Network traffic analysis for the computing cluster at IHEP

A. A. Kotliar[a], V.V. Kotliar[b]

National Research Center "Kurchatov Institute" State Research Center of Russian Federation Institute for High Energy Physics, Protvino, Russia

E-mail: [a] Anna.Kotliar@ihep.ru, [b] Viktor.Kotliar@ihep.ru

A task for analysis of network traffic flows on the high performance computing network for the computer cluster is very important and allows to understand the way of complicated computing and storage resources usage by different software applications running on the cluster. As soon as all these applications are not managed by the cluster administrators they need a tool to understand usage patterns to make then an appropriate tuning for the core cluster software to achieve more effective usage for the cluster resources. The paper presents the development of such system for the IHEP cluster.

Keywords: sflow, netflow, flox, pmacct, R programming language, network traffic analysis, packet capture, traffic classification, GRID-computing

Introduction

A task for analysis of network traffic flows on the high performance computing network for the computer cluster is very important and allows to understand the way of complicated computing and storage resources usage by different software applications running on the cluster. As soon as all these applications are not managed by the cluster administrators they need a tool to understand usage patterns to make then an appropriate tuning for the core cluster software to achieve more effective usage for the cluster resources. Open source software for performing such analysis out of the box does not exist. Also the problem is that very complex and specific computing cluster infrastructure need to be taken into account. The more general way is to use different software components clued together in a system witch fulfill requested properties. In this paper such system based on flow collector software, relational data base software, web services and a programming language for statistic analysis is described.

System for flow data analysis

The system architecture for network traffic analysis is presented on the figure 1. It consists of the following software components:

- many network devices which generate sFlow data [Claise, Trammell, …, 2013];
- pmacct system for collecting this data, filtering them, splitting to the data blocks per date and time;
- MySQL relation DB to store prepared by pmacct data block;
- FLOX web-server software for simple analysis of flow data tables;
- R interface to the MySQL DB for sophisticated and statistical analysis.

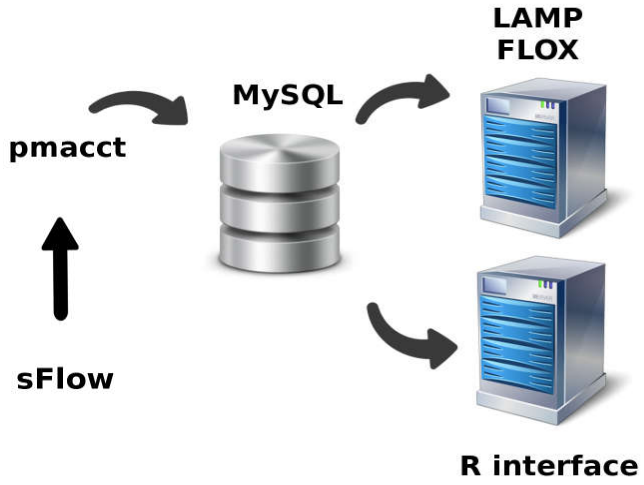

Figure 1. System architecture for collecting and analyzing flow data

For the data source sFlow (it is a short for "sampled flow") is used. It is an industry standard for packet export at Layer 2 of the OSI model. sFlow uses sampling to achieve scalability and is applicable to high speed networks such as the network of the computing cluster. By monitoring traffic flows on all ports continuously, sFlow can be used to instantly highlight congested links, identify the source of the traffic, and the associated application level conversations [Hofstede, Celeda, …, 2014].

All data comes to the pmacct system. It is a small set of multi-purpose passive network monitoring tools. It can account, classify, aggregate, replicate and export forwarding-plane data, ie. IPv4 and IPv6 traffic. It collects data in memory tables and then store it persistently to MySQL DB. pmacct is able to perform data aggregation, offering a rich set of primitives to choose from; it can also filter, sample, re-normalize, tag and classify at L7.

FloX (Flow eXplorer) is a simple PHP tool to examine large tables of flow data in a SQL database. It is easy extendable and a allow to use SQL like requests to the netflow database. It is used as a first tool for real-time data analysis.

For more complex analysis R language can be used. It is a language and environment for statistical computing and graphics. R provides a wide variety of statistical (linear and nonlinear modeling, classical statistical tests, time-series analysis, classification, clustering, …) and graphical techniques, and is highly extensible. R can easy produce well-designed publication-quality plots.

Integration to IHEP infrastructure

Computing cluster at HEP has two kind of networks: external network attached to the campus LAN with an access to the research network in Internet and internal high throughput network which serves for the inter-cluster communications. Described system for traffic analysis is applied to the internal network. The network core of the cluster is built on top of the HP ProCurve switches 5406zl. All switches bound together by pairs into logical switches with the distributed trunking technology. Connection bandwidth between cores switches is 2x10Gb/s when all computer hardware connected by 2x1Gb/s or 2x4Gb/s links with using bonding technique over link aggregation protocol. Figure 2 shows the implemented schema for sFlow analysis.

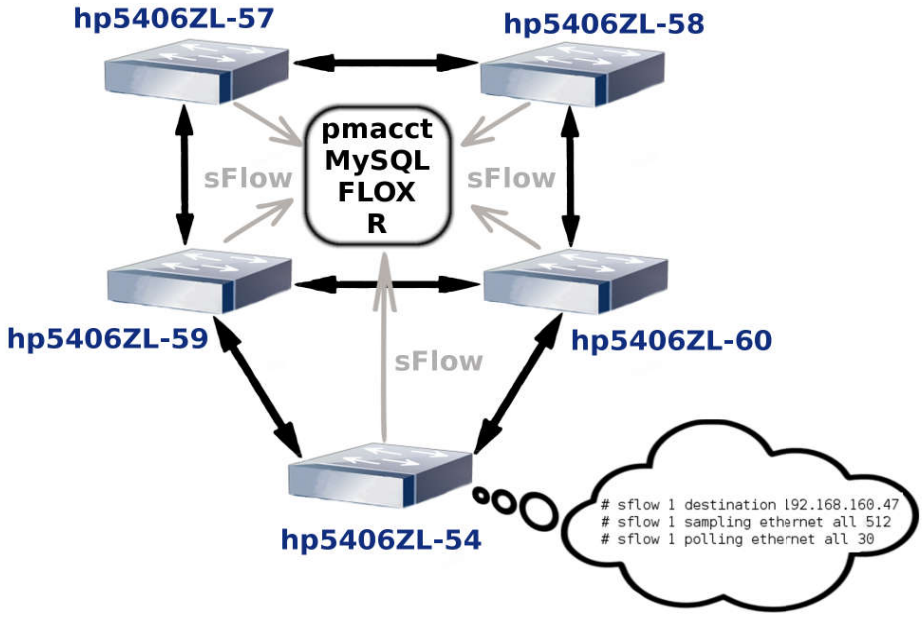

Figure 2. Implemented schema for sFlow analysis

Data analyze

For data analysis stateless packet inspection technologies were used [Getman A.I., 2015]. Flow data analysis consist of two part. First of all it is an interactive real-time analysis of traffic collected from network devices in MySQL database. FloX web-interface is used for it. This interface allows easily navigate through simple flow table where stored only fields like: source and destination IP addresses, traffic type, tcp/udp ports, and number of packets transmitted. In the web interface it is also possible to use complex SQL queries if needed. Many helper tables were created inside MySQL to allow to be made Union queries where some types of network devices need to be grouped. For instance there were created lustreOSSNodes , seAlicePoolNodes, seAtlasPoolNodes, seCMSPoolNodes tables with ip_int, ip_ext fields. These tables allow create queries like "how much traffic was sent to Lustre storage file system" or "how much network packets were transmitted from Atlas storage system to nodes". All helper tables are specific for IHEP computing cluster but the idea is general and could be used anywhere. As the result of such analysis it was discovered a misconfiguration on the IHEP cluster where incorrectly set storage nodes sent traffic through external interfaces to the internal cluster nodes. So additional setup was performed to mitigate such problem. It shows that it is the only way to understand traffic flows on the complex cluster system with dual network setup where we even do not know exactly how each program works as soon as these programs still under constantly development in the Grid community.

Second and more complicated analysis is done by using R programming language for statistical computing and graphics. In this analysis we try to use statistical methods to analyze unstructured data to understand how flows of the networks traffic could be used for describing usage patterns of the computing cluster for later modification in the setup environment to minimize additional traffic between different working zones means maximize computation usage. Increase effectivity of the usage of computing resources is a primary goal. Maybe using some machine leaning techniques which allows R language we even will be able to predict abnormal behavior of the cluster which depends on flow connections and trigger alarms or we will be able to self heal our system. The simplicity of using R for analysis is shown on figure 3.

```
acc<-dbGetQuery(conn = con, statement = "select
sum(bytes),(UNIX_TIMESTAMP(stamp_updated)-
UNIX_TIMESTAMP(DATE('20160330'))) from acct_20160330 where ip_src in (select
ip_int from lustreOSSNodes) group by stamp_updated;")
plot(acc[,2]/60/60,acc[,1]/1024/1024,main="Cluster from Lustre
",xlab="Time",ylab="MiB",col='blue')
```

Figure 3. Generate traffic plot with R

Here we can see two lines for generating graphical plot for traffic from the Lustre storage within computing cluster (figure 4).

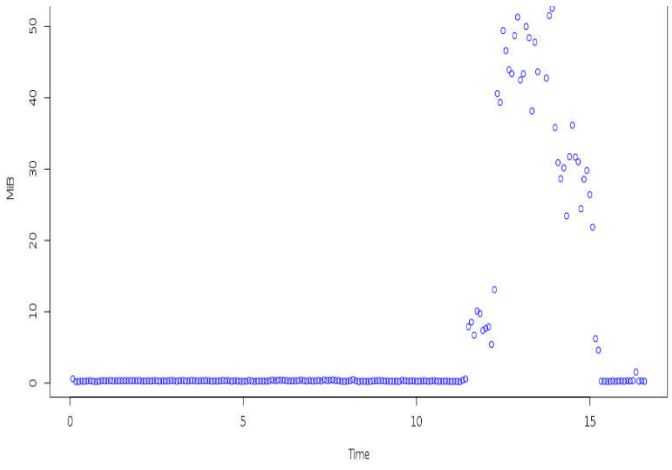

Figure 4. R plot for network traffic

Conclusion

The way of creation and usage flow tools for analysis network traffic for high-throughput network for high-performance and Grid computing cluster was presented in the described work. As soon as all these tools based on open source it is possible to easy extend them for any particular usage. Main achievement is a creation of the very simple architecture for performing almost any kind of network data analysis. By combining statistical computing language for programming and simple query language for DB it is possible to get helpful results in a short period of time.

The system is implemented on the production high-performance computing cluster at IHEP and allowed to find anomalies and misconfigurations in the cluster environment.

As future works it is planned to use R for detecting anomalies in traffic patterns with statistical and machine learning techniques which could help to implement some principles of autonomic computing [White paper…, 2006] such as self-optimization, self-healing, self-protection for the IHEP computer cluster.

References

Claise B., Trammell B., Aitken P. Specification of the IP Flow Information Export (IPFIX) protocol for the exchange of flow information. // RFC 7011 (Internet Standard), Internet Engineering Task Force. — 2013. [Electronic resource]. URL: http://www.ietf.org/rfc/rfc7011.txt.

Hofstede R., Celeda P., Trammell B., Drago I., Sadre R., Sperotto A., Pras A. Flow Monitoring Explained: From Packet Capture to Data Analysis With NetFlow and IPFIX // IEEE communications surveys & tutorials. — 2014. — Vol. 16, No. 4. — P. 2037.

White Paper. An architectural blueprint for autonomic computing // IBM. — 2006.

Getman A.I., Evstropov E.F., Markin Y.V. Wirespeed network traffic analysis: survey of applied problems, approaches and solutions. // Preprint ISP RAS. — 2015. [Electronic resource]. URL: http://www.ispras.ru/preprints/docs/prep_28_2015.pdf

http://www.sflow.org/about/sampling_history.php