메모리 증가가 없는 실시간 네트워크 트래픽 모니터링 구현 연구

이양원

호남대학교 정보통신공학과

A Study of Implementation of Real-Time Network Traffic Monitoring without Memory Increase

Yang-Weon Lee

Dept. of Information and Communication Eng. of Honam University

요약문

본 논문에서는 인터넷을 이용한 응용분야 중에서 원격으로 네트워크의 트래픽을 모니터링함에 있어서 시간이 흐름에 따라서 메모리의 증가를 요구하지 않는 Round Robin Database 방식을 이용한 방법에 대한 응용 연구 방법을 기술하였다. 먼저 기본적인 RRDtool을 이용하여 트래픽 모니터링 데이터베이스 구조 설계를 구현한 과정을 기술하였고, 데이터의 모니터링을 위한 모니터링 프로그램으로서 Perl 스크립트 언어를 이용한 작업 과정을 보였다. 마지막으로 본 연구결과를 통하여 구축된 트래픽 모니터링 시스템을 이용하여 실제 트래픽을 모니터링한 결과를 실험 결과 에 제시하였다.

1. 서 론

인터넷을 이용한 계측시스템의 구현에서 가장 문제가 되는 것은 시간의 경과에 따라서 많은 양의 데이터가 누적되어 상대적으로 큰 하드디스크의 용량이 요구되는 점이다. 따라서 장시간 모니터링이 필요한 시스템에서는 새로운 개념의 데이터베이스의 설계가 필요하게 되었다. 따라서 이러한 시간에 비례한 메모리의 용량 증가를 없애면서 연속적으로 데이터베이스를 구축할 수 있는 기법으로 개발되어진 것이 RRD 방식의 RRDtool이다. RRDtool은 MRTG의 개발자인 Swiss Federal Institute of Technology의 Tobias Oetiker에 의해 개발되었다. RRDtool은 Round Robin Database tool을 의미하는데 여기에서 Round robin은 고정된 양의 데이터와 현재 요소를 가리키는 포인터와 같이 동작하는 것을 말한다. RRDtool은 연속적으로 들어오는 데이터들을 고정된 양으로 데이터베이스에 효율적으로 저장하고 그래픽 파일을 생성하는 CLI(command line interface)방식 툴의 일종이다. 즉 RRD에는 일정 시간 동안 몇 가지 어떤 값을 측정하여 이 정보를 임의 기간동안 RRDtool DB에 저장하고 있다. 저장되는 값은 일반적으로 숫자들이 되지만 반드시 그러한 제한이 있지는 않다. 많은 예제들에서 SNMP를 사용하여 장비가 가지고 있는 MIB의 값을 추출하여 이 값을 RRD에 저장하는 것을 볼 수 있다. RRDtool은 데이터베이스를 만들고, 그곳에 데이터를 저장하고, 이 데이터를 추출하여 웹 브라우저를 통해 볼 수 있도록 GIF(혹은 PNG) 형식으로 그래프를 만든다. 이 이미지들은 수집된 데이터에 의존한다. 예를 들어 평균 네트워크 사용률이나 최대 사용률 등을 계속 모니터링하면서 볼 수 있게 한다. RRDtool의 원리를 이용한 것으로서 널리 알려진 것은 MRTG-3(Multi Router Traffic Grapher)이다. 이는 RRDtool을 이용하여 디자인된 트래픽 모니터링 프로그램의 일종이다. 이것은 네트워크 장비나 서버의 온도, 전압, 전송속도, 메모리 사용률, CPU 점유율 등의 정보를 수집하여 만들어진 이미지를 웹 상으로 실기간으로 보여주는 프로그램이다. 이 외에도 조수흐름, 태양방사열, 소모전력, 전시회관람자, 공항주변의 소음, 날씨 등에 사용이 된다. 이는 데이터 수집을 위해 센서를 필요로 한다. 기본적인 구조는 C로 구현이 되었으며 Perl로써 핸들링 된다. 현재 Unix 와 NT 버전의 1.0.33이 나와 있다.

본 논문에서는 기존에 나와 있는 MRTG와는 별도로 간단하게 데이터를 모니터링할 수 있는 시스템의 데이터베이스 설계를 구현하고, 이를 이용하여 서버나 라우터, 스위칭 허브, 등의 인터페이스 단위별 트래픽을 모니터링하는 시스템 프로그램의 개발의 연구결과를 보이고 있다.

2. RRD의 구조

RRD 데이터베이스의 구조는 레코드, 필드를 갖는 일반 데이터베이스의 형식과 흡사하다. 먼저 수집이 시작될 시간을 정하게 되며, 다음으로 데이터베이스의 소스 타입을 정하고, 저장 될 데이터의 저장공간을 확보하게 한다. 생성되는 파일은 rrd확장자를 가지며 XML(Extensible Stylesheet Language) 파일로의 변환과

복원이 가능하다. 저장되는 데이터는 반드시 숫자일 필요는 없으며 상태나 상황의 저장이 가능하다. 입력 데이터의 저장포맷은 바이너리로써 데이터 접근에 훨씬 더 빠른 처리를 속도를 나타낸다.

다음은 RRD데이터베이스 구조 정의의 예를 보인 것이다.

```
rrdtool create test.rrd
            --start 920804400
            DS:speed:COUNTER:600:U:U
            RRA:AVERAGE:0.5:1:24
            RRA:AVERAGE:0.5:6:10
## Test.rrd라는 이름의 DB파일을 만듦
## 시작시간
## COUNTER값인 데이터소스 네임이 speed인 것을 만듦
## 업 데이트 최대시간간격 600초, 최소치와 최대치 정의 없음
## 5분동안의 평균데이터를 24회(24회5분-2시간)
## 30분 동안의 평균데이터를 10회(5시간) 평균을 구함
```

다음 그림 1은 데이터베이스의 구조를 나타낸 것이다.

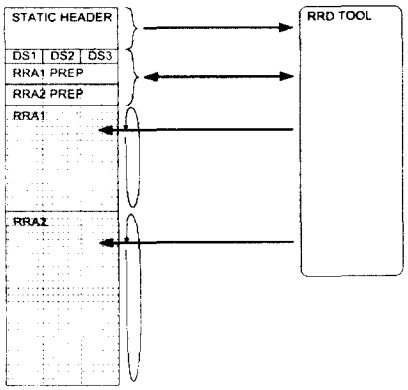

그림1. 데이터 베이스 구조

DS(Data Sources)로 데이터 구조를 정의 하며 RRA(Round Robin Archives)로 저장데이터의 정의와 크기를 정의한다. 나머지 공간은 실질적으로 데이터가 저장되는 공간을 확보하게 된다. 이는 처음 데이터 베이스 생성시 그 데이터베이스의 크기가 확정됨을 의미하고 주기적으로 반복되어 저장된다. 데이터 입력은 멀티 스텝 프로세싱으로 저장이 된다. 입력 값이 들어오면 RRDtool은 샘플링 하고 RRA 정의를 참고 하여 다시 샘플링 과정을 거쳐 RRA 저장공간에 저장하게 된다. 이는 그림 1에서 처럼 동시에 두개 이상의 입력과정이 처리됨을 의미한다. 다음 그림2는 RRDtool의 원리를 이용한 MRTG-2의 로그 파일이 주기적으로 업 데이트 되는 과정을 나타낸 것이다.

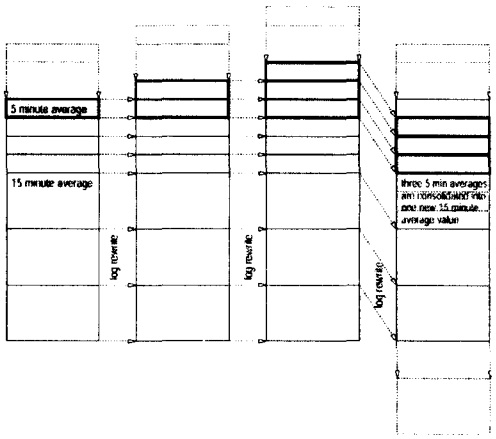

그림 2. MRTG Log 데이터 저장 구조

5분 간격으로 데이터를 수집하여 5분으로 정의된 DB 테이블에 저장하게 되고 15분이 지나게 되면 이 5분간의 데이터 3개의 평균을 구해 15분으로 정의 된 DB 테이블에 데이터 입력하게 된다. 이런 방법으로 30분, 2시간, 1일 단위의 평균을 구하여 일간, 주간, 월간, 년간의 평균적인 데이터 입력이 가능하다. 저장구조는 스택형식과 유사하다. RRD 데이터베이스 리샘플링은 입력 받은 데이터를 RRDtool에서 시간단위로 나누어 저장하는 기법을 말한다. 다음 표1은 300초 단위의 RRD리 샘플링을 실제 입력되어 계산되는 과정과 비교하여 설명한 것이다.

<표 1 >RRD 리샘플링

입력 계산 데이터		RRD 계산 데이터		
0	U	Time+000	0	U
300	300	Time+300	300	300
603	303	Time+600	300	300
900	297	Time+900	300	300

RRD 의 300초 단위로 입력된 카운터 데이터를 수집하여 다시 샘플링 하는 과정을 보면 그림 3과 같다.

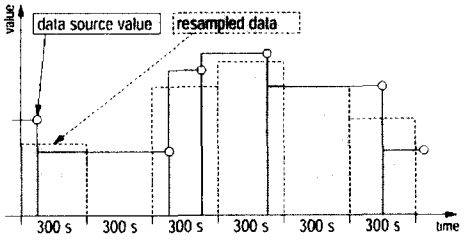

그림 3. 데이터 리샘플링 예

3. 데이터베이스 설계와 그래프 표현

데이터베이스 설계는 표현하는 그래프 파일과 밀접한 관계를 가진다. 예를 들어 12시간동안의 5분 평균 그래프와 일주일 동안의 15분 평균의 그래프 표현에 있어 데이터베이스의 구조는 다음과 같다.

RRA:AVERAGE:0.5:1:144 ##(12*12)
RRA:AVERAGE:0.5:3:672 ##((4*24*7)

하지만 12시간보다 긴 24시간 동안의 5분 평균의 그래프와 일주일 동안의 30분 평균의 그래프를 표현하고자 한다면 다음과 같은 구조를 가져야 할 것이다.

RRA:AVERAGE:0.5:1:288 ##(12*24)
RRA:AVERAGE:0.5:6:336 ##((2*24*7)

따라서 최종 표현하고자 하는 그래프의 시간보다 더 큰 데이터베이스의 설계를 필요로 한다. RRDtool의 그래프는 기본적으로 저장된 데이터에 의존하며 생성 파일종류, 크기, 표현시간, 등 매우 유연한 그래프를 만들어낸다. 생성된 그래픽 파일은 별도의 수동조작이 없으면 디폴트 값을 적용하여 최대한의 상태를 적용하여 생성한다. 뿐만 아니라 저장된 데이터에 몇몇 수학적 연산을 적용시켜 나타나는 값들을 적용시킬 수 있다. 다음은 RRDtool을 이용한 그래프 생성의 표현을 보인 예제와 몇 가지 생성된 그래프를 보인 것이다.

```
rrdtool graph test.gif                          \
        --start 920804400 --end 920808000       \
          --vertical-label km/h                 \
          DEF:myspeed=test.rrd:speed:AVERAGE    \
          "CDEF:kmh=myspeed,3600,*"
           CDEF:fast=kmh,100,GT,100,0,IF                \
          CDEF:over=kmh,100,GT,kmh,100,-,0,IF   \
          CDEF:good=kmh,100,GT,0,kmh,IF                \
          HRULE:100#0000FF:"Maximum allowed"    \
          AREA:good#00FF00:"Good speed"                \
          AREA:fast#550000:"Too fast"                   \
          STACK:over#FF0000:"Over speed"

## test.gif 이름의 그래픽 파일을 생성## 범위는 20:00~21:00까지
## 버티컬 라벨은 km/h
## test.rrd에서 speed:AVERAGE을 받아 myspeed 에 저장
## CDEF는 계산 정의. myspeed를 km/h로 계산
## myspeed가 100이 넘는 값은 fast
## Fast의 크기를 계산(over)
## 100을 기준으로 하는 수평선
## good을 area type으로 표현
## over는 stack type으로 표현 (100 이상의 값 표현)
```

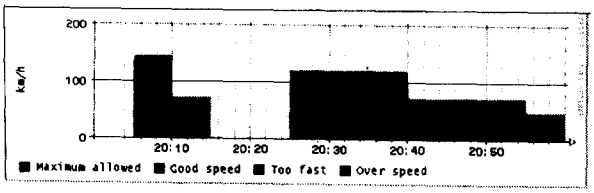

그림4. 일정시간 단위의 속도 표현 방식

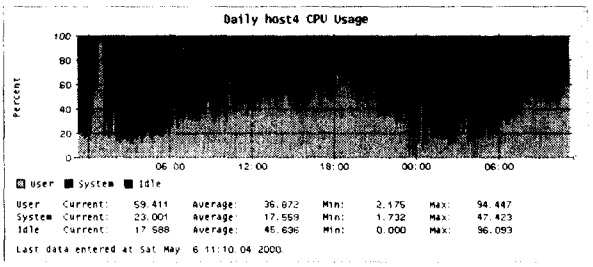

그림5. 하루단위의 CPU사용률 표현 방식

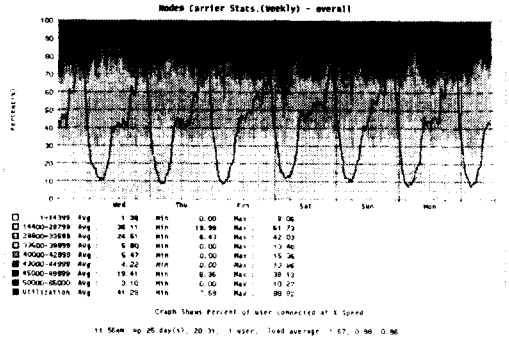

그림6. 주간단위의 모뎀상태 표현 방식

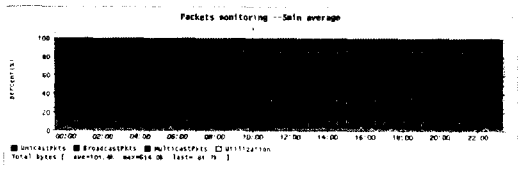

그림7. 하루단위의 패킷 모니터링 표현 방식

4. 라우터 인터페이스 적용 실험

RRDtool은 Perl 바인딩 모듈과 Perl CGI모듈을 지원하여 웹 상에서 실시간 그래프 확인이 가능하게 한다. 다음 예제는 네트워크 장비로부터 5분 간격으로 Packet과 traffic정보를 수집하여 이것을 Perl를 이용하여 웹 상으로 실시간으로 모니터링 하는 과정 중 NT system의

내부구성을 나타낸 것이다.

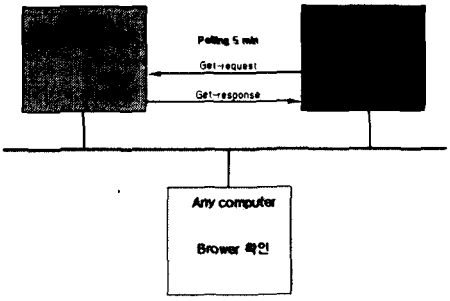

그림8 . NT 시스템의 구성

NT Server는 Perl의 SNMP 모듈을 이용하여 네트워크장비의 정보를 가져오게 되고 이 수집된 데이터를 Perl shared module를 통해 업 데이트하고 cgi 파일을 생성한다. 다음으로 IIS4를 이용하여 웹 서비스를 하게 된다.

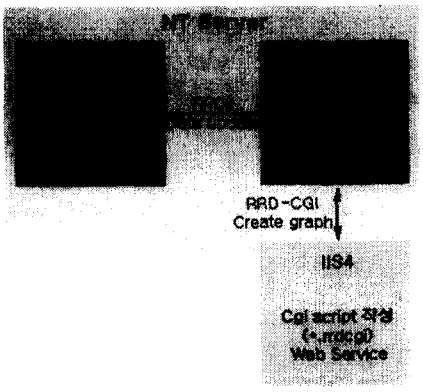

그림9 . NT 시스템 내부 구성도

아래의 결과는 5분마다 업 데이트를 자동적으로 보여주는 모니터링의 화면이다.

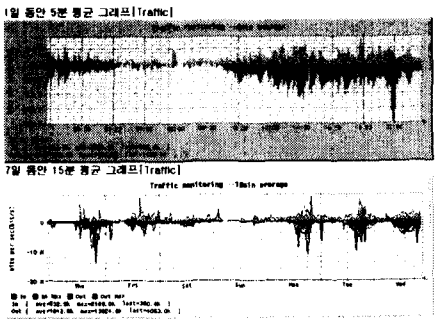

그림10. 5분 간격 갱신 그림

다음 그림은 약 3주간의 log 데이터와 4주간의 RRD 데이터의 파일 크기를 비교한 것이다. 시간이 지남에 따라 log 파일은 계속해서 늘어나게 되고 rrd 파일은 변함이 없다.

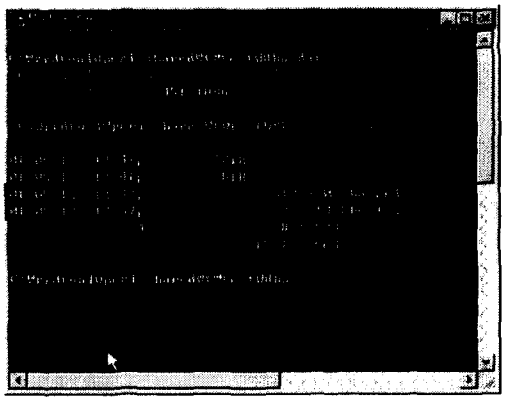

그림11. log파일과 RRD의 파일크기 비교

5.결 론

인터넷에서 RRDtool을 이용한 모니터링 시스템을 설계하는 과정을 보이고, 이의 구현을 위한 실험 내용을 보였다. 실험 결과 기존의 MRTG을 이용함에 있어서 그래픽 파일의 디자인과 수집 데이터(2개)의 제한에서 벗어나 간단하게 설치하여 유용하게 모니터링할 수 있음을 확인할 수 있었다. 이는 간단하게 RRDtool에 대한 개념만 파악한다면, 보다 쉽게 라우터의 리소스 사용률이나 인터페이스에 대한 traffic을 RRDtool을 이용하여 구성할 수 있게 해주는 프로그램들을 만들 수 있다는 것이다.

6. 참고 문헌

[1] Internet h/p, RRDtool,
[2] Internet h/p, MRTG, [3] Internet h/p, SNMP support Perl 5,
[4] Internet h/p, usenix,
[5] Allan Leinwand & Karen Fang Conroy, "Network Management A Practical Perspective", ADDISON WESLEY, 1996