

## SOLUTION BRIEF

# Enabling Modern Service Analytics

Improve customer satisfaction while controlling costs.

Ensuring a positive customer experience is no longer as simple as providing a well-made product and having a trained staff to answer questions. You need to be one step ahead of problems, predicting and resolving *minor* issues before they become *major* headaches. Cloud-based analytics are a powerful tool to solve service issues, but they can be expensive, especially as the number of data sources and tests you need to run grow. To ensure your customers are happy and your budget isn't overwhelmed, you need a service analytics strategy powered by a hybrid cloud infrastructure that delivers speed, scale, and simplicity.

## Cloud-based Analytics and Cost Challenges

The heart of a good service-analytics strategy is data. Whether gathered via batch or real-time acquisition, the more data you store, the greater is your ability to:

- Correlate data sources to accurately identify root cause.
- Plot trends to accurately identify current and future issues.
- Mine the install base to find other customers that may need support.

These help proactively resolve issues and ensure customers stay happy.

In a cloud model, you acquire data via compute instances running stream processing software (e.g. Apache Kafka). Once acquired, cloud storage holds the data. Analytic pipelines then read the data through ETL processes to relational or NoSQL databases for alerting and use by support visualization tools.

The challenge is that reading data into cloud-based pipelines has a cost. It may be nominal at first, but as you use more data for root causes, investigate your install base for similar issues, or do advanced customer satisfaction tests cloud costs can quickly grow by five to ten times.



### Utilize Cloud Analytics

Harness the power of hybrid cloud for high speed analytics while dramatically reducing costs.



### Scale Without Compromise

Leverage 3-dimensional scaling to retain growing data sets and run concurrent analysis on data sets without slowdowns.



### Minimize Risk

Enterprise resiliency, intelligent fault recovery and non-disruptive upgrades ensure you're always ready to serve customers.

## Modeling and AI: Exponential Value and Cost

Cost escalation becomes a major consideration as you expand the way you use logs and other data sources. Some common examples of these escalating costs are:

- **Ad Hoc Support Queries:** When your support team receives a call from a customer, systems must read, analyze, and review data. Having access to historical customer data as well as the ability to see if anyone else has seen similar issues on other systems enables your teams to resolve customer issues faster, but this ability can also carry additional read costs.
- **Simulation Modeling:** Using hardware and workload modeling, service teams can enable customers to proactively determine if and when they may need to upgrade their environment. Simulation modeling can be a powerful tool for helping customers do what-if planning, but it requires historical data to build trend lines, account for seasonality, and determine the benefit of moving to a range of system types. Analyzing this data using the cloud carries a cost.
- **AI Development:** No service algorithm is static. To deliver the best customer experience, you must constantly test and re-train your AI models. Accomplishing this requires data intake and processing to account for new things that customers may be seeing and account for updates in a your offerings.

Combined, advanced services analytics like these can enable your support teams to deliver customer sat scores exponentially higher than before. But since they require on average 10 reads per log, cloud costs can grow quickly.

## The Pure1® Solution: Hybrid Cloud with FlashBlade®

Pure1 is Pure Storage's cloud-based monitoring, fault management, and forward planning service. Similar to the discussion above, it acquires analytics from customer-deployed FlashArray and FlashBlade systems via encrypted transfer of logs, performance statistics, and alerts.

When you build platforms and add new service capabilities—ad hoc queries, simulation modeling, AI development, etc.—costs can escalate quickly. Building new services for customers should not come with massive monthly cloud provider bills. To eliminate exponential cost increases, Pure Storage developed a hybrid cloud strategy that leverages the scalable compute capabilities of public cloud with the powerful data processing capabilities of FlashBlade.

In this hybrid cloud model, Pure1 acquires array data via compute instances running stream processing software. As Pure 1 receives analytics, it streams and filters them (via Apache Kafka and Spark) for immediate analysis to determine if Pure1 needs to issue an alert. At the same time, the Pure1 servers pass the data via Direct Connect to Pure FlashBlades at a nearby Equinix facility. Data passes directly between the EC2 instances and FlashBlade via an S3 mount point, eliminating cloud read operations.

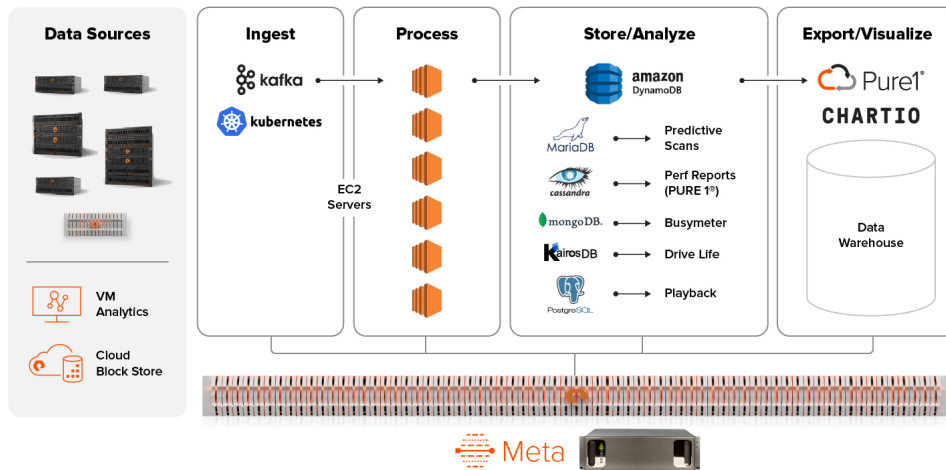


Figure 1: Cloud analytics with Pure Storage

Once stored on FlashBlade, the system makes the data available to a broad range of analytic pipelines including:

- Predictive scanning applications running on MariaDB
- Resiliency testing applications (looking at wear levels and soft errors) running on KariosDB
- Visualization tools used by support staff and running on ChartIO
- Machine learning tools used to build modeling and recommendation engines

Some of the most latency-sensitive applications run at Equinix, but for efficiency and cost control, most of these pipelines run in AWS. Based on Pure1 experiences, this does not affect latency, even as the number of concurrent service processes increase. There are two primary reasons for this. First, the Amazon S3 connection to the FlashBlade uses high-bandwidth connectivity to prevent latency. Second, we built FlashBlade around a non-blocking, scale-out architecture that automatically distributes workloads across different blades for maximum performance. We also designed multi-blade architecture for concurrency to readily deliver up to 1.5TB per second of throughput across tens of thousands of parallel requests.

Combined, this strategy has enabled Pure Storage to reduce service analytic costs by 50% while doubling the number of value-added support services we offer to customers. At the same time, we have been able to increase the frequency of AI/ML model re-training. This has enabled Pure to improve the accuracy of support analysis and accelerate the time to resolve issues by two times or more.

### Additional Resources

- Find out more about [FlashBlade](#) and Pure Storage [solutions for analytics](#).
- See how Pure1 helps [proactively resolve support issues](#) and empower customer decision making.

[purestorage.com](http://purestorage.com)

800.379.PURE

