

[인터넷] 가상화된 대규모 데이터 센터 망에서의 ARP/ND 확장성 이슈

인터넷 프로토콜 표준화 기구인 IETF에서는 최근 들어 가상화(virtualization) 기술을 사용하는 대규모 인터넷 데이터 센터 망에서 발생할 수 있는 이슈들에 대한 논의를 ARMD(Address Resolution for Massive number of hosts in Data Center) BOF를 중심으로 활발히 진행하고 있다. 2010년 11월 초 북경 IETF 79차 회의에서 있었던 첫 번째 ARMD BOF 회의에서는 ARMD 관련 problem statement, VM migration에 따른 문제, 이들 문제를 해결하는 것과 관련해서 ARP 감소, Moose, L2 VPN, virtual subnet 등에 대한 발표가 있었다. BOF는 워킹그룹(WG) 형성을 목표로 임시 형성된 모임으로 IETF 79차 회의에서 있었던 ARMD BOF 미팅에서도 ARMD WG 형성 가능 여부에 대한 열띤 토론이 있었다.

본 고에서는 먼저 가상화된 대규모 데이터 센터 망의 특성에 대해 살펴보고, 이와 관련해서 ARMD WG을 통해 다루고자 하는 이슈에 대해서 소개한다. 또한 ARMD WG 형성을 위해 현재 논의되고 있는 내용에 대해서도 간략하게 소개한다.

가상화된 대규모 데이터 센터 망의 특성

대규모 데이터 센터에는 다수의 서버와 호스트, 클라이언트 서브넷이 존재하며, 데이터 센터 자원을 효율적으로 사용하기 위해 호스트를 가상 머신(virtual machine, VM)화 해서 서버에 할당하는 서버 가상화(virtualization) 기술이 최근 들어 많은 관심을 받고 있다. 가상화된 데이터 센터 환경에서도 목표(target) 호스트의 IP 주소를 획득하기 위해서는 ARP/ND(Address Resolution Protocol/Neighbor Discovery) 메시지를 사용해야 한다. 서버 가상화로 인해 서버 한 대가 수백 대 이상의 VM을 지원할 수 있으며, 이로 인해 빈번한 ARP/ND 메시지의 브로드캐스트/멀티캐스트 또는 unknown flooding이 발생하는 문제가 생길 수 있다. 호스트의 VM이 과부하된 서버에 할당되어 있는 경우 이 VM을 부하가 적은 서버로 재할당할 수 있으며, 이것을 VM migration이라고 한다. VM migration으로 인해 가상 서브넷이 네트워크 전반에 걸쳐서 존재할 수 있고, 그 결과 ARP/ND 트래픽이 네트워크 상으로 퍼져나가는 문제가 생긴다. 또한 하나의 서브넷에 속한 호스트들이 여러 TOR(Top of Rack; 또는 access switch) 스위치에 걸쳐서 존재하는 경우 TOR 스위치의 FDB(Filtering Data Base)를 재갱신(refresh)하기 위해 ARP/ND 메시지를 빈번하게 브로드캐스트/멀티캐스트 하는 문제도 발생한다.

ARMD WG에서 다루고자 하는 이슈들

호스트가 목표 호스트의 MAC 주소를 찾기 위해 ARP request 메시지를 브로드캐스트 하는 오버헤드를 줄이기 위해 호스트에 MAC 주소와 IP 주소 간의 매핑 정보를 유지하는 ARP 캐시가 있다. 그러나 ARP 캐시에 있는 MAC-IP 주소 간의 매핑 정보가 일정 시간(예, 2분) 이내에 사용되지 않는 경우 이 정보를 버리며 또한 만일 사용되고 있더라도 일정 시간(예, 10분)마다 ARP 메시지를 브로드캐스트 한다. 이러한 ARP 기능은 서브넷(브로드캐스트 도

메인) 규모가 작은 경우는 문제가 없으나, 서브넷 규모가 커지게 되면 그에 비례해서 ARP 메시지 전송 오버헤드가 증가하는 문제가 생긴다. 하나의 서버가 관리하는 VM 수가 수 백대 이상이 될 것으로 예상되는 가상화된 데이터 센터 망에서 서버가 처리해야 하는 ARP 메시지 오버헤드는 상당할 것으로 예상된다.

여기에 VM이 한 서버에서 다른 서버로 이동하는 경우 ARP 브로드캐스트 문제는 더 심각해질 것이다. VM migration은 VM이 동일한 2계층 네트워크 내의 한 서버에서 다른 서버로 이동하는 것을 뜻하며 이를 통해 서버 효율 및 데이터 센터 적응성(agility)을 향상시킬 수 있다. 이때 VM은 장소 변경과 무관하게 IP 주소와 MAC 주소를 동일하게 유지하며, 따라서 VM이 새로운 장소로 이동하면 이 VM이 할당된 서버의 TOR이 이 VM과 같은 브로드캐스트 도메인에 속해 있는 호스트들이 할당된 서버의 목표 TOR이 어디에 있는지 알지 못하기 때문에 gratuitous ARP 메시지를 모든 TOR 스위치로 브로드캐스트 하게 된다.

또한 하나의 데이터 센터에 4095개 이상의 VPN이 존재하는 경우 상이한 VPN에 속하는 호스트들을 분리하기에는 VPN의 수가 부족하며, 이를 해결하기 위해 MAC-in-MAC 인캡슐레이션 등을 사용해서 추가로 분리해주어야 한다. 이 경우 VM이 이동하면 이동한 곳의 새 TOR이 인캡슐레이션된 데이터 프레임의 외부 헤더에 있는 목적지 주소를 모르기 때문에 이 데이터 프레임을 플러딩하는 문제가 생긴다.

ARP 브로드캐스트 문제를 해결하는 가능한 방법으로는 TOR이 ARP 캐싱을 하고 프록시 역할을 하는 방법 또는 디렉토리 기반 방법 등이 있을 수 있으며, TOR이 프록시 역할을 하는 경우 VM migration시 ARP 브로드캐스트 문제를 해결하지 못하는 단점이 있다.

ARMD BOF 이후의 ARMD WG 형성을 위한 노력

IETF 79차 회의의 ARMD BOF에서는 VM migration에 따른 이슈에 대한 다양한 해결방안들에 대한 발표가 있었고, ARMD에서 해결하고자 하는 이슈에 대한 충분한 분석 없이 해결방안을 제시하는 것은 문제가 있다는 의견이 많았다. 회의 마지막에 ARMD WG 형성에 대한 참석자들의 의견을 물은 결과 워킹그룹 형성의 필요성은 인정하나 problem scope에 대한 조정 및 분석이 선행되어야 한다는 의견이 대다수였다.

이 ARMD BOF 이후 ARMD 메일링 리스트를 통해 ARMD WG 형성을 위한 논의가 활발히 진행되어 왔으며, ARMD BOF에서 지적되었던 scope 조정을 위한 논의가 있었고 그에 따라 수정된 ARMD WG의 charter(안)에서는 가상화된 데이터 센터 환경에서 ARP/ND의 확장성 특성을 분석 등을 통해 파악하고 이러한 확장성 문제로 인한 제약사항을 식별하며 이러한 제약사항을 극복할 수 있는 망 설계와 관련된 권고를 제시하는 것으로 되어 있다. 수정된 ARMD WG charter(안)에서는 ARP/ND 확장성 문제 해결방안 도출을 표준화 항목에서 제외시켰으며, 향후 상황에 따라 추가적으로 고려할 수도 있는 것으로 여지를 남겨 놓았다.

안상현 (서울시립대학교 컴퓨터과학부 교수, ahn@uos.ac.kr)