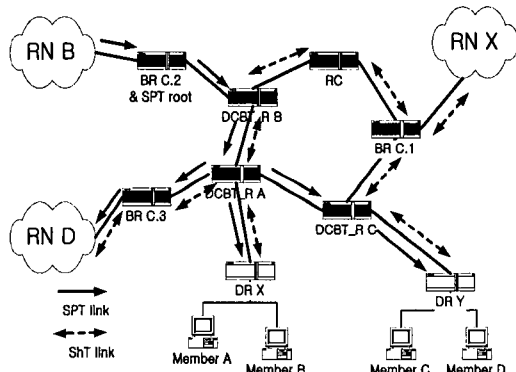


그림 5 RBTS의 결과로 형성된 RN C내의 멀티캐스트 트리

A를 제외한 나머지 지역망으로부터의 트래픽이 흐르는 트리이고, 실선의 단방향성 트리는 RN B로부터 전송되는 RN A의 트래픽이 흐르는 멀티캐스트 트리를 의미한다.

RBTS에 의해서 병목의 원인이 되던 RN A로부터의 멀티캐스트 데이터가 RN B를 거쳐 RN C로 흘러들게 되면, C.1은 그 처리부하가 줄게 되고 동시에 IP 패킷 손실율과 라우터 버퍼에서의 시간 지연을 동시에 감소시킬 것이다. 이 효과는 RN C에 의해 멀티캐스트 트래픽을 수신하는 RN D의 통신품질에도 분명히 영향을 미칠 것이다.

3.4.2 트리전환 시 경로결정 방안

본 절에서는 전체 네트워크에 이미 부하가 걸려있는 상황을 추가로 가정하여, 앞서 3.3.1절에서의 가정에 현실성을 부여하고자 한다. 예를들어 B.2-C.2 링크에 이미 멀티캐스트와는 무관한 트래픽이 흘러 그 가용 용량이 줄어들어 있다고 가정하면, RBTS에 의해 무조건 RN B로 트리를 전환한다면, 트리전환의 효과가 오히려 역으로 일어날 수도 있다. 따라서, 트리전환을 위한 조건적 수용방안이 필요한데, 이를 위한 해법으로 두가지가 가능하다. 첫째, 각 지역망의 자원상태를 관리하는 자원관리 에이전트(Resource Management Agent)를 두고 이들간의 상호작용을 통한 트리전환 여부를 판단하는 방법과 둘째, 프로토콜을 통해 경계 라우터들간의 상호작용을 통한 조건적 트리전환하는 방법이다. 자원관리 에이전트를 이용할 경우, 보다 융통성이고 다양한 효과를 얻을 수 있으나 이들을 위한 독립적인 망을 재구축해야 하고 관리해야 하기 때문에 차선책으로서 고려한다. 본 논문에서는 후자를 통해 문제를 해결한다. 이 방법은 자원탐색이 간단하고 비교적 빠르게 이루어지는 장점이 있는 반면 자원의 탐색이 경계 라우터에 의해서만 이루어지므로 지역망 내부적인 상황을 고려할 수 없다는 단점이 있다. 그러나, 최근 망들간의 링크에 비해 지역망이 고속화되는 경향이 있으므로 큰 문제는 아니라고 사료된다.

3.5 라우팅 루프 제거 방안

멀티캐스트 라우팅 프로토콜에서 주의해야할 설계기 준들 중 하나가 라우팅 루프(routing loop)를 제거하거나 원천적으로 발생하지 않도록 설계하는 것이다. 그런데, EDCBT에서는 RBTS를 수행할 경우, 경계 라우터를 루트로하는 SBT와 공유트리가 하나의 지역망 내에 공존하므로 망 내의 EDCBT 라우터들은 트리전환한 특정 소스로부터의 트래픽과 그 외의 공유트리를 통해 전송되어야하는 트래픽을 구별하지 못하게 되어 라우팅

루프가 발생할 가능성이 생긴다. 망 내부의 EDCBT 라우터들이 라우팅 정책을 수행한다면, 트리전환 시 해당 소스주소를 라우팅 테이블에 기록하여 문제를 쉽게 해결할 수 있다. 그러나, 불행히도 대부분 지역망 내의 라우터들은 라우팅 정책을 쓰지 않는다. 예를들어, 그림 5에서 EDCBT_R A는 EDCBT_R C로 공유트리 트래픽을 보내서는 않되나, SBT 트래픽과 공유트리 트래픽을 구별할 수 없기에 EDCBT_R C로 모든 소스로부터의 트래픽을 보내고 이는 다시 BR C.1을 거쳐 계속해서 루프를 돌게 되어 망에 큰 손실을 가한다.

따라서, 이를 해결할 방법은 현재 IETF에서 표준작업을 진행하고 있는 차등 서비스(Differential Service)를 멀티캐스트 라우팅에 확장 이용하는 것이다 [18]. IP 헤더의 ToS(Type of Service) 필드 중 6비트를 차등 서비스에서는 재정의하였고, 나머지 2비트는 현재 이용되지 않고 있다. 이 여분의 2비트를 이용하여 트리전환에 의해 형성된 지역망내 SBT의 루트 경계라우터(그림 5의 C.2)에서는 트리전환된 소스 지역망으로부터 들어오는 트래픽에 "01"을 삽입하여, 내부 EDCBT 라우터들이 공유트리 트래픽과 SBT 트래픽을 구별할 수 있게 한다. 이로써, 트리전환된 망 내부에서 유일한 라우팅 루프 발생 원인을 제거할 수 있다.

4. OPNet 모의실험

본 장에서는 4장에서 설계한 EDCBT 멀티캐스트 라우팅 프로토콜의 구현 및 검증을 OPNet 네트워크 시뮬레이터를 통해 보이고자 한다. 실험에 사용된 OPNet은 3.5A 유무선 통합 버전이다.

4.1 모의실험 모델

OPNet 시뮬레이터를 이용하여, 그림 6의 네트워크를 구성한다. 그림 6(a)는 전체적인 네트워크 구성을 보여주고 있으며, 각 지역망(RN으로 표기)은 그림 6(b)와 같은 구조를 갖는다. 배경으로 보이는 전세계 지도의 스케일은 실제 거리를 나타내고 있다. 즉, 각 지역망이 전세계의 5대륙에 걸쳐 분포하도록 구성했다. 네트워크에 병목을 유발시키기 위해 `rte_r3_p` 경계라우터의 성능과 링크 속도를 줄였다. 이에 대한 자세한 내용은 다음 절에서 상술한다. 루트 망으로는 RN1이 설정되었다. 따라서, 각 지역망은 RN1으로 향하는 최단 거리를 유니캐스트 라우팅 정보에 따라 판단하여 지역망간 공유트리를 구성하게 된다.

앞서 기술한 EDCBT망의 4개 구성요소를 모두 구현했다. 이를 그림 7에 도시하였다. 그림 7(b)는 지명 라우터의 구조이며 IGMP는 IGMPv2를 구현하였다.

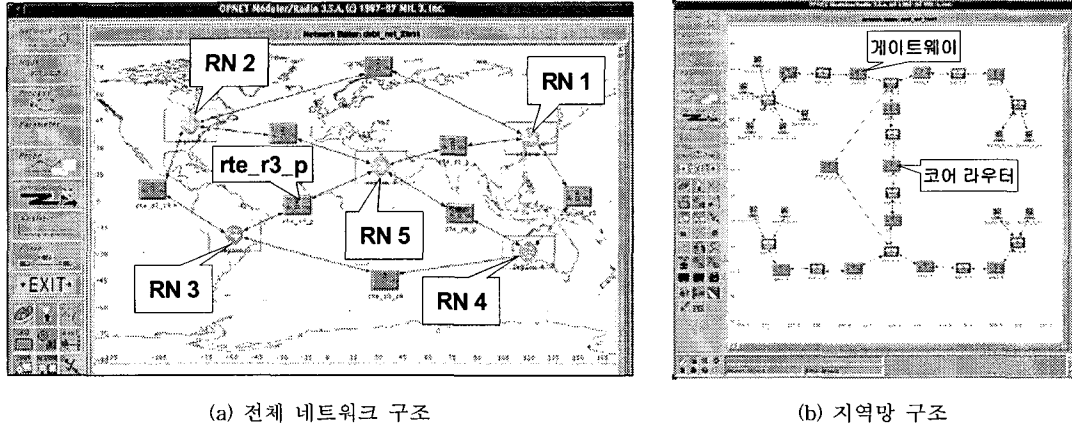


그림 6 OPNet 네트워크 모델 구성

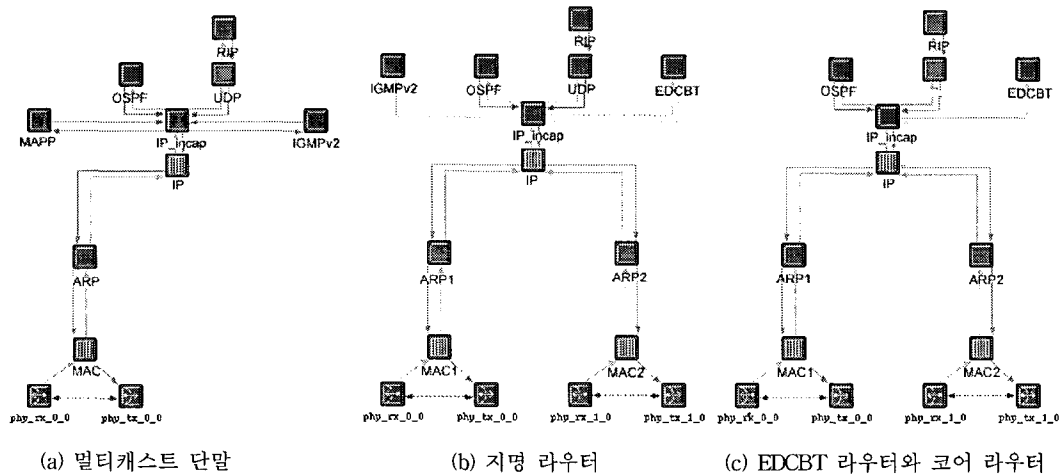


그림 7 OPNet 노드 모델

EDCBT 라우팅 모듈은 기본적으로 코어 라우터 기능을 포함하며, 모의실험 과정에서 코어라우터로 동작할 것인지 단순 EDCBT라우터로 동작할 지를 지정할 수 있도록 구현하였다. 멀티캐스트 단말의 경우, MAPP (Multicast Application)는 멀티캐스트 응용프로그램 모듈로 그룹에 가입/탈퇴를 확률적으로 결정하며 IGMPv2를 이용하고, 지명라우터와 연동하여 IETF의 IGMPv2 표준에 기술된 대로의 그룹관리 동작을 수행한다. 단말은 두가지 모드로 동작하는데, 첫째는 단순한 송신자로

서 멀티캐스트 트래픽을 발생시켜 EDCBT에 의해 형성된 멀티캐스트 트리에 트래픽을 보내는 모드이고, 둘째는 그룹에 가입되어 있는 경우에 트리를 통해 흘러드는 멀티캐스트 트래픽을 수신하고 동시에 송신할 수도 있는 모드이다. 망은 FDDI를 기반으로 구성되어 있으므로 MAC 계층은 Token 방식의 FDDI MAC이 사용된다. 유니캐스트(Unicast) 라우팅 프로토콜로는 OSPF와 RIP이 모두 지원되나, 본 실험에서는 RIP을 사용했다. 그러므로, 라우팅 메트릭으로 홉(Hop) 카운트를 기준으

로 하게 된다. 단말과 라우터들을 포함하는 모든 네트워크 노드들은 IP주소를 가지고 있고, IP 처리는 노드모델의 IP 모듈에서 수행한다. 이렇듯, OPNet에서는 각 프로토콜의 초기화 동작, 완벽에 가까운 프로토콜 처리 기능 및 물리적인 전송지연시간 등의 섬세한 모의실험 처리가 이루어져 실제 네트워크 환경에 매우 유사한 신뢰성있는 결과를 제공하게 된다.

4.2 모의실험 목적 및 파라미터

본 실험의 목적은 다음 4가지로 정리할 수 있다.

- ① EDCBT 프로토콜의 동작 여부를 검증한다.
- ② 라우팅 루프의 발생 여부를 검사한다.
- ③ 지역망간 트리전환의 효율성을 종단간 시간지연과 패킷 소실율에 입각하여 평가한다.
- ④ 코어 라우터의 지역망 내 위치에 따른 통신품질의 차이를 조사한다.

표 1 모의실험 파라미터

실험 파라미터		설정 값
MAPP	패킷 전송율	평균 0.8 패킷/초 Poisson process
	패킷 크기	1512 byte
	평균 그룹 유지 시간	400 초, Constant PDF
	그룹 재가입 시간	탈퇴후 200초, Constant PDF
	그룹의 수	1
링크 전송 속도	RN1-RN5-RN3 링크	T1(1.544Mbps)
	기본 링크	T3(45Mbps)
라우터의 패킷처리 성능	rte_r3_p	50 패킷/초
	기본 EDCBT 라우터	500 패킷/초
각 지역망 내 노드 수	EDCBT 라우터 수	7, 총: 35개
	코어 라우터 수	1, 총: 5개
	지명 라우터 수	4, 총: 20개
	멀티캐스트 단말 수	8(Default), 11(RN 1) 총: 43개

모의실험 파라미터를 표 1에 정리한다. 표 1에서 보이는 바와 같이, 전체망에 존재하는 그룹의 수는 하나로 지정하였다. 다중의 멀티캐스트 그룹을 동적으로 발생시키는 것도 가능하나, 모의실험 상 큰 의미는 없기 때문에 멀티캐스트 그룹의 수는 하나로 고정시켰다. 단일 MAPP에서 발생시키는 패킷의 발생율은 초당 0.8 패킷이다. 이론적으로는 34.4 패킷/초(=0.8패킷/초 x 43 단말)가 단말에서 멀티캐스트 트리를 통해 네트워크로 주

입되지만, 네트워크 노드들에서의 다중큐잉지연과 전송 지연등을 고려하면 수신하는 단말에서는 이와는 다른 매우 불특정한 수신율로 받게 된다. 이를 바탕으로 RN3와 RN5 사이의 경계라우터인 rte_r3_p를 다른 라우터의 패킷처리용량의 1/10 (50패킷/초)로 설정하였다. 즉, 공유트리를 통해 트래픽이 RN5로부터 RN3로 흘러들어갈 때, 저성능의 rte_r3_p에서 버퍼오버플로우를 발생시켜 상대적으로 많은 트래픽을 발생하는 RN1으로 RBTS를 유발시키기 위함이다. 마찬가지로 rte_r3_p의 출력링크 속도도 T1으로 설정되었으며, 이는 최대 초당 120 패킷정도를 전송할 수 있는 용량이다. 이처럼 rte_r3_p에서는 다른 네트워크 노드에 비해 과부하가 걸리도록 설정되었다.

4.3 실험 결과

그림 6의 계층적 네트워크 구조에서 표 1의 파라미터들을 가지고, 약 200초 동안 모의실험을 수행한 결과 그림 8에서 14까지의 흥미로운 결과를 도출했다. 그림 8은 경계 라우터인 rte_r3_p에서의 버퍼 오버플로우에 의한 3% IP 패킷 소실량이 발생하고 있음을 보여 준다. 본 모의실험에서는 일정 시간동안 미리 지정된 RBTS 임계치 이상 오버플로우가 발생할 경우 트리스위칭이 일어나도록 설정되었다. 실제적인 구현 시에는 이를 망관리자가 망관리 정책에 따라 경계라우터의 EDCBT모듈에 직접 혹은 간접적으로 설정할 수 있도록 해야한다. rte_r3_p에서의 오버플로우에 의해 RBTS가 RN3로부터 RN1으로 일어나고, 그 결과 두 지역망간에 단방향성 우회경로가 형성된다. 그림 9과 그림 10은 RN 1으로부터 RN 3로 유입되는 트래픽에 대한 RBTS 이전과 이후의 패킷소실량과 종단간 지연시간을 비교한 결과이다. RBTS 이후에는 패킷소실이 거의 일어나지 않고 있음을 알 수 있고, 종단간 지연시간은 1/40 수준으로 줄어든 것을 알 수 있다. 두 지역망간 종단간 지연시간이 획기적으로 줄어들 수 있었던 원인은 병목이 발생하고 있는 RN 3의 rte_r3_p로부터 다른 경계라우터로 RN1으로부터의 멀티캐스트 트래픽을 우회시키므로써 얻어지는 로드감소효과와 본 모의실험에서 인터넷 백그라운드 트래픽을 무시하고 전체망에 오직 멀티캐스트 트래픽만이 흐른다고 가정한 데도 있을 수 있다. 인터넷 백그라운드 트래픽이란, 본 모의실험에서 사용하는 멀티캐스트 트래픽을 제외한 일반적인 유니캐스트용 트래픽을 말한다. 하지만, 실제적인 인터넷 트래픽을 수학적으로 모델링하는 것은 사실상 힘들고, 임의로 발생시킨다는 것도 의미 없는 일이므로 본 실험에서는 이를 무시하였다. 그러나, 그림 11의 망 처리율 1.2배 증가한 결과를 통해

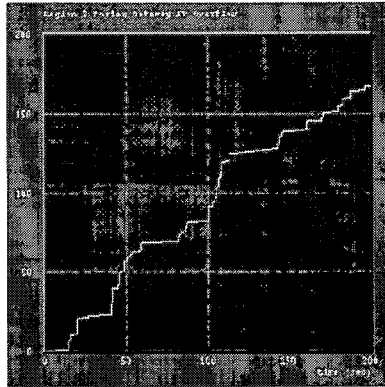


그림 8 rte_r3_p에서의 IP Buffer Overflow

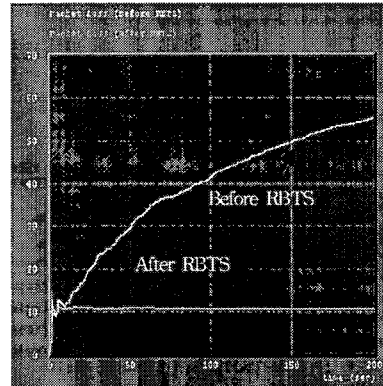


그림 9 RN1->RN3, Packet Loss

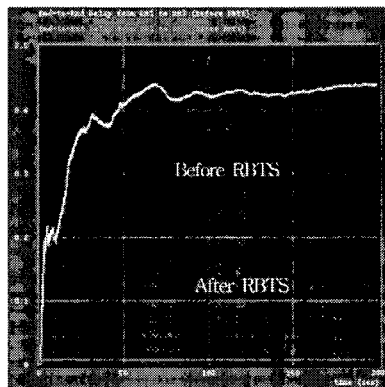


그림 10 RN1->RN3, 종단간 지연시간

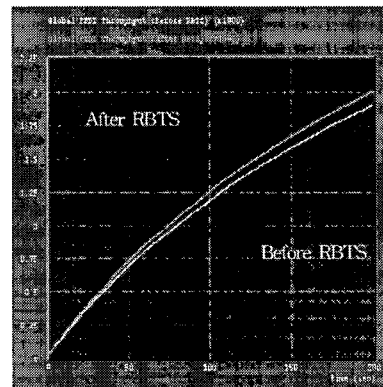


그림 11 전체망의 평균처리율

볼 때, 망의 처리능력 향상에 RBTS가 기여할 수 있다는 것을 알 수 있다. 또한, 2.2절에서 기술한 바대로 인터넷에 어느 정도 일정한 패턴을 갖는 백그라운드 트래픽이 흐른다고 가정하면, 3.4.2에서 기술한 다양한 스위칭 경로탐색 기법을 이용하여 통신품질 향상을 기대할 수 있다. 그림 12에서 그림 14은 망 전체에서 발생된 패킷소실량과 종단간 지연시간을 구한 것이다. 전체망에 대한 종단간 지연시간 역시 큰 폭으로 줄어들었음을 볼 수 있고, 전체 패킷 소실율은 50%가량 줄어들었다. 특히, 그림 13과 그림 14은 코어 라우터의 위치를 망 중앙에서 게이트웨이로 이동시키기 이전과 이후 그리고 RBTS 이전과 이후를 동시에 비교한 결과이다. 기존의 코어 라우터를 중심으로 멀티캐스트 트리를 구성하는

공유트리방식에서 많은 논쟁거리가 되었던 코어의 위치 결정 문제가 본 논문에서 제안하는 수정 계층적 망 구조에서는 그리 큰 비중을 차지 않음을 입증하였다. 모의 실험한 망의 형상이 그물형이거나 토폴로지없이 구성되는 임의적 구조라면 [4], 코어의 위치에 따른 통신품질의 변동이 매우 크리라 예상된다. 그러나, 본 저자의 망 형상에 관한 고찰 결과 2.1절에서 제안한 계층구조형이 일반 인터넷의 타당한 모델이라 생각되므로 코어의 위치에 관련된 모의실험의 결과는 신뢰할 만하다 하겠다.

5. 결론

본 논문에서는 현재의 인터넷구조를 고려한 수정 계층적 네트워크 모델을 설계하고, 이를 기반으로 지역망

내부와 지역망 사이의 멀티캐스트 트리를 포괄 수용할 수 있는 통합 멀티캐스트 라우팅 프로토콜을 제안했다. EDCBT는 분산코어 구조이며, 양방향성 공유트리틀을 기반으로 한다. 지역망 내부에서는 CBT와 유사한 방법으로 공유트리틀을 구성하고, 지역망들 사이는 루트 지역망을 트리의 정점으로하여 경계라우터들과 각 코어 라우터들간의 상호공조로서 거대 트리가 구성된다. 그러므로, 현재의 멀티캐스트 라우팅 프로토콜들이 안고 있는 확장성 결여 문제를 해결할 수 있다. EDCBT의 핵심적 기능 중, RBTS는 효과적으로 네트워크 병목을 해결하여, 전반적인 네트워크 자원 활용률을 향상시켰다. 결과적으로 종단간 지연시간, 패킷 소실을 및 망처리율등의 통신품질을 획기적으로 개선시켰다. RBTS에 의해 형성된 SBT와 공유트리틀을 구분하지 못하는 라우팅 루프 문제는 IP의 ToS 필드를 재정의하여 간단히 해결했다. 마지막으로, 지역망 내에서의 코어 라우터의 위치가 본 논문의 망 구조에서는 통신품질에 영향을 미치지 않는다는 사실을 입증했다.

참고문헌

- [1] S. Deering, "Host Extensions for IP Multicasting," IETF RFC1112, Aug. 1989.
- [2] P. Bagnall, R. Briscoe and A Poppitt, "Taxonomy of Communication Requirements for Large-scale Multicast Applications," IETF Draft, May 1998.
- [3] W.T. Kim and Y.J. Park, "DCBT: An Efficient Multicast Architecture for Wide Scale and Large Group Multimedia Communications," the proceedings of ICC'99, Vancouver, Canada
- [4] B. M. Waxman, "Routing of multipoint connections," IEEE JSAC, vol. 6, no. 9, pp. 1617-1622, 1988
- [5] E. W. Zegura, K. L. Calvert and M. J. Donahoo, "A Quantitative Comparison of Graph-Based Models for Internet Topology," IEEE/ACM Trans. on Networking, Vol. 5, NO. 6, Dec. 1997
- [6] Y. Rekhter and T. Li, "A Border Gateway Protocol 4 (BGP-4)," IETF RFC1771, Mar. 1995.
- [7] A. Viswanathan, N. Feldman, Z. Wang and R. Callon, "Evolution of Multiprotocol Label Switching," IEEE Communications Magazine, May, 1998
- [8] P. Newman, G. Minshall, T. Lyon, and L. Huston, "IP Switching & Gigabit Routers," IEEE Communications Magazine, Jan. 1997
- [9] R. G. Addie, M. Zukerman and T. D. Neame, "Broadband Traffic Modeling: Simple Solutions to Hard Problems," IEEE Communications Magazine, Aug. 1998
- [10] K. Thompson, G. J. Miller, and R. Wilder, "Wide-Area Internet Traffic Patterns and Characteristics

Observations on the patterns and characteristics of wide-area Internet traffic," IEEE Network, Nov. 1997

- [11] C. Shields and J.J. Garcia-Luna-Aceves, "The ordered core-based tree protocol," in Proc.IEEE INFOCOM'97, Kobe, Japan, pp. 884-91, April 1997.
- [12] M. Handley, J. Crowcroft, and I. Wakeman, "Hierarchical Protocol independent multi-cast(HPIM)," University College London, Nov. 1995.
- [13] A. Ballardie, "Core Based Trees (CBT version 2) Multicast Routing: Protocol Specification," IETF RFC2189, Sep. 1997.
- [14] S. Deering, "The PIM Architecture for Wide Area Multicast Routing," IEEE/ACM transactions on Networking, Apr. 1996.
- [15] T. Billhartz, J. B. Cain, E. F. Goudreau, D. Fieg and S. G. Batsell, "Performance and Resource Cost Comparisons for the CBT and PIM Multicast Routing Protocols," IEEE JSAC VOL. 15 NO. 3, April 1997
- [16] W. Fenner, "Internet Group Management Protocol, Version 2," IETF RFC 2236, Nov. 1997.
- [17] S. Kumar, P. Radoslavov, D. Thaler, C. Alaettinoglu, D. Estrin, and M. Handley, "The MASC/BGMP Architecture for Inter-domain Multicast Routing," in Proc. ACM SIGCOMM 98, September 1998, Vancouver, Canada.
- [18] K. Nichols, S. Blake, F. Baker and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers," IETF RFC2474, Dec. 1998.



김 원 태

1994년 한양대학교 전자공학과 졸업.
1996년 한양대학교 전자공학과 석사.
1996년 ~ 2000년 현재 한양대학교 네트워크 컴퓨팅 연구실에서 박사과정 수학 중. 1999년 한국학술진흥재단에서 신진연구인력으로 선정. 관심분야는 차세대 인터넷 기술, 이동인터넷 통신 및 분산가상현실



박 용 진

1969년 와세다대학교 전자통신공학과 졸업. 1971년 와세다 대학원 석사. 1978년 와세다 대학원 박사(공학박사). 1979년 ~ 현재 한양대학교 전기전자공학부 교수. 1983년부터 1년간 Univ. of Illinois, Urbana 전산학과 방문 부교수. 1991년부터 1년간 영국 Kent 대학 방문교수. 관심분야는 컴퓨터 통신, 분산 시스템, 초고속 통신, 이동데이터 통신.