Feeling the enterprise pulse:Unified monitoring with Beats and ELK stack

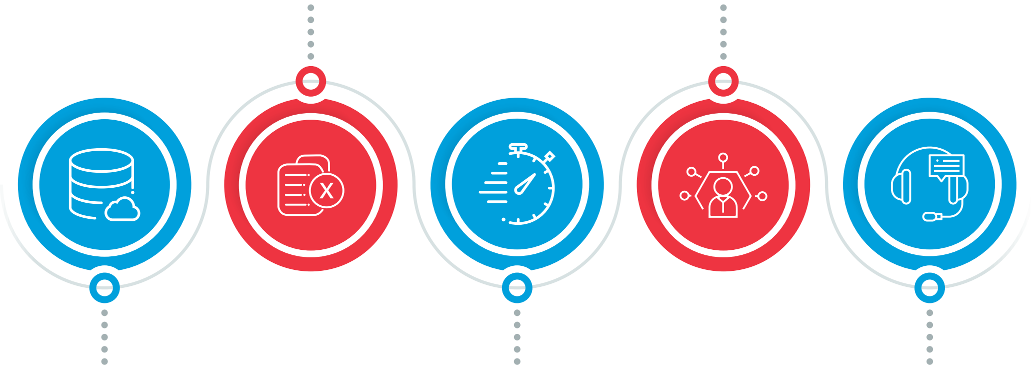
Organizations generate huge volumes of business and technical data in the form of logs and metrics, collectively called knowledge data. These are invaluable sources of information if mined properly. Timely analysis of this data can help businesses improve their decision-making process, resulting in better performance and profit making. There are ways to mine this data in the form of clickstream analysis, customer sentiment analysis, customer 360-degree analysis and other such methods. Such analyses can help in real-time monitoring, performing root cause analysis for system failures, predictive monitoring with anomaly detection of distributed platforms and the like.

In any enterprise, there exists a plethora of logs and metrics that aid in end-to-end monitoring and providing business intelligence. These (logs and metrics) aided with Unified Log Monitoring Frameworks can amplify business performance through proactive detection of system issues and anomalies and agility in decision making from the business perspective

Just as log monitoring is a key part of organizational decision making, there are technical challenges related to defining and keeping such a framework for that. We shall talk about those challenges and the potential solutions for log analysis and monitoring.

The sticky issues

Log analysis and monitoring are fraught with several challenges that prevent organizations from using logs and metrics to the best advantage.



Data is generated from disparate sources and systems

All logs are not in structured format.

Unlimited time series data is generated from these systems, making them difficult to mine in real time

Metrics and data are highly distributed

The advent of microservices has further compounded the problems related to traceability, collating and aggregating data/metrics from different sources to provide meaningful information to businesses

Unified monitoring to the rescue

So, what's an ideal monitoring framework? A framework capable of capturing data, logs from any source, efficiently storing them, utilizing historical data to create machine learning models for predicting system issues and utilizing distributed tracing for Root Cause Analysis.

Elastic releases and components, such as Beats, Elasticsearch, Logstash and Kibana (BELK), part of an open source stack, have a huge pool of easily pluggable light-weight edge data shippers. These shippers are capable of collecting system and application-related metrics, reading log files and forwarding the time series data and metrics to data pipelines for further processing. The data pipelines are the ingestion workhorses that collect, parse, transform and store data. This solution also provides the ability to search and analyze large data volumes, apart from supporting extensive reporting features.

There are additional capabilities to enhance the power of the Elastic Stack.

In a highly distributed architecture, correlation of data across multiple layers is a challenge. With microservice implementations, the challenge multiplies. The answer to this problem lies in Distributed Tracing. Distributed tracing frameworks, like Zipkin, provide uniform and consistent tracing capabilities and are a force to be reckoned with.

Discovering a problem is the first step in resolving it. Eliminating the delay between when the problem occurs and the time the problem is detected, immediately brings one closer to identifying the root cause. Predictive Monitoring using anomaly detection comes in handy here.

Anomaly detection with an open framework such as Apache Spark, helps predictive monitoring in finding outliers in an otherwise ordinary set of data. Machine Learning algorithm such as Support Vector Machine, Bayesian Networks and k-means clustering are widely used in anomaly detection. Based on the use case and data attributes, either supervised unsupervised learning techniques can learn from historical data and easily detect outliers in the present data. Anomaly detection can be used for cases such as network security analysis, fraud detection and error detection.

A predictive monitoring solution built on top of open source frameworks, such as BELK, OpenTracing and Apache Spark, leverages the flexibility of Beats and Logstash, the power of an OpenTracing framework like Zipkin, along with the intelligence of the Spark ecosystem. It provides a cost-effective yet flexible monitoring platform that can be used for better decision making. It can also be applied across a variety of business use cases.

About the authors

Nishant Sahay
Senior Architect, Wipro

Nishant Sahay is a senior architect at Wipro, with extensive experience in data analysis, design and visualization. Nishant works at the Open Source COE lab at Wipro, where he is responsible for research and solution development in machine learning and deep learning. He has written articles on technology in online forums and presented at multiple open source conferences and platforms across the world.

Bhavani Ananth
Architect, Wipro

Bhavani is an architect at Wipro and is a part of the Open Source CoE. She is currently working on predictive log analysis and monitoring using ELK. In the past, she has worked on Java based technologies and middleware solutions.

Wipro Limited
Doddakannelli, Sarjapur Road,
Bangalore-560 035, India

Tel: +91 (80) 2844 0011
Fax: +91 (80) 2844 0256
wipro.com

Wipro Limited (NYSE: WIT, BSE: 507685, NSE: WIPRO) is a leading global information technology, consulting and business process services company. We harness the power of cognitive computing, hyper-automation, robotics, cloud, analytics and emerging technologies to help our clients adapt to the digital world and make them successful. A company recognized globally for its comprehensive portfolio of services, strong commitment to sustainability and good corporate citizenship, we have over 160,000 dedicated employees serving clients across six continents. Together, we discover ideas and connect the dots to build a better and a bold new future.

For more information, please write to us at **info@wipro.com**