IPFIX
Internet Draft
Intended status: Experimental
Expires: July 2013
January 18, 2013

R. Krishnan
D. Meyer
Brocade Communications
Ning So
Tata Communications

Flow Aware Packet Sampling Techniques

draft-krishnan-ipfix-flow-aware-packet-sampling-00.txt

Status of this Memo

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Abstract

The demands on the networking infrastructure and thus the switch/router bandwidths are growing exponentially; the drivers are bandwidth hungry rich media applications, inter data center communications etc. Using sampling techniques, for a given sampling rate, the amount of samples that need to be processed is increasing exponentially. This draft suggests flow aware sampling techniques for handling various scenarios with minimal sampling overhead.

Table of Contents

1. Introduction

Packet sampling techniques in switches and routers provide an effective mechanism for approximate detection of various types of flows (long-lived and short-lived) with minimal packet replication bandwidth overhead. A large percentage of the packet samples comprise of long-lived large flows and a small percentage of the packet samples comprise of other flows. The long-lived large flows aka top-talkers consume a large percentage of the bandwidth and small percentage of the flow space. The other flows, which are the typical cause of security threats like Denial of Service (DOS) attacks, Scanning attacks etc., consume a small percentage of the bandwidth and a large percentage of the flow space. This draft explores light-weight techniques for automatically detecting the top-talkers in real-time with a high degree of accuracy and sampling only the other flows - this makes security threat detection more effective with minimal sampling overhead.

1.1. Conventions used

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

The following acronyms are used:

DOS: Denial of Service

MPLS: Multi Protocol Label Switching

NVGRE: Network Virtualization using Generic Routing Encapsulation

TCAM: Ternary Content Addressable Memory

VXLAN: Virtual Extensible LAN

2. Flow Aware Packet Sampling

The steps in flow aware packet sampling are described below

1) Any flow which exceeds minimum flow duration and a minimum bandwidth would be characterized as a long-lived large flow. For identifying long-lived large flows, use the techniques described in Section 2.1. This helps in identifying the long-lived large flows aka top-talkers in real-time with a high degree of accuracy.

2) The identified long-lived large flows can be broadly classified into 2 categories as detailed below; these flows can be sampled at a low rate or need not be sampled.

   a. Well behaved (steady rate) long-lived large flows, e.g. video streams

   b. Bursty (fluctuating rate) long-lived large flows e.g. Peer-to-Peer traffic

3) The other flows (excluding the long-lived large flows) can be sampled at a normal rate. The other flows can be examined for determining security threats like DOS attacks, Scanning attacks etc. [LANCOPE]

For packet sampling, it is recommended to use PSAMP [RFCs 5474-5477] or sFlow [RFC 3176].

2.1. Long-lived Large Flow Identification

A flow (long-lived large/small flow or short-lived large/small flow) can be defined using one or more of the following suggested formats as described below

- IP 5 tuple: IP Protocol, IP source address, IP destination address, TCP/UDP source port, TCP/UDP destination port
- IP 3 tuple: IP Protocol, IP source address, IP destination address
- MPLS Labels
- VXLAN, NVGRE
- IP source address, IP destination address and IPv6 flow label (RFC 6437)
- Other formats

The techniques described in this document are agnostic to the format of the flow.

2.1.1. Automatic identification

Automatic identification of long-lived large flows can be implemented in ingress and/or egress processing elements of switches and routers. The characteristics of such an implementation would be

- Inline solution
- Minimal system resources

- Maintain line-rate performance

- Perform accounting of long-lived large flows with a high degree of accuracy

The advantages and disadvantages of automatic identification are detailed below.

Advantages of Automatic Identification

- Accurate identification of long-lived large flows

- Real-time identification of long-lived large flows

Disadvantages of Automatic Identification

- Not supported in many switches and routers

The implementation of automatic identification of long-lived large flows is vendor dependent. Below is a suggested technique.

Note: Netflow learns all the flows (long-lived and short-lived) and has scalability issues in terms of flow-cache size and CPU utilization.

2.1.1.1. Programmable parameters in Switches and Routers for Automatic Identification

- Minimum measurement interval for determining a candidate long-lived large flow (for e.g. 60 seconds)

- Minimum bandwidth of long-lived large flow (for e.g. 100 Mbps)

- Policy specification (for e.g. flows from a given IP source and/or destination address)

2.1.1.2. Suggested Technique for Automatic Identification

Step 1) If the long-lived large flow exists in a flow-table (e.g. TCAM), increment a per flow counter. Else, proceed to Step 2.

Step 2) There are multiple hash tables, each with a different hash function. Each hash table entry has an associated counter. On packet arrival, a new flow is looked up in parallel in all the hash tables and the corresponding counter is incremented. If the counter exceeds a programmed threshold in a given time interval in all the hash table entries, a candidate long-lived-flow is learnt and programmed in a flow-table.

There may be some false positives due to multiple other flows masquerading as a long-lived large flow; the amount of false positives is reduced by parallel hashing using different hash functions.

This technique is also suggested in [draft-krishnan-opsawg-large-flow-load-balancing].

3. Acknowledgements

The authors would like to thank Juergen Quittek for all the support and valuable input.

4. IANA Considerations

This memo includes no request to IANA.

5. Security Considerations

This document does not directly impact the security of the Internet infrastructure or its applications. In fact, it proposes techniques which could help in identifying a DOS attack pattern.

6. Data Model Considerations

In Section 2, for exporting the identified long-lived large flows to an external entity, it is recommended to use one of the protocols recommended in evaluation of candidate protocols for IPFIX [RFC 3955]. For any packet formats (for e.g. VXLAN, NVGRE) which are not covered by the above RFCs, a flow export data model needs to be defined - IETF could potentially consider a standards-based activity

around this.

Section 2.1.1.1 defines programmable parameters in switches and routers for automatic identification. IETF could potentially consider a standards-based activity around defining a data model for moving this information from a central management entity to the switch/router.

7. References

7.1. Normative References

[1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[2] Crocker, D. and Overell, P.(Editors), "Augmented BNF for Syntax Specifications: ABNF", RFC 2234, Internet Mail Consortium and Demon Internet Ltd., November 1997.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2234] Crocker, D. and Overell, P.(Editors), "Augmented BNF for Syntax Specifications: ABNF", RFC 2234, Internet Mail Consortium and Demon Internet Ltd., November 1997.

7.2. Informative References

N. Duffield et al., "A Framework for Packet Selection and Reporting", RFC 5474, March 2009.

T. Zseby et al., "Sampling and Filtering Techniques for IP Packet Selection", RFC 5475, March 2009.

B. Claise, Ed. et al., "Packet Sampling (PSAMP) Protocol Specifications", RFC 5476, March 2009.

T. Dietz et al., "Information Model for Packet Sampling Exports", RFC 5477, March 2009.

S. Leinen "Evaluation of Candidate Protocols for IP Flow Information Export (IPFIX)", RFC 3955, October 2004

P. Phaal et al. "InMon Corporation's sFlow: A Method for Monitoring Traffic in Switched and Routed Networks", RFC 3176, September 2001

[LANCOPE] Benefits of Flow Analysis Using sFlow: Network Visibility, Security and Integrity
http://www.lancope.com/files/Lancope_Generic_sFlow_WP.pdf

[draft-krishnan-opsawg-large-flow-load-balancing] R. Krishnan et al., "Best Practices for Optimal LAG/ECMP Component Link Utilization in Provider Backbone Networks", January 2013

Authors' Addresses

Ram Krishnan
Brocade Communications
San Jose, 95134, USA

Phone: +001-408-406-7890
Email: ramk@brocade.com

David Meyer
Brocade Communications
San Jose, 95134, USA

Phone: +001-408-333-4193
Email: dmm@1-4-5.net

Ning So
Tata Communications
Plano, TX 75082, USA

Phone: +001-972-955-0914
Email: ning.so@tatacommunications.com