제18회 한국정보처리학회 추계학술발표대회 논문집 제9권 제2호 (2002.11)

한국전력 백본망 실시간 트래픽 측정 및 분석

오도은*, 이진기*
*한전 전력연구원
e-mail:hifive@kepri.re.kr

Realtime Measurement and Analysis of KEPCO's Backbone Traffic

Do-Eun Oh*, Jin-Kee Lee*
*Information & Telecommunication Group, KEPRI

요 약

네트워크 기반의 응용 서비스들이 다양하게 개발되고 업무 전산화로 인한 분산 컴퓨팅 환경이 확대됨에 따라 네트워크 트래픽은 증가 일로에 있다. 특히, 많은 트래픽이 집중되는 백본망에서는 제한된 자원을 효율적으로 사용하기 위하여 효과적인 트래픽 관리 방법이 요구되고 있다. 이를 위해서는 우선적으로 네트워크 트래픽의 총 양에 의한 정량적이고 단순 평면적인 정보가 아닌 네트워크 트래픽을 정밀 분석하고 특성을 파악할 수 있는 방법이 필요하며 또한 어느 한 순간의 데이터나 일괄처리 방식에 의한 분석이 아닌 실시간으로 장기적인 트래픽 측정 및 분석이 필요하다.

본 논문은 트래픽 측정방법 및 특성과 한국전력 백본망을 대상으로 네트워크 트래픽의 발원지로부터 목적지까지의 트래픽을 플로우별로 수집, 분석할 수 있는 트래픽 측정 및 분석 시스템에 대하여 기술하였다. 본 시스템은 백본망 전체의 트래픽을 지속적이며 실시간으로 이용자별, 시간대별, 프로토콜별, 응용 서비스별로 정보 추출이 가능하며 그래픽 환경과 웹 기반의 사용자 환경을 제공한다.

1. 서론

인터넷의 폭발적인 증가와 함께 최근의 업무 환경이 클라이언트/서버 모델의 분산 컴퓨팅 환경으로 변화함에 따라 네트워크 트래픽이 증가 일로에 있다. 이에 따라 어느 곳에서, 어떤 형태의, 얼마나 많은 트래픽이 유발되고 소모되는지를 알아내는 일은 한정된 네트워크 자원을 효율적으로 활용하기 위한 네트워크 관리자들에게 당면한 중요한 과제가 되었다. 이를 위해서는 먼저 네트워크 트래픽 측정 및 분석 작업이 필수적이다. 특히, 많은 트래픽이 집중되는 백본망에서는 다양한 품질을 요구하는 여러 가지 응용 서비스들이 동시에 이용되고 있기 때문에 네트워크 트래픽 측정 및 분석이 반드시 필요하다.

네트워크 트래픽 측정은 그 사용목적에 따라서 측정방법이 다양하게 분류될 수 있다. 현재 네트워크 운영 관리에는 네트워크관리시스템에서 사용되는 SNMP 기반 Passive 측정방식이 가장 널리 사용되고 있다.[1] 하지만 이 방식은 SNMP MIB를 통한 네트워크 트래픽의 총 양에 대한 정보를 제공해 줄 뿐 어떤 호스트나 응용 서비스에서 어느 정도의 트래픽을 발생시키는지, 어떤 프로토콜이 이용되고 있는지 등의 정보는 제공하여 주지 못한다.[2] 이러한 문제를 해결하기 위해 패킷 캡처나 RMON Probe를 통한 트래픽 분석 도구들이 사용되고 있지만 대규모의 네트워크에서는 적합하지 못할뿐더러 백본망이 아닌 서브넷 단위의 네트워크를 대상으로 하고 있다. 따라서 백본망을 대상으로 네트워크 트래픽의 총 양에 의한 정량적이고 단순 평면적인 정보가 아닌 네트워크 트래픽을 정밀 분석하고 특성을 파악할 수 있는 방법이 필요하며 또한 어느 한 순간의 데이터나 일괄처리 방식이 아닌 실시간으로 장기적인 트래픽 측정 및 분석이 필요하다.

본 논문은 한국전력 백본망을 대상으로 트래픽을 지속적이며 실시간으로 이용자별, 시간대별, 프로토콜별, 응용 서비스별로 측정 및 분석할 수 있는 시스템에 대하여 설명한다. 본 논문은 먼저 Passive 트래픽 측정방법 가운데 네트워크 장치에 기반을 둔 Netflow 기반 트래픽 측정 방식에 대하여 살펴보고

이를 바탕으로 개발된 트래픽 측정 및 분석 시스템에 대하여 소개한 후 결론으로 맺는다.

2. 플로우 기반 트래픽 측정 방법

2.1 플로우

IP 플로우는 그림 1과 같이 연속된 IP 패킷들로 정의될 수 있다. 즉, IP 플로우는 응용의 주소 쌍(송신자 주소, 송신자 포트 번호, 수신자 주소, 수신자 포트 번호), 호스트 쌍(송신자 호스트 주소, 수신자 호스트 주소), 네트워크 주소 쌍(송신자 네트워크 주소, 수신자 네트워크 주소), AS 번호 쌍(송신자 AS 번호, 수신자 AS 번호) 등으로 주어지는 명세를 만족시키는 제한된 시간 내에 도착하는 IP 패킷들의 흐름으로 정의된다.[3]

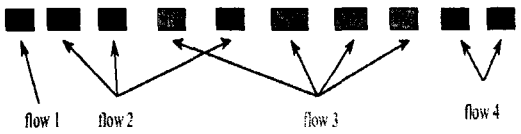

그림 1. 플로우(flow)

플로우 스펙이라고 부르는 일정한 조건들을 만족하면서 도착하는 패킷들은 일정한 간격으로 플로우 타이머에 의해 검사되어 지정한 시간동안 플로우 스펙을 만족하는 패킷이 도착하지 않으면 플로우의 종료를 선언한다. 트래픽 측정에서 사용되는 플로우 스펙은 패킷들의 정보를 다양한 측면에서 종합할 수 있도록 하는 것이다. 이러한 플로우 스펙은 다음과 같은 인자들로 구성된다.

● 플로우 방향성

플로우는 단방향 또는 양방향의 패킷 전송으로 구분할 수 있다. 예를 들어 TCP 연결에서는 A에서 B로의 데이터 전송은 B에서 A로의 데이터 전송을 수반한다.

● 플로우 종단의 수

플로우는 기본적으로 양쪽의 종단 사이에서의 패킷 흐름이지만, 특정 목적지 종단으로 가는 모든 패킷들의 흐름을 하나의 종단만으로 정의할 수 있다.

● 플로우의 종단 타입

포트, 호스트, 서브넷, 네트워크, AS 및 공통 경로 등의 여러 가지 종단 종류에 따라 패킷들을 다양한 측면에서 플로우로 정의할 수 있다.

2.2 Netflow 기반 트래픽 측정 방식

2.2.1 고려사항

Netflow 기반 트래픽 측정에 앞서 다음과 같은 점들을 사전에 고려하여야 한다.

- 네트워크 장치가 Netflow를 지원 하는가
- Forwarding 가능한 Netflow version은 얼마인가
- Netflow를 이용함으로써 네트워크 장치의 성능에 어느 정도의 영향을 줄 것인가
- Netflow를 이용함으로써 어느 정도의 트래픽을 유발하는가
- 측정시스템이 어느 정도의 트래픽 양을 처리할 수 있는가
- 어느 인터페이스를 측정할 것이며 그 때 얻을 수 있는 정보는 무엇인가
- 어느 기간동안 측정할 것인가
- 측정 데이터의 보존은 어떻게 할 것인가
- 보존할 데이터는 raw 데이터인가, flow 데이터인가, 가공된 결과인가

2.2.2 Netflow 기반 트래픽 측정 도구

● FlowScan

Netflow 정보를 바탕으로 원하는 기간만큼의 프로토콜별, 응용 서비스별 네트워크 사용에 대한 시간축 상의 그래프를 그려낸다. 또한 Custom graph, Top AS usage, Top user, Raw flow dump 등 대부분의 유용한 분석이 가능한 것이 장점이다. 하지만 시각화된 정보 이외의 통계 데이터를 얻기 힘들다.

● Cflowd

Netflow로부터 얻은 플로우 정보를 수집하여 Arts++ 형태의 파일로 저장한다. 기본적으로 수집부분만 담당하며, 분석은 Arts utility를 이용해 몇 가지 정해진 통계 정보를 텍스트 형태로 얻어낼 수 있다.

● MADAS

MIRNET-STARTAP 사이의 트래픽 측정을 위해 러시아에서 개발한 트래픽 측정 도구로 실시간 그래프 기능과 쿼리 인터페이스가 가능한 이상적인 도구이다. MADAS는 국가간 트래픽 측정에 초점이 맞추어져 있다. 이는 다른 측정 도구에는 없는 특이한 기능인데 AS를 바탕으로 국가간 트래픽을 분석해 낸다는 점에서 백본망에 어울리는 기능이라 할 수 있다. 쿼리의 결과가 그래프로 나타내지는 장점이 있긴 하지만, 응용 서비스별 분석이 불가능 한 점은 취약한 부분이다.

● Flowtools

Cflowd & ARTS와 비슷한 기능을 한다. Netflow로부터 정보를 모아 미리 정해진 20가지의 통계 정보를 얻어낼 수 있다. CGI를 사용해 웹에서 쿼리가 가능하지만 기본적으로 커맨드 기반의 프로그램이다.

● NetFlow FlowCollector

CISCO에서 개발한 상업용 Netflow 기반 트래픽 측정 도구이다. Netflow 트래픽을 export device로부터 수집하고 가공하여 저장하는 역할을 한다. 저장된 데이터는 Netflow DataAnalyzer와 같은 다른 Netflow 응용프로그램에 의해 사용될 수 있다. Thread와 Filter라는 기능에 의한 집합에 의해 사용자가 원하는 정보를 정확하고 간결하게 저 용량으로 저장할 수 있고, 플로우를 다양한 방법으로 집합할 수 있다. 또한 포트, 호스트, AS, 프로토콜 등 다양한 방법으로 데이터를 집합해서 저장할 수 있다. 그러나 상용 트래픽 측정 도구로 매우 고가이며, Raw data 분석이 용이하지 못한 단점이 있다.

● NetFlow DataAnalyzer

CISCO에서 개발한 상업용 NetFlow 기반 트래픽 측정 도구로 Netflow FlowColletor에 의해 수집, 집합, 저장된 Netflow 데이터를 받아와서 시각화한다. Netflow FlowCollector에게서 데이터를 받아오는 형식이므로 Netflow FlowCollector에서 어떻게 설정을 해서 데이터를 수집했는지가 중요하다. 여러 가지 집합으로 데이터를 볼 수 있게 하며 Time Slider, AS Drilling Down, Search 등의 특수한 기능도 제공 한다. 그래프에서 시간 축으로 전체 트래픽이 표시가 안 되고, 웹이 지원이 안 되며 느리다는 단점이 있다.

2.2.3 비교 분석

각 Netflow 기반 트래픽 측정 도구들을 항목별로 비교하면 표 1과 같다.[4]

2.2.4 Netflow 기반 플로우 분류 기준

Netflow 기반 트래픽 분석에서는 Source Address, Destination Address, Source Port, Destination Port, Layer 3 Protocol, Type Of Service(TOS) Byte, Input Interface 등 7가지 기준으로 패킷을 플로우로 분류하고 이에 대한 정보를 저장한다.

2.2.5 Netflow 기반 트래픽 분석 정보

다음 표 2는 Netflow 기반 트래픽 분석 정보를 보여준다.

표 2. Netflow 기반 트래픽 분석 정보

Usage	- Packet Count - Byte Count
Time of Day	- Start Time Stamp - End Time Stamp
Port Utilization	- Input Interface Port - Output Interface Port
QoS	- Type of Service - TCP Flags - Protocol
From/To	- Source IP Address - Destination IP Address
Application	- Source TCP/UDP Port - Destination TCP/UDP Port
Routing and Peering	- Next Hop Address - Source AS Number - Destination AS Number - Source Prefix Mask - Destination Prefix Mask

3. 트래픽 측정 및 분석 시스템

기존의 Netflow 기반 트래픽 측정 도구들 가운데 쿼리 인터페이스와 시각화 기능을 동시에 가지고 있는 도구는 존재하지 않는다. 따라서 본 시스템은 두 가지 기능을 모두 가지도록 기존의 Cflowd를 개

표 1. Netflow 기반 트래픽 측정 도구 비교

비교 항목	FlowScan	Cflowd	MADAS	Flowtools	FlowCollector	DataAnalyzer
Input	Cisco Netflow	Cisco Netflow	Cisco Netflow	Cisco Netflow	Cisco Netflow	FlowCollector
Output	Web(Graph)	Flow Dump	Web(Graph)	Text	Flow Collect	Graph/Table
Realtime Analysis	○	×	○	×	해당사항 없음	○
Query Interface	×	×	○	○	해당사항 없음	○
User Interface	Command	Command	GUI	Command	Command	GUI
Raw Data Collect	○	○	○	○	○	해당사항 없음
Data Analysis	○	○	○	○	해당사항 없음	○
Visualize	○	×	○	×	해당사항 없음	○
Time Series Graph	○	×	○	×	해당사항 없음	○
Raw Flow Dump	○	○	-	-	○	해당사항 없음

선하였다. 이 때 Cflowd의 텍스트 기반 저장 형태를 데이터베이스 저장 형태로 변경함으로써 그래픽을 포함한 웹 기반의 사용자 환경을 구축할 수 있게 하였다. 이를 통하여 어느 한 순간의 데이터나 일괄처리 방식이 아닌 실시간으로 장기적인 트래픽 측정 및 분석이 가능하게 되었다.

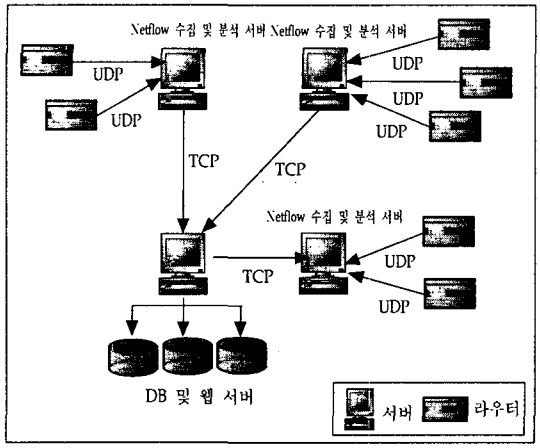

그림 2. Netflow 기반 트래픽 분석 시스템 구성도

그림 2에서 각 라우터는 Flow export 패킷을 cflowdmux와 cflowd를 실행하고 있는 Netflow 수집 및 분석 서버에 보낸다. Netflow 수집 및 분석 서버의 cflowdmux는 UDP 데이터그램인 Flow export 패킷을 수신하여 공유메모리 버퍼에 쓰고, Netflow 수집 및 분석 서버의 cflowd는 공유메모리에 쓰여진 패킷을 읽어서 파일에 저장한다. 이 때 서브넷별 프로토콜 사용 분포, AS별 포트 사용 분포 등의 정보를 얻어낼 수 없는 cflowdmux와 cflowd의 단점을 보완하기 위하여 flowmediator를 개발하였다. flowmediator는 5분 간격으로 저장된 파일을 읽어서 이를 인터페이스별로 파일에 저장하고 DB 및 웹 서버의 요청에 따라 저장된 파일을 전송하며 생성된 후 일정 시간이 지난 데이터 파일은 삭제한다. DB 및 웹 서버는 Netflow 수집 및 분석 서버로부터 저장된 파일을 전송 받아 이를 분석해 Raw Flow 데이터, 포트 데이터, 프로토콜 데이터, AS 데이터, 서브넷 데이터 등의 필요한 플로우 정보를 데이터베이스에 저장한다. 또한 Netflow 수집 및 분석 서버와 통신하여 cflowd의 설정을 변경하는 기능을 제공하며 dataviewer 모듈을 두어 Servlet내에서 사용자의 요청에 따라 해당 정보를 분석해 HTML, Applet, Image, Table 등의 형태로 반환하는 기능을 수행한다.

4. 결론

네트워크 기반의 다양한 응용 프로그램 개발에 따른 급속한 네트워크 트래픽 증가로 네트워크 트래픽 측정 및 분석 작업이 필수적인 요소로 인식되는 가운데 특히 많은 트래픽이 집중되는 백본망에서는 다양한 품질을 요구하는 여러 가지 응용 서비스들이 동시에 이용되고 있기 때문에 그 필요성이 더욱 증가하고 있다. 하지만 현재까지 백본망을 대상으로 네트워크 트래픽의 총 양에 의한 정량적이고 단순 평면적인 정보가 아닌 네트워크 트래픽을 정밀 분석하고 특성을 파악할 수 있는 방법이 제시되지는 못하고 있다. 본 논문은 한국전력 백본망을 대상으로 효율적인 네트워크 자원 활용을 위한 네트워크 트래픽 측정 및 분석 시스템을 개발하였다. 본 시스템은 백본망을 대상으로 어느 한 순간의 데이터나 일괄처리 방식이 아닌 실시간으로 장기적인 트래픽 측정 및 분석이 가능하다.

참고문헌

[1] 정태수, 윤승현, 양지호 "인터넷 트래픽 측정 시스템 개발", 한국정보처리학회 추계 학술발표논문집 제 8권 제 2호, 2001년 11월

[2] 홍순화, 김재영, 조범래, 홍원기 "분산 시스템 환경에서의 로드 밸런싱을 통한 웹기반 네트워크 트래픽 모니터링 및 분석", 통신망운용관리학술대회 논문집, 2001년 5월

[3] 옥도민 "플로우 분류기를 이용한 인터넷 트래픽 측정 및 특성 분석", 서울대학교 석사논문, 2000년 2월

[4] KAIST "Passive Measurement Tool의 분석과 적용" 중간보고서, 2001년 9월