Background

Zeek is a powerful framework for taking network traffic and extracting relevant metadata from it. The tool generates a series of logs that provide a variety of features both at a session level and individual protocol transactions.

Zeek provide a simple mechanism for correlating across logs. The Zeek *conn.log* represents high-level TCP/UDP/ICMP sessions, while other logs provide insight into individual protocols, files, and behaviors.

Dataset

https://mcfp.felk.cvut.cz/publicDatasets/CTU-Normal-27/bro/

Problem Description

Design a simple query interface that allows a user to analyze a set of Zeek logs. The utility needs only support *conn.log* and the *network protocol* logs. The query interface can be implemented as a command line utility, and does not rely on a database backend. Use the above dataset link to test your utility.

Query Modes

The utility must operate in the following query modes:

- Free-text query - Return all entries matching the provided string. (E.G "`google.com`")
- log.column query - Return all entries within a single log matching the provided query (E.G `conn.id_orig_h = "192.168.0.1"`)

Log Correlation

The utility must provide an option to include correlated logs in search results. When the option is included, the utility will also include additional entries in the data-set matching the **uid** field.

Exporting

The utility must provide an export mechanism and will write the results of a search to one or more CSV files. For example, if a search query results in multiple log entry types being returned, generate separate CSV output files for each unique log type (E.G `dns_matches.csv, http_matches.csv, conn_matches.csv, etc`).