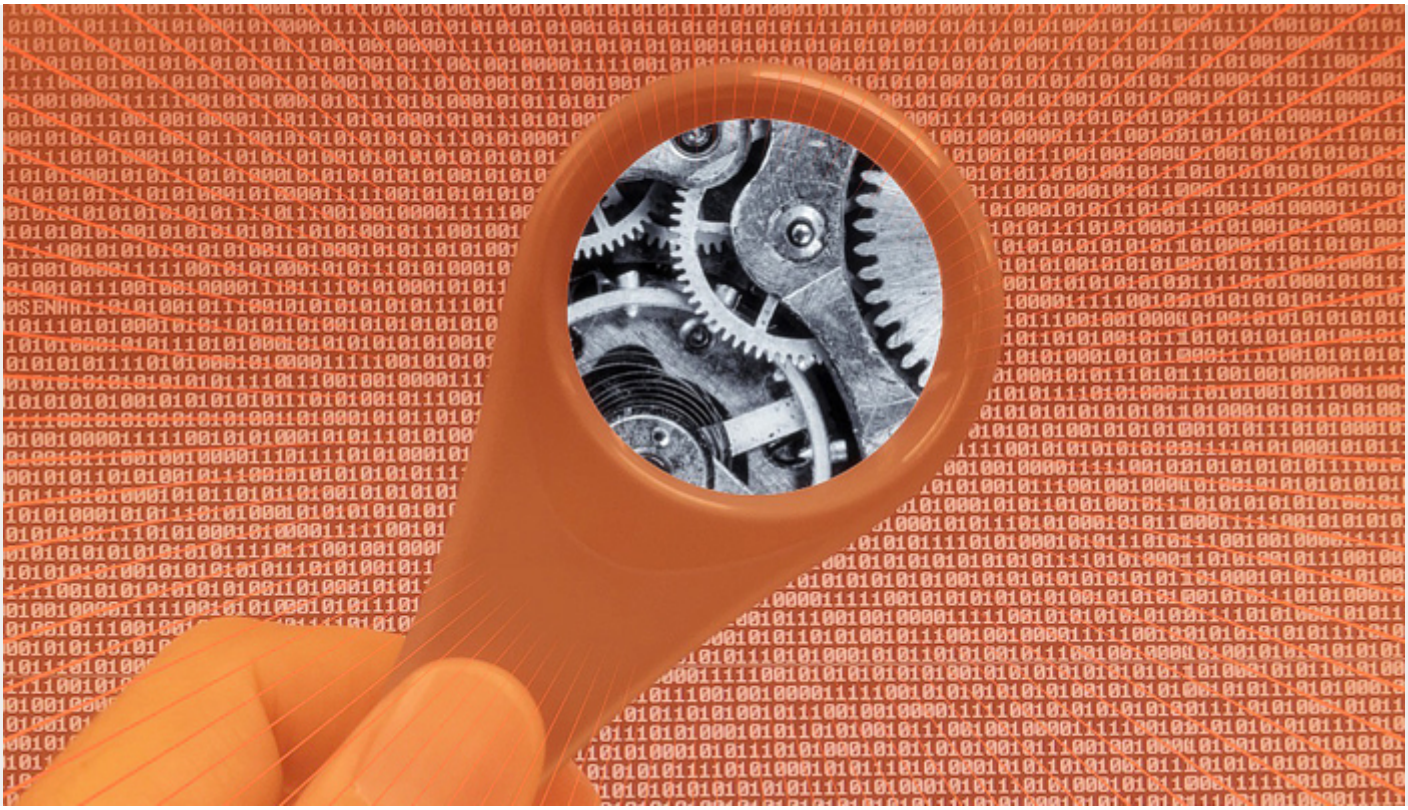
PREDICTIVE LOG ALERTING WITH ML ANOMALY DETECTION



Logging is vital to the success of any IT project. With a solid logging practice, you can troubleshoot errors, find patterns, calculate statistics, and provide diagnostics information easily. Given the size and complexity of many modern systems and the fact that they're always on with 24x7 availability, logs can rapidly become difficult to manage. This, combined with aggregating logs from multiple systems, makes it infeasible to manually process logs.

Log data also contains anomalies that represent potential system faults, which makes them critical to debugging application performance and errors. However, if you look at the logs, most of the entries simply say that "an event occurred." What we want is a way to detect when things aren't following the normal pattern, which means that the automated analysis needs to look at individual lines and groups of entries to determine whether they're expected or indicate any deviation. This can help you proactively find concerns before they become a problem and help troubleshoot errors when they arise.

BMC Helix Log Analytics provides automated log analysiswith machine learning (ML)-based anomaly detection to process log contents and find abnormal entries and behavior patterns in logs.

Why anomaly detection is important

Imagine you log in into your system to find that an application you manage has been running slowly. Your team updated a few patches in the last release, but that was over a week ago. There's no reason why anything should be different now. Maybe it's an integration that's causing problems. Or

maybe the server has a hardware issue. Whatever the case, you're going to have to look at the logs. Without a log analytics solution, you need to go through the raw log file with Ctrl+F and some regular expression (regex).Maybe you will modify that script you tried to make last time. It didn't quite work, but you think it sped the process up.

Keep in mind that the log has been recording a tremendous number of messages, which may take hours and days of effort to search and troubleshoot. You don't even know what you're looking for. Or where to look first. The problem might not even be in the error and warning messages. It might be hidden in success messages that are fired too quickly or out of order. No amount of regex will find that.

To overcome this, you will need an automated log analytics solution that can identify the entries and behaviors that don't look like they fit. This approach may not find the problem immediately, but it's going to give you a subset to work with—things that you can investigate further without having to dive into those 600,000+ entries manually. That's where BMC Helix Log Analytics and its ML-based anomaly detection capability can come to the rescue.

ML anomaly detection by BMC Helix Log Analytics

BMC Helix Log Analytics anomaly detection uses ML to detect anomalies from logs and allows you to generate events that quickly alert you to impending problems in your application or system. It incorporates an unsupervised deep-learning model which is based on an artificial neural network technique and involves the following steps:

- Data pre-processing
- Anomaly detection
- Evaluation

The first step is data pre-processing, where the raw and unstructured log data is transformed into features that can be ingested into the anomaly detection algorithm. It parses the raw data to extract key value pairs and remove extraneous or execution-specific details, and the output is used as input for the anomaly detection stage. At this stage, the ML model looks at every incoming log record, finds patterns and behaviors from log messages, their similarity and frequency of occurrence, calculates the anomaly score, and identifies records which are anomalies. Finally, these anomalies are also cross validated with the domain expert labeled list of anomalies for each dataset to identify false positives, false negatives, true positives, and true negatives in order to derive precision and recall. This helps to auto-tune, optimize, and improve the overall accuracy of the ML anomaly detection algorithm.

When the log alert policy with anomaly detection is defined, the ML model is trained on the incoming logs that are categorized as training data, per the matching criteria of the anomaly alert policy, and then a threshold value is calculated. After the ML model is built, it calculates the anomaly score for every new log record that comes in, and when its value exceeds the threshold, the log record is flagged as an anomaly and an anomaly event is generated. The ML model keeps training continuously and auto-updates if it finds new patterns or behavior in logs.

You can keep track of rare and anomalous log patterns and generate events using the anomaly alert policy, as shown below.

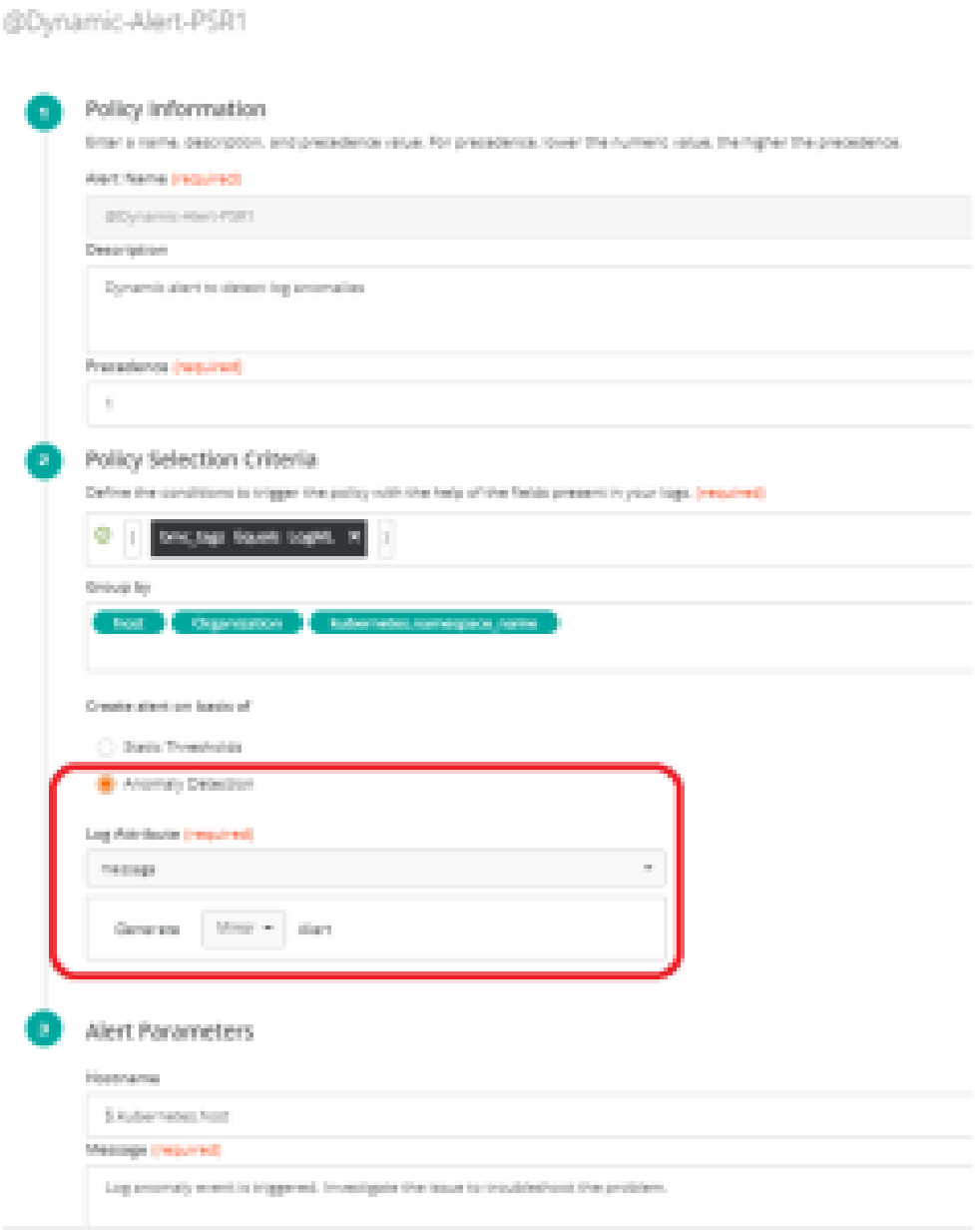


Figure 1. Log alert policy to generate anomaly events.

These anomaly log events are acted upon in the BMC Helix Operations Management events console and further correlated in the context of a given service for root cause isolation with BMC Helix AIOps capabilities. You can cross-launch into log analytics in the context of an anomaly event to see associated logs and troubleshoot for more information.

You can also visualize the anomalies in the log explorer and troubleshoot the probable cause, and apply a filter based on the anomaly score or query on specific conditions and further slice and dice log records.

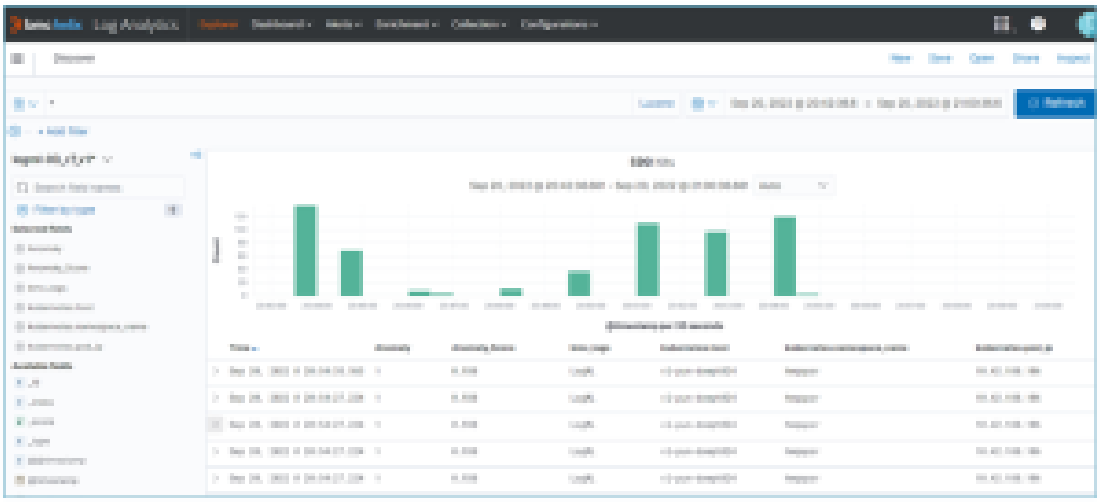


Figure 2. Log explorer to analyze log anomalies.

To summarize, BMC Helix Log Analytics allows you to use ML-based anomaly detection on log files to help troubleshoot why processes are failing, identify whether you have any security concerns, and perform a check on your software. It's best suited to large, complex systems that include access, runtime, development, and security logs and which generate tons of logs every minute. BMC Helix Log Analytics can run on any log at any time, including regularly in the background, to proactively find and solve concernsbeforethey become problems, and increase the likelihood of finding the root cause of a problem. This also helps increase uptime, reduce errors, and improve system design, all of which are key to the success of your business.

Related Content

- Observability with Logs to Accelerate MTTR