학내망에서 발생하는 IPv6-over-IPv4 트래픽 분석

박하늘, 박준상, 김명섭

고려대학 컴퓨터정보학과

{skzizix, junsang_park, tmskim}korea.ac.kr

Traffic analysis of IPv6-over-IPv4 in campus

요 약

IPv4 주소체계를 IPv6 로 변환하기 위해서는 장기간의 노력과 많은 비용이 요구되기 때문에 IPv6 로의 완전한 체계 변화까지는 상당 시간이 필요하다. 따라서 일부 IPv6 기반 네트워크의 지원을 위해서 IPv4 에 기반한 IPv6 트래픽이 증가하고 있는 추세이다. 하지만 현재의 트래픽 분석 시스템은 IPv4 트래픽에 중점을 두고 분석을 수행하고 있다. 현재 네트워크의 안정적인 관리와 효과적인 IPv6 체계로의 변환을 위해서는 IPv6 트래픽에 대한 분석과 관리가 요구되는 시점이다. 본 논문에서는 학내망에서 발생하는 IP protocol Number 41 인 IPv6-over-IPv4 트래픽 양에 대한 트랜드를 분석하고, 다양한 실험을 통하여 IPv6 트래픽의 특징을 분석하여 기술하였다.

1. 서 론

IP(Internet Protocol) 트래픽의 매년 성장률이 높아지고 있고 인터넷 사용자의 급증과 브로드밴드의 속도 향상, 인터넷 접속이 가능한 단말기의 증가로 인해 IP 트래픽은 앞으로도 지속 적인 성장률을 보일 것이다. 이러한 네트워크 환경 변화를 고려했을 때 네트워크 관리에 대한 중요성은 커져가고 있다.

트래픽 분석에 있어서 현행 IPv4 기반 인터넷 주소자원의 고갈은 예상보다 빨리 진행되고 있다. 차세대 인터넷 주소(IPv6) 전환이 필요한 시점에서 이미 IPv6 의 트래픽은 발생 하고 있고, 네트워크 안정적인 관리에 있어서 취약점으로 작용하고 있다. 따라서 네트워크상에서 IPv6 트래픽 발생 여부를 확인하고, 트래픽의 사용 목적을 판단하여 비정상적인 트래픽에 대한 차단이 요구된다.

학내망에서 발생하는 트래픽을 수집하여 트래픽을 분석한 결과, 2012.01.06 하루 동안 발생하는 트래픽 중 TCP, UDP 를 제외한 나머지(Other)의 대부분은 IP Protocol Number 41 트래픽이 차지했다. 또한, Total 트래픽에 대한 IP Protocol Number 41 의 비율은 평균 Flow 0.35%, Packet 3.77%, Byte 2.91%를 차지 했다. 이러한 관점을 비추어 볼 때 IP Protocol Number 41 트래픽에 대한 분석이 필요하다.

본 논문에서는 학내망에서 발생하는 IP Protocol Number 41 트래픽의 트랜드를 분석하고 특징을 분석하여 기술 하였다.

본 장에 이어서 2 장에서는 IPv6-over-IPv4 에 대한 전환기술과 관련연구를 소개하고 3 장에서는 IP protocol Number 41 인 IPv6-over-IPv4 트래픽양에 대한 트랜드를 나타내 보이고, 4 장에서는 IPv6-over-IPv4 트래픽이 갖는 특징을 분석하고 기술한다. 마지막으로 5 장에서는 결론 및 향후 연구를 기술한다.

2. 관련 연구

기존 네트워크와의 호환성을 위해 IPv4 를 IPv6 로 전환하는 기술로 듀얼스택, 터널링, 주소 변환이 있다[1]. 변환기법은 일반적으로 게이트웨이 상에서 계층에 따른 변환을 수행하는 기법으로서 헤더 변환, 주소 매핑, 프로토콜 변환과 같은 동작을 수행하고 터널링 기법을 적용 할 수 없는 상황에서 사용된다. 터널링 기법은 기존의 IPv4 인프라를 활용하여 IPv6 트래픽을 전송하는 방법을 제공하는 기법이다. 위의 두 가지 연동 기법 중에서 일반적으로 많이 사용되는 것은 터널링 기법이며 IPv6-over-IPv4 의 프로토콜이 41 번을 사용한다.

IPv4 를 사용한 터널링 기술에 관한 연구는 진행되어 왔지만 터널링 과정에서 발생되는 IP Protocol Number 41 트래픽에 대한 분석은 이루어지지 않고 있다[2]. IPv6 트래픽에 대한 모니터링과 분석연구 또한 많이 진행 되어 왔지만, 대부분 IPv4 기반의 도구이기 때문에 실제 사용되고 있는 네트워크에서 발생하는 IPv6 트래픽이 아닌, 임의로 IPv6 트래픽을 발생하였거나 시나리오 기반으로 연구가 진행되어 왔다[3].

* 이 논문은 정부(교육과학기술부)의 재원으로 2010 년도 한국연구재단-차세대정보컴퓨팅기술개발사업(20100020728) 및 2012 년도 한국연구재단(2012R1A1A2007483)의 지원을 받아 수행된 연구임.

따라서 본 논문에서는 현재 네트워크에서 발생하는 트래픽의 특징을 정확하게 분석하기 위해서 실망에서 발생하는 IP Protocol Number 41 트래픽 양에 대한 트랜드와 특징을 분석하였다.

3. IPv6-over-IPv4 트래픽양의 트랜드 분석

본 장에서는 논문에서 이용한 트래픽 트레이스에 대한 설명과 학내망에서 발생하는 IPv6-over-IPv4 의 트래픽 양의 변화 추이를 보여준다.

본 논문의 이 후의 내용에서는 IP Protocol Number 41 인 IPv6-over-IPv4 을 Prot.41 이라 명명하고 기술하도록 하겠다.

3.1 트래픽 트레이스

본 절에서는 실험에 사용되는 트래픽을 수집하기 위한 환경과 트래픽 구성에 대하여 기술한다.

그림 1 은 학내망에서 발생하는 트래픽의 수집 및 분석 구조를 나타내고 있다.

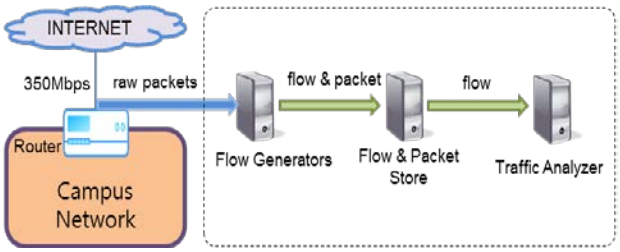

그림1. 학내망의 트래픽 수집 및 분석 구조

학내망과 인터넷을 연결하는 링크의 모든 패킷을 수집하여 플로우와 패킷 형태로 구성하고, 수집된 플로우와 패킷은 트래픽 트레이스 저장소에 저장되고 분석된다. 캡쳐 지점의 대역폭은 업링크과 다운링크를 포함하여 350Mbps 이며, 3000 여대의 호스트가 존재한다.

트래픽양의 변화 추세의 신뢰도를 향상 시키기 위해서 2012 년 1 년동안 학내망에서 발생한 전체 트래픽을 수집하고 분석하였다. 학내망 특성상 방학, 네트워크 장비 교체, 정전등의 원인으로 네트워크를 이용하지 않아서 트래픽양이 적은 기간이 있지만 데이터의 트랜드 분석에는 영향을 미치지 않을 것으로 판단된다. 2012 년 1 년 동안 수집된 트래픽 중 8 월 같은 경우가 학내 네트워크 공사로 인해 트래픽양이 적은 기간이다.

표 1 은 2012 년 학내망에서 발생한 총 트래픽을 Flow, Packet, Byte 를 기준으로 월별로 나타낸 것이다. 본 논문에서 Flow 는 트랜드 분석과 호스트간의 트래픽 발생 특징 분석에 용이하게 3-tuple(SrcIP, DstIP, Transport Layer Protocol)이 동일한 패킷의 집합으로 정의하였다. 본 논문의 3-tuple Flow 는 트래픽 분석에서 port 번호로 구분하며 일반적으로 사용하는 5-tuple 기반 Flow 여러 개로 나뉠 수 있다.

2012 년동안 월 평균적으로 Total 트래픽은 Flow 959M, Packet 52G, Byte 36TB 가 발생하였고, Prot.41 트래픽은 Flow 2,816K, Packet 736M, Byte 630GB 가 발생하였다.

표 1. 트래픽 트레이스 구성

Month	Total			Prot.41		
	Flows	Packets	Bytes	Flows	Packets	Bytes
1	765M	39G	34TB	1,818K	818M	565GB
2	844M	43G	38TB	1,806K	743M	609GB
3	1,107M	56G	48TB	2,068K	800M	732GB
4	745M	38G	31TB	1,274K	462M	423GB
5	486M	27G	22TB	1,102K	366M	325GB
6	1,272M	62G	51TB	2,789K	675M	589GB
7	651M	34G	24TB	2,775K	629M	525GB
8	235M	11G	9TB	1,093K	289M	252GB
9	1,581M	49G	40TB	4,855K	889M	763GB
10	1,341M	58G	46TB	4,820K	866M	740GB
11	1,388M	57G	46TB	5,093K	1,367M	1,230GB
12	1,094M	53G	43TB	4,295K	933M	805GB
Total	11,513M	532G	439TB	33,793K	8,843M	7,564GB

3.2 트래픽 양의 변화 추이

본 절은 Prot.41 트래픽양의 추이 변화를 보인다. 그림 2 는 1 년동안 발생한 전체 트래픽 대비 Prot.41 트래픽 변화량을 월별로 나타내고 있다.

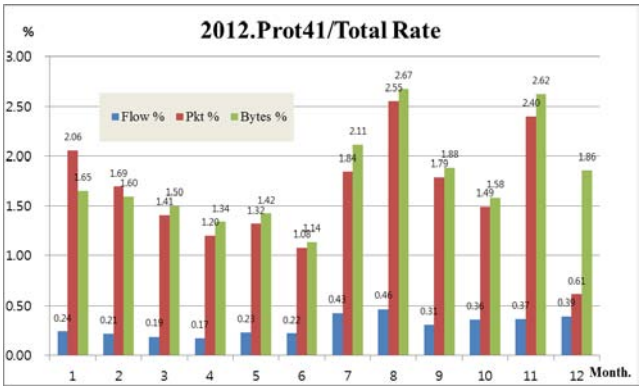

그림2. 전체트래픽에 대한 Prot. 41 비율

전체 트래픽 중 Prot.41 트래픽이 가지는 Packet 과 Byte 의 비율은 평균 1.79% 였다. 또한 Flow 를 기준으로 1~6 월은 약 0.21%였고, 7~12 월의 Flow 는 약 0.38%로 7 월부터 Prot.41 트래픽의 Flow 가 증가 추세를 보였다.

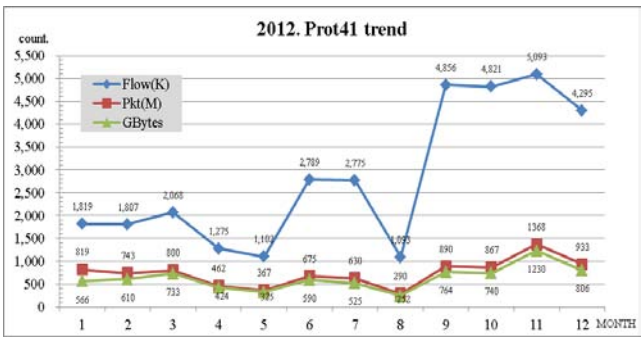

그림 3. 2012년 Prot.41 트래픽의 추이

그림 3 은 Prot. 41 트래픽의 Flow, Packet, Byte 양의 변화 추이를 나타내고 있다. Prot.41 트래픽에 대한 Flow 수는 6 월달부터 증가 추세를 보였고, 9 월달은 급격히 증가하였다. 또한 Packet 과 Byte 역시 9 월달부터는 증가하는 추세이다. 8 월달은 네트워크 공사로 인해 트래픽 양이 적기 때문에 추이를 반영하지 못하고 있다. 이러한 추세를 볼 때, 2013 년에는 더 많은 상승세를 보일 것으로 예상된다.

4. Prot. 41 트래픽 특징 분석

본 장에서는 Prot.41 트래픽의 특징 분석을 Local&Remote Host 간에 전송되는 상위 호스트의 Byte 분포 특징, Prot.41 트래픽의 Inbound&Outbound 특징, Prot.41 트래픽이 가지는 데이터의 유무로 총 3 가지 측면에 대하여 분석하였다.

4.1 상위 호스트의 Byte 분포 분석

분석의 신뢰성을 높이기 위하여 2012 년을 1 학기, 여름방학, 2 학기, 겨울방학 총 4 분기로 나누고 해당 분기에서 임의의 달을 선택하여 1 주차 동안 실험을 하였다.

3.1 절에서 정의한 Flow 의 출발지 및 목적지 주소를 Local 과 Remote 로 분류하고 하루 동안의 Flow 를 모아 Bytes 를 기준으로 하여 Top Flow 를 선정하였다. Top10 / Byte 는 Flow 의 상위 10 개 이며 Top 10% / Byte 는 상위 10%개 이다.

표2. Top10 / Mbyte

	일	월	화	수	목	금	토
5월	97.87%	96.93%	77.33%	94.24%	89.61%	94.45%	98.96%
	1433/1464	6002/6192	6017/7781	4388/4657	2673/2983	6179/6541	1718/1736
7월	79.16%	34.45%	35.12%	60.93%	34.02%	90.56%	52.70%
	5816/7348	6691/19423	7326/20862	10136/16636	6825/20061	2397/2647	11218/21286
9월	55.94%	35.53%	37.41%	49.93%	40.59%	42.15%	56.36%
	7430/13284	7457/20989	11184/29893	20074/10022	8538/21035	6081/14427	5298/9401
1월	89.48%	51.03%	51.02%	64.07%	66.65%	87.76%	56.95%
	6118/6837	15175/29740	7115/13945	9636/15041	7562/11346	8824/10054	11259/19770

표 2 는 분기 별로 나누고 임의의 달 5, 7, 9, 1 월을 선택 하여 일주일 동안 Top10 / Byte 의 실험 결과 이며, Total Byte 에 대한 Top10 Byte 비율과 각각의 Byte 양을 알 수 있다. 표 3 는 표 2 와 같은 데이터를 가지고 Top10% Byte 에 대한 실험 결과 이다.

표3. Top10% / Mbyte

	일	월	화	수	목	금	토
5월	99.75%	99.92%	99.86%	99.85%	99.71%	99.93%	99.73%
	1461/1464MB	6187/6192	7770/7781	4650/4657	2974/2983	6537/6541	1731/1736
7월	99.68%	99.61%	99.83%	99.87%	99.60%	99.39%	99.92%
	7324/7348	19348/19423	20827/20862	16615/16636	19982/20061	2630/2647	21271/21286
9월	99.36%	99.07%	99.77%	99.69%	99.67%	99.75%	99.77%
	13201/13284	20794/20989	29824/29893	20014/20074	20967/21035	14392/14427	9380/9401
1월	99.85%	99.91%	99.75%	99.82%	99.83%	99.90%	99.94%
	6828/6837	29715/29740	13910/13945	15014/15041	11327/11346	10045/10054	19759/19770

표 2 를 보면 Top10 Byte 는 Total Byte 의 최대 98% 의 Byte 를 차지한 날이 있었고, 대부분이 Total Byte 의 많은 비율을 지녔다고 볼 수 있다. 또한 표 3 Top10% Byte 를 통해서 Top 10%는 Total Byte 의 약 99.8%로 Byte 의 거의 대부분을 차지하고 있다는 것을 알 수 있다.

4.2 Prot.41 트래픽의 Outbound, Inbound 분석

4.2 절은 Prot.41 트래픽의 Outbound 와 Inbound 분석을 총 3 단계로 나누어 기술하였으며 1 단계는 Local&Remote IP, 2 단계는 단방향&양방향, 3 단계는 1, 2 단계를 조합하여 Local&Remote 에 대한 단방향&양방향의 트래픽을 분석하여 기술하였다.

Prot.41 트래픽의 Local&Remote IP 분류

학내망에서 발생하는 Prot.41 트래픽의 출발지 및 목적지 주소를 Local 과 Remote 로 분류하고 Packet, Byte 기준으로 Inbound&Outbound 를 분석하였다.

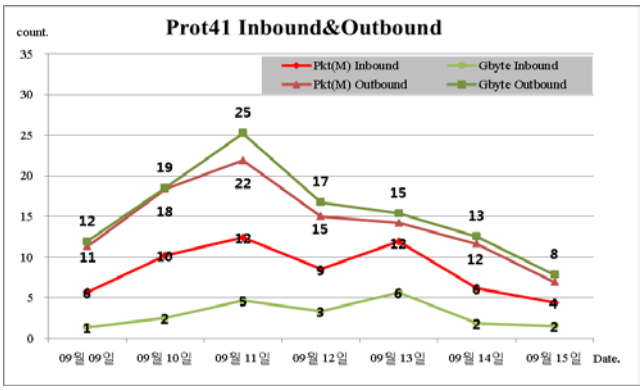

그림4. Prot.41 Inbound&Outbound

그림 4 를 보면 Prot.41 트래픽을 Local 과 Remote 로 분류하였을 때, Inbound 가 Outbound 보다 많은 트래픽을 발생하였다. 2012 년 9 월 1 주차동안 Inbound 트래픽은 Packet 59M, Byte 21GB 였으며 Outbound 트래픽은 Packet 99M, Byte 108GB 로 평균적으로 Inbound 의 Packet 은 Outbound Packet 보다 약 2 배가 높았고, Byte 는 약 5 배가 높았다.

Prot.41 트래픽의 단방향&양방향 분류

학내망에서 발생하는 Prot.41 트래픽을 단방향&양방향 트래픽으로 분류하여 Flow, Packet, Byte 기준으로 분석하였다.

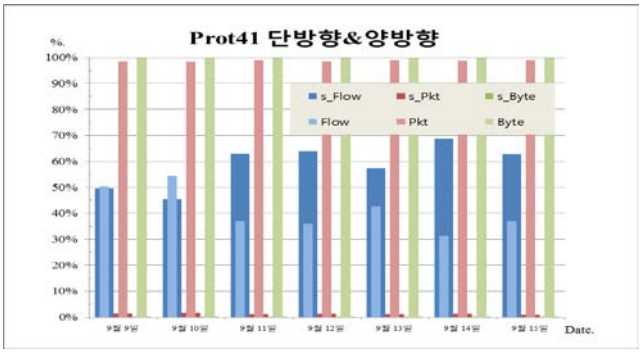

그림5. Prot.41 트래픽의 단방향&양방향

그림 5 를 보면 단방향의 Flow 가 양방향에 비해 대부분 많은 Flow 를 가지고 있었고, 2012 년 09 월 1 주차동안 단방향은 평균 Flow 56.13%, Packet 1.30%, Byte 0.19%를 차지했고 양방향의 경우 Flow 43.87%, Packet 98.70%, Byte 99.81%로 적은 양의 Flow 에도 불구하고 Packet, Byte 의 비율을 대부분 차지하는 것으로 보아 비교적 큰 데이터를 차지하고 있다는 것을 알 수 있다. 또한 단방향의 경우 Flow 양에 비해 너무 적은 Packet 과 Byte 를 지닌다는 점은 비정상적인 트래픽 일 수도 있다.

Prot.41 단방향&양방향 트래픽의 Local&Remote IP

학내망에서 발생하는 Prot.41 트래픽의 출발지, 목적지를 Local 과 Remote 로 분류하고 분류된 트래픽을 다시 단방향과 양방향으로 분류하여 Flow, Packet, Byte 를 분석하였다. 2012. 9 월 1 주차 트래픽을 가지고 실험을 하였고, 실험 결과는 다음과 같다.

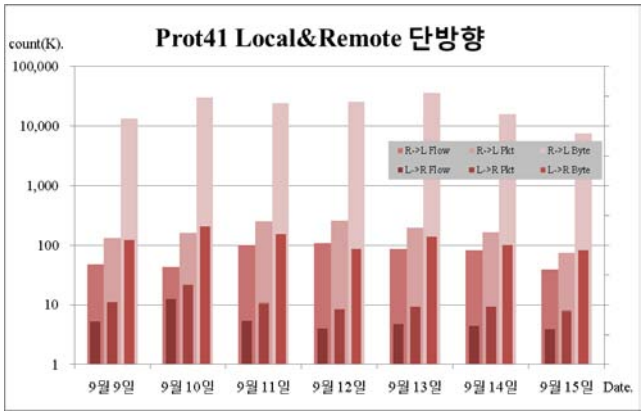

그림6. Prot.41 단반향 트래픽의 Local&Remote IP

그림 6 은 Prot.41 단방향 트래픽을 Local&Remote 로 분류하였을때의 Flow, Packet, Byte 양이다. R->L 은 Remote 에서 Local 방향을 나타내며, L->R 은 Local 에서 Remote 방향을 나타낸다.

표4. Prot.41 단방향 트래픽의 비율

Date	Local -> Remote			Remote -> Local			Total Simplex		
	Flow	pkt	Byte	Flow	pkt	Byte	Flow	pkt	Byte
9월 9일	37.43%	47.47%	53.16%	62.57%	52.53%	46.84%	76K	252K	28M
9월 10일	78.14%	73.66%	59.23%	21.86%	26.34%	40.77%	197K	625K	74M
9월 11일	22.85%	31.40%	49.77%	77.15%	68.60%	50.23%	130K	360K	47M
9월 12일	13.00%	21.90%	22.98%	87.00%	78.10%	77.02%	125K	325K	32M
9월 13일	20.32%	31.05%	35.19%	79.68%	68.95%	64.81%	108K	285K	54M
9월 14일	18.73%	34.99%	38.14%	81.27%	65.01%	61.86%	102K	254K	26M
9월 15일	28.27%	45.56%	48.50%	71.73%	54.44%	51.50%	54K	135K	14M

그림 6 을 보면 분류되어진 Prot.41 단방향 트래픽은 Local 에서 Remote 방향이 Remote 에서 Local 방향 보다 Flow, Packet, Byte 의 양이 모두 낮다는 것을 알 수 있다. 또한, 표 4 에서 볼 수 있듯이 Total 에 대한 비율을 보면 Remote->Local 의 비율이 Flow 기준 약 70%, Packet, Byte 기준 약 60%를 나타냈다. 결과적으로 Prot.41 단방향 트래픽은 외부에서 들어오는 트래픽이 대부분 이라는 것을 알 수 있다.

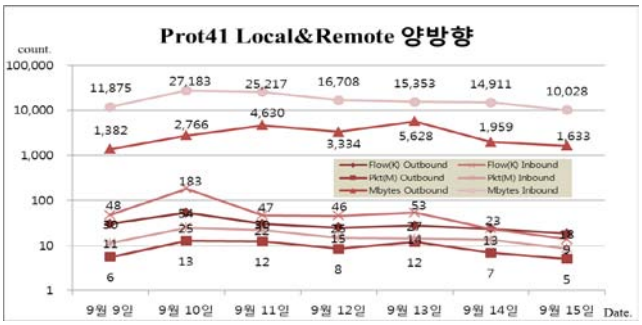

그림7. Prot.41 양방향 트래픽의 Local&Remote IP

그림 7 은 그림 6 과 같은 실험데이터로 Prot.41 양방향 트래픽을 Local&Remote 로 분류하였을 때 의 Flow, Packet, Byte 양이다. Prot.41 양방향 트래픽의 경우에도 Inbound 의 Flow, Packet, Byte 양이 더 많다는 점을 알 수 있다. 결과적으로 Prot.41 트래픽은 단방향과 양방향 모두 Outbound 보다 Inbound 의 Flow, Packet, Byte 양이 많았으며 양방향 트래픽은 상대적으로 적은 Flow 에도 불구하고 많은 Byte 를 발생 시킨다는 점을 알 수 있다. 따라서 양방향 트래픽의 경우 특정 어플리케이션이 사용 될 확률이 높고 단방향 트래픽 경우에는 공격성을 가진 트래픽일 수도 있다.

4.3 Prot.41 트래픽의 데이터의 유무

4.3 절은 Prot.41 트래픽에 대한 데이터의 유무를 판단하기 위해 Flow/Byte 의 분포와 Flow 의 Duration 비율에 대해 2012.09 1 주차동안 실험을 하였고 분석 내용을 기술하였다.

Flow / Byte 의 분포

4.2 절에서 분류되어진 트래픽을 바탕으로 Flow 당 Byte 의 분포를 구하여 분석하였다.

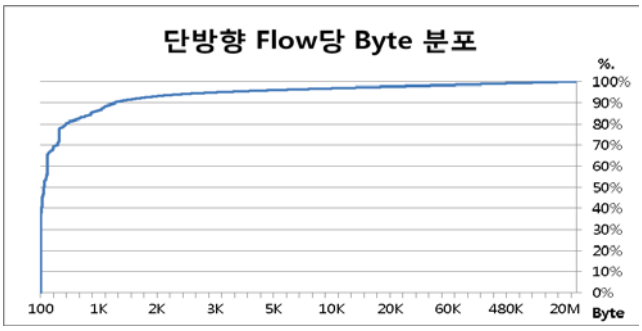

그림8. 단방향 Flow의 Byte 분포

2012.09 1 주차 동안 발생한 Prot.41 단방향 트래픽의 Byte 분포를 보면 78Byte ~ 232Byte 구간에 50%를 넘고 78Byte ~ 1400Byte 구간에는 90%를 차지 하는 것을 볼 수 있다. 결과적으로 Prot.41 이 발생하는 단방향 트래픽의 Flow 는 대부분이 최소 Byte 만을 가지는 트래픽이고 70% 이상이(표 4 단방향 비율 참조) Remote 에서 들어오는 트래픽이다.

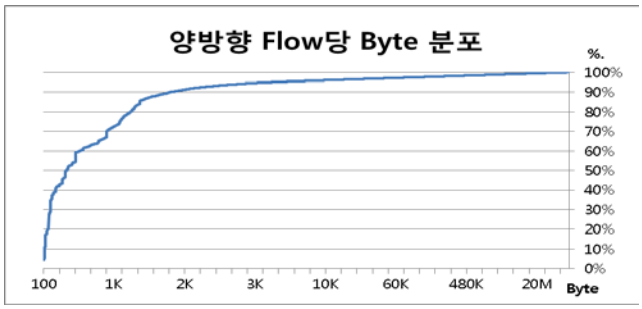

그림9. 양방향 Flow의 Byte 분포

그림 9 는 그림 8 과 같은 실험 데이터로 양방향 Flow 의 Byte 분포이다. 단방향과 비교해보면 좀 더 많은 Byte 를 가지는 Flow 가 많다는 점을 알 수 있고 따라서 양방향 Flow 는 데이터를 지닐 확률이 높다.

Flow / Duration 비율

2012.09 1 주차 동안 Flow 의 단방향과 양방향 트래픽 Flow Duration 비율을 분석하였다.

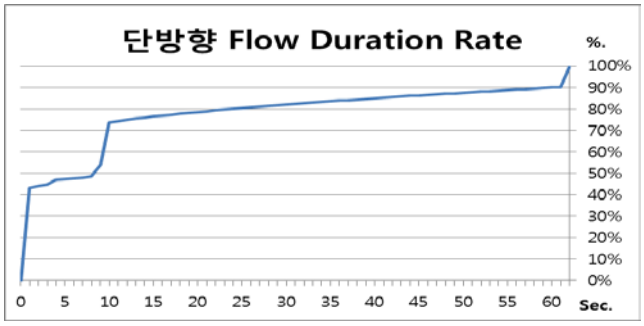

그림10. 단방향 Flow Duration 비율

그림 10 은 단방향 Flow Duration 을 누적분포로 나타낸 것이다. 그림 10 을 보면 1 초 미만에서 약 47%를 차지하며 8~10 초 구간에선 약 15%를 차지하는 급격한 변화를 나타냈고 10 초 미만 구간에서는 약 75%를 차지했다. 60 이상 구간은 60 초 ~ 60,967 초에 해당하고 그 분포가 산발적이며 비율 또한 크지 않기 때문에 그림 10 에서는 이러한 구간에 대하여 60 초 이상으로 표현하였다. 60 초 이상의 비율은 약 10%를 나타냈다.

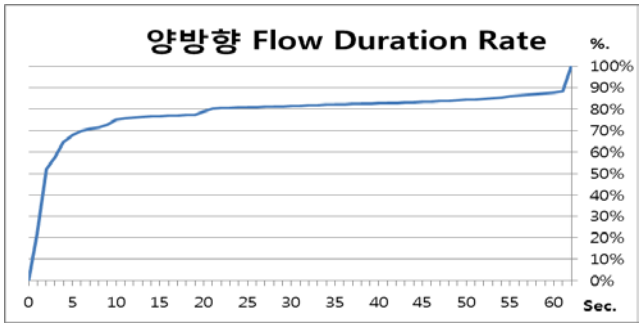

그림11. 양방향 Flow Duration 비율

그림 11 은 양방향 Flow Duration 의 누적분포 그림이며, 단방향과 비교하여 보면 10 초 미만의 구간에서 약 75%로 비슷한 양상을 나타냈지만 단방향 Flow Duration 8~10 초 구간에서 나타난 급격한 변화는 없었다. 또한 60 초 이상의 비율은 약 12%로 2% 더 많은 비율을 가지고 있었다.

5. 결론 및 향후 연구

Prot.41 트래픽의 Trend 를 보면 점점 증가하는 추세이고, 전체 Byte 에 대해서도 약 2%로 간과해서는 안될 트래픽을 가지고 있다.

단방향 트래픽은 Prot.41 이 발생하는 트래픽 중 많은 Flow 를 차지하지만 Packet, Byte 양은 미미 하였고 그 중 외부에서 들어오는 Flow 양이 약 70%를 차지한 점과 Flow 당 Byte 의 분포와 Duration Rate 를 비추어 볼 때, 단방향 트래픽의 Flow 는 많은 Byte, 긴 Duration, 혹은 두가지 모두 지닌 Flow 가 하나 이상은 존재한다는 점에서 비정상 트래픽일 확률이 있다.

양방향 트래픽은 Prot.41 트래픽의 절반 이하로 나타났는데 비교적 적은 Flow 에도 불구하고 Packet, Byte 의 양이 거의 대부분을 차지한다는 점과 단방향 트래픽보다 Flow 의 Byte 분포를 볼 때 데이터를 지닐 확률이 높다는 점이다. 따라서 이러한 추세와 특징을 비추어 볼 때 Prot.41 이 발생하는 트래픽에 대한 연구는 필요하다.

향후 연구로 Prot.41 이 발생하는 트래픽에 대한 상세한 분석이 요구되고 만약 발생한 트래픽 중 응용이 사용됐다면 그 특정 응용이 무엇인지, 사용되지 않았다면 Prot.41 트래픽이 가지는 데이터가 무엇인지를 판별해야 한다. 또한 비정상 트래픽이 발생한다면 이런 트래픽에 대해서 어떻게 대처 해야 할지도 생각해 보아야 할 것이다.

6. 참고 문헌

[1] 김미영, 문영성, “IPv4/IPv6 변환기 보안 위협”, 한국통신논문지, 제23 권 제 9호, pp. 53-64, 2006.9.
[2] 이정남, 장주욱, “방화벽에 호환성을 갖는 IPv6 터널링 기법 및 구현”, 한국정보과학회 학술발표논문지, 제29 권 제 2 호, pp. 382-384 , 2002.10.
[3] 김선영, 이홍규, 오승희, 서동일, 오창석. “IPv6 기반 트래픽 분석 도구 설계”. 한국콘텐츠학회 논문지, 제 5 권 제 2 호, pp.115-121, 2005.4.
[4] “Generic Packet Tunneling in IPv6 Specification” http://tools.ietf.org/html/rfc2473.
[5] B. Carpenter and K. Moore, “Connection of IPv6 Domains via IPv4 Clouds”, RFC 3056, Feb. 2001.
[6] E. Chen, et al, “Analysis of IPv6 Network Communication UsingSimulation”, in Research and Development, SCOReD 2006, pp. 11-15, 2006.