

IPv6 에서 멀티호밍을 위한 BGP 확장

우종욱*, 류승협*, 이재용*

정재훈**, 박정수**

*연세대학교 High Performance Multimedia Network Lab.

**ETRI

e-mail : jongwook@alumni.usc.edu

BGP Extension for Multihoming in IPv6

Jongwook Woo*, Seunghyup Ryu*, Jaiyong Lee*

Jaehoon Chung**, Jungsoo Park**

*Dept. of Electronic Engineering, Yonsei University

**ETRI

요 약

기존의 인터넷 프로토콜 IPv4 의 32-bit 주소 field 는 미래의 인터넷 관련 제품의 수요 증가에 미치지 못하는 한계성을 내포하고 있다. 따라서, 128-bit 주소 field 를 가진 IPv6 에 관한 연구가 IPv6 포럼을 중심으로 진행되고 있다. 본 논문은 IPv6 의 한 관심분야인 멀티호밍 특히 Site Exit 라우터들 간의 멀티호밍 기법에 대해 서술하였다. 또한, IPv4 및 IPv6 에서 기존의 Site Exit 라우터들 간에 널리 쓰이는 BGP Routing 알고리즘을 멀티호밍에 적합하도록 수정한 알고리즘을 제안하였다. 그러므로, 이 알고리즘을 통한 Redundancy 와 Load-Sharing 의 이점을 살릴 수 있으며 BGP 의 수정은 적으므로 기존의 라우터를 어렵지 않게 멀티호밍에 적합하도록 변환 시킬 수 있다.

1. 서론

차세대 인터넷 주소방식인 IPv6 는 Address 주소를 기존 IPv4 의 32 bit 에서 128 bit 로 늘렸다. 따라서, IPv4 의 주소 부족 문제를 해결하였고 Computer Host 나 주변기기 뿐 아니라 다양한 단말들 - 핸드폰 이나 가전제품 - 의 인터넷 접속을 가능하게 한다. 또한, 자동 주소 할당등 많은 장점을 가진다. 그러나, IPv4 를 대상으로 연구 중인 여러 문제들이 IPv6 에서도 연구되어야 할 필요가 있다. 그 중의 하나인 멀티호밍 (Multihoming) 문제를 본 논문에서 서술한다.

멀티호밍은 어떠한 Site 나 Server 또는 AS (Autonomous System) 등이 복수의 인터페이스를 가지고, 각각의 인터페이스가 ISP 나 AS 로 연결되어 복수의 ISP 나 AS 를 가지게 되는 방식이다. 따라서, 이러한 특징을 이용하여 하나의 ISP 혹은 AS 로 가는 연결이 문제가 생기더라도, 다른 ISP 나 AS 를 통해 인터넷에 접근할 수가 있다. 반대로, 외부에서 복수의 연결 중 1 개를 통하여 Multihomed Site 혹은 Server 에 접근할 수 있게 된다. 정상시에는 복수의 인터페이스, ISP 와 AS 를 가지므로, 이들 모두를 통해 인터넷에

접근할 수가 있게 되는데, 여기에서 효율적인 Load Sharing 을 제공하거나, 유사시를 대비한 Redundancy 를 고려 하여 네트워크 효율을 높일 수 있다.

본 논문에서는 멀티호밍을 한 AS 에 있는 Exit 라우터 Group 에 복수의 외부 AS 가 있는 방식 또는 한 외부 AS 에 복수의 라우터들이 있는 방식으로 정의한다. 따라서, 외부 AS 간에 가장 널리 쓰이는 BGP Routing 알고리즘에 대하여 분석한다. 현재, IPv6 을 이용한 멀티호밍에 대한 연구가 여러 측면에서 이루어지고 있다. IPv6 멀티호밍이 갖추어야 할 요구사항등에 대한 draft 가 IETF working group 에서 제시 되고 있으며, Load Sharing, Address Selection, 라우터에 대한 사항, Simplicity 및 그 외에 필요한 사항등을 포함 하고 있다. 또한, 멀티호밍에서의 Default 라우터 Selection, Backbone 이 복수일 경우의 Backbone Selection 등의 제안 및 연구결과도 발표가 되고 있다 [1, 2, 4, 5, 6, 7, 8, 9].

현재의 멀티호밍은 주로 단순히 유사시를 대비한 수단으로 이용되고 있다. 따라서 정상시에 복수의 연결을 모두 사용할 수 있을때에도, 특별한 관리 없이

Traffic 을 주고 받는 방식을 사용하고 있다. 이는 네트워크 자원 측면에서 효율적이지 못하다. 효율적인 자원 관리를 위하여 연결의 단절 등의 문제를 대비하는 Redundancy 뿐 아니라 Load Sharing 등의 이용이 필요하다.

본 논문에서는 Redundancy 를 위한 기존의 멀티호밍 방식 [1, 2] 과 Load Sharing 에 관하여 소개한다. 그리고, AS 들 내의 Site Exit 라우터들 간의 BGP routing 알고리즘을 설명하고 멀티호밍을 구현하기 위하여 BGP Routing 알고리즘을 수정한 BGP 를 제안한다. 제 2 장은 Redundancy 구현을 위한 Site Exit 라우터들 간의 기존의 멀티호밍 구조와 Load-Sharing 에 대하여 서술한다. 제 3 장은 기존의 BGP4+ Routing 알고리즘을 소개한다. 제 4 장은 본 논문에서 제안하는 수정된 BGP4+ Routing 알고리즘을 설명한다. 마지막으로 결론을 나열한다.

2. IPv6 에서의 멀티호밍 방식

2.1. Site Exit 라우터들 간의 멀티호밍 구조

RFC 2260 [1] 에서는 Site Exit 라우터들 간의 멀티호밍 구조를 크게 2 가지로 제시하고 있다. 하나는 Auto Route Injection 방식이고, 다음은 Non-direct eBGP Peering 방식이다. 이 둘은 복수의 Site Exit 라우터와 복수의 ISP 를 가지는 상황을 보여주고 있다. 본 논문에서는 ISP 대신 보다 광범위 한 개념인 AS 로서 멀티호밍 구조를 설명한다.

가. Auto Route Injection

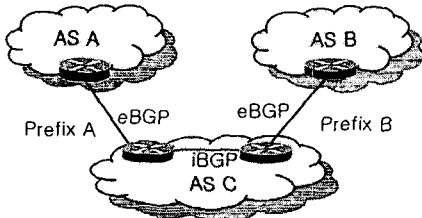


그림 1. Auto Route Injection

그림 1 에서 AS A, B, C 에 있는 각 Site Exit 라우터는 각각의 이웃 AS 의 Site Exit 라우터들과 연결되어 있다. AS C 는 AS A 와 B 의 각 Site Exit 라우터를 통하여 prefix A, B 에 대하여 접근을 할 수 있다. AS C 내 에 연결된 Local Site Exit 라우터 들은 iBGP (Interior BGP) peering 을 이용하여 서로의 prefix 정보, 즉 prefix A 와 B 를 공유한다.

만약, 그림 2 와 같이 AS B 를 담당하는 Site Exit 라우터에 문제가 생겨서 AS B 로 가는 route 가 끊어졌다고 하자. 이때 iBGP 로 연결된 다른 라우터는 끊어진 prefix B 에 대한 정보를 AS A 에 보내준다. 이를 Auto Route Injection 방식이라고 한다.

즉, AS C 에서 각각의 AS 와 연결되어 있던 Site Exit 라우터들은 서로가 담당하고 있는 prefix 의 정보를 iBGP peering 으로 교환하고 있다가, 어느 한쪽에 문제가 생기면, 자신의 Local Site Exit 라우터들에게 문제가

생긴 부분의 prefix 정보를 보냄으로써, Fault Tolerance 를 제공한다.

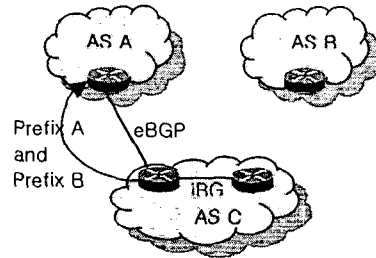


그림 2. Auto Route Injection에서의 Redundancy

나. Non-direct eBGP Peering

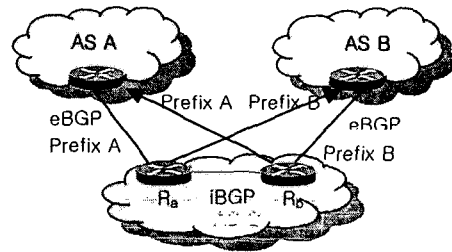


그림 3. Non-direct eBGP Peering

그림 3 은 Non-direct eBGP Peering 구조이다. 이 구조는 앞의 Auto Route Injection 과는 달리 각 Site Exit 라우터가 자신이 담당하는 Primary AS 뿐만 아니라 Secondary AS 도 eBGP Peering 을 하고 있다. 보통 때에는 자신이 담당하는 Primary AS 에 대해서 자신이 사용하는 Primary prefix 에 대한 것만을 담당한다; 라우터 Ra 의 primary prefix 정보는 prefix A 이고 secondary prefix 정보는 prefix B 이다; 라우터 Rb 의 primary prefix 정보는 prefix B 이고 secondary prefix 정보는 prefix A 이다. 그림 4 에서처럼 AS B 를 전담하는 Site Exit 라우터에 문제가 생겼을 경우에 AS A 를 전담하던 Site Exit 라우터를 통해서 Traffic 이 전달될 수 있게 된다.

Auto Route Injection 방식은 한쪽의 Local Site Exit 라우터에 문제가 생기거나 외부 연결이 끊기면, 그림 2 에서처럼 Traffic 이 다른 하나의 AS 를 통해 들어오게 된다. 이는 Traffic 이 하나의 AS 로 집중될 수 있음을 말한다. 따라서, AS 와 연결된 대규모 Network 에 - Internet 같은 경우 - 부담을 줄 수 있다. 하지만 Non-direct eBGP Peering 방식은 한쪽의 Site Exit 라우터에 문제가 생겼더라도, 각각의 primary 와 secondary AS 를 통해 Traffic 을 받기 때문에 AS 들이 연결된 대규모의 Network 에 부담을 주지 않는다.

2.2. Load Sharing

기존의 Static Load Sharing 방식은 BGP 와 IGRP protocol 에서 사용하는 방식을 예로 들 수 있다 [9, 10, 16]. 이러한 것들은 관리자의 설정에 의한 방식이라고

볼 수 있다. 즉, 관리자가 Network 상황을 고려하여 Bandwidth, Link Cost 등과 같은 여러 가지 Parameter 들을 설정한다. 그리고 이러한 Parameter 들을 가지고 라우터들은 최적의 경로를 찾아서 routing 을 하고, Load Sharing 을 하게 된다. Cisco 에서는 기존의 BGP4+ Protocol 을 사용하여 Site Exit 라우터들 간의 Load Sharing 을 Static 하게 구현할 수 있게 하였다.

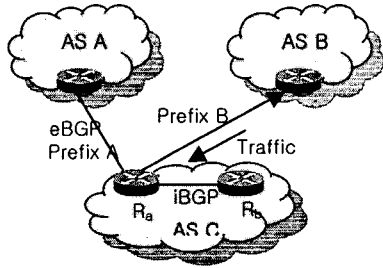


그림 4. Non-direct eBGP Peering에서의

즉, 라우터를 Configure 할 때 Prefix 정보를 분할하여 Multihomed 된 각 Exit 라우터에 Load 가 분산되도록 하였다. 이러한 방식은 단순하다는 장점이 있으나, Network 상황이 변화하게 되면, 설정된 Parameter 들을 가지고서는 그 라우터가 최적의 성능을 발휘하지 못하게 되며, 전체적인 Network 의 성능 저하를 가져오게 된다는 문제점이 있다.

이러한 문제를 해결하는 방법중의 하나는 관리자에 의해 설정되던 Parameter 들을 라우터 혹은 Server 가 Network 상황을 감지하여서, 자동으로 빠르게 변화시켜 주는 Dynamic 방식을 사용하는 것이다. 라우터가 스스로 상황을 파악하여 자신의 Parameter 를 변화하는데 걸리는 시간은 관리자의 대처 시간보다는 짧다. 이는, 손실되는 Packet 이나 Delay 되는 Packet 이 줄어들 수 있음을 시사한다. 또한, 전체적인 Network 성능의 향상을 기대해 볼 수 있게 된다.

그림 1 Auto Route Injection 과 그림 3 Non-direct eBGP Peering 구조에서 한 Site Exit 라우터 Group 에서의 Multihomed 된 Site Exit 라우터들 중 Load Sharing 대상이 될 라우터를 선택하는 Policy 를 생각해 보면 다음과 같은 임의의 라우터 선택, 최적의 라우터 선택으로 생각해 볼 수 있다.

- 임의의 라우터 선택
 - 라우터의 상태와 상관 없이 Round-Robin 또는 Random 하게 Site Exit 라우터를 순차적으로 선택하여 Load 를 분산한다. 가장 단순하게 Load Sharing 을 구현할 수 있으나 각 라우터의 work load 를 고려하지 않음으로써 delay 등의 문제를 일으킬 수 있다.
- 최적의 라우터 선택
 - 앞에서의 각 고려사항에 기반하여 적합한 Router 를 선택하는 방법은 다음과 같다.
 - a. Run-Time 시 Multihomed 된 라우터들의 상태를 고려하여 Idle 한 라우터에 Load 를 분산한다.
 - b. Multihomed 된 라우터들에 Multicast Request 를 보내서 Response Time 이 가장 빠른 라우터에 Load 를 분산한다.

- c. Run-Time 시 Site Exit 라우터 Group 에 저장된 Load 분산 정보를 분석하여 라우터 선택 Weight 을 둔다.
- d. 그 밖에 Metrics, Preference Value, Prefix Information 등을 조합하여 라우터를 선택하는 factor 로 이용할 수 있다.

3. BGP4+ Algorithm

EGP (Exterior Gateway Protocol) 중에서 현재 가장 많이 사용되는 것은 BGP4 이다. BGP 는 여러 버전을 거쳐 BGP4 까지 발전되었다. 이의 가장 중요한 기능은 AS 간에 BGP peer 라우터사이의 네트워크 연결성 (reachability)를 교환하는 것이다. BGP 는 신뢰성을 보장하기 위하여 TCP 를 사용하며, 현재 많이 사용되고 있는 BGP4 는 IPv4 만을 지원하고 있다. RFC 2283 [12]에서는 이러한 BGP4 를 확장하여 다른 여러 프로토콜의 정보를 교환할 수 있도록 추가 사항을 정의하고 있다. 예를 들어 IPX 나 IPv6 에 대한 정보를 교환할 수 있도록 한 것이다. 따라서, IPv6 망에서 BGP 로 AS 간의 Routing 정보를 교환하기 위해서 BGP4+ 규격을 정의 하게 된다.

BGP4 는 한 목적지에 한 Path 만을 Best Path 로 선택한다 [11]. 그 Path 가 정해지면 BGP4 는 그 Path 를 Routing Table 에 삽입한다. 그리고 그 Path 를 neighbor 들에 전달한다. 이 Best Path 선택 알고리즘은 BGP4+ 에서도 마찬가지이다. 따라서, 본 논문에서는 BGP4+ 와 BGP4 의 Multihomed Network 에서의 Load Sharing 해법을 동일하게 본다. 다음의 Decision Process 는 BGP4 Best Path 선택 알고리즘에서 최적의 Route 를 선택하는 부분이다.

한 BGP 라우터를 위한 BGP Decision Process [11]

Phase 1. Preference Degree 계산: 한 Neighboring AS 의 한 BGP Speaker 로 부터 받은 각 Route 의 Preference Degree 를 계산한다. 그리고, local AS 의 다른 BGP Speaker 들에게 각 Destination 에 대한 Highest Degree Preference Route 를 advertising 한다.

Phase 2. Route 선택: local AS 에서 각 Destination 에 대한 Best Route 를 선택하고 Loc-Rib 에 설치한다. 즉, Adj-RIBS-In 에 한 가능한 Route 들이 있는 각 Destination 에 대하여, local BGP Speaker 는 다음의 route 를 선택한다.

- (a) 같은 Destination 들에 대한 Highest Preference 를 가진 Route
 - (b) 한 Destination 에 대한 유일한 Route
 - (c) Tie-breaking rule 에서 선택된 Route
- 선택된 Route 는 각 Speaker 의 Loc-RIB 에 같은 Destination 을 가진 Route 와 교체된다.

Phase 3. Route 유포: PIB (Policy Information Base) 의 policy 에 따라 각 Neighboring AS 의 각 BGP peer 에 Loc-Rib 의 Route 들을 유포한다. Route Aggregation 과 Information Reduction 이 선택적으로 수행된다.

4. BGP4+ Algorithm 의 확장

본 장에서는 IPv6 에서 멀티호밍 Network 의 Exit 라우터들 간의 Load Sharing 을 위하여 Routing Table 과 BGP Path 선택 알고리즘을 확장한 해법을 제시한다. 먼저 한 AS 의 한 Exit 라우터 a 가 다른 한 AS 또는 여러 AS 들의 n 개의 Exit 라우터와 연결 되어 있으면 그 라우터 a 는 한 목적지 d 에 대하여 최대한 n 개의 Next Hop 을 가진다. 따라서, BGP Routing table 은 최소한 n 배로 증가한다. 또한, Route 선택 알고리즘은 한 목적지 d 에 대하여 최대한 n 개의 Next Hop 을 가진 여러 Path 들을 선택하도록 수정된다.

이를 바탕으로 다음은 기존의 BGP Decision Process 에서 Phase 2 를 확장 및 수정한 본 논문에서 제시한 알고리즘이다. 받은 Route 들과 기존의 최적 Route 비교로 하나의 최적 Route 를 결정하는 대신 한 destination 에 대한 각 Next-Hop 별 최적의 Route 를 선택한다.

한 Multihomed BGP 라우터를 위한 BGP Decision Process

Phase 2. Route 선택: local AS 에서 각 Destination 에 대한 다른 next-hop 중에서 Best Route 를 선택하고 Loc-Rib 에 설치한다. 즉, Adj-RIBs-In 에 한 가능한 Route 들이 있는 각 Destination 에 대하여, local BGP Speaker 는 다음의 Route 들을 선택한다.

- (a) 같은 Destination 들에 대한 다른 next-hop AS Path 중에서 Highest Preference 를 가진 Route 들 \square 즉, 각 interface 당 Best Route 들
- (b) 한 Destination 에 대한 유일한 Route
- (c) 같은 Destination 들에 대한 다른 next-hop AS Path 중에서 Tie-breaking rule 에서 선택된 Route 들 선택된 Route 들은 각 Speaker 의 Loc-RIB 에 같은 Destination 을 가진 Route 와 교체된다 \square 다른 next-hop AS Path 일 경우는 교체되지 않는다.

기존의 BGP Best 선택 알고리즘에 본 논문의 Phase 2 를 적용하면 멀티호밍의 Load Sharing 이 가능하다. 물론, Dynamic 한 Traffic Load Sharing 을 구현하는 연구는 본 논문에서 다루지 않는다. 이 알고리즘을 통하여 한 Destination 에 대한 모든 Path 들 중 각 multihomed network 의 복수의 Next Hop eBGP 라우터 Path 들중에서 Next Hop 당 최적의 Path 를 선택할 수 있다.

이 수정된 BGP 알고리즘을 분석하기 위해서는 이 알고리즘이 영향을 미치는 한 BGP 라우터의 Loc-Rib Table 을 분석해야 한다. 즉, 각 Destination D 에 대하여 Loc-Rib Table 의 Space Complexity 는 기존의 BGP 알고리즘이 $O(l \times D)$ 이며 본 알고리즘에서는 $O(n \times D)$ 이다; n 은 한 AS 에 있는 interface 갯수이다. 마찬가지로, 각 Destination D 에 대하여 Loc-Rib Table 의 Time Complexity 는 기존의 BGP 알고리즘이 $O(l \times D)$ 이며 본 알고리즘에서는 $O(n \times D)$ 이다. 본 논문의 BGP 알고리즘 Time 과 Space 효율이 기존의 알고리즘보다 떨어지지만 전체 멀티호밍 구조에서 자원 및 Time 효율성은 높으리라 기대한다.

5. 결론

본 논문은 IPv6 에서의 Multihomed Site Exit 라우터들 간의 Load Sharing 구현에 대하여 서술 하였다. 각 Multihomed Network 구조는 RFC 2260 [1] 과 RFC 3178 [2] 에서 언급된 Auto Route Injection 과 Non-Direct EBGP Peering 구조를 기반으로 한다. 본 논문에서는 Load Sharing 구현을 위해 요구되는 고려 사항과 한 Site 의 Exit 라우터 Group 에서 Multihomed 된 Site Exit 라우터들을 선택하는 확장된 BGP 알고리즘을 제안하였다. 본 연구는 SSFNet Simulator [17] 를 이용하여 자원의 효율성과 Performance 에 관한 실험중이다.

참고문헌

- [1] T. Bates and Y. Rekhter, Scalable Support for Multihomed Multi-Provider Connection, RFC 2260
- [2] J. Hagino and H. Snyder, IPv6 Multihoming Support at Site Exit Routers, RFC 3178
- [3] Yong Kim, Myung Shin, Jung Park, Seung Lee, Next Generation Internet Protocol IPv6, 2002, Da-Sung publication
- [4] Steven Deering, Overview of IP(v6) Multihoming Issues, IPng working, September 29 - October 1, 1999 in Tokyo
- [5] B. Black, V. Gill, J. Abley, Requirements for IP Multihoming Architectures, draft-ietf-multi6-multihoming-requirements-01, Network Working Group, June 25, 2001
- [6] B. Black, V. Gill, J. Abley, IPv4 Multihoming Motivation, Practices and Limitations, draft-ietf-multi6-v4-multihoming-00, Network Working Group, June 25, 2001
- [7] R. Hinden, IPv6 Host to Router Load Sharing, draft-ietf-ipv6-host-load-sharing-00.txt, Jan 4, 2002
- [8] Kris Bubendorfer, Nomad: Towards an Architecture for Mobility in Large Scale Distributed Systems, PhD Thesis, Victoria University of Wellington, December 2001
- [9] Syed Faraz Shamim, How Does Unequal Cost Path Load Balancing (Variance) Work in IGRP and EIGRP, <http://www.cisco.com/warp/public/103/19.html>, CISCO Systems
- [10] Praveen Akkiraju, Kevin Delgadillo, Yakov Rekhter, Enabling Enterprise Multihoming with Cisco IOS Network Address Translation (NAT)
- [11] Y. Rekhter, T. Li, A Border Gateway Protocol 4, RFC 1771
- [12] T. Bates, R. Chandra, D. Katz, Y. Rekhter, Multiprotocol Extensions for BGP4, RFC 2283
- [13] An Introduction to Interdomain Routing and BGP, Timothy G. Griffin, <http://www.research.att.com/~griffin/interdomain.html>
- [14] ARIN, BGP Tutorial, Avi Freedman, <http://avi.freedman.net>
- [15] An Application of the BGP Community Attribute in Multi-home Routing, E. Chen and T. Bates, RFC 1998
- [16] Using the Border Gateway Protocol for Interdomain Routing, Cisco Systems, <http://www.cisco.com/univercd/cc/td/doc/cisintwk/ics/icsbgp4.htm>
- [17] Scalable Simulation Framework Network Models, <http://www.ssfnet.org/>