Managing aperiodic data in Splunk InfraMon

Some of the metrics that are used in applications and infrastructure monitoring are aperiodic in nature. In other words, successive datapoints from a given emitter for that metric are not regularly spaced in time. For example, you’re measuring the latency of transactions for your application, and your instrumentation captures values only when a transaction has occurred and then sends them once per minute to Splunk Infrastructure Monitoring. Additionally, your transactions occur more often during the day than the night, such that there are some minutes-long stretches during the night where you have only one or two transactions. For those minutes where there are no transactions, you don’t send any datapoints to Splunk Infrastructure Monitoring, so the analytics engine perceives null values for those times. You still want to be alerted when latency is trending above a certain threshold.

Your detector will need to account for the fact that the data is more regular during the day. For example, you could create a detector with a static threshold with a percent of duration (for example, above 5 seconds for 80% of the last 15 minutes) and it would work when the metrics are being received by Splunk Infrastructure Monitoring regularly. However, during the night, such a detector would likely never fire, as you would rarely have enough full sets of data.

To compensate for this, you can use analytics to express a percentage of datapoints received during some window above (or below) a static threshold and alert on that.

When the detector is based on the actual metrics data, you can sometimes use an extrapolation policy (typically Last Value) along with a percent of duration condition. However, this may lead to the triggering/clearing behavior being determined largely by extrapolated values.

Another approach is to use window transformations. Especially for aperiodic data, alerting when the error rate over a 15-minute window is above, for example, 10% will have much better properties, and is more reasonable, than alerting when the error rate, calculated at every timestamp, is over 10% for 15 minutes.

Finally, consider using the aperiodic module in the SignalFlow library.

Next steps

These additional Splunk resources might help you understand and implement these recommendations:

- Signal Flow Library: aperiodic.flow