

QoS를 고려한 RMCP 기반의 응용 계층 멀티 캐스트

탁광옥^o 박시용 정기동

부산 대학교 컴퓨터 공학과

{kotak^o, sympark, kdchung}@melon.cs.pusan.ac.kr

The application level multicast based RMCP considering QoS

Kwangok Tak^o Siyong Park Kidong Chung

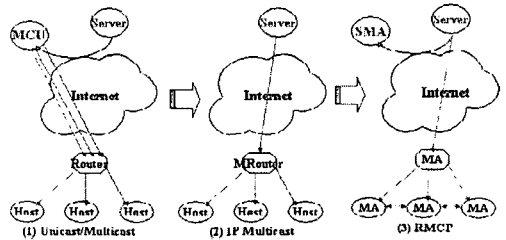
Dept. of Computer Engineering, Pusan University

요 약

본 논문에서는 응용 계층 멀티캐스트 방식을 이용하여 미디어 스트림 데이터를 실시간으로 전송할 수 있는 서비스망을 구현하고자 한다. 일반적으로 IP 멀티캐스팅 서비스를 구현하기 위해서는 멀티캐스팅 라우터를 이용한다. 그러나 멀티캐스팅 라우터를 이용하여 멀티캐스트 서비스를 구현할 때, 사업적, 기술적 측면에서 많은 문제점을 가지게 된다. 이러한 문제점을 해결하기 위하여 애플리케이션 레벨에서 중계 기능을 수행하는 모듈을 이용한다. 기존에 있는 네트워크망을 그대로 이용하면서, 멀티 캐스트 서비스를 제공한다. 본 논문에서는 애플리케이션 레벨 멀티캐스트 서비스를 지원하기 위하여 세션망 가임, 세션망 탈퇴, 각 프로토콜에 대해서 정의한다. 정의된 방식을 이용하여 토폴로지를 형성하고, IP 멀티캐스팅 서비스와 응용 계층 멀티캐스트를 이용 하였을 때 패킷 손실의 차이를 NS2를 이용하여 알아본다.

1. 서 론

멀티캐스트 서비스는 일대다, 다대다 방식으로 스트림 데이터를 전송하는 방식이다. 이러한 전송 방식들은 주로 비디오 컨퍼런싱, 분산 소프트웨어, 인터넷 TV, 공유 화이트 보드와 같은 프로그램에 사용된다. 이와 같은 프로그램들은 다수의 수신자들이 동시에 접속하여 서비스를 제공받기 때문에 많은 대역폭 할당량을 요구하기 때문에 서비스를 제공하는 링크상에 데이터 폭주로 인한 오버헤드가 발생한다. 이러한 문제점을 해결하기 위하여 멀티캐스팅 라우터를 이용해 전달하는 방식과 내부적으로 유니캐스트 방식을 이용하여 여러 수신 노드들에게 전송하는 방식을 이용한다. 멀티캐스팅 라우터를 이용할 경우, 멀티 캐스트 서비스에 대한 수익 모델 부재로 인한 사업적 측면의 문제점과 멀티 캐스트 라우터의 불안정성, 트래픽 제어의 불안정성에 관한 기술적 문제점을 가지고 있다. <그림 1. (2)> 그리고, 내부적으로 유니캐스트를 이용하여 멀티 캐스팅 서비스를 이용할 경우, 동시 접속자가 너무 많아 송신측망의 게이트웨이를 통하여 나가는 부분에서 트래픽 혼잡으로 인한 많은 패킷 손실이 발생하게 된다.<그림 1. (1)> 이러한 문제점을 해결하기 위하여 응용 계층 멀티캐스트(Application level multicast) 전송 방식을 이용한다. Relayed Multicast 또는 Overlay Multicast라고 칭하기도 한다. 이 전송 방식은 멀티캐스팅 라우터에서 처리하는 데이터 복사, 경로 배정 같은 업무를 중간 노드의 애플리케이션 계층에서 처리하는 방식이다. <그림 1.(3)>



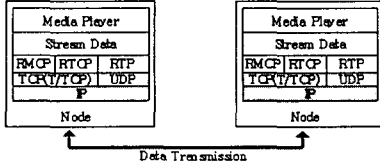
<그림 1> 멀티캐스트 지원 방식

현재 응용 계층 멀티캐스트를 관한 연구들이 활성화 되고 있다. 중간 노드간의 연결을 구성하는 방식에 따라서 크게 메쉬, 트리 방식으로 나눌 수가 있다. 메쉬 방식에는 Narada[1], ALMI[2], Scattercast 방식이 있으며, 트리 방식에는 TBCP[3], TAG[4], Peercast, Host Multicast, Overcast, Yoid 방식이 있다. 그리고, 메쉬, 트리 방식이 아닌 독창적인 방식으로 중간노드들을 구성하는 NICE, CAN[5] 방식이 있다. 위와 같은 연구들은 소규모 단위 그룹 통신에서 신뢰적으로 서비스가 이루어질 수 있도록 고안되었다. 그리고, 데이터의 송, 수신을 담당하는 Data Plane과 세션내 멤버 구성, QoS를 모니터링하는 Control Plane이 통합적으로 구성되어 있다. 본 논문에서는 Data Plane과 Control Plane을 분리하여 독립적으로 데이터 스트림, 제어 메시지를 운영할 수 토폴로지를 구성한다. 이러한 구성을 위해서 RMCP[6] 프로토콜을 이용하고 중간 노드간의 구성은 QoS를 고려한 트리를 기반으로 구성한다.

본 논문의 구성은 다음과 같다. 2장에서 RMCP 프로토콜에 대하여 알아보고, 3장에서 멀티캐스트 그룹 세션망 관리를 소개한다. 4장에서 실험 및 향후 연구 방향에 대해 알아본다.

2. RMCP 프로토콜

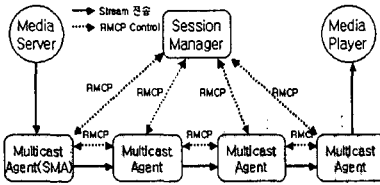
본 장에서는 멀티캐스팅 토폴로지를 형성하기 위하여 사용하는 RMCP 프로토콜에 대하여 간단히 소개한다. RMCP는 Data Plane과 Control Plane이 구분되어 기존의 미디어 전송 프로그램에 별도의 조작 없이 사용할 수 있다.



<그림 2> 프로토콜 스택

2.1 RMCP의 기본 구조

<그림 3>에서는 RMCP의 기본 전송 과정을 나타낸다. 멀티미디어 전송을 위하여 SMA(Sender Multicast Agent)가 서비스를 제공하고 중간 노드인 MA(Multicast Agent)가 중계 역할을 하면서 미디어 재생기까지 스트림을 전송하는 과정을 보여준다.



<그림 3> RMCP를 이용한 스트림 전송

2.2 RMCP 프로토콜 메시지

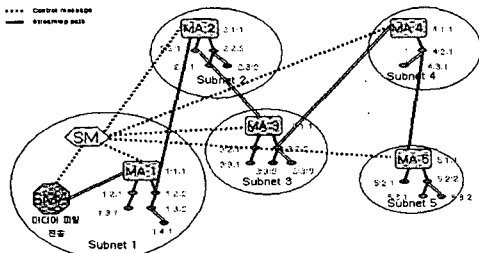
RMCP 세션 관리를 위하여 <표 1>과 같은 Control 메시지를 이용한다. 각 메시지는 세션 참가 (JR, JC), 미디어 스트림 전송 요청 (RR, RC), 및 QoS 관리 (SR, SC)를 위해서 사용된다.

Message	RMCP Operation	From	To
Join Request(JR)	Operation	MA	Session Manager
Join Confirm(JC)	Session Join	Session Manager	MA
Relay Request(RR)	Data Channel	DownStream MA	UpStream MA
Relay Confirm(RC)	Control	UpStream MA	DownStream MA
Status Report(SR)	Session	MA	Session Manager
Status Confirm(SC)	Monitoring	Session Manager	MA

<표 1> RMCP 제어 메시지

3. Session 관리

본 논문에서는 <그림 4>와 같은 망구조를 배경으로 설명한다. 멀티캐스트 서비스의 QoS를 고려한 확장성, 결합 허용, 신뢰성을 관리하기 위하여 SM과 DMA(Designated MA: Subnet에서 가장 상위 노드)는 각각 <표 2>, <표 3>과 같은 정보 테이블을 가진다.



<그림 4> RMCP를 이용한 미디어 전송 구조

From	To	Gateway	Tree ID	Dest IP addr	subnetmask	aver bw	last node
SMA	MA1	164.125.164.1	1:1:1	### ### ###	255.255.252.0	325kb/s	1:2:1
MA1	MA2	2:1:1	### ### ###	255.255.255.0	294kb/s	2:2:1
MA2	MA3	3:1:1	### ### ###	255.255.255.0	265kb/s	3:2:2
MA3	MA4	4:1:1	### ### ###	255.255.0.0	243kb/s	4:2:1
MA4	MA5	5:1:1	### ### ###	255.255.128.0	229kb/s	5:2:1

<표 2> SM의 정보 테이블

Tree ID	From	To	Gateway	IP addr	subnetmask	bw
1:2:1	1:1:1	1:2:1	164.125.164.1	### ### ###	255.255.252.0	473kb/s
1:2:2	1:1:1	1:2:2	### ### ###	255.255.252.0	323kb/s
1:3:1	1:2:1	1:3:1	### ### ###	255.255.252.0	312kb/s
1:3:2	1:2:2	1:3:2	### ### ###	255.255.252.0	264kb/s
1:4:1	1:3:2	1:4:1	### ### ###	255.255.252.0	253kb/s

<표 3> MA1의 정보 테이블

정보 테이블에서 aver bw는 각 subnet망에서 사용되고 있는 평균 대역폭을 나타낸다. <수식 1>을 이용하여 구한다.

$$\text{average bw} = \sum_{i=1}^n N_{bw} / N_n \quad \text{<수식 1>}$$

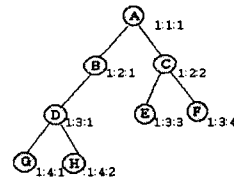
N_{bw} : subnet망에서 각 노드간의 link의 bandwidth

N_n : subnet망에 속해 있는 노드 수

이 평균 bw는 새로운 노드가 세션 가입시 적절한 Subnet망을 찾기 위해서 사용된다. 자세한 내용은 3.2절에서 설명한다.

3.1 Tree 구성

본 논문에서는 'RMCP' 세션에 참가하는 노드들을 식별하기 위하여 Tree ID를 제안한다. Tree ID는 노드의 식별성을 높여주고, Tree 재구성이 쉽게 이루어질 수 있도록 한다. <그림 5>는 하나의 Subnet 구성을 Tree로 나타낸 것이다. Tree ID 부여 방식은 아래의 규칙에 따른다.



<그림 5> Tree ID

Tree ID =
subnet number : level number : node number

▶ Initial node number = $F_n \times (N_{upper} - 1) + 1$

▶ Final node number = $F_n \times N_{upper}$

예를 들어, E노드의 경우 subnet number가 1, level number가 1, node number는 첫번째 노드이기 때문에 initial node number의 수식에 의해서 3이다. 그러므로 C노드의 첫번째 자식 노드는 1:3:3의 Tree ID를 부여하게 된다. F_n 은 부모 노드가 가질 수 있는 자식 노드의 수를 나타내며, N_{upper} 는 부모 노드의 node number를 나타낸다.

3.2 Session 가입

세션내에서 미디어 스트림 서비스를 이용하기 위해서는 세션 멤버로 가입을 해야 한다. 본장에서는 <그림 4>에서 새로운 노드가 가입하는 방법을 설명한다. 세션에 가입하는 방법은 2가지로 나뉜다. 먼저 새로운 노드가 접근하기 위한 하부망을 찾기 위한 inter connection 접근 방식과 하부망에서 부모 node를 찾기 위한 intra connection 방식이 있다.

▶ Inter Connection 접근 방식

1) 현재 Active한 MA들 중 동일한 subnetmask가 있는지를 검색한다.

2) 동일한 하부망이 있을 경우, Intra Connection 접근 방식으로 접근 시도한다. 만약 동일한 하부망이 없을 경우, dynamic, static 접근 기법을 이용하여 새로운 노드가 연결할 하부망을 검색한다.

dynamic 접근 방식은 <표 2>에서 현재 서비스하고 있는 미디어의 비트 전송률(프레임 전송율)과 근접한 Best-fit 방식으로 하부망을 찾는 방식이다.

static 접근 방식은 <표 2>에서 최상위 정보가 새로 진입하는 하부망으로 지정하는 방식이다.

네트워크가 형성되어 있는 Active한 환경에서라면 static방식보다 dynamic방식이 좀 더 QoS를 고려한 방식이라 할 수 있다. 그래서 본 논문에서는 dynamic 접근 방식을 이용한다.

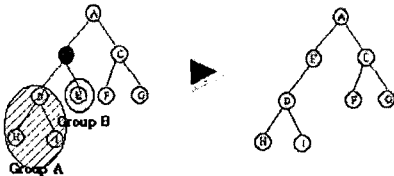
▶ Intra Connection 접근 방식

Inter Connection에서는 새로 진입할 노드가 연결될 subnet을 선택한다. 그리고, 진입 노드는 다시 subnet의 DMA에게 가입 요청을 하고, subnet내에서 적절한 부모 노드를 찾아야 한다. 부모 노드를 찾기 위한 조건은 아래와 같다.

- subnet내에서 최대의 bw값을 가진 노드를 검색한다.
- child를 허용하는 노드이어야 한다.
- fan-out(자식 노드 허용 범위) 노드를 체크한다.
- Tree level의 depth가 최소인 부모 후보 노드를 찾는다.

3.3 Session 탈퇴

앞장에서 세션망에 가입하는 절차에 대하여 알아보았다. 본 장에서는 세션망에서 탈퇴하는 과정에 대하여 알아보겠다. <그림 6>에서 B 노드가 탈퇴하는 과정을 나타낸다.



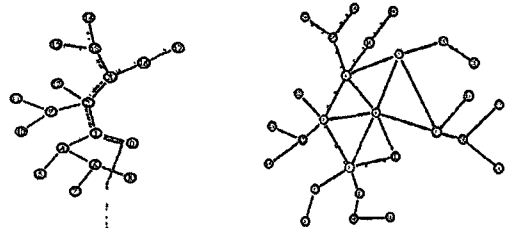
<그림 6> 세션 탈퇴

▶ 탈퇴 과정

- 1) B노드는 A,D,E(부모, 자식 노드)노드에게 leave signal을 보낸다.
- 2) B 노드는 미디어 재생을 멈추고, 스트림의 라우팅 역할을 지속적으로 진행한다.
- 3) B노드를 기준으로 대역폭의 합이 작은 그룹의 최상위 노드를 찾는다. <그림 6>에서 E노드의 그룹이 대역폭의 합이 작으므로 E노드를 B노드의 자리로 대체한다.
- 4) E노드는 D,H,I를 오른쪽 그림과 같이 차례로 연결한다. <그림 6>에서 만일 Group A가 level의 depth가 깊고, fanout의 조건에 위배된다면 3,4항목을 반복적으로 수행한다.

4. 실험 및 향후 연구 방향

지금까지 RMCP 프로토콜 기반으로 망구조를 형성하고, 미디어 스트림을 위한 세션 참가, 탈퇴 그리고, 트리 형성에 대하여 알아보았다. NS-2를 이용하여 IP 멀티캐스트 전송과 본 논문에서 제시한 망구조상에서의 전송을 아래의 그림과 같이 실험을 하였다. 각 망구조에서는 하나의 서버(<그림 7>상의 파란색 노드)에서 8개의 client에게 3가지 스트림의 전송 서비스를 제공하여 준다. 그림에서 나타내듯이 IP 멀티캐스트 서비스에서는 패킷 dropping으로 인한 데이터 손실이 일어나고 있는 반면에 응용계층 멀티캐스트에서 데이터 손실이 없이 서비스를 제공하고 있음을 알 수 있다.



(a) IP 멀티캐스트 (b) 응용 계층 멀티캐스트

<그림 7> 실험 구조

본 논문에서는 실시간 미디어 전송을 전제로 서비스를 형성하였다. 즉, 미디어 데이터를 받는 즉시 사용자에게 플레이 하여 주는 방식을 말한다. 이러한 서비스는 하나의 문제점을 가지게 된다. 세션 탈퇴시 트리 재구성으로 인한 재생 끊김 현상이 발생하게 된다. 이 문제점을 보완하기 위해서 일정한 크기의 버퍼를 이용하는 방법에 대한 연구를 진행 할 것이다.

참고 문헌

- [1] Y.-h. Chu, S. G. Rao, and H.Zhang, "A Case for End System Multicast", ACM SIGMETRICS 2000, June 2000
- [2] D. Pendarakis, S. Shi, D. Verma and M. Waldvogel, "ALMI: An Application Level Multicast Infrastructure", the 3rd Usenix Symposium on Internet technologies & Systems(USITS 2001), March 2001
- [3] L. Mathy, R. Canonico and D. Hutchison, "An Overlay Tree building Control Protocol", NGC 2001, November 2001
- [4] M. Kwon and S. Fahmy, "Topology Aware Overlay networks for Group Communication", Proceeding of NOSSDAV'02, May 2002
- [5] Sylvia Ratnasamy, Mark Handley, Scott Shenker, Richard Karp, "Content Addressable Network", ACM SIGCOMM2001, August 2001
- [6] Seok Joo Koh, "Specification of Relayed Multicast Control Protocol", Working in Progress in Question 8 of ITU-T SG17. ITU-T SG17 X.rmcp. November 2002