

# Layer 3 이더넷 스위치 성능 시험 방법론 연구

김용선

한국전자통신연구원

## A Methodology for Performance Testing of Ethernet Switch

Yong-Sun Kim

ETRI

E-mail : doori@etri.re.kr

### Abstract

This paper covers the performance testing for layer 3 Ethernet switch based on various methodologies by which we can measure essential metrics such as throughput, latency, frame loss rate, and back to back frames. In the first place, layer 2 and layer 3 switch evolution is introduced followed by description of IP packet switching in layer 3 switch. And then, the above test metrics and test methodologies are illustrated as well. At last, we conduct the performance testing for layer 3 switch in case of transmitting packets of 64, 128, 256, 512, 1024, 1280, and 1518 byte size and analyze the results.

### I. 서론

Layer 2 스위칭은 MAC(Media Access Control) 프레임에 기준으로 동작하며, 하드웨어기반의 병렬처리 스위칭을 제공하기 때문에 single CPU bridge의 순차적인 소프트웨어 기반 스위칭의 비효율성을 해소해 줄 수 있다. 또한 layer 2에서 동작하기 때문에 여러 가지 다른 프로토콜에 독립적일 수 있다. 그러나 브로드캐스트로 인해 불필요한 트래픽이 전체 스위치 포트에 전송되는 문제점이 있다. 비록 VLAN이 이런 문제를 어느 정도는 경감시켜 줄 수 있지만, 서로 다른 VLAN간의 통신을 위해 라우터를 사용해야만 하며 특히, VLAN으로 분리된 서로 다른 subnet간의 트래픽이 많이 발생하는 경우, 라우터를 경유해야하는 트래

픽이 많아지고 이는 라우터로의 트래픽 집중을 초래하여 성능의 손실을 가져온다. 그리고 일반적으로 라우터는 소프트웨어 기반의 스위칭을 하므로 단순히 스위치 내의 통신보다 속도가 느리게 된다. 라우터가 이와 같이 WAN 트래픽 전송을 위해 사용되지 않고 VLAN간의 통신을 위해 동일한 포트에 트래픽을 송수신하게 되는 경우 이를 one-armed router라고 부른다[1]. 바로 이런 시점에서 layer 3 스위치가 등장하게 된다. Layer 3 스위치는 라우터를 스위치 자체에 구현하고 forwarding을 하드웨어적으로 처리하여 VLAN간의 트래픽에 의한 속도저하를 방지할 수 있다. Layer 3 스위치는 layer 2 스위치 기능을 제공하는 경우도 있고 그렇지 않은 경우도 있다. Pure layer 3 스위치의 경우 여러 layer 2 스위치에 연결되어 VLAN간의 연결을 제공하는, 일반적인 라우터와 같은 역할을 수행하므로 layer 2 스위치 기능은 필요하지 않으며 라우터를 inter-VLAN processing으로부터 해방시켜줄 수 있다. 본 논문에서는 먼저 layer 3 이더넷 스위치의 스위칭 과정에 대해 좀 더 자세히 고찰하고, 이를 바탕으로 성능시험요소 및 성능시험방법론에 대해 알아본다. 또한 성능시험방법론을 실제 layer 3 이더넷 스위치에 적용하여 시험을 수행하고 그 결과를 분석한다.

### II. Layer 3 이더넷 스위치 성능시험요소 및 시험방법론

#### 2.1 Layer 3 이더넷 스위치의 스위칭 과정

라우터를 스위치 자체에 구현하고 forwarding을 하드웨어적으로 처리하기 위해서는 몇 가지 제약점이 따른다. Layer 2 스위치는 단지 이더넷 MAC 프레임에

대해서만 동작하면 되기 때문에 하드웨어 상에 쉽게 구현 가능한 반면, layer 3 스위치의 경우 IP, IPX, AppleTalk 등 다양한 프로토콜이 존재하고, 이와 같은 프로토콜의 forwarding decision이 layer 2 forwarding decision에 비해 훨씬 복잡하다는 문제점이 있다. 예를 들어 IP option processing은 layer 3 forwarding decision을 보다 복잡하게 만드는 요인이 된다. 그러므로 일반적으로 layer 3 스위치는 IP option field를 포함하지 않는 IP 패킷만을 하드웨어를 통해 fast forwarding을 수행하고 나머지는 layer 3 스위치의 CPU에 의해 처리하도록 구현되어 있다. 그림 1은 일반적인 layer 3 스위치의 패킷 스위칭 순서도이다.

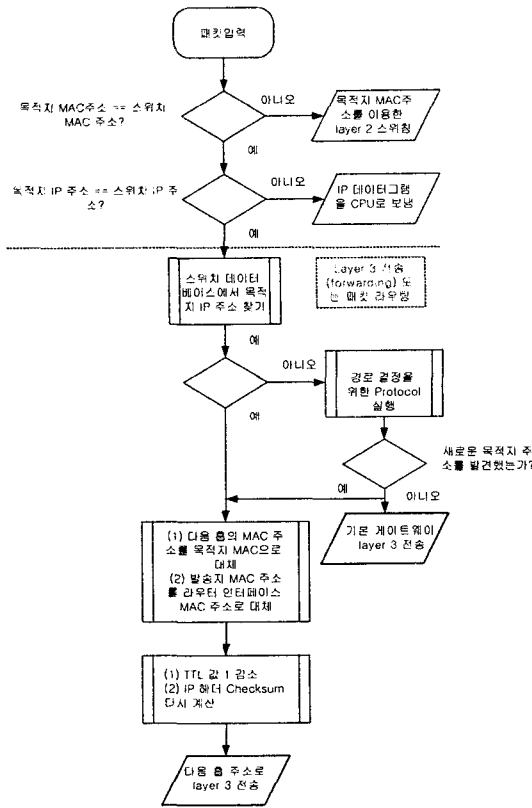


그림 1. IP 패킷 스위칭 순서도

2.2 Layer 3 이더넷 스위치의 성능시험요소

본 논문에서 측정하고자 하는 시험대상장비(DUT: Device Under Test)의 주요 성능 요소와 시험방법론은 IETF의 BenchMarking Working Group(BMWG)의 RFC1242[2], RFC2285[3], RFC2544[4], RFC2889[5]에서 정의되어있다. 주요 시험요소를 간략히 나타내면 다음과 같다:

- Lossless Throughput (Maximum Forwarding Rate)
- Latency
- Frame Loss Rate
- Back-to-Back Frame

위와 같은 성능요소를 unidirectional 또는 bidirectional test traffic 과 같이 다양한 traffic orientation 방법을 이용하여 측정함으로써 스위치의 성능을 시험한다.

시험 데이터의 분산 측면에서 살펴보면, 시험대상장비인 이더넷 스위치의 입력과 출력 port로 test frame이 송수신 되는 방식에 따라 다음과 같이 시험 traffic을 분류한다:

- Non-meshed traffic
- Partially meshed traffic
- Fully meshed traffic

상기한 traffic 분산에 의한 스위치의 처리 능력을 주요 성능시험 요소를 중심으로 시험한다.

2.3 이더넷 스위치 성능시험방법론

본 절에서는 2.2절의 이더넷 스위치 성능시험요소를 측정하기 위한 시험방법에 대해 설명한다. 본 논문에서는 non-meshed traffic을 인가하여 측정하는 maximum forwarding rate, latency, frame loss rate, back-to-back frame의 성능시험요소만을 측정한다.

2.3.1 Maximum forwarding rate

그림 2와 같이 1대1 traffic mapping을 갖는 N쌍의 DUT port-pair에 traffic 전송을 시도하여, 프레임 손실 없이 전송 가능한 maximum forwarding rate를 측정한다.

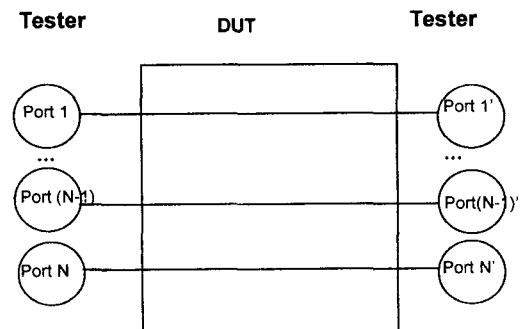


그림 2. N 1-to-1 non-meshed traffic mapping

Test traffic을 전송하고 최대 forwarding rate를 측정하는 방법은 다음과 같다.

전송되는 traffic의 InterFrame Gap(IFG)은 topology와 송신 SmartCard의 전송속도에 따라서 최소 IFG로 전송된다. 전송되는 frame의 수는 설정된 시험기간 동안 전송된 frame 갯수이며, keep-alive와 routing update frame 등은 시험결과에 포함되지 않는다. 초기 전송률 100%에서 시험을 시작하여, 송신을 한 결과 frame의 손실이 발생하면 전송률을 20% 낮추어 시험을 수행한다. 시험결과 손실된 frame이 없으면 binary search algorithm에 따라 전송률 증가를 결정(가장 최근 실패한 전송률과 성공한 전송률의 1/2 값)하여 시험을 반복 수행한다. 반복된 시험결과 허용된 frame 손실율을 만족하면 시험을 종료한다.

### 2.3.2 Latency

그림 2와 같이 1대1 traffic mapping을 갖는 N쌍의 DUT port-pair에 traffic 전송을 시도할 때, burst traffic을 설정된 초기 전송률부터 최대 전송률까지 변화시키면서 tester에서 발생시킨다. Latency를 측정하기 위하여 burst stream 중간의 한 frame에 trigger를 설정한다. Trigger가 설정된 frame에만 timestamp를 부착하여 전송하고 tester가 이 frame을 수신하면 수신 timestamp를 부착하여 아래와 같이 latency를 구한다.

$$\text{Latency} = (\text{수신 Timestamp}) - (\text{송신 Timestamp})$$

Forwarding 방법 (Cut-Through 또는 Store and Forward)에 따라서 Timestamp를 설정하는 방법이 상이하다(RFC1242의 section 3.8 참조).

### 2.2.3 Frame loss rate

그림 2와 같이 1대1로 연결된 N 쌍의 DUT port-pair에 다양한 frame size(64, 128, 256, 512, 1024, 1280, 1518bytes)를 가진 traffic을 부가하여 frame loss rate를 측정한다. Frame loss rate는 각 frame size와 offered load에서 초당 transmitted frame에 대한 lost frame의 비를 의미한다.

### 2.2.4 Back-to-back frames

그림 2와 같이 1대1 traffic mapping을 갖는 N쌍의 DUT port-pair에 traffic 전송을 시도할 때, 주어진 전송 속도에서 가능한 최소 InterFrame Gap을 유지하면서 하나의 burst에 전송할 수 있는 최대 frame의 수를 센다. 주어진 load에서 frame loss가 발생하게 되면 test duration을 binary search algorithm에 따라서 감소시키고 frame loss가 발생하지 않으면 증가시키면서 시험을 반복한다. Frame size와 load를 변화시키면서 시험을 반복하고 평균치를 구한다.

## III. 이더넷 스위치의 성능시험 수행 및 결과 분석

본 장에서는 II장에서 제시한 각 성능시험요소의 시험방법론에 따라 실제로 성능시험을 수행하고 그 결과를 분석한다. 본 성능시험을 위해 시험대상장비로서 C사의 10/100 layer 3 이더넷 스위치를 사용하였고 시험장비로서 Netcom Systems 사의 SmartBits system을 사용하였다[6].

### 3.1 이더넷 스위치 성능시험 환경구성

본 논문에서 수행한 성능시험의 환경구성은 그림 3과 같다.

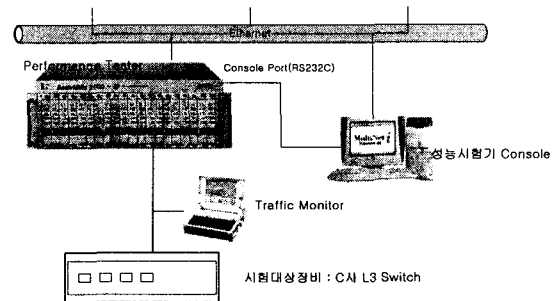


그림 3. 이더넷 스위치 성능시험 환경구성

### 3.2 Maximum forwarding rate 시험결과

그림 4는 각각의 frame size에 대해 throughput을 %단위로 보여준다. 그림 4에서 보듯이 Frame size가 64byte인 경우 최대 throughput은 “99.41%”, 128 byte인 경우 “99.66%” 이었으나 다른 모든 frame size에 대해서는 100%의 throughput을 나타내었다.

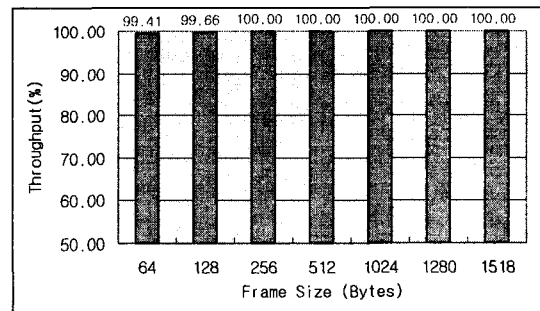


그림 4. Maximum Forwarding Rate in N 1-to-1 Configuration (Throughput(%) vs. Frame Size)

### 3.3 Latency 시험결과

각 port pair에 unidirectional mode로 test traffic을 부가하였고, DUT가 내부적으로 store and forward 방식으로 test traffic을 전송하므로 latency를 LIFO방식으로(RFC1242의 section 3.8 참조) 측정된 결과, 그림 5에서 알 수 있듯이 모든 frame size에 대하여 load가 증가할수록 latency가 다소 증가하였다. 또한 frame size가 클수록 latency가 증가함을 알 수 있다.

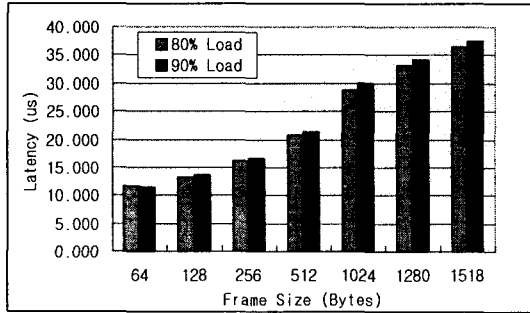


그림 5. Latency N 1-to-1 Configuration (Latency vs. Frame Size)

### 3.4 Frame loss rate 시험결과

Frame loss rate는 64 bytes frame을 100% load로 전송했을 경우에만 0.001%의 frame loss가 발생하였을 뿐, 다른 모든 frame size에 대해 load에 관계없이 0.000%로 frame loss없이 traffic을 전송함을 알 수 있었다.

### 3.5 Back-to-back 시험결과

표 1의 시험결과에서 알 수 있듯이 64byte, 128 byte frame을 제외한 모든 frame이 test duration인 60초와 동일한 back-to-back burst duration을 나타내었다. 이것은 DUT가 60초 동안 frame과 frame사이에 minimum interframe gap만을 가지고 전송할 수 있음을 의미한다. 여기서 burst duration은 back-to-back frame의 수(burst size)를 시간으로 환산하여 나타내었다. 그러나 64 byte와 128 byte frame의 경우 100% load에서 각각 19.1932 초와 33.8458 초의 burst duration을 나타내었다. 64 byte frame을 예로들어 설명하면 이것은 DUT가 19.1932 초(2,856,137개의 패킷에 해당)동안 minimum interframe gap을 가지고 전송한 후, interburst gap만큼 backoff 한 후 다시 19.1932 초 동안 minimum interframe gap을 가지고 전송하며 계속해서 이 과정을 반복하게 됨을 의미한다.

표 1. Back-to-back (100% offered load일 때)

Frame Size (bytes)	Burst Size (# of frames)	Burst Duration (sec)
64	2,856,137	19.1932
128	2,858,579	33.8458
256	2,717,400	60.0000
512	1,409,760	60.0000
1024	718,380	60.0000
1280	576,900	60.0000
1518	487,620	60.0000

## IV. 결론

본 논문에서는 먼저 layer 3 이더넷 스위치를 소개하고 IP패킷 전송 방법 등을 설명하였고 layer 3 스위치의 성능시험요소를 추출하고 이의 시험방법을 분석하였다. 또한, 이를 이용하여 직접 시험을 수행하여 그 결과를 분석하였다. 이를 바탕으로 추후에는 Layer 4 이더넷 스위치와 Router 등의 성능시험으로 확장해 나갈 예정이다.

## 참고문헌

- [1] Thayumanavan Sridhar, Future Communications Software, "Layer 2 and Layer 3 Switch Evolution," The Internet Protocol Journal, Volume 1, Number 2, September 1998.
- [2] RFC 1242, "Benchmarking Terminology for Network Interconnection Devices," IETF, July 1991.
- [3] RFC 1944, "Benchmarking Methodology for Network Interconnect Devices," IETF, May 1996.
- [4] RFC 2285, "Benchmarking Terminology for LAN Switching Devices," IETF, Feb. 1998.
- [5] RFC 2889, "LAN Switch Benchmarking Methodology," IETF, August 2000.
- [6] <http://www.netcomsystems.com>