

Intelligent SSD Performance Anomaly Detection



Presenter: *Vijila N, Sagar Kumar Sahu, Mradul Kumar, Satisha Atmakuri, Dhananjay Pal*

Samsung Semiconductor India

Disclaimer

This presentation and/or accompanying oral statements by Samsung representatives collectively, the “Presentation” is intended to provide information concerning the SSD and memory industry and Samsung Electronics Co., Ltd. and certain affiliates (collectively, “Samsung”). While Samsung strives to provide information that is accurate and up-to-date, this Presentation may nonetheless contain inaccuracies or omissions. As a consequence, Samsung does not in any way guarantee the accuracy or completeness of the information provided in this Presentation.

This Presentation may include forward-looking statements, including, but not limited to, statements about any matter that is not a historical fact; statements regarding Samsung’s intentions, beliefs or current expectations concerning, among other things, market prospects, technological developments, growth, strategies, and the industry in which Samsung operates; and statements regarding products or features that are still in development. By their nature, forward-looking statements involve risks and uncertainties, because they relate to events and depend on circumstances that may or may not occur in the future. Samsung cautions you that forward looking statements are not guarantees of future performance and that the actual developments of Samsung, the market, or industry in which Samsung operates may differ materially from those made or suggested by the forward-looking statements in this Presentation. In addition, even if such forward-looking statements are shown to be accurate, those developments may not be indicative of developments in future periods.

Agenda

- ❖ Introduction
- ❖ SSD Sustained Performance Evaluation Key Challenges
- ❖ Solution Overview
- ❖ Solution Architecture and Flow Diagram
- ❖ OpenDistro Anomaly Detection Procedure
- ❖ Results
- ❖ Benefits

Introduction

Objective:

- To meet with user sustained performance demand, identifying unintended performance bottlenecks during SSD evaluation phase by analyzing huge set of performance/system log and metric is tedious.
- To automate the analysis we introduced powerful unsupervised machine learning mechanism for performance anomaly detection, which helps us to identify unintended events during sustained performance evaluation.
- SSD performance anomaly is an unintended behaviour like huge drop in read/write or spike in latency or system resources during course of usage.

Key Challenges

SSD sustained performance evaluation key challenges:

- Performance bottleneck identification on varying operating conditions without targeted QoS.
- Maintain and perform real-time analysis of large set of data generated out of performance evaluation result and system logs.
- Discovering and alerting anomalies using conventional methods by creating visualizations and dashboards can be challenging.
- Alert based on static threshold require prior domain knowledge.
- Real-time application workload performance evaluation without known access patterns
- Defect triaging & classification
- Derive use case out of anomalous behaviour for failure reproduction and future evaluation.
- Automatic correlation between time series QoS data, defect history, system metric and debug log.

Solution Overview

This solution introduces powerful unsupervised machine learning mechanism for performance anomaly detection, which helps us to identify unintended events during sustained performance workload evaluation with automatic log correlation.

- **Analytics Solution**

- ✓ Centralized live log stashing
- ✓ Real-time performance analysis and trend

- **Random Cut Forest(RCF) Based Unsupervised Learning Solution**

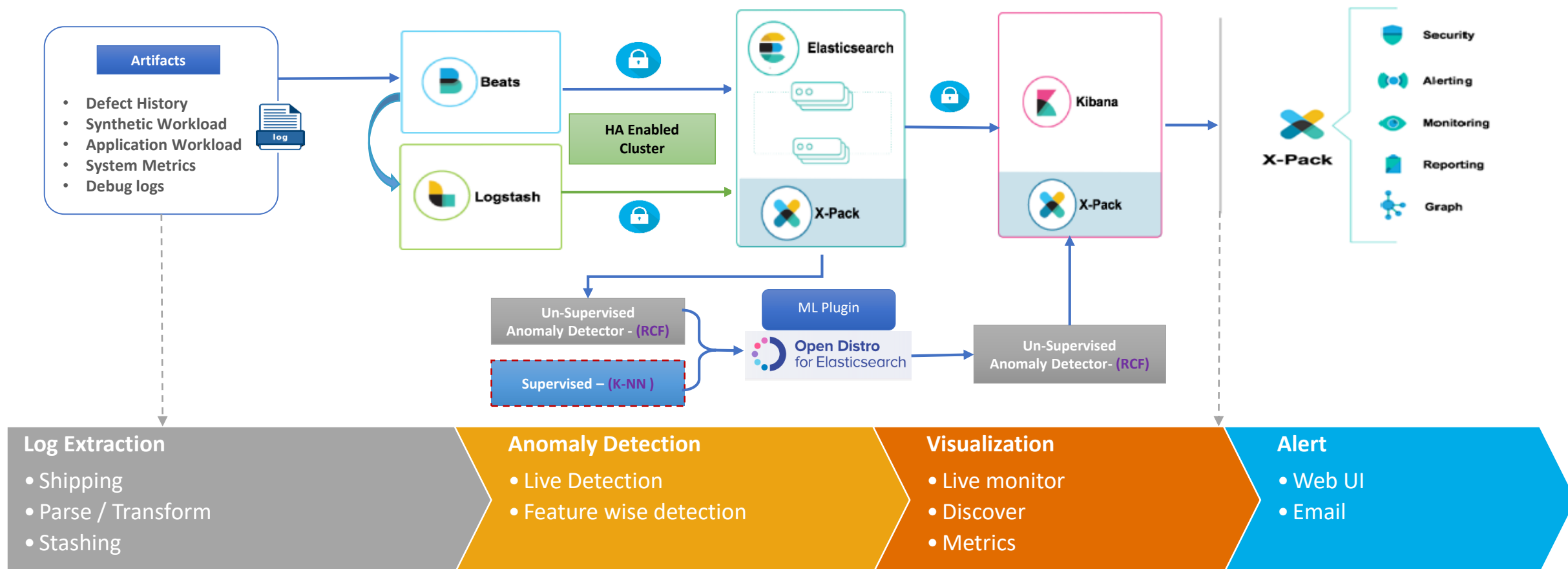
- ✓ Performance anomaly detection & alert through WebUI, e-Mail

- **K-Nearest Neighbors(KNN) Based Supervised Learning Solution(*Future Work*)**

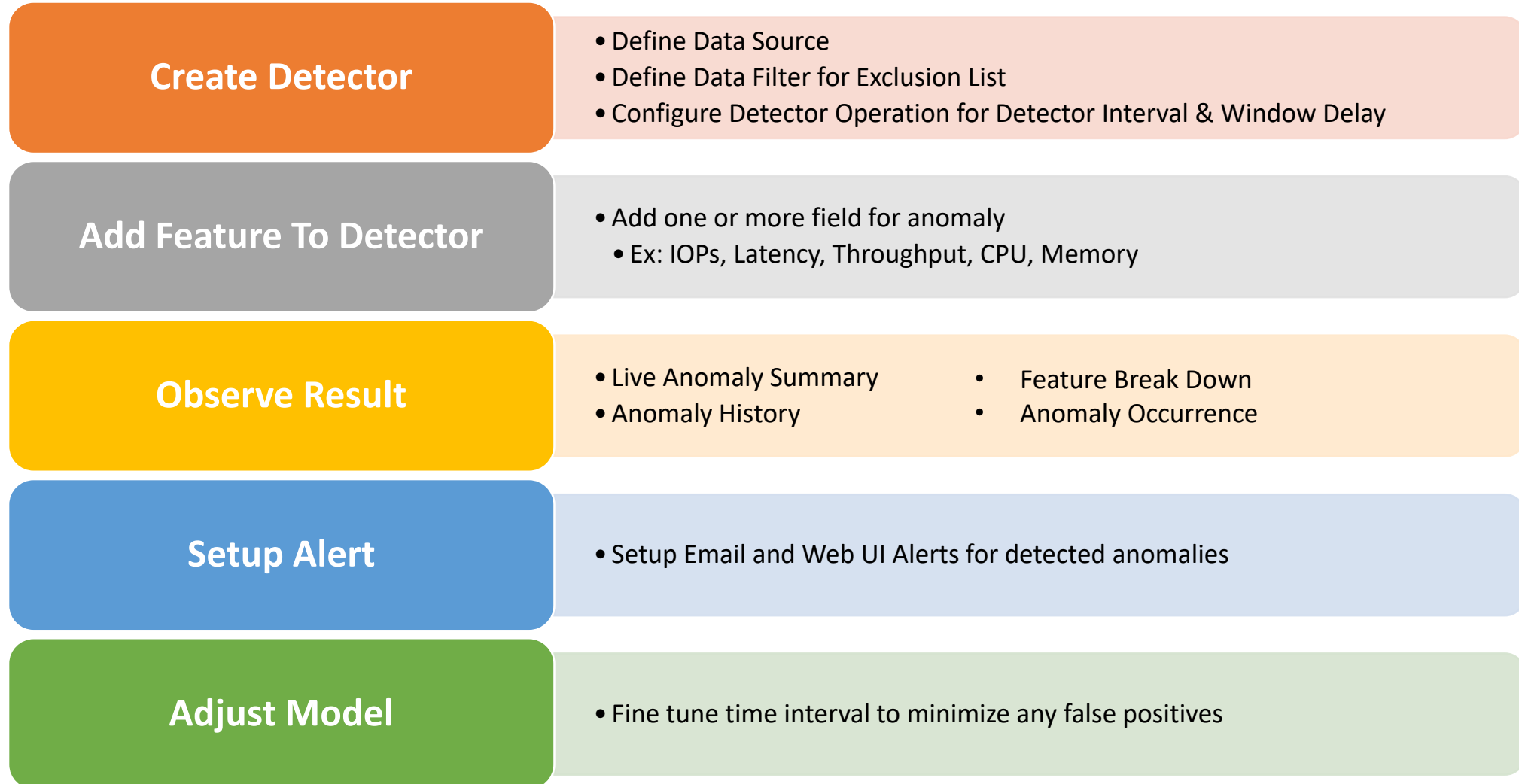
- Defect Triaging & Classification

Solution Architecture & Flow Diagram

- Performance Log Extraction, Transformation and Analysis using Elastic Stack
- RCF Machine learning methodology for performance anomaly detection
- Alerting on performance outliers detection



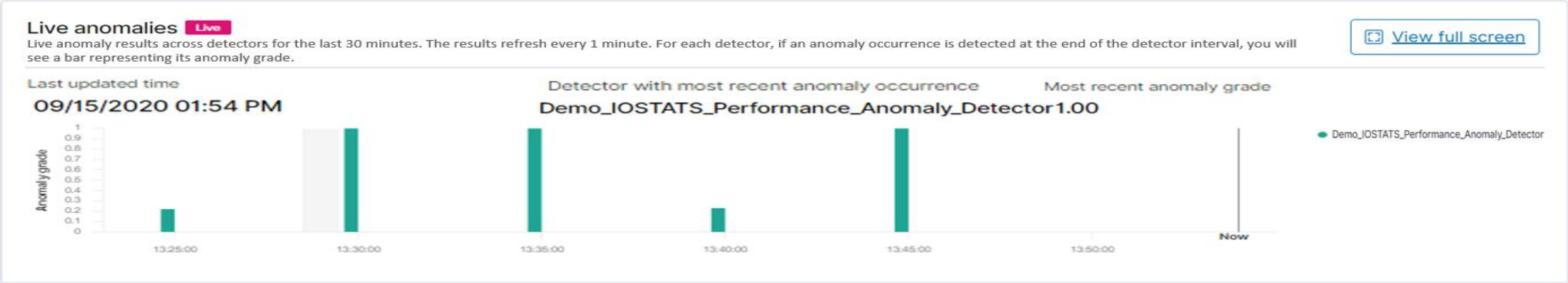
OpenDistro Anomaly Detection Procedure



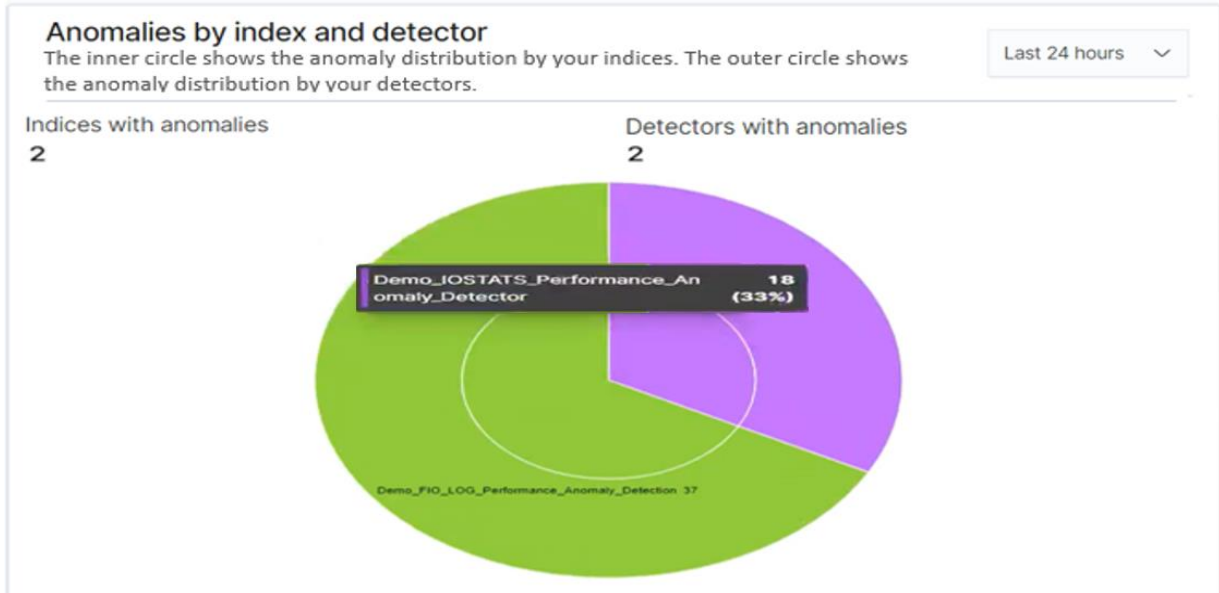
Result – Live Anomaly Summary



- ❑ Displays live anomaly result across detectors for last 30 minutes along with bar representing anomaly grade with respect to each anomaly.



- ❑ Displays anomalies with respect to each data source and detectors.



Result – Anomaly History



Live anomaly result for Demo_IOSTATS_Performance_Anomaly_Detector





Result – Live Exploration Metrics

User can perform live exploration of different matrices to root cause identified anomalies.

Dashboard / Demo_Performance_Data_Discover

Full screen Share Clone Edit

Search KQL Last 2 days Show dates Refresh

+ Add filter

FIO Performance Result Details

FIO Workload Access Specification & Performance Test Result

| Time | jobname | disk_name | host | queue_depth | error | read_bandwidth(KiB/sec) | write_bandwidth(KiB/sec) | read_clat_mean(usec) | write_clat_mean(usec) | cpu_sys(%) | cpu_user(%) | read_iops | write_iops |
|-------------------------------|---------|-----------|-----------|-------------|-------|-------------------------|--------------------------|----------------------|-----------------------|------------|-------------|-----------|------------|
| > Sep 16, 2020 @ 18:11:35.963 | seqread | nvme0n1 | elkserver | 16 | 0 | 670,511 | 0 | 90.273 | 0 | 57.92 | 21.357 | 167,627 | 0 |
| > Sep 16, 2020 @ 18:11:35.962 | seqread | nvme0n1 | elkserver | 16 | 0 | 670,543 | 0 | 90.27 | 0 | 57.915 | 21.358 | 167,635 | 0 |
| > Sep 16, 2020 @ 18:11:34.957 | seqread | nvme0n1 | elkserver | 16 | 0 | 670,048 | 0 | 90.336 | 0 | 57.915 | 21.353 | 167,512 | 0 |

FIO Process & System Info Details

System and Process specific CPU, Memory usage info

| Time | process.name | system.process.cgroup.blkio.total.bytes | system.process.cgroup.blkio.total.ios | system.process.cpu.total.pct | system.process.cpu.total.norm.pct | system.process.memory.size | system.process.cgroup.memory.mem.usage.bytes | system |
|-------------------------------|--------------|---|---------------------------------------|------------------------------|-----------------------------------|----------------------------|--|--------|
| > Sep 16, 2020 @ 18:11:35.770 | fio | 137,158,915,575,808 | 5,033,587,143 | 0.049 | 0.001 | 677,421,056 | 8,414,109,696 | 12,56 |
| > Sep 16, 2020 @ 18:11:34.770 | fio | 137,157,920,051,200 | 5,033,582,071 | 0.04 | 0.001 | 677,449,728 | 8,552,996,864 | 12,56 |
| > Sep 16, 2020 @ 18:11:34.770 | fio | 137,157,920,837,632 | 5,033,582,075 | 0.823 | 0.017 | 610,357,248 | 8,553,521,152 | 12,56 |

Disk IO Stats Details

Disk IO Stats info

| Time | metricset.name | system.diskio.name | system.diskio.iostat.read.request.per_sec | system.diskio.iostat.write.request.per_sec | system.diskio.iostat.read.per_sec.bytes | system.diskio.iostat.write.per_sec.bytes | system.diskio.iostat.read.await | system.disk |
|-------------------------------|----------------|--------------------|---|--|---|--|---------------------------------|-------------|
| > Sep 16, 2020 @ 18:11:36.222 | diskio | nvme0n1 | 2,918.551 | 0 | 382,248,517.503 | 0 | 0.349 | 0 |
| > Sep 16, 2020 @ 18:11:35.223 | diskio | nvme0n1 | 5,757.94 | 0 | 754,417,804.635 | 0 | 0.334 | 0 |
| > Sep 16, 2020 @ 18:11:34.222 | diskio | nvme0n1 | 5,835.113 | 0 | 764,533,282.572 | 0 | 0.34 | 0 |

Benefits

- Minimize analysis time
- Performance anomaly detection with high accuracy
- Quick feedback on performance results
- Assist in performance anomaly triaging
- Reduce turn around time for Performance Qualification
- Derive anomalous performance use case

THANK YOU