SOLUTION BRIEF

Accelerated CI/CD Software Delivery

Enhance developer productivity and support expanding test requirements and aggressive delivery schedules.

Building software releases that deliver on functional requirements and user expectations is no easy task. You need to continually revise code to ensure it behaves as expected and stands up to demanding loads. As your software becomes more advanced, you must test more features and use cases, and do it in less time. To ensure you meet aggressive timelines while delivering quality release candidates, your CI/CD processes need to be backed by log analytics pipelines built on an infrastructure that supports exponential growth and rapid changes to test environments.

Analytic Pipelines and the Need for Modern Storage

Delivery of high-quality software requires continuous analysis of error logs and performance metrics pulled from functional and stress testing. These logs help developers identify bugs and determine if the issue is new, or one already being tracked. Traditionally, testing of code was done by manually parsing logs, correlating data, and analyzing issues. This testing required numerous test engineers to spend hours or even days to root cause issues.

Log analytics pipelines improve developer productivity by streamlining and automating the analysis process so code issues can be identified and resolved in less time. The challenge is that these pipelines need high-speed access to large amounts of data to precisely identify issues. As the number of log sources grows, the throughput required to pull logs can quickly scale to multiple GBs per second and TBs of data stored per day.

This issue becomes more severe when enrichment and data tagging are added to improve the precision of analysis. Tagging can cause capacity needs to grow by four, five, or even more times while dramatically increasing the performance load. With legacy storage, development teams are forced to choose between doing less analysis or suffering delays in code releases due to analysis taking longer and longer to complete.

Enhance Productivity

- Quickly see if bugs are known or new, enabling developers to act sooner.
- Store more data and re-filter it faster to speed code readiness.

Accelerate Analysis

- Process data through pipelines in less time to alert on issues sooner.

Increase Agility

- Leverage containers with Pure Storage Orchestrator to dynamically adjust test environments in a fraction of the time.

FlashBlade: Accelerating Pipeline Performance and Agility

Pure Storage® FlashBlade® provides the scale, speed, and simplicity necessary to support the aggressive data acquisition requirements of log analytics pipelines, even as the number of features, uses cases, and active code branches increase. The platform's three-dimensional scaling enables growth to multiple PBs of capacity and over 24 million MBps of throughput. This allows more data from more sources to be stored longer. It also allows the number of concurrent tests to increase exponentially without adding latency that delays delivery. As needs grow, FlashBlade makes capacity and performance scaling simple through a unique design that manages load balancing and scale-out networking inside the FlashBlade. This minimizes risks of misconfiguration and dramatically reduces the amount of time you'll spend between install and giving resources to testers working on tight schedules.

Accelerated Filtering and Indexing Processes

As testing becomes more complex, your developers need to be able to adjust their DevOps and test strategies. Often that involves reading data back, re-filtering, and then indexing data stores to build a more complete view of the environment. Here FlashBlade offers multiple advantages over legacy NAS systems. In addition to scalability and read/write optimization, the NFS file system has been optimized to reduce latency. The FlashBlade NFS client sequentially walks the file system structure to accelerate data discovery and skips the kernel for high speed, parallelized discovery. Reads are buffered to improve throughput and lower latency for up to 75% faster re-filtering and rebuilding of indexes.

Ensuring Testing Agility with Container Support

Today's testing environments are highly dynamic (Figure 1). The number of applications in a pipeline and the number of resources allocated to individual tests can shift quickly and developers do not have time to wait for legacy provisioning to occur, especially as the number of tests runs per day grow by ten times or more and delivery timelines shrink.

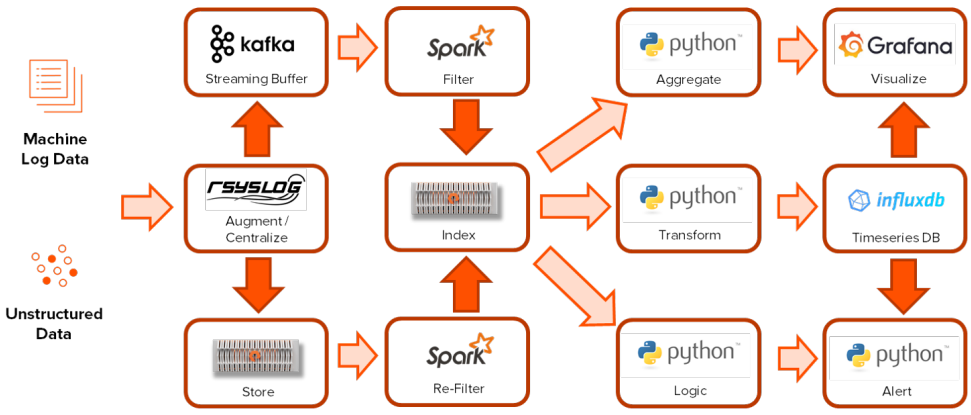

Figure 1. Log Analytics Pipeline with Different Pipeline Functions and Applications That Can Be Used by Function.

To ensure stability testing occurs and teams can execute on last-minute bug fixes, development teams are turning to containers for dynamic provisioning of compute and storage resources in a fraction of the time. This allows test teams to pivot quickly and deploy new application instances without waiting before tests can resume. To meet these needs, Pure Storage offers Pure Service Orchestrator. Pure Service Orchestrator virtualizes FlashBlade systems into a single resource, eliminating manual array management with Kubernetes and providing intelligent provisioning based on system loads as well as Quality of Service needs. This can easily reduce the time to spin up or adjust a development environment by two times or more.

Pipelines in Production: Pure Storage FlashBlade Testing

At Pure, our development team recognized similar issues with traditional functional and stress testing methodologies. Over the last five years, we have rapidly expanded the FlashBlade development program, growing from a few hundred tests per day to over 120,000 tests per day running across 2,500 simulated environments with 1,000 load generators managed by 20 Jenkins instances.

The log and tertiary data set ingested from these tests also grew from roughly 500,000 events per second to over 4.5 million events per second. That concurrent load dramatically increased throughput needs as well as storage requirements, with over 24TB of raw data accumulated per day—72TB when enriched with metadata. The net was a capacity needs growth of 50x.

As the data levels grew, so did the index size and the frequency at which data needed to be re-filtered so the index could be rebuilt to enable aggregation and transformation of data. This was critical to ensure visualization tools (e.g. Grafana) and time-series databases (e.g. InfluxDB) used in the bug-analysis process provided the most accurate picture of what was happening and why. At the same time, the team needed a way to quickly spin up Python, Jenkins instances, and more using Docker.

To meet the performance, capacity, and container requirements, the team implemented FlashBlade. The move to FlashBlade eliminated bottlenecks and enabled the team to:

- Scale concurrent testing from a few hundred tests per day to more than 120,000 tests per day.
- Increase data retention by 36 times to improve the accuracy of log analysis across tests.
- Accelerate bug analysis per day by five times while reducing issue triaging times.

The net result is that developers are more productive, release cycles finish with fewer delays, and testing is more thorough, ensuring a better customer experience.

Additional Resources

- Find out more about FlashBlade.
- Learn how Pure Service Orchestrator simplifies Kubernetes deployments.
- Learn more about using Pure Professional Services to accelerate your strategy.

©2020 Pure Storage, the Pure P Logo, and the marks on the Pure Trademark List at https://www.purestorage.com/legal/productenduserinfo.html are trademarks of Pure Storage, Inc. Other names are trademarks of their respective owners. Use of Pure Storage Products and Programs are covered by End User Agreements, IP, and other terms, available at: https://www.purestorage.com/legal/productenduserinfo.html and https://www.purestorage.com/patents

PS1924-01 11/2020