Google Cloud

Query tools: RegEx and YARA-L

As you've learned, regular expression (RegEX) is a sequence of characters that form a pattern. You can use RegEx to search through log data and extract the relevant information you need. You've also learned about YARA-L. YARA-L is a computer language that is used to create detection rules for searching through ingested log data. In this reading, you'll take a deeper dive into RegEx, and you'll learn about best practices to use with YARA-L.

YARA-L

In Google Cloud's Chronicle, YARA-L is a computer language you can use to create detection rules for searching through ingested log data. Yara-L lets you hunt for threats and other events across large volumes of data. It works in conjunction with the Chronicle Detection Engine.

Note: This reading relates to version 2.0 of Yara-L. Be sure to use the current version with the Detection Engine. Older versions may cause problems.

Yara-L best practices

- **Filter out any zero values:** If you equate two omitted fields (they may be automatically omitted in the events you run rules against) they default to zero values. This will lead to unintended matches. For example, if you specify a rule for two fields, `$e1.field1 = $e2.field2` and they are empty, your search will result in an unintended match. To filter out any zero values, you can instead specify your rule as `$e1.field = ""`. You'll prevent a match since the fields won't contain any data.
- **Add an event type filter:** If your YARA-L rule only does its detection on Unified Data Model (UDM) events, you can add an event-type filter. This will reduce the number of events the rule needs to evaluate.

RegEx

When working with logging systems in a cloud environment, you'll likely encounter large amounts of data from multiple sources. Knowing how to effectively query log data can help you filter, sort, and manipulate log data to find specific information. To make this process easier, you can use RegEx to search through log data and extract relevant information. RegEx works with query tools to provide pattern matching for more focused querying. Patterns define the parameters of a search. If a piece of text matches a given pattern, RegEx retrieves it for you.

RegEx is used as part of the expressions in YARA-L. For instance, RegEx is one of many pattern matching functions within Yara-L used to create detection rules. Besides YARA-L, RegEx is also used in other areas within Chronicle, like searches.

Key takeaways

YARA-L, a computer language used to create detection rules for searching through ingested log data, has some best practices to follow, like filtering out zero values. Using YARA-L lets you identify threats that may be present in large amounts of data. You can also create and use a RegEx detector to help sensitive data protection by finding matches based on a RegEx pattern.