

With the explosion of the number of distributed applications, a new dynamic server environment emerged grouping servers into clusters, whose utilization depends on the current demand for the application. Detecting and fixing erratic server behavior is paramount for providing maximal service stability and availability. Using standard techniques to detect such behavior is yielding sub-optimal results. We have collected a dataset of OS-level performance metrics from a cluster running a streaming distributed application and injected artificially created anomalies. We then selected a set of various machine learning algorithms and trained them for anomaly detection on said dataset. We evaluated the algorithms performance and proposed a system for generating notifications of possible erratic behavior, based on the analysis of the best performing algorithm.