

논문 2008-45TC-3-7

# BGP에서의 End-to-End 트래픽 엔지니어링 방안

( An Improved BGP Routing Algorithm for End-to-End Traffic Engineering )

공 현 민\*, 권 영 미\*\*

( Hyonmin Kong and Youngmi Kwon )

## 요 약

BGP는 인터넷에서 망사업자 및 대규모 네트워크에서 네트워크 간에 라우팅 정보를 교환하는 라우팅 프로토콜이다. 라우터가 end-to-end 트래픽에 대해 트래픽 엔지니어링을 위한 정책을 설정할 때 neighbor AS 이외의 AS들에게는 원하는 정책이 영향을 미치지 않기 때문에 원하는 트래픽 엔지니어링을 할 수 없는 한계가 있다. 본 논문에서는 global 정책의 설정을 통해 원하는 트래픽 엔지니어링의 특성을 반영시킬 수 있도록 함으로써 BGP에서 End-to-End간 트래픽 엔지니어링을 위한 라우팅이 실현될 수 있도록 하는 확장된 BGP 라우팅 알고리즘을 제안한다. 새로운 BGP 라우팅 알고리즘은 전통적인 BGP 라우팅 테이블의 구조적인 변경을 하지 않고 새로운 메시지 형식을 추가하지 않으며 기존의 UPDATE message를 이용하여 global 정책 설정을 하게 한다. 또한 새로운 BGP 라우터가 전통적인 BGP 라우팅 알고리즘으로 동작하는 BGP 라우터들과 함께 망을 구성하고 있는 경우에도 확장성 문제를 야기하지 않는다. 새로운 BGP 알고리즘과 전통적인 BGP 알고리즘의 성능을 비교하기 위해 패킷 loss와 평균 delay를 시뮬레이션을 통해 비교 분석하여 그 결과를 확인하였다.

## Abstract

BGP is an exterior routing protocol which exchanges routing information between ASs in ISP or large networks. It is necessary for the BGP to provide traffic enforcement among ASs, for an End-to-End traffic engineering capability in the Internet. However, traditional BGP-4 is lack of end-to-end traffic engineering capability because the policy-based routing affects the traffic only between directly-connected neighbor ASs. This paper proposes an extended BGP routing algorithm to achieve an End-to-End traffic engineering capability. This new method does not require an additional BGP message type, nor change a structure of existing routing table. It only extended the operation of UPDATE message with AS\_PATH attribute, which is well-known mandatory. Simulation shows that the new BGP algorithm improves the packet delay and packet loss without the problem of backward compatibility.

**Keywords :** BGP, Routing, Traffic Engineering, Policy

## I. 서 론

BGP (Border Gateway protocol)는 인터넷에서 AS

(Autonomous System) 간 정보를 교환하는 라우팅 프로토콜이다. 서로 다른 망간에 Routing 정보를 주고받는 Exterior Gateway Protocol로 현재는 BGP-4가 사용되고 있다<sup>[1]</sup>. BGP는 정책에 의한 라우팅 프로토콜로서 네트워크 경로가 네트워크를 운영하는 운영자의 정책에 의해서 결정되는 라우팅 프로토콜이다. 라우팅 테이블에는 라우터들의 목록과, 접근할 수 있는 네트워크 주소, 그리고 최적의 경로를 선택하기 위해 각 라우터까지의 경로와 관련된 비용 정보가 담겨있다. BGP는 다른 라우팅 방법들에 비하여 흡 카운트의 제한을 받지 않아 규모가 큰 망을 지원할 수 있는 Path-vector 라우

\* 학생회원, 충남대학교 정보통신공학과  
(Electronics, Radio, Information Communications Engineering, Chungnam National University)  
\* 정회원, 충남대학교 전기정보통신공학부  
(Electronics, Radio, Information Communications Engineering, Chungnam National University)  
※ 본 연구는 한국과학재단 목적기초연구(R04-2003-000-10041-0)지원으로 수행되었음.  
접수일자: 2008년2월1일, 수정완료일: 2008년3월19일

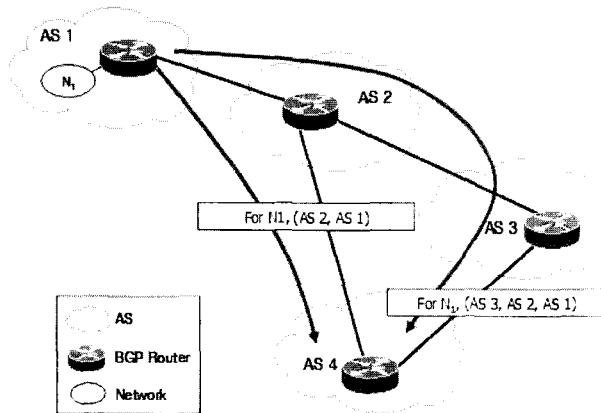


그림 1. BGP의 UPDATE 메시지 정보에 의한 라우팅  
Fig. 1. BGP Routing with UPDATE message.

팅 프로토콜로서 이웃 라우터 간에 연결을 유지하면서 경로 정보를 교환한다.

인터넷에서 source-to-destination 연결은 IGP 뿐 아니라, EGP를 거쳐 설정되며, 따라서 end-to-end간의 QoS를 보장하기 위해서는 BGP의 QoS 또한 향상시켜야 하고<sup>[3]</sup>, end-to-end traffic engineering을 실현하기 위해서는 BGP의 end-to-end traffic engineering 또한 보장하여야 한다.

BGP에서는 반드시 최소 흡 방식으로 경로를 설정하지는 않고 미리 설정된 정책에 따라 경로를 설정한다. 예를 들어, 그림 1에서 N1에 대한 경로 정보가 AS4에 있는 라우터에게 다른 경로를 거쳐 두 개의 BGP 패킷으로 전달되면, 이들을 수신한 AS4의 BGP 라우터는 흡 수가 아닌 미리 설정된 정책에 따라 경로를 선택하는 것이다.

비용이나 보안 또는 정치적인 정책에 의해 BGP는 자신의 경로정보를 다른 라우터에게 알리거나 수집된 정보들을 선택적으로 버리거나 무시할 수 있다.

그러나 한 AS의 BGP 정책이 영향을 미치는 범위는 바로 이웃한 AS까지이다. 이러한 제한을 확장시키기 위해 source-to-destination 간에 정책이 enforce 되도록 하려는 방안이 최근 발표되었는데, 이는 RFC 4360<sup>[4][5]</sup>에 정의된 community 값을 활용하여 그림 2와 같이 path를 end-to-end로 제어하고자 하는 것이다<sup>[5]</sup>. 그러나 이러한 방법은 AS 대 AS의 전체 트래픽은 제어하지만, AS에서 발생하는 모든 트래픽이 하나의 경로만을 사용함으로써 전체적인 망 링크를 활용하지 못하는 제약이 있다.

본 논문에서는 하나의 AS를 구성하는 network의 단위로 end-to-end의 다른 링크들을 사용하도록 엔지니

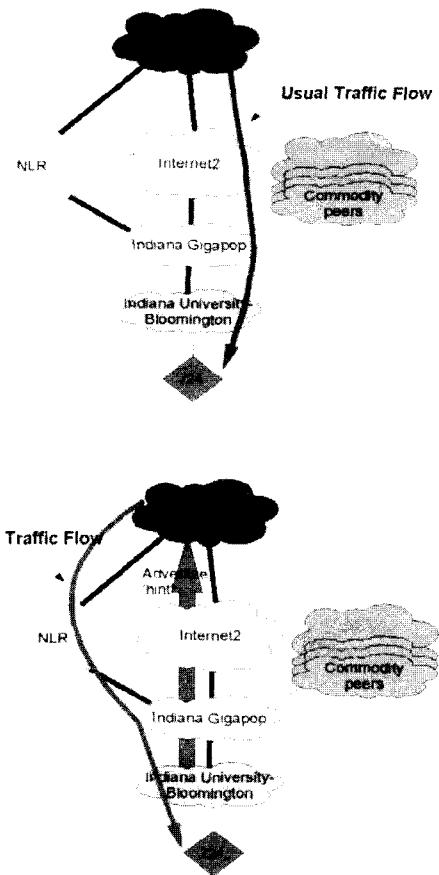


그림 2. Community 속성을 이용한 트래픽 엔지니어링  
Fig. 2. Traffic Engineering with Community Attribute.

어링 함으로써 전체적인 BGP 망의 이용도를 높일 수 있도록 하는 방안을 제안한다.

본 논문의 구성은 II장에서 제안하는 새로운 BGP 알고리즘을 설명하고, III장에 시뮬레이션 결과를 보이며, IV장에서 결론을 맺는다.

## II. 새로운 BGP 라우팅 알고리즘

### 1. 기존 BGP 라우팅의 제한사항

BGP는 동일한 목적지에 대해 거쳐 가는 AS의 수가 적은 경로 혹은 대역폭이 높은 경로 등, 최적의 경로를 선택하도록 정책을 정한다. 그림 3과 같은 네트워크 환경에서, AS1으로부터 AS100과 AS200으로 트래픽이 송신된다고 할 때, 일반적으로 관리자는 throughput이 좋은 AS3을 경유해서 보내도록 설정을 할 것이다. 관리자는 AS1의 경계라우터(border router) RTA의 LOCAL\_PREF나 MED를 조정하여 원하는 방향으로 트래픽을 엔지니어링한다<sup>[6]</sup>.

그림 3에서 AS1의 N1 망이 N100, N200과 연결을 갖

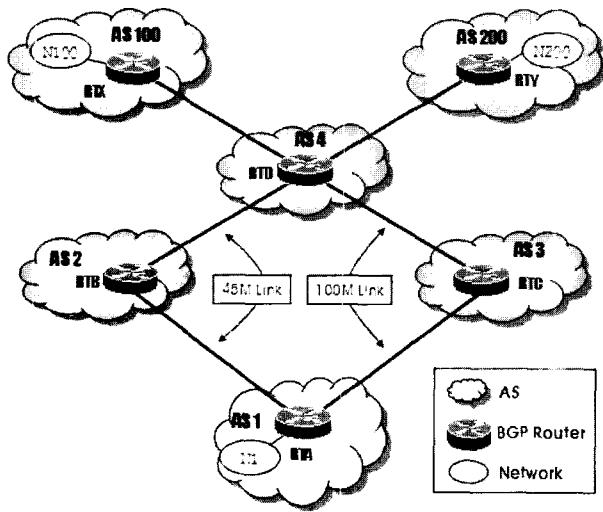


그림 3. BGP 망 구성도 예

Fig. 3. Example of BGP Network.

표 1. 라우터 RTA의 라우팅 테이블

Table 1. Routing Table of Router RTA.

Destination Network	NEXT HOP	LOCAL_PREF	Etc.
N100	RTB	200	
N200	RTC	100	...
N100	RTB	100	
N200	RTC	200	
...			

고 있고, 트래픽 분산 등 여러 이유로 서로 다른 경로 (N100-AS4-AS2-AS1과 N200-AS4-AS3-N1)를 통해 송수신 하고 싶은 경우를 가정하자.

AS1의 경계라우터 RTA의 라우팅 테이블을 표 1과 같이 설정하면, RTA는 N100으로 보내는 트래픽을 LOCAL\_PREF 200인 RTB(=AS2)를 next hop으로 하여 보낼 것이고, N200으로 전송하는 트래픽은 RTC(=AS3)를 next hop으로 분산하여 보낼 것이다.

그러나 N100, N200으로부터의 트래픽은 AS4의 경계 라우터 RTD의 라우팅 테이블(표2)에 의해 N1을 destination으로 하는 최선의 next hop인 RTC(=AS3)만을 통해 수신하게 된다. 이는 AS1의 정책이 바로 이웃하지 않은 AS4에는 전달되지 않음으로 인해 생기는 한계이다. 따라서 동일한 목적지 네트워크로 향하는 트래픽 일지라도 트래픽 목적지 AS의 망 사용 정책에 따라 트래픽을 보낼 수 있도록 인접하지 않은 AS에게 정책을 알리는 방안이 필요하다 하겠다.

즉, 현재의 BGP에서는 LOCAL\_PREF나 MED와 같은 Metric들을 이용하면 이웃한 AS와는 BGP 라우터들의 Incoming/Outgoing traffic을 정책에 의해 트래픽 경로를 설정할 수 있지만 이웃하지 않은 AS의 BGP와의

global policy 설정은 불가능하다. 본 논문에서 제안하는 새로운 BGP 라우팅 알고리즘은 동일한 목적지로 향하는 트래픽 일지라도 목적지의 망 사용 정책에 따라 트래픽을 분산하여 전달할 수 있게 한다.

## 2. BGP 알고리즘의 확장

BGP는 경로 정보를 UPDATE 메시지를 이용해 이웃 AS들에게 알린다. UPDATE 메시지를 받은 BGP 라우터들은 그 메시지의 정보들을 이용해 자신의 라우팅 테이블을 갱신한다.

앞 장에서 지적한 기존 BGP의 한계를 넘기 위해서는 UPDATE 메시지를 보낼 때 그 UPDATE 메시지를 보내는 source network의 정보를 같이 보내어, 같은 경로 정보라도, 입력받은 링크에 따라 다른 라우팅 엔트리로 취급하면 된다. 표 2의 라우팅 테이블에 source network 정보를 포함하도록 확장한 것을 표3에 보였다.

표 3의 라우팅 테이블을 갖는 라우터 RTD는 N100으로부터 N1으로 보내는 트래픽은 next hop RTB(=AS2)로 보내고, N200에서 N1으로 보내는 트래픽은 next hop RTC(=AS3)로 보내게 되어 같은 N1으로 보내는 트래픽이라도 분산하여 전송할 수 있다.

그러나 BGP 라우팅 테이블의 구조를 변경하는 경우, 기존의 BGP 와의 연동에 문제가 생길 수 있어, 본 논문에서는 별도의 source network 항목을 사용하지 않고, 기존의 BGP 라우팅 테이블의 한 항목인 AS\_PATH 정보에 source network 정보가 들어갈 수 있게 함으로써 동일한 효과를 갖게 하는 방법을 제안한다. 이를 위해 AS\_PATH에 AS number를 추가해서 보낼

표 2. 라우터 RTD의 라우팅 테이블

Table 2. Routing Table of Router RTD.

Destination Network	NEXT HOP	LOCAL_PREF	Etc.
N1	RTB	100	
N1	RTC	200	...
...			

표 3. Source network 항목이 포함된 라우팅 테이블

Table 3. Extended Routing Table with Source Network Information.

Source Network	Destination Network	NEXT HOP	LOCAL_PREF	Etc.
N100	N1	RTB	200	
N100	N1	RTC	100	...
N200	N1	RTB	100	
N200	N1	RTC	200	
...				

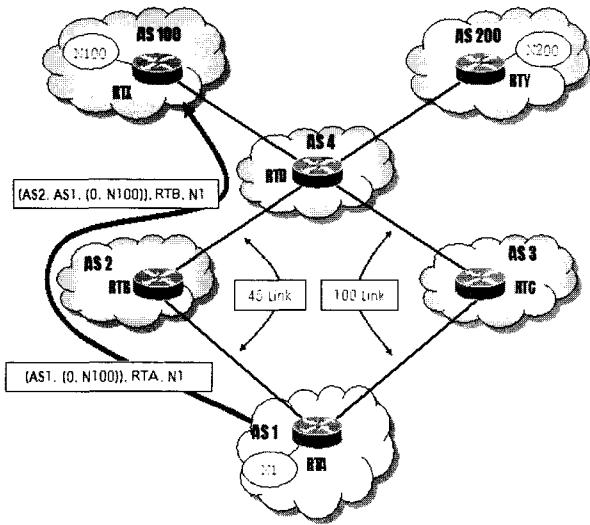


그림 4. Global UPDATE message  
Fig. 4. Global UPDATE message.

때, AS number로 사용되지 않고 있는 0번을 특수 제어 정보로 사용하여 source 정보를 넣을 수 있도록 하였다.

현재 AS number는 1부터 65535까지 사용하도록 권고되어 있다<sup>[1]</sup>. 인접하지 않은 AS에게 global policy 설정을 요청하고자 하는 AS는 UPDATE 메시지의 AS\_PATH 필드에 자신의 AS number를 넣기 전에 zero와 특별한 policy 설정을 원하는 network 정보를 넣는다.

BGP에서 UPDATE 메시지를 받은 경계라우터는 unreachable 하다고 withdrawn routes 정보로 explicit하게 주어지지 않는 네트워크 정보 이외에는 라우팅 테이블에서 삭제하지 않는다. 따라서 하나의 네트워크 A에 대해 가능한 라우팅 정보 여럿을 라우팅 테이블에 유지하고 있게 되며, next hop 경로를 결정해야 하는 시점에 여러 변수들을 참조해 best route를 선택해 사용한다.

그림 4는 N100과의 통신에 AS2의 경로를 사용하도록 global policy를 사용하고 싶은 AS1이 UPDATE 메시지를 보낼 때 각 hop에서의 {AS\_PATH, NEXT\_HOP, NLRI(Network layer reachability information)} 내용을 보인 것이다.

제안한 알고리즘을 적용하는 BGP 라우터를  $BGP_{new}$ 라고 적용하지 않은 라우터를  $BGP_{old}$ 라고 표기하겠다. 그림 4에서 라우터 B는  $BGP_{old}$ 이고 라우터 D는  $BGP_{new}$ 라고 가정할 때, 라우터 A에서 global UPDATE message가 보내진 후 라우터 RTB와 라우터 RTD의 라우팅 테이블은 각각 표 4 및 5와 같다.

AS number 0이 존재하는 경우에는 다른 필드의

표 4. Router RTB의 BGP 라우팅 테이블  
Table 4. BGP Routing Table in Router RTB.

Destination Network	NEXT HOP	AS_PATH	Etc.
N1	RTA	AS1	
N100	RTD	AS4, AS100	
N200	RTD	AS4, AS200	...
...			

표 5. Router RTD의 BGP 라우팅 테이블  
Table 5. BGP Routing Table in Router RTD.

Destination Network	NEXT HOP	LOCAL_PREF	AS PATH	Etc.
N1	RTB	100	AS 2, AS 1, (0, N100)	
N1	RTC	100	AS 3, AS 1	...
N1	RTB	100	AS 2, AS 1	
N1	RTC	200	AS 3, AS 1	
N100	RTX	100	AS 100	
N200	RTY	100	AS 200	
...				

LOCAL\_PREF 값에 상관없이 가중치를 가장 높게 적용한다. 만약 우선순위가 같을 때는 AS\_PATH의 길이가 짧은 것을 우선한다.

라우터 RTD는 N100으로부터 N1으로 향하는 트래픽이 들어왔을 때 AS\_PATH에 zero 값이 있고, 그 뒤의 network 값이 N100과 일치하는 정보를 LOCAL\_PREF 값과 상관없이 가장 높은 우선순위로 사용하며, 따라서 LOCAL\_PREF = 200인 4번째 엔트리를 사용하지 않고 1번째 엔트리인 next hop RTB를 best route로 선택해 forward하게 됨으로써, AS1이 의도했던 global policy대로 링크를 사용하게 된다. N200으로부터 N1으로 향하는 트래픽에 대해서는 가장 높은 LOCAL\_PREF를 갖는 4번째 엔트리 정보를 사용해 next hop RTC(=AS3)로 forward 하므로 AS1과 AS2 간에 global load balancing 효과를 갖게 된다.

Global UPDATE message는 기존의 UPDATE message를 이용하여, 이 메시지는 필요한 라우터들만을 업데이트를 한다. 메시지의 의미를 인식하지 못하는  $BGP_{old}$  라우터의 경우도 그냥 이웃에게 메시지를 전달해 버리므로, 전통적인 BGP 라우팅 알고리즘으로 동작하는 BGP 라우터들과의 연동에 있어서도 문제가 없다.

BGP 라우터에 트래픽이 들어오면 라우터는 다음의 알고리즘을 수행하여 다음 목적지로 트래픽을 보낸다.

- 1) 라우팅 테이블에서 도착지 Network 항목과 트래픽의 도착지를 비교하여 일치하는 항목이 있는지

찾는다.

- 2) AS\_PATH 항목에 AS number zero가 있는지 확인한다. 만약 zero가 없다면 BGP<sub>old</sub> 라우팅 알고리즘과 동일하게 동작한다.
- 3) AS number zero가 있다면, zero 다음의 네트워크 주소와 패킷의 source가 일치하는지 확인한다. 만약 일치하는 것이 없다면 BGP<sub>old</sub> 라우팅 알고리즘과 동일하다.
- 4) 만약 source가 일치하면 경로설정에 있어 다른 metric들을 무시하고 가장 먼저 우선권을 부여한다. 만약 일치하는 것이 2개 이상이면 일반적인 BGP의 다음 우선순위 결정원칙에 따라 결정한다.

### III. 시뮬레이션 및 결과

BGP<sub>new</sub> 알고리즘의 성능을 알아보기 위해 traffic loss와 평균 delay 값을 시뮬레이션을 통해 측정하여 트래픽 QoS의 성능 정도를 알아보았다. 시뮬레이션은 java 기반의 SSFNet<sup>[8]</sup>를 사용하였으며 그림 5와 같이 망을 구성하였다.

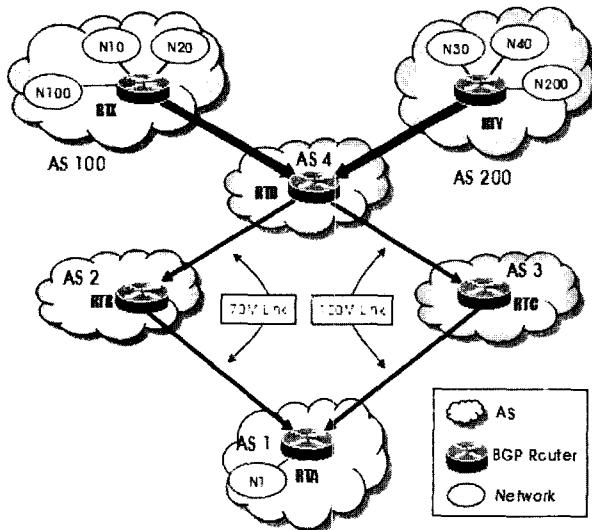


그림 5. 시뮬레이션 네트워크 구성

Fig. 5. Network configuration for simulations.

제안한 방법의 시뮬레이션을 위한 가정은 다음과 같다

- 1) 네트워크 N1은 N100, N200과의 통신에 AS1-AS2-AS4를 사용하려고 하며, 그 이외의 트래픽에 대해서는 용량이 높고 path 길이가 짧은 경로를 사용한다.

- 2) AS1-AS2-AS4의 링크는 70Mbps이고 AS1-AS3-AS4의 링크는 100Mbps이다.
- 3) AS\_PATH 항목에 AS number zero가 있는지 확인하고 만약 zero가 없다면 BGP<sub>old</sub> 라우팅 알고리즘과 동일하게 동작한다.
- 4) 트래픽은 self-similar traffic<sup>[7]</sup>이고 transmitter의 buffer size는 10Mbyte이다.

모든 라우터들이 BGP<sub>old</sub> 일 경우, 라우터 D의 테이블은 표 6과 같다. 만약 모든 라우터에 BGP<sub>new</sub> 알고리즘이 적용되어 있고 RTA에서 global UPDATE message를 보내어 RTD가 이를 반영한다면 라우팅 테이블은 표 7과 같게 될 것이다.

이와 같은 상황에서 traffic loss와 delay를 측정하기 위해 N10, N20, N30, N40, N100, N200의 트래픽을 1Mbps ~ 20Mbps로 증가시키면서 시뮬레이션을 실행하였다.

그림 6이 traffic loss와 end-to-end packet delay 결과를 보인 것이다. BGP<sub>old</sub>의 경우, 모든 트래픽이 AS4 - AS3 - AS1의 경로만을 선택하게 되고, 각 망에서 생성되는 트래픽 양이 증가함에 따라 packet loss가 매우 심해지고 이와 더불어 end-to-end delay도 급격히 증가하는 것을 볼 수 있다.

BGP<sub>new</sub>의 경우는 트래픽이 각 경로를 따라 분산되게 되어 BGP<sub>old</sub>의 경우보다 packet loss와 end-to-end delay에서 훨씬 나은 결과를 보인다.

Delay는 N10, N20, N30, N40, N100, N200의 트래픽이 10Mbps 이하일 때는 BGP<sub>old</sub>와 BGP<sub>new</sub>에서 거의 동일한 수준이지만 트래픽이 10Mbps 이상으로 증가하면서 BGP<sub>old</sub>에서 그 증가폭이 급격히 상승하였다.

표 6. BGP<sub>old</sub> 일 때 라우터 RTD의 라우팅 테이블  
Table 6. BGP<sub>old</sub> Routing Table of Router RTD.

Destination Network	NEXT HOP	LOCAL_PREF	AS PATH	Etc.
N1	RTB	100	AS 2, AS 1	
N1	RTC	200	AS 3, AS 1	
...				

표 7. BGP<sub>new</sub> 일 때 라우터 RTD의 테이블  
Table 7. BGP<sub>new</sub> Routing Table of Router RTD.

Destination Network	NEXT HOP	LOCAL_PREF	AS PATH	Etc.
N1	RTB	100	AS 2, AS 1, (0, N100)	
N1	RTB	100	AS 2, AS 1, (0, N200)	...
N1	RTB	100	AS 2, AS 1	
N1	RTC	200	AS 3, AS 1	
...				

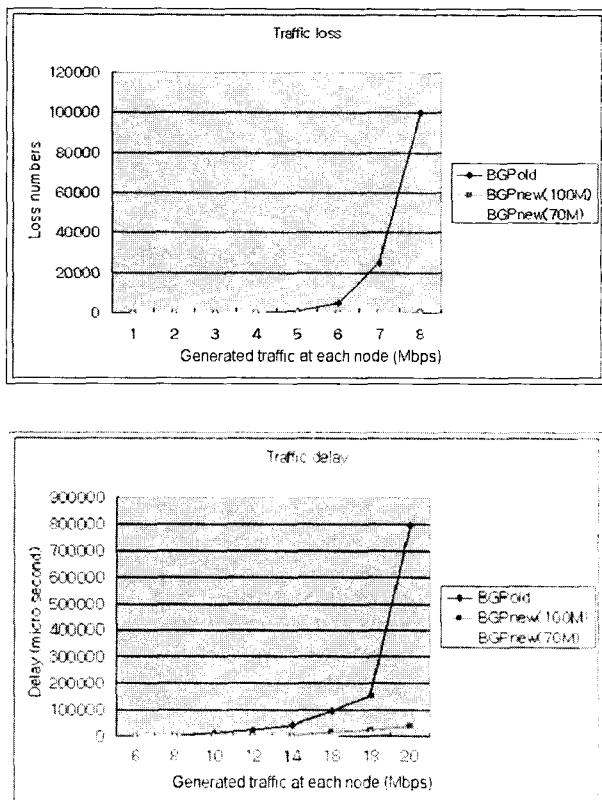


그림 6. 시뮬레이션 결과

Fig. 6. Simulation Results.

#### IV. 결 론

본 논문에서는 BGP에서 end-to-end Policy에 의한 트래픽 엔지니어링이 가능하도록 개선한 BGP 알고리즘을 제안하였다. 제안한 BGP 확장 알고리즘은 기존 BGP의 라우팅 테이블의 구조적인 변경 없이 동작하고 라우터들 간의 협상을 위해서 새로운 형식의 메시지가 아닌 기존의 UPDATE message를 사용하였으며, 기존 BGP 프로토콜만을 인식하는 다른 라우터들과 연동되는 데에 문제를 일으키지 않는다.

시뮬레이션 결과 packet loss나 traffic 평균 delay에 있어서 제안한 방법이 기존의 방법보다 모두 더 나은 성능을 보였다.

본 논문에서의 traffic engineering granularity는 AS내의 network 대 network 단위이다. Customer 대 customer 단위의 granularity를 갖도록 하기 위해서는 BGP가 갖는 여러 가지 현실적인 문제들을 해결해야 할 것이다.

#### 참 고 문 헌

- [1] Y. Rekhter and T. Li, RFC 1771 "A Border Gateway Protocol 4 (BGP-4)", <http://www.ietf.org/rfc/rfc1771.txt?number=1771>, March 1995.
- [2] P. Traina, RFC 1965 "Autonomous System Confederations for BGP", <http://www.ietf.org/rfc/rfc1965.txt?number=1965>, June 1996.
- [3] L. Xiao, K. Lui, J. Wang and K. Nahrstedt, "QoS Extension to BGP", Proceedings of the 10th IEEE International Conference on Network Protocols, pp. 100-109, Paris, France, November 2002.
- [4] S. Sangli, D. Tappan and Y. Rekhter, BGP Extended Communities Attribute, <http://ietf.org/rfc/rfc4360.txt?number=4360>, February 2006.
- [5] B. Sweeny, "BGP Path hinting update," Joint Techs Workshops, January 2008.
- [6] S. Halabi and D. McPherson, Internet Routing Architectures, 2nd ed., ch.6, Cisco Systems, 2000.
- [7] A. Adas, "Traffic models in broadband networks," IEEE Communications Magazine, vol 35, no. 7, pp.82-89, Jul 1997.
- [8] Homepage of Networking Modeling, Simulation and Analysis with SSFNet, <http://www.ssfnet.org>

---

저자소개

---



공 현 민(학생회원)  
 2001년 목원대학교 컴퓨터공학과  
 학사.  
 2003년~현재 충남대학교 정보  
 통신공학과 석사과정  
 재학중.  
 2006년~지에스인스트루먼트  
 (주) 기반연구소 연구원.

<주관심분야 : Computer Network, Traffic Engineering, Femtocell Networking>



권 영 미(정회원)  
 1986년 서울대학교 컴퓨터공학과  
 학사.  
 1988년 서울대학교 컴퓨터공학과  
 석사.  
 1996년 서울대학교 컴퓨터공학과  
 박사.

1993년~1995년 한국전자통신연구원 연구원  
 1996년~2002년 목원대학교 컴퓨터공학과 조교수  
 2002년~현재 충남대학교 전기정보통신공학부  
 부교수  
 <주관심분야 : Traffic Engineering, QoS Routing, Sensor Networks, Internet Protocols>