SAND2019-8169D

The Center for Cyber Defenders

Expanding computer security knowledge

Low Resolution Indexing

Fast searches on full text indexing systems at petabyte scale

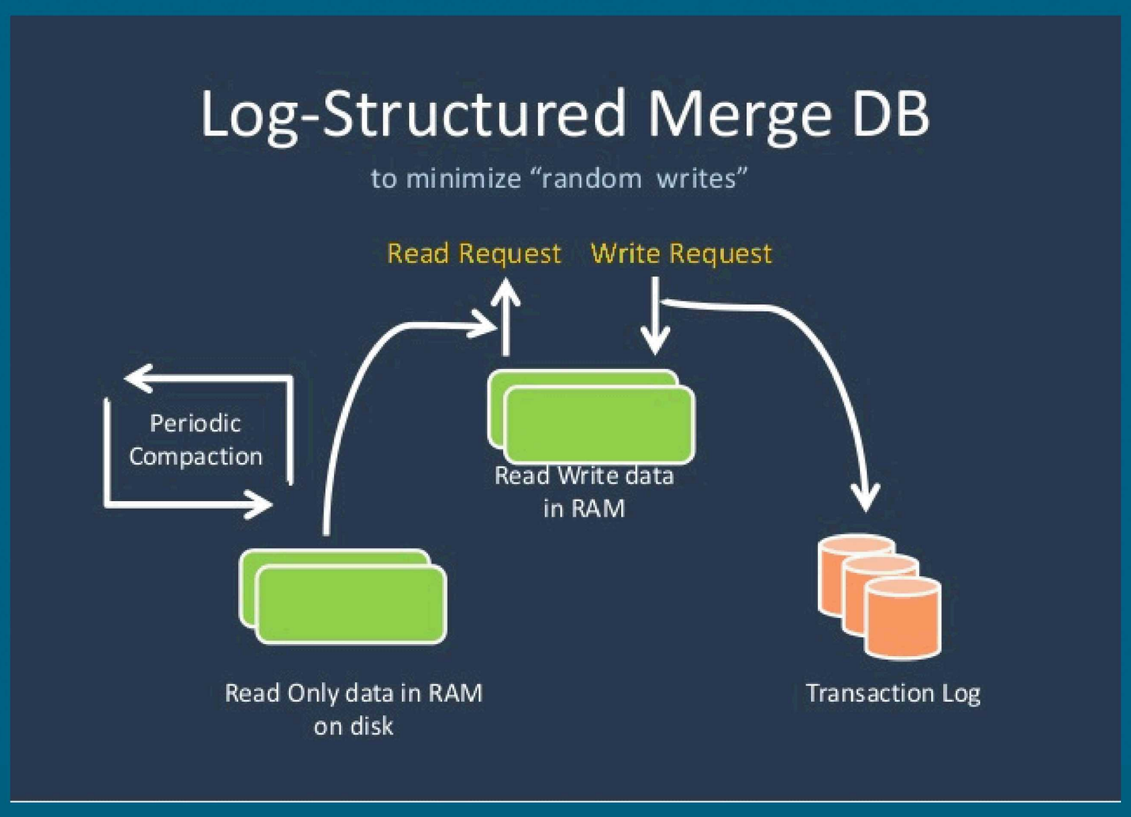
Adam Fasulo , New Mexico Institute of Mining and Technology
Eric Richardson, Prairie View A&M
Charles Smutz, Org. 9312

Problem Statement:

Very broad searches over large datasets can be prohibitively slow on common full text indexing systems. Broad searches such as, “where have we seen activity from this IP?” or “where did we first see this DOMAIN?” are common and necessary for effective incident response and building threat intel in the (IC) intelligence community. Searches can take an inferior amount of time on typical Splunk/ELK deployments at PetaByte scale.

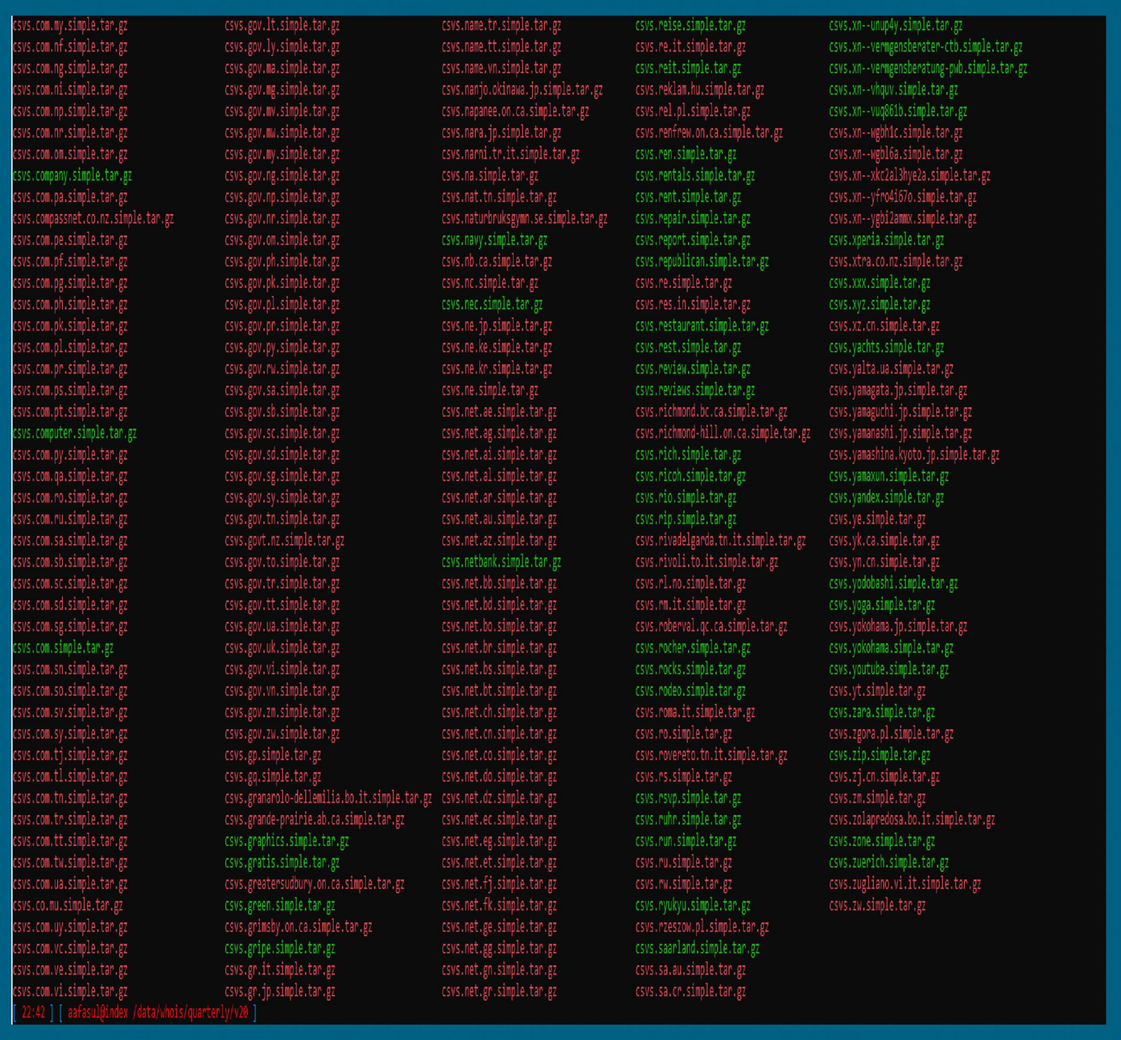
Objectives and Approach:

Use a fast and embeddable persistent key value store for a low resolution index that allows range scans using bloom filters, which uses a Log Structured Merge Tree data structure design.



Results

- Millisecond searches on a varying range of large and unique datasets.



Impact and Benefits:

- The benefit of our research is that it offers faster querying times for Splunk’s results.
- The impact of this research is that blue teams and intrusion detection users will be able to find malicious activities on lots of intrusion detection data at faster speeds.
- This makes it harder for malicious hackers to consistently infiltrate systems, because scanning for malicious activity is now ameliorated.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525. SAND2017-XXXXC.