SC19 Network Research Exhibition: Demonstration Abstract

Tracking Network Events with Write Optimized Data Structures

Tom M Kroeger, Brian J Wright, Evan West, Cynthia A Philips

Sandia National Labs

{tmkroeg, bjwright, ewest, caphil}@sandia.gov

Abstract

The basic action of two IP addresses communicating is still a critical part of most security investigations. Typically security tools log events and send them to a variety of traditional databases. Unfortunately, when faced with indexing billions of events such databases are usually unable to keep up with the rate of network traffic. As a result, security monitors typically log with little indexing.

Write-optimized data structures (WODS) provides a novel approach to traditional data structures (e.g. B-trees). WODS are able to ingest data 10 to 100 times faster while answering queries in a timely manner. Our Diventi project uses a write optimized B-Tree known as a B^e tree to index layer 3 network activity either from bro conn logs or netflow data. In 2017 & 2018 our tool was able to track all bro-ids monitored traffic & netflow indexing at rates above 160,000 events per second, and typically answering queries in milliseconds.

This year we intend to grow that rate and focusing on netflow data as the primary source for monitoring of IP activity. In addition, we believe the ability to answer queries in sub-second timeframe will enable analytic scripts to integrate the data in this index into their automated monitoring and alerting.

Goals

This effort will provide practical hands on insights into how WODS can be used for security monitoring of high speed networks.

Our goals here will be:

1. Understanding the growth of our data structure as it works to keep up with high data rates.
2. Understanding uses for security monitoring and how our structure responds to queries.
3. Gain a preliminary understanding around expiration of data within our tool.
4. Gain knowledge around the use of such indices to support automated alerting.

Resources

Our Diventi system will be running on a single dedicated machine that will require 2U slot in a rack. For data input we need to get a copy of the security team's tap of netflow data. This will enable our system to observe and index top rates of layer 3 activity. The system would need one IP address for management and a 10Gbps link to the systems generating the netflow logs. The management interface should be modest in its traffic load.

The two members doing this work (Brian and Tom) are also members of the SCinet network security team so we shouldn't need any addition work space or access.

Involved Parties

This is a collaborative effort between researchers at Stony Brook University, UC Santa Cruz, VM Ware labs and Sandia Nation labs.

- Tom M Kroeger, Brian Wright, Evan West, Cynthia A Phillips, Jon Berry, Sandia National Labs, {tmkroeg, bjwrigh, ewest, caphil, jberry} @sandia.gov
- Michael Bender, Stony Brook University, bender@cs.stonybrook.edu
- Rob Johnson, VMWare Labs, rob@cs.stonybrook.edu