CERIAS

The Center for Education and Research in Information Assurance and Security

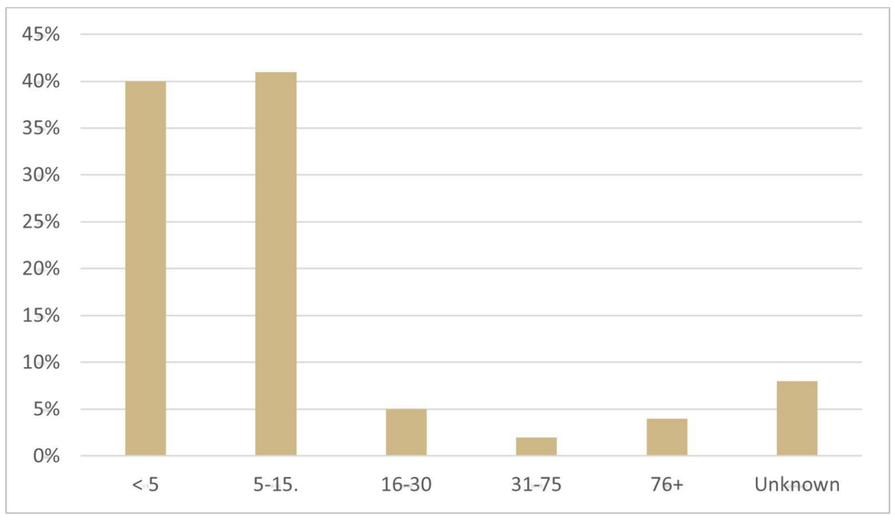
eBPF-based APM and Observability for Cloud-native Infrastrcture

Bhavye Sharma, sharm609@purdue.edu and Deepak Nadig, nadig@purdue.edu

Identifying application anomalies & performance bottlenecks has become difficult due to:

- Microservice architecture with different programming languages and tooling
- Ephemeral container-based workloads
- Large scale distributed applications in the cloud

Average # of observability tools used[1]



Observability tools collect 3 types of data i.e., **Metrics, Traces and Logs**

Traditional APM tools fail to provide sufficient visibility into cloud-native workloads.

Limitations of traditional APM tools[2]

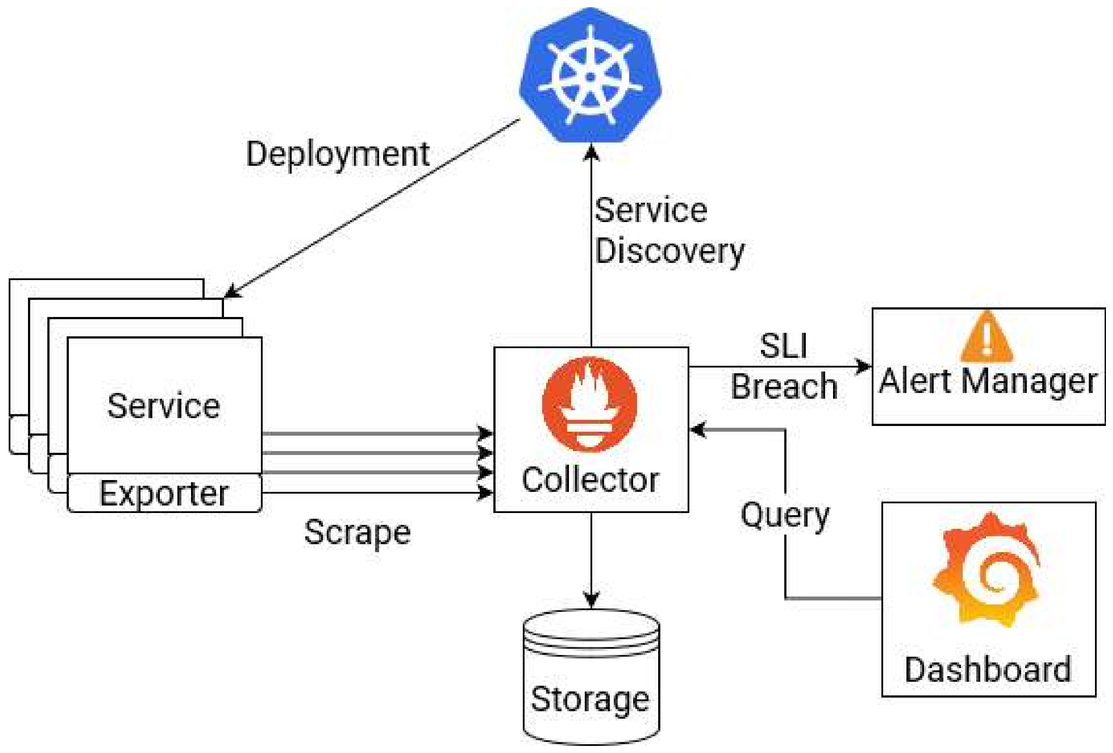
Complexity of root cause analysis	**Vast amount of data**	**Correlation issues across environments**
Speed of modern app development	**Vendor lock in with multiple tools**	**Security risks during peak traffic**

Our Research Questions:

- Can we export high-accuracy observability data without instrumenting the application layer?
- Can we use ML and DL to classify application behavior based on observability data?

Our solution exports Linux host and container level metrics to a time-series database. The data is used to train our machine learning models.

Most organizations that build observability with a **centralized monitoring and scraping approach**



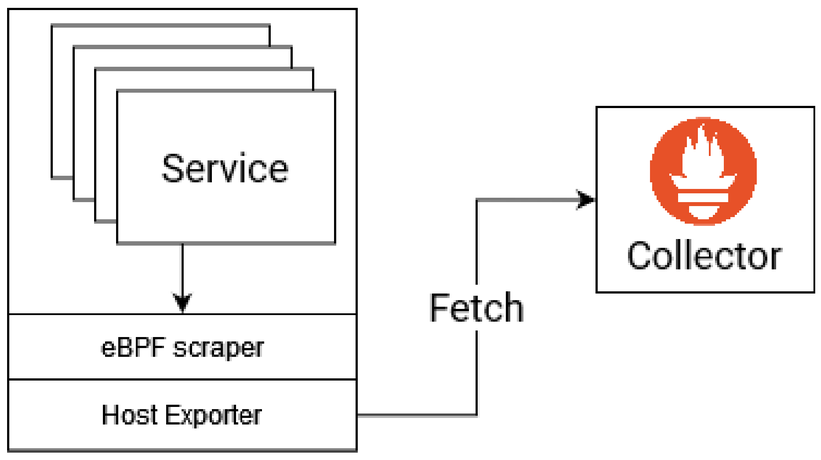
Our Proposed Methodology:

Build open-sourced unified pipeline for monitoring and collecting baseline performance data.

Simulate production workloads like Data Mining, Machine Learning, Web Servers and Databases.

Analyze the subset of critical hardware metrics, OS metrics, application logs, kernel traces for productive workloads.

Train one-class anomaly detector neural network[3] to identify anomalies in our workloads.



[1]: Grafana Lab's 2023 Observability Survey
[2]: Simform's Traditional APM to Enterprise Observability Guide
[3]: Anomaly Detection using One-Class Neural Networks

Purdue Cloud-native, Cyberinfrastructure and Networks

CYAN

PURDUE UNIVERSITY

CERIAS