

NetFPGA를 이용한 멀티캐스트 오버레이 네트워크 설계

유태완, 최훈규, 손석신, 권태경, 최양희
서울대학교, 컴퓨터 공학부

{twyou, hgchoi, ssson, tkkwon, yhchoi}@mmlab.snu.ac.kr

The Design of Multicast Overlay Network based on NetFPGA

Taewan You, Hoongyu Choi, Seokshin Son, Taekyong Kwon, Yanghee Choi

School of Computer Science & Engineering, Seoul National University

{twyou, hgchoi, ssson, tkkwon, yhchoi}@mmlab.snu.ac.kr

요약

1990년대 Multicast protocol이 제안된 뒤 IP multicast는 이후 10년 이상 많은 연구가 이루어졌다. 대부분의 multicast 연구는 데이터들을 효율적으로 그룹에게 전송하기 위한 optimization 및 utilization 관점에서 제안된 multicast tree 구축 알고리즘과 같은 것이었으며 이로 인해 실제 망에서 multicast를 운용하기에는 복잡하고 관리가 용이 하지 않는 deployment에 약점을 가지고 있다. 그러나 이 multicast는 미래 인터넷의 중요한 communication 형태로 다시 주목 받고 있다. 앞으로의 인터넷의 대부분의 traffic은 비디오와 오디오와 같은 멀티미디어 데이터가 주를 이루는 것으로 예측되며, IPTV와 Mobile IPTV, 그리고 YouTube 등과 같이 다양한 멀티미디어 데이터를 그룹간 공유하는 웹 2.0의 출현과 함께 multicast는 다시 한번 중요성을 인정 받게 되었다. 본 논문은 미래인터넷의 platform으로 사용되기 위해 제안된 programmable router인 NetFPGA를 이용하여 이러한 미래 인터넷에 적합한 multicast를 개발 및 검증을 위한 테스트베드로써 구축하기 위한 오버레이 네트워크의 구축과 관련된 설계 기법을 소개하고, 실제 구축된 멀티캐스트 오버레이 네트워크를 소개한다.

1. 서론

IP Multicast는 그룹간 통신을 위해서 효율적인 link 사용을 위해 S. Deering에 의해 1990년에 제안된 프로토콜이다 [1]. 이 IP Multicast의 개발 이후 10년 이상 많은 multicast 연구들이 이루어졌으며, IGMP, PIM [2][3] 등 관련 많은 protocol들이 개발되었다. 현재 대부분 라우터에는 multicast protocol이 탑재되어 있으며 호스트의 대부분 OS의 경우 IGMPv3 등이 구현되어 있는 상황이다. 그러나 이러한 multicast 서비스는 대부분 라우터 등에서 활성화 되어 사용되지 않고 있으며, IPTV와 같은 특정 서비스를 위해 한 ISP 도메인 안에서 public 인터넷 망과 분리되어 운영되고 있는 실정이다. 대부분의 multicast 연구는 데이터들을 효율적으로 그룹에게 전송하기 위한 optimization 및 utilization 관점에서 제안된 multicast tree 구축 알고리즘과 같은 것이었으며 이로 인해 실제 망에서 multicast를 운용하기에는 복잡하고 관리가 용이 하지 않는 deployment에 약점을 가지고 있다 [4].

그러나 이 multicast는 미래 인터넷의 중요한 communication 형태로 다시 주목 받고 있다. 앞으로의 인터넷의 대부분의 traffic은 비디오와 오디오와 같은 멀티미디어 데이터가 주를 이루는 것으로

예측되고 있으며, 특히 multicast와 같이 그룹 간 communication 형태의 지원을 받을 때 가장 효율적으로 서비스 될 수 있는 애플리케이션들이 많이 등장하고 있다. IPTV, 모바일 IPTV 기반의 개인 웹캐스팅, 비디오회의, MMORPG, 파일배포 등이 그 예들이다. 특별히, IPTV [5]는 미래 멀티캐스트를 활용한 킬러 애플리케이션(Killer application)이 될 것으로 전망된다. 모바일 IPTV는 서버나 콘텐츠 공급업자가 아닌 일반인들이 소규모 수신자들을 위한 방송의 소스가 되는 환경을 가정하고 있으며, 특별히 소스가 되는 사용자가 이동기기(Mobile devices)를 서버로 활용하면서 무선연결을 활용하여 이동하는 경우까지를 전망하고 있다.

또한, 일대일 통신이 주류를 이루는 현재와 달리 가까운 미래에는 소규모 그룹내의 그룹통신, 사회적 네트워크의 등장과 단 방향의 정보게시가 아닌 상호작용을 핵심으로 하는 웹2.0 [6] 등이 광범위하게 활용되면서, 통신의 기본 패러다임 자체가 multicast와 같이 그룹간 communication이 기본이 되는 특성을 필요로 하는 사회적 요구가 급격히 증가하고 있다.

이렇듯 미래 인터넷의 대부분 서비스가 multicast 형태의 서비스로 예측되기 때문에 실제 deployment 관점에서 새롭게 multicast protocol을

설계하는 등의 접근방법이 필요한 상황이다. 그러나 대부분 제안된 IP multicast의 경우 inter-domain 간의 scalability 문제와 여러 제약들을 가진 application을 이용한 overlay multicast 등이 실제 deployment 가능한 프로토콜로 여겨지고 있다.

따라서 본 논문은 실제 미래 인터넷과 같은 환경에서 새로운 형태의 multicast protocol들을 제안하기 위한 platform으로 사용을 위한 multicast overlay network를 설계 및 실제 구축하였다. 먼저 미래인터넷의 platform으로 사용되기 위해 제안된 programmable router인 NetFPGA [7]를 사용하였으며 이 NetFPGA 라우터를 multicast overlay를 구축할 수 있도록 proxy 기능을 추가하였다. Multicast proxy 기능을 구현하기 위해 ID-Locator 분리 아키텍처를 사용하였고 [8], multicast를 가능하게 하기 위한 전송 메커니즘을 설계 및 구현하였다. 결론적으로 본 논문을 통해 제안 및 구현된 NetFPGA 기반의 proxy overlay multicast 테스트 베드는 추 후 multicast 형태의 미래인터넷 서비스 개발 시 실제 테스트 베드로 활용 할 수 있으며, proxy 기능을 새롭게 추가하여 새로운 형태의 multicast protocol 제안 및 활용 할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서 NetFPGA 및 터널링 기반의 ID-Locator 분리 기술에 대해 소개하고, 3장에서 proxy 기반 multicast overlay network 설계에 대해 살펴본다. 4장에서는 실제 구축된 multicast overlay network를 소개하고 테스트된 성능에 대해 언급한다. 마지막으로 5장에서 앞으로의 작업에 대해 언급하고 결론을 맺는다.

2. NetFPGA와 ID-Locator 분리 기반 터널링 아키텍처

2.1 NetFPGA 소개

NetFPGA[12]는 학생 또는 연구자가 자신이 설계한 시스템을 하드웨어를 이용하여 구축할 수 있도록 해주는 플랫폼이다. FPGA (Field Programmable Gate Array)가 포함된 NetFPGA 카드를 표준 PC에 장착하면 하나의 NetFPGA 시스템이 완성되기 때문에 저렴한 가격으로 실험 환경을 구축할 수 있다. 또한 공개 소스를 기초로 하여 새로운 설계를 Verilog 코드로 작성하여 FPGA에 로딩함으로써 쉽고 빠르게 새로운 시스템의 프로토타입을 만들어 낼 수 있다.

NetFPGA 시스템은 하나의 라우터로써 동작한다. 라우터의 일반적인 Datapath 구조는 그림 8과 같다. 들어온 패킷은 대해 헤더 프로세싱을 거쳐 Output Port로 나가게 된다. Datapath 코드에서 Delay를 조절하여 라우터의 기능을 강화할 수 있다. Verilog 소스 코드의 확장자는 .v 파일로 되어 있으며

Emacs와 같은 텍스트 편집기를 이용하여 수정가능하다. 주의할 점은 Verilog 코드를 수정하고 이를 반영하기 위해서는 합성(Synthesis) 과정을 반드시 거쳐야 한다.

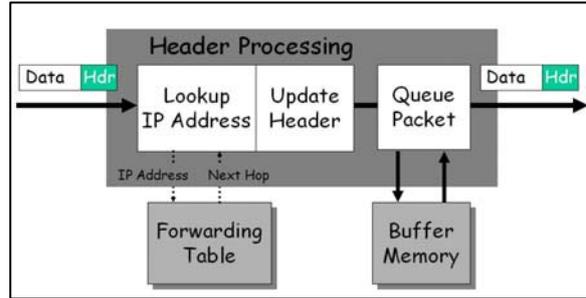


그림 1 라우터의 일반적인 Data path 구조

본 연구는 Forwarding table을 위해 제안된 control plane인 scone [7]에서 새롭게 ID-Locator 분리 기술을 적용하기 위해 map table과 이를 활용한 IP-in-IP 터널링 [8] 기능을 구현하였다. 이 터널링 기법은 현재 IETF를 중심으로 인터넷 확장성 문제를 해결하기 위해 연구되고 있는 ID-Locator 분리 기술 [9]을 활용하여 개발되었다. ID-Locator 분리 기술은 다음 절에서 소개한다.

2.2 ID-Locator 분리 기반 터널링 아키텍처

ID-Locator 분리를 위해서는 user network에서 사용되는 Identifier와 Locator 사이의 mapping을 담당하는 mapping table을 유지해야 한다. 그림 2는 실제 NetFPGA를 통해 ID-Locator 분리 네트워크 설계하여 구축한 네트워크이다.

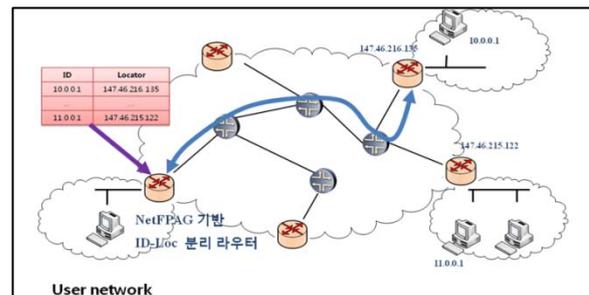


그림 2 NetFPGA기반 ID-Locator 분리 네트워크

각 NetFPGA 라우터는 통신 하고자 하는 user network와 터널링 end-point가 되는 라우터의 주소 - 이를 Locator로 사용한다 - 들을 유지하는 map table을 가지고 있다. 이 map table은 global network에 적용시키기 위해서는 LISP (Locator Identifier Separation Protocol)의 ALT (Alternative Topology) [9]와 같이 update 또는 distribute 하는 기능을 가져야 한다. 그러나 본 설계에서는 간단히 manual 하게 map table을 구축한다. 아래 표는 간단히 작성한 map table의 예이다.

표 1 Map Table

ID	Locator	RTT
10.0.0.1	147.46.216.135	84
...

이 map table은 먼저 forwarding 테이블에서 match가 되지 못하도록 ID를 private 하게 할당하고, 이를 외부로 announce 하지 않도록 하였다. 이 match 되지 못한 ip 패킷은 process_ip_packet()함수에 도달하여 처리를 받게 되는데, 이때 map table을 참조하여 encapsulation과 decapsulation이 되어 전송되게 된다.

3. Proxy Multicast Overlay Network 설계

본 연구는 앞 장에서 살펴본 것과 같이 ID-Locator 분리 기술로써 터널링을 사용하여 라우터와 목적지 라우터 사이를 1홉과 같이 오버레이 네트워크로 구축할 수 있다. 특히 ALT와 같이 ID와 Locator 사이의 mapping 정보를 global network에 분산 및 업데이트 시키는 별도의 시스템을 구축하면 쉽게 오버레이 네트워크를 구축할 수 있다.

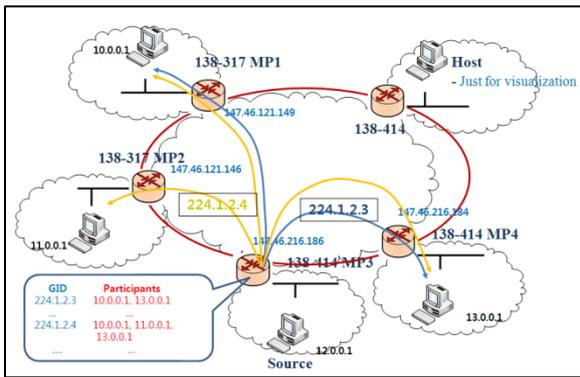


그림 3 NetFPGA 기반 Proxy multicast overlay network

이러한 오버레이 네트워크에 multicast와 같이 그룹 통신을 할 수 있는 형태의 라우팅 기능을 위해서는 위의 그림 3과 같이 multicast routing table을 유지해야 한다. 즉 멀티캐스트 서비스를 위해 ID로 사용되는 multicast address와 함께 그 그룹 address에 참가하는 참가자의 EID 정보들을 유지하고 있어야 한다. 각 EID는 map table을 통해 locator로 변환되어 전송될 수 있다. 본 연구에서는 NetFPGA의 그림 1의 packet 처리 알고리즘 중 control plane을 확장하여 각 group ID에 참여하고 있는 ID의 수만큼 packet을 복사하여 전송할 수 있는 multicast 기능을 구현하였다. 또한 router와 host 간의 multicast 처리를 위한 기능 역시 구현하였다.

3.1 Multiple packet 전송 구현

NetFPGA 라우터는 기존map table인 maplocal과 multicast를 위해 동적으로 생성되는 multicast table인 mapremote라는 두 가지의 테이블을 유지 관리한다.

maplocal은 static하게 scone시작 시점에서 라우터에 map table을 입력할 때 사용한다. 본 연구에서 map table을 텍스트파일에 미리 저장해 두고, scone 가동 시 maplocal로 읽어 들이게 함으로써 map table을 생성했다.

반면 mapremote는 동적으로 multicast table을 유지할 때 사용된다. receiver가 join메시지를 보내면 receiver proxy는 이 메시지를 encapsulation하여 source proxy로 보내게 되는데, source proxy는 mapremote에 동적으로 변화하는 GID와 참가자의 ID를 기록한다. 마찬가지로 라우터에 leave 메시지가 도달한 경우, 라우터는 mapremote에서 해당 entry를 삭제한다.

maplocal과 mapremote는 linked-list로 구현하였다. 하나의 GID에 해당하는 참가자 ID는 여러 개가 될 수 있으며, 라우터는 자신의 maplocal을 활용하여 encapsulation을 한 후 참가자의 ID 개수만큼 패킷을 만들어 전송한다.

3.2 Multicast session 관리 기능 구현

End node와 NetFPGA 라우터 사이에서 multicast를 지원하기 위해서는 IGMP [2] 기능이 라우터에 구현되어야 한다. 기본적으로 end node의 대부분 OS에서는 IGMPv3를 지원하므로 라우터는 multicast group에 대한 query 또는 정보를 알리기 위한 제어 메시지와 timer를 구현하여 주기적인 query 또는 reply 메시지를 보낼 수 있는 기능을 구현하였다.

4. Proxy multicast overlay network 테스트 베드 및 성능 평가

본 장은 앞서 설명한 NetFPGA를 이용한 Proxy 기반의 multicast overlay network를 실제 서울대학교 학내 망에 연결하여 구축한 테스트베드를 소개한다.

아래 그림 4는 서울대학교 학내 망에 설치된 NetFPGA 기반의 proxy 멀티캐스트 라우터이다. 그림 3과 같이 5대의 proxy 라우터를 서울대학교 138동 연구동의 두 연구실의 학내 스위치에 연결하였으며, 각각 5개의 호스트를 연결하여 vlc [10]를 이용하여 video stream 테스트를 수행하였다.

테스트베드를 위해서 3Com Switch 2824에 직접 NetFPGA 라우터를 연결하였으며, 학내 망의 안정성을 위해 147.46.216.0 서브넷과 147.46.215.0 서

브넷만을 이용하였다. 실험은 200Kbps 의 video file을 vlc 서버모드를 활용하여 multicast address 를 목적으로 하여 UDP 스트리밍을 시작하였고, 최종 4개의 host가 join 하는 상황에서도 별다른 문제 없이 잘 동작함을 볼 수 있었다.

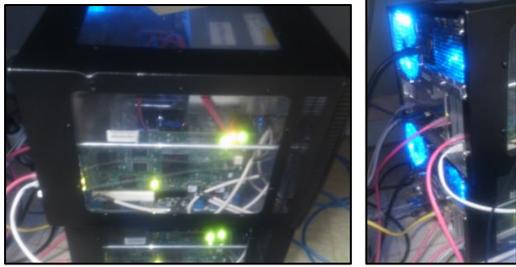


그림 4 NetFPAG proxy multicast 라우터

NetFPAG가 사용한 link는 10Gbyte Ethernet link로써 hardware forwarding을 사용하였을 경우 약 7.5 ~ 8.2 Gbps 의 성능을 보였으나, 본 논문에서 구현된 multicast proxy 라우터는 software 부분을 활용하여 작성되어 대부분의 packet들이 Linux의 user level application과 같이 처리되었다. 따라서 실제 10GB의 link를 사용하였으나 측정된 link throughput은 1000배 이상이 느린 성능을 보였다. 즉 터널링을 위해 encapsulation과 decapsulation을 하는 process가 hardware 처리가 아닌 Linux의 user 레벨로 packet 처리가 됨으로 급격한 성능저하가 발생됨을 알 수 있었다. 아래 그림은 200Kbps로 받은 UDP packet을 나타낸다.

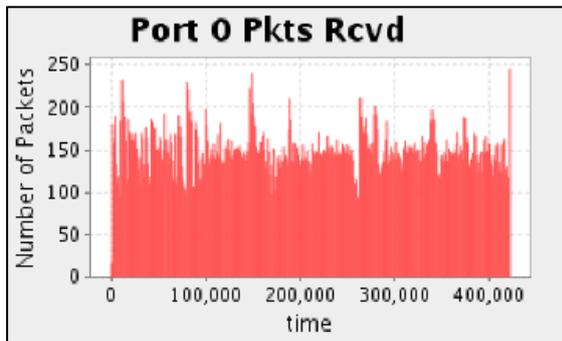


그림 5 NetFPAG 라우터의 received data

5. 결론 및 향후 계획

본 논문은 실제 미래 인터넷과 같은 환경에서 새로운 형태의 multicast protocol들을 제안하고 검증하기 위한 platform으로 사용하기 위해 multicast overlay network를 설계 및 구축하였다. 특히 미래 인터넷을 설계하기 위해 제안된 programmable router인 NetFPGA를 사용하였고, ID-locator 분리 기법을 활용한 오버레이 네트워크 설정과 multicast 기능을 위한 proxy 기능 구현을 통해 손쉽게 multicast 테스트 베드를 구축할 수 있게 하였다.

본 논문을 통해 제안된 NetFPAG proxy는 서울대 학내 망에 연결하여 테스트베드를 구축하고, vlc 를 활용하여 쉽게 multicast가 가능함을 보였다.

그러나 현재 구현된 proxy는 software를 기반으로 하였기 때문에 1000배 이상의 link 성능 저하를 보이는 치명적인 단점을 가지고 있으나, 추 후 hardware 기반의 forwarding 기능을 구현한다면 약 7-8G의 link 속도를 내 실제 활용 가능한 platform으로 사용될 수 있을 것을 예상할 수 있다.

감사의 글

본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었고 (NIPA-2009-C1090-0902-0006)", 지식경제부 및 정보통신연구진흥원의 IT산업원천기술개발사업의 일환으로 수행하였음 [2007-F-038-03, 미래 인터넷 핵심기술 연구]. 이 연구를 위해 연구장비를 지원하고 공간을 제공한 서울대학교 컴퓨터연구소에 감사 드립니다.

참고문헌

- [1] S. Deering, "Host Extensions for IP Multicasting," IETF RFC 1112, Aug. 1989.
- [2] B. Cain, etc, "Internet Group Management Protocol, Version 3," IETF RFC 3376, Oct. 2002
- [3] B. Fenner, etc, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)," IETF RFC 4601, Aug. 2006. IGMP
- [4] S. Ratnasamy, A. Ermolinskiy and S. Shenker, "Revisiting IP Multicast," SIGCOMM'06, 2006.
- [5] Y. Xiao, etc, "Internet Protocol Television (IPTV): The Killer Application for the Next-Generation Internet," IEEE Communications Magazine, November 2007.
- [6] T. O'Reilly, "What Is Web 2.0 - Design Patterns and Business Models for the Next Generation of Software,"
- [7] NetFPGA homepage, <http://www.netfpga.org>
- [8] W. Simpson, "IP in IP Tunneling," IETF RFC 1853, Oct. 1995.
- [9] V. Fuller, etc, "Locator/ID Separation Protocol (LISP)," IETF draft-ietf-lisp-05.txt, Sep. 2009.
- [10] Video LAN, <http://www.videolan.org/>