Passive Performance Monitoring and Traffic Characteristics on the SLAC Internet Border

Connie Logg and Les Cottrell

Passive Performance Monitoring and Traffic Characteristics on the SLAC Internet Border

Connie Logg, Network Analyst, Stanford Linear Accelerator Center (SLAC)
Les Cottrell, Assistant Director SLAC Computing Services, Stanford linear Accelerator Center (SLAC)

Abstract

Understanding how the Internet is used by HEP is critical to optimizing the performance of the inter-lab computing environment. Typically use requirements have been defined by discussions between collaborators. However, later analysis of the actual traffic has shown this is often misunderstood and actual use is significantly different to that predicted. Passive monitoring of the real traffic provides insight into the true communications requirements and the performance of a large number of inter-communicating nodes. It may be useful in identifying performance problems that are due to factors other than Internet congestion, especially when compared to other methods such as active monitoring where traffic is generated specifically to measure its performance. Controlled active monitoring between dedicated servers often gives an indication of what can be achieved on a network. Passive monitoring of the real traffic gives a picture of the true performance. This paper will discuss the method and results of collecting and analyzing flows of data obtained from the SLAC Internet border. The unique nature of HEP traffic and the needs of the HEP community will be highlighted. The insights this has brought to understanding the network will be reviewed and the benefit it can bring to engineering networks will be discussed.

Keywords: Passive monitoring, Flow monitoring, SNMP, Inter-lab HEP traffic

1   INTRODUCTION

The SLAC Internet border is instrumented with a Cisco [1] Catalyst 6506 containing a Cisco MSFC module for routing. SLAC is connected to the Internet via a 155 megabit/second link to ESNet and a gigabit link to Stanford University. There are two types of passive monitoring that are performed on this border: SNMP monitoring of port utilization and network support device loads, and, monitoring of the traffic characteristics (via Cisco's Netflow Version 7 on the Catalyst 6506). These two different types of passive monitoring complement each other to provide a perspective on SLAC's communications with the outside world.

2   SNMP MONITORING

Detailed SNMP monitoring of SLAC's switches and routers [2] has been in production for several years. Every morning the entire SLAC network infrastructure is modeled, and configuration files are created from the model for use in the SNMP monitoring of all the active ports in the network. The modeling process entails determining the infrastructure connectivity (composed of 133 switches and 22 routing devices, and, including how and where the thousands of nodes in the network are connected to the network) by reading out the Cisco Discover Protocol, the ARP, bridge tables, and other standard SNMP MIBs. There are approximately 8700 total ports in the network, of which about 4600 are active currently. The connections to the world via ESNet and Stanford are also monitored by this system. Monitored traffic parameters include (but are not limited to) the in and out byte and in and out packet counts which are analyzed and plotted to indicate interface utilization. These readouts are performed every 5 minutes. The code for the SNMP data collection and analysis has been developed in house and uses Perl [3], SNMP Research's Brass Utilities [4], and Gnuplot [5].

The SNMP monitoring of the traffic provides statistics on the total traffic through the connections. Figure 1 is a graph of 1 point per day of the average and maximum in and out throughput for the 6 month period February-July 2001. It can be seen that the traffic volume is highly variable. Note the change around June 1. This was due to the installation of a 155 megabit link between SLAC and IN2P3, a major SLAC collaborator in Lyon, France.
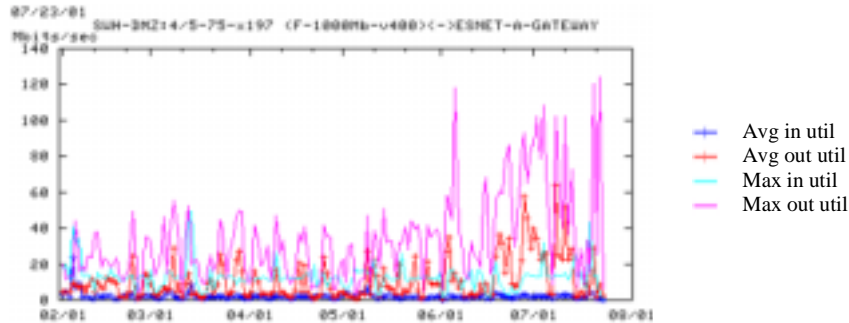
Figure 1: 6 Month Throughput for ESNet link

3   NETFLOW MONITORING [6]

Cisco's Netflow technology is used to monitor the type of traffic between SLAC and the Internet. Netflow Version 7 is enabled on the Catalyst 6506 and the records are sent to a Solaris system where they are aggregated via Cisco's Netflow Flowcollector Release 3 (aggregation type CALLREC). An aggregation file is created every 10 minutes. A Netflow record contains the source IP address, the destination IP address, the source port, the destination port, the protocol, the type of service, the number of packets, octets and flows the record represents, and, the starttime, endtime, and activetime (elapsed time) that the record represents. The aggregation by flows compared to saving all packet headers is a good compromise of data volume versus the granularity of information. The compressed Netflow aggregation files require about 100 megabytes/day. The analysis code has been developed in house. It is written in Perl and uses RRDTool [7] and Gnuplot. The Netflow monitoring provides information on the protocols and applications in use, as well as the conversations.

The Netflow statistics provide 3 different measures of the type of traffic: number of bytes, number of packets, and number of flows. As can be seen from Table 1, these 3 different measures present 3 very different views of the traffic. The following table is an overall average of the daily averages for the number of bytes, packets, and flows for the period February-July 2001.

Type of Measure	TCP	UDP	ICMP
% of total bytes	92%	7%	.4%
% of total packets	88%	9%	2%
% of total flows	76%	15%	9%
Average megabytes/day	180000	8600	470

Table 1: Traffic Volume by 3 Different Measures

4   TRAFFIC CHARACTERIZATION

For the purposes of traffic characterization, the units of Megabytes/day (as opposed to packet or flows) will be used, since this actually measures the bandwidth utilization of the traffic.

4.1 TCP Traffic

The Netflow analysis shows that the 3 main components of the TCP traffic are: SSH, FTP, and WWW. Table 2 shows the overall average of the daily averages for the period February-July 2001.

Type of Measure	FTP	SSH	WWW
Average megabytes/day	94000	30000	15000
% of TCP bytes	39%	25%	13%
% of TCP packets	33%	27%	15%
% of TCP flows	1%	2%	84%

Table 2: Primary Components of TCP Traffic

The quantity of FTP and SSH traffic are highly variable while the WWW traffic is relatively stable. WWW traffic regularly ranges between 10000 and 25000 megabytes/day. FTP varies between 6500 and 660000 megabytes/day and SSH varies between 3000 and 70000 megabytes/day.

4.2 UDP Traffic

The identifiable UDP traffic is overwhelmingly composed of AFS and Real (realaudio ports [8] are 6970,6971,6972,6973,7070).

Type of Measure	AFS	REAL
Average Megabytes/day	5400	1700
% of total UDP bytes	62%	18%
% of total UDP packets	44%	11%
% of total UDP flows	50%	.2%

Table 3: Primary Components of UDP Traffic

5   SOURCES AND DESTINATIONS OF THE TRAFFIC

Figure 2: Apps for top 10 sites for Jul. 20, 2001
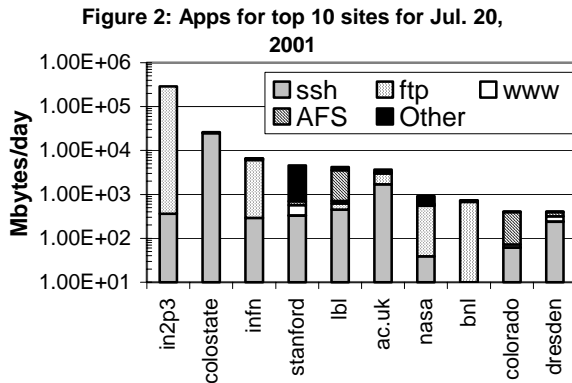
SLAC's network architecture is based on virtual lans (VLANs). Analysis of the SLAC destined and originated traffic (measured in megabytes) indicates that (on a daily basis) 90% of the time the top 2 VLANs are those associated with the BABAR compute farm and its supporting servers.

Analysis of the external sources and destinations (see for example figure 2) reveals that the primary users are SLAC's HEP collaborators with differing application requirements. The biggest applications are bulk file copying (STP, SSH/SCP) and file access (AFS).

6   FLOW CHARACTERISTICS

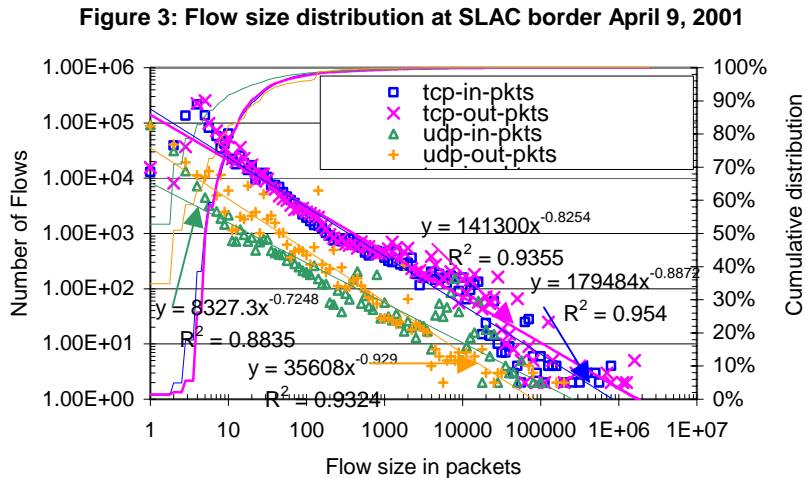**Figure 3: Flow size distribution at SLAC border April 9, 2001**

Figure 3 shows an example of the flow size distributions in packets for TCP *in* (to SLAC) and *out*, and UDP *in* and *out* flows. The *in* distributions have roughly the same behavior as the *out* distributions. It is apparent that there is roughly an order of magnitude more TCP than UDP flows. From the cumulative distributions it is also apparent that the UDP flows contain fewer

$y = 141300x^{-0.8254}$
$R^2 = 0.9355$
$y = 179484x^{-0.8872}$
$R^2 = 0.954$
$y = 8327.3x^{-0.7248}$
$R^2 = 0.8835$
$y = 35608x^{-0.929}$
$R^2 = 0.9324$

packets than the TCP flows. The flow distributions have heavy tails that can be fitted with power law fits of the form $y=a^{-b}$ which show up as a straight lines on a log-log plot. Typical values of the parameter *b* are in the range 0.6 – 0.9 and stay fairly constant from day to day. Further study of the distributions reveals that 75% of the TCP-*in* flows are < 5Kbytes, 75% of the TCP-*out* flows are < 1.5Kbytes (< 10 packets), for UDP-*in* 80% of the flows are < 600 Bytes (< 3 packets). The peak in the UDP *out* flows at about 100 packets is due to SNMP, and the peak at about 2000 packets is due to Real traffic [8]. Looking at the flow lengths, we find 60% of the TCP flows are < 1 second.

7   SECURITY

In addition to facilitating traffic characterization, the Netflow data is also useful for security purposes. Since the data provides detailed information on the source, destination, and type of packets in a well defined format, it can be easily by mined for information on attempted breakins and unauthorized computer facility utilization (such as running game servers). Each day a file is created for each TCP and UDP application port with one line per source and destination pair that provides the count of the bytes, packets and flows in each direction between the pair. If questionable use is identified, the raw data files for the day can be easily mined for the records of all communications between the pairs. The raw data files can also be mined further for communications information on IP addresses that have been identified as security concerns.

8   CHALLENGES

Several challenges (addressed in detail in [6]) encountered in the course of this analysis have included:

- Processing the large volume of data.   Between 2.5 – 3 million records a day are generated by the switch on which we have Netflow running..
- Identifying the application ports (mapping number to name). There are inconsistencies between the various reference lists.
- Deciding whether to use the source port or destination port to categorize a record. The algorithm we are using is described in [6].
- Translating IP addresses to names – Given that 50% of the IP addresses are non-SLAC IP addresses, the DNS lookups can take quite a while (several minutes and often timeout if an address if non-translatable), since they can involve world wide DNS nameserver searches. Several shortcut techniques (and internal script caching) have been developed and are discussed in detail in [6].
- Categorizing IP fragment packets. Packet capture analysis has shown that the IP fragments (which amount to about 39% of the total UDP bytes per day), are actually AFS packets fragments. This is not obvious from the Netflow records themselves.

9   CONCLUSIONS

SNMP and flow monitoring provide detailed information on the utilization of SLAC's connections to the rest of the world. This information facilitates greater understanding of the requirements of HEP and provides information needed to justify present and future bandwidth capability. Although SNMP monitoring provides information on the bandwidth utilization, flow analysis complements and completes the utilization information with details on precisely how the bandwidth is being used.

Flow monitoring can also provide detailed information for computer security purposes.

10   REFERENCES

[1] http://www.cisco.com - Provides information on Cisco's products and technologies
[2] http://www.slac.stanford.edu/comp/net/quick-guide.html - Detailed description of SLAC's SNMP, network connectivity, and services monitoring.
[3] http://www.perl.com - PERL information and source
[4] http://www.snmp.com - Provides information on SNMP Research International's products
[5] http://www.gnuplot.org - GNUPLOT information and source
[6] http://www.slac.stanford.edu/~cal/netflow/SLAC-Netflow.html - Describes SLAC's Netflow Analysis
[7] http://people.ee.ethz.ch/~oetiker/webtools/rrdtool - RRDTool information and source
[8] http://www.caida.org/tools/measurement/coralreef - Points to their Master Applications Ports list.