IPFIX 기반 실시간 플로우 분석 도구의 설계 및 구현[1)]

신성호°, 이영석, 권택근
충남대학교 전기정보통신공학부 컴퓨터전공
{shshin, lee, tgkwon}@cnu.ac.kr

Design and Implementation of an IPFIX-based Real-time Flow Analysis Tool

S. Shin°, Y. Lee, and T. Kwon
Dept. of Computer Science and Engineering, Chungnam National University

요 약

인터넷 사용자의 증가와 함께 인터넷 트래픽 역시 기하급수적으로 증가하고 있다. 이에 따라 네트워크 보안, 트래픽의 특성 분석, 사용자별/응용별 과금, 및 트래픽 엔지니어링을 수행하기 위해서는 네트워크 모니터링 도구가 필수적이다. 그러나, 기존 연구에 많이 사용되고 있는 Cisco NetFlow v5는 포트번호를 기반으로 응용 트래픽을 분류하여 분석할 수 있지만, IPv6와 MPLS 등의 다양한 프로토콜에 대한 트래픽을 분석할 수는 없다. 따라서, IETF에서는 IPFIX표준을 제정하여 확장성 있고 SCTP/TCP 전송 프로토콜을 이용하여 신뢰성있는 플로우 측정 프레임워크를 제시하였다. 본 논문은 Cisco NetFlow v5와 NetFlow v9(IPFIX의 기본 프로토콜)를 모두 지원하여 IPv4/IPv6 트래픽을 실시간으로 분석할 수 있는 플로우 기반 트래픽 모니터링 도구를 구현하였고, 실험을 통하여 IPv6 트래픽을 신뢰성있게 분석할 수 있다는 것을 보였다.

1. 서 론

인터넷에서는 많은 사용자가 다양한 응용 프로그램을 사용하고 있다. 웹, 이메일, 원격접속, 파일 다운로드 등의 전통적인 응용 뿐만 아니라, P2P, 멀티미디어 스트리밍, 인터넷 폰, 바이러스 등이 최근에 등장하였다. 응용별 트래픽 특성과 인터넷 트래픽의 패턴 등에 대한 분석, 보안 트래픽 차단, 응용별 과금 및 트래픽 엔지니어링을 수행하기 위해서는 네트워크 모니터링을 위한 도구가 필수적이다. 대표적으로 간단한 SNMP[1]를 기반으로 네트워크 주요 구성요소의 입출력 바이트/패킷 수를 관찰하는 다양한 도구가 개발되어 사용되고 있다. 하지만, 세부적인 응용 포트별 통계 정보를 생성할 수는 없다. 따라서, 좀 더 다양한 응용별 분석이 필요하다. 대표적으로 Cisco사의 라우터에서 패킷들의 흐름을 플로우로 분류하여 플로우별 트래픽 모니터링 기능을 NetFlow[2]로 제공하고 있다. 라우터에서의 트래픽 모니터링 기능은 라우터 자체의 성능을 저하시키지 않기 위하여 고속의 링크를 모니터링할 경우에는 패킷 샘플링 기법을 이용하여 플로우별 통계를 제공하고 있다. 이와는 별도로 링크계층의 프레임을 읽어들여 패킷별 트래픽 측정기능을 수행하고 있는 다양한 도구들이 있는데, 대표적으로는 pcaplib를 기반으로 하는 tcpdump[3]이다. 소프트웨어기반의 패킷 측정 도구는 고속 링크에서는 성능이 저하되기 때문에 전용 하드웨어로 구현된 것들이 제품화되어 출시되고 있다[4].

라우터에서 제공하는 트래픽 모니터링 표준은 IETF IPFIX(IP Flow Information eXport)로 제정되고 있는데, 이는 Cisco사의 NetFlow v9을 기초로 하고 있다. 현재는 NetFlow v5를 이용하는 도구들이 많이 사용되고 있다. 하지만, NetFlow v5는 IPv6와 MPLS와 같은 다양한 프로토콜의 트래픽을 모니터링할 수가 없고, 수집된 플로우별 통계 정보를 UDP로 전송하기 때문에 트래픽 측정 정보의 손실이 발생할 수 있다. 또한, 고정된 플로우 통계 필드만을 사용하기 때문에 다양한 분석 기능을 구현하기 어렵다. NetFlow v9기반의 IPFIX에서는 템플릿 구조를 이용하여 IPv6와 MPLS 등의 트래픽 측정이 가능하고, SCTP[4]와 TCP를 기본 전송프로토콜로 사용한다.

본 논문에서는 차세대 라우터에서의 트래픽 모니터링 표준으로 제안된 IPFIX(IP Flow Information eXport)를 지원하는 플로우별 트래픽 분석 도구의 프로토타입을 구현한 결과를 제시하도록 한다.

본 논문의 구성은 다음과 같다. 2장에서는 라우터/스위치에서 트래픽 모니터링 기능의 표준으로 제공될 IPFIX의 표준을 소개하고, 3장에서는 IPFIX을 기반으로 한 실시간 플로우 분석도구의 설계 및 구현에 대해 상세히 기술하고, 4장에서는 실험결과를 기술하며 마지막으로 결론 및 향후 과제를 5장에서 다룬다.

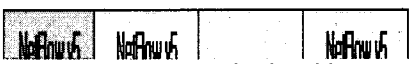
1) 본 연구는 정보통신부 대학 IT연구센터 육성사업과 한국전산원에서 지원을 받았음.

2. IPFIX 표준

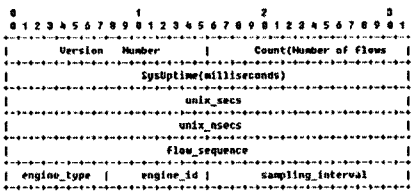
NetFlow는 패킷 헤더의 다섯가지 필드들(src IP addr, src port, dst IP addr, dst port, proto)을 이용하여 플로우를 생성하며 일반적으로 버전 5가 많이 사용되고 있다. 따라서, NetFlow v5와 NetFlow v9을 기반으로 한 IPFIX 표준을 설명하도록 한다.

2.1 NetFlow v5 구조

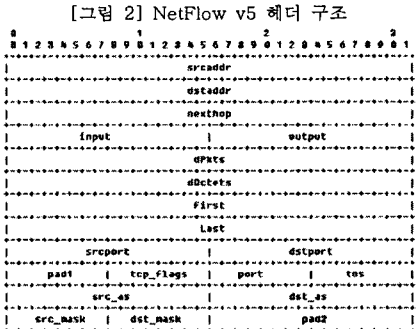
[그림 1]은 NetFlow v5에 대한 헤더 구조와 데이터 레코드 구조를 보여주고 있고, 헤더와 데이터 레코드는 각각 [그림 2]와 [그림 3]과 같다.



[그림 1] NetFlow v5 네트워크 전송 구조



[그림 2] NetFlow v5 헤더 구조



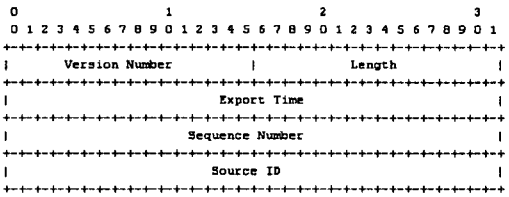
[그림 3] NetFlow v5 데이터 레코드 구조

[그림 1]에서 처럼 하나의 UDP패킷에는 하나의 NetFlow 헤더(24 바이트)와 최대 30개(30*48=1,440 바이트)의 데이터 레코드를 포함할 수 있다. 하나의 데이터 레코드는 하나의 플로우 정보를 가지고 있다.

2.2 NetFlow v9 기반의 IPFIX 구조

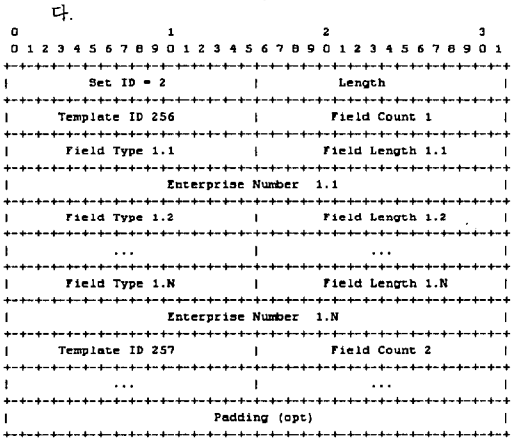
Cisco NetFlow v9을 바탕으로 표준화 작업이 진행 중이며 IPFIX 헤더[그림 4], IPFIX 템플릿 세트[그림 5], IPFIX 데이터 세트[그림 6]의 3가지 포맷으로 구성되어있다.

- IPFIX 헤더



[그림 4] IPFIX 헤더 구조

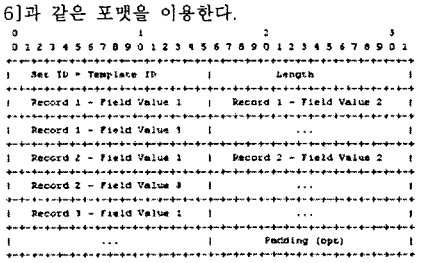
- IPFIX Template Set: IPFIX의 큰 장점 중 하나는 플로우 정보를 전달할 때 가변적으로 필요한 정보만으로 데이터 레코드를 구성하여 보낼 수 있도록 설정하는 것이다. Template Set은 가변적으로 데이터 레코드를 구성하기 위해 플로우 정보를 전달하기 전에 플로우 정보가 어떻게 구성되어있는지 알려주는 정보이다. 주요 필드들은 다음과 같다.



[그림 5] IPFIX Template Set 구조

1. Template ID [식별번호]: Data Set에 해당되는 Template 식별값.
2. Field Count [flow 순번]: NetFlow 데이터 패킷으로 전송할 Field type의 개수.
3. Field Type [flow 순번].[field 순번]: Flow 특성. (e.g. bytes, packets, protocol, addr...)
4. Field Length [flow 순번].[field 순번]: Field Type에 해당하는 값의 저장 공간 크기

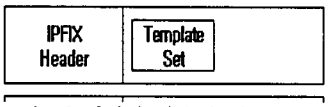
- IPFIX Data Set: 플로우 정보를 실제로 전달할 때 [그림 6]과 같은 포맷을 이용한다.



[그림 6] IPFIX Data Set

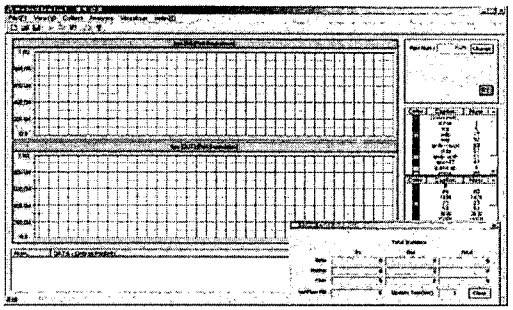
1. Record [flow 순번]
2. Field Value [flow 순번]: 측정 값이 저장됨.

- 네트워크 전송 패킷 구조: 본 논문에서 사용된 IPFIX 전송 구조는 [그림 7]과 같다.



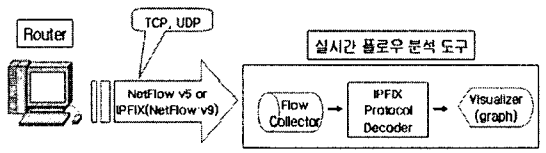
[그림 7] 실험에 사용된 레코드 구조

3. IPFIX 기반 실시간 플로우 분석 도구



[그림 8] 실시간 플로우 분석 도구

[그림 8]에서 윈도우 운영체제에서 구현된 IPFIX기반의 실시간 플로우 분석 도구의 전체적인 모습을 보여주고 있다.



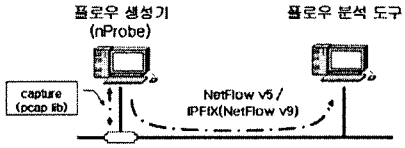
[그림 9] 실시간 플로우 분석도구 구조

본 논문에서 구현한 실시간 플로우 분석 도구의 구조는 [그림 9]와 같고, 주요 구성요소는 다음과 같다.

- 플로우 수집(Flow Collector) 모듈: UDP나 TCP로 전송되는 플로우 데이터를 받아서 NetFlow v5/v9(IPFIX) 데이터를 수집할 수 있도록 하였다.
- IPFIX 프로토콜 디코더(IPFIX Protocol Decoder) 모듈: NetFlow v5/v9 플로우를 디코딩 및 분석하여 시각화 모듈에서 그래프로 출력할 수 있도록 입출력 링크의 비트/패킷/플로우 전송률을 각 프로토콜과 포트번호로 구분하여 통계정보를 생성한다.
- 시각화(Visualizer) 모듈: 프로토콜 디코더에서 생성된 통계 정보를 전달 받아 주기적인 시간 단위로 그래프를 업데이트하여 네트워크 관리자가 쉽게 트래픽 경향을 파악할 수 있도록 하였다. 트래픽 정보를 비트 수, 패킷 수, 플로우 수 통계에 대하여 전체, 프로토콜별, 및 포트 별로 구분한 그래프로 볼 수 있도록 하였다.

4. 실험결과

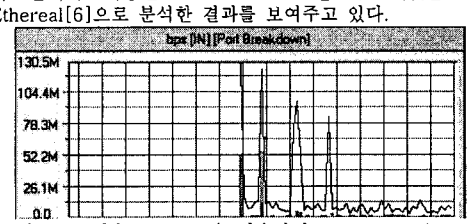
플로우 분석 도구의 실험을 위해서 tcpdump로 저장된 파일을 nProbe[6]라는 도구를 이용하여 NetFlow v5/v9 플로우를 생성한 후, 플로우 분석도구가 설치된 호스트로 TCP/UDP 프로토콜을 이용하여 전송하였다. 실시간 플로우 분석 도구에서는 전송된 NetFlow v5/NetFlow v9 플로우 패킷을 수신하여 그래프로 분석 결과를 나타내도록 하였다[그림 10]. 본 실험에서는 IPv6 트래픽에 대한 처리결과와 TCP/UDP 프로토콜을 이용할 때의 성능을 비교하였다.



[그림 10] 실험 환경

4.1 NetFlow v9를 이용한 IPv6 트래픽 분석

IPv6 트래픽(Secure copy 프로그램으로 파일 복사)을 생성한 후 이에 관한 플로우의 통계 정보를 NetFlow v9 포맷의 플로우 패킷을 생성하여 TCP 프로토콜을 이용하여 분석한 결과를 [그림 11]에서 보여주고 있다. [그림 12]에서는 [그림 11]의 트래픽 분석에 사용된 NetFlow v9 플로우 패킷을 캡쳐하여 Ethereal[6]으로 분석한 결과를 보여주고 있다.



[그림 11] TCP 전송 프로토콜을 이용하여 NetFlow v9(IPFIX) 플로우 분석

[그림 12]에서는 템플릿 세트(Template Set) 패킷 내용을 보여주고 있다.

```
0000  00 0f ea 34 e4 d0 00 03  47 72 9a f0 08 00 45 00
0010  00 6c 11 2c 40 00 40 06  7a d6 a8 bc 2e 8d a8 bc
0020  2e 84 ed 49 1f 2b fc 3b  b2 81 70 63 5b 54 80 18
0030  00 17 51 e1 00 00 01 01  08 0a 1e f2 ab ca 00 00
0040  00 00 00 09 00 01 00 00  0b 52 43 26 d4 0a 00 00
0050  00 1e 00 00 00 00 00 00  00 24 01 01 00 07 00 01
0060  00 04 00 02 00 04 00 04  00 01 00 07 00 02 00 1b
0070  00 10 00 0b 00 02 00 1c  00 10
```

[그림 12] NetFlow v9 Template Set

1. 00 09 : NetFlow version(NetFlow 데이터 패킷의 시작)
2. 00 24 : Template Flow 길이
3. 01 01 : Template ID
4. 00 07 : Template Field Count
5. 이후 4바이트 단위로 {필드 식별 코드(2), 필드 크기(2)} 쌍을 이룬다.

[그림 13]은 레코트 세트(Data Set) 패킷으로 밑줄 표시의 내용은 순서적으로 템플릿의 BYTE(4), PKCKET(4), PROTOCOL(1), SRC_PORT(2), IPV6_SRC_ADDR(16), DST_PORT(2), IPV6_DST_ADDR(16) 필드들을 보여주고 있다.

```
0000  00 0f ea 34 e4 d0 00 03  47 72 9a f0 08 00 45 00
0010  05 dc 11 2e 40 00 40 06  75 64 a8 bc 2e 8d a8 bc
0020  2e 84 ed 49 1f 2b fc 3b  b2 b9 70 63 5b 54 80 10
0030  00 17 33 23 00 00 01 01  08 0a 1e f2 ab cb 00 00
0040  00 00 00 09 00 01 00 00  0b 52 43 26 d4 0a 00 00
0050  00 1e 00 00 00 00 01 01  05 4a 00 00 01 68 00 00
0060  00 03 06 c3 c7 20 01 02  20 08 04 00 20 00 00 00
0070  00 00 00 00 01 00 16 20  01 02 20 08 04 00 20 00
0080  00 00 00 00 00 00 02 00  00 3c c0 00 00 00 1e 06
```

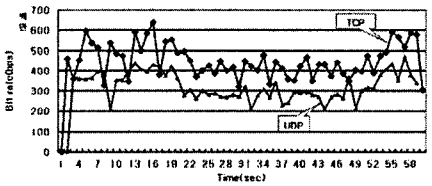
[그림 13] NetFlow v9 Data Set

[그림 13]에서 밑줄 표시의내용은 다음과 같이 순서적으로 설명한다.

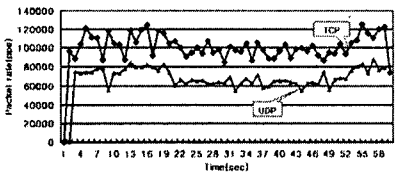
1. 00 00 01 68 : 측정된 Byte 수.(360)
2. 00 00 00 03 : 측정된 Packet 수.(3)
3. 06 : flow에 해당하는 프로토콜 번호.(TCP)
4. C3 C7 : 소스 호스트 포트번호.(50,119)
5. 20 01 … 00 01 : IPv6 소스 호스트 주소. (2001:220:804:20::1)
6. 00 16 : 목적지 호스트 포트번호.(22)
7. 20 01 … 00 02 : IPv6 목적지 호스트 주소. (2001:220:804:20::2)

4.2 NetFlow v9 플로우 전송 프로토콜의 성능 비교

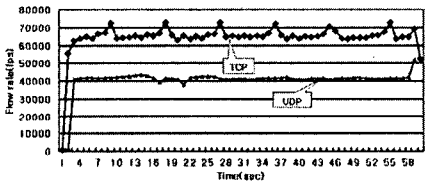
NetFlow v9 플로우의 전송 프로토콜을 TCP와 UDP로 이용했을 때 모니터링 트래픽의 비트/패킷/플로우 전송률은 크게 차이가 발생한다는 것을 [그림 14-16]에서 확인할 수 있다. 기존의 NetFlow v5는 UDP를 기본적인 전송 프로토콜로 사용하고 있기 때문에 IPFIX 표준에서 사용해야하는 SCTP/TCP와 같은 프로토콜을 이용하게 되면 보다 정확한 트래픽 통계를 생성할 수 있다.



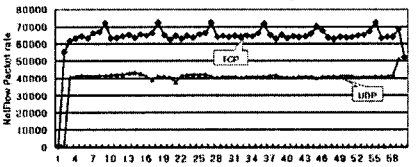
[그림 14] Bit rate(bps)



[그림 15]Packet rate(pps)



[그림 16] Flow rate(fps)



[그림 17] NetFlow Packet rate

5. 결론

본 논문에서는 IPFIX 표준을 지원하는 실시간 플로우 분석 도구를 설계하고 구현한 결과를 제시하였다. 구현한 실시간 플로우 분석 도구는 기존에 사용되는 NetFlow v5 뿐만 아니라 IPFIX 플로우를 분석할 수 있다. IPFIX를 지원하기 때문에 IPv4와 IPv6의 트래픽을 모두 처리할 수 있으며 전송 프로토콜로 TCP를 이용하도록 하여 플로우 수신률을 향상시키도록 하였고, 실험을 통하여 기본적인 기능과 성능을 보였다.

6. 참고문헌 :

[1] J. Case, M. Fedor, M. Schoffstall, J. Davin, "A Simple Network Management Protocol (SNMP)", IETF RFC 1157, May 1990.
[2] Cisco NetFlow, http://www.cisco.com/warp/public/cc/pd/iosw/ioft/netflct/tech/napps_ipfix-charter.html
[3] Tcpdump, http://www.tcpdump.org
[4] R. Stewart, Q. Xie, K. Morneault, C. Sharp, H. Schwarzbauer, T. Taylor, I. Rytina, M. Kalla, L. Zhang, V. Paxson, "Stream Control Transmission Protocol(SCTP)", IETF RFC 2960, October 2000.
[5] nProbe, http://www.ntop.org/nProbe.html
[6] ethereal, http://www.ethereal.com/