

UNLEASH MODERN ANALYTICS WITH FLASHBLADE

SUMMARY

Big data is no longer static, predictable data. In the modern era of analytics, infrastructure must be architected for the unknown, delivering real-time performance with cloud-like agility for any unstructured data. Modernize with Pure Storage® FlashBlade™, the industry's first dynamic data hub purpose-built for today's analytics workloads.

LEGACY, DISK-BASED STORAGE: THE OLD ERA

Data from a decade ago is different than data today. Hadoop File System (HDFS) and its distributed direct-attached storage (DDAS) architecture was designed for data that is static and predictable. While it worked for batched, sequential workloads, it led to complex application silos that lack the agility and real-time performance expected from modern applications like Apache® Spark™ and machine learning.



DECADE AGO

ASSUMPTIONS ABOUT DATA IN HADOOP FILE SYSTEM

- Typical File is **Large**
- Access is **Sequential**
- Hardware Failure** is the Norm
- Data is **Batched**
- Network is **Slow**

MODERN DATA CHANGES EVERYTHING

The industry is in a new era. While Hadoop was the only widely available analytics tool a decade ago, data scientists have a plethora of tools at their disposal today. Spark, Kafka™, and Elastic all push the limits of data in drastically different ways. And with the advent of machine learning and AI, truly unstructured, random data is enabling new insights, driving innovations once considered science fiction. A new class of data platform is needed, one that is built for the needs of modern, dynamic data – and architected for the unknown. This is FlashBlade.



TODAY

DATA IS NOW DIFFERENT

- Small** to Large Files
- Random** to Sequential Access
- Real-time** or Batched
- Apps & Data **Evolve Quickly**
- Elastic** Infrastructure

CUSTOMER HIGHLIGHT

10X FASTER APACHE SPARK FOR MAN AHL, A LEADING INVESTMENT FIRM

Man AHL is the world's leading public hedge fund company, with more than \$19 billion in assets under management. The firm's investment decisions are based on mathematical models running on computers, without human involvement.

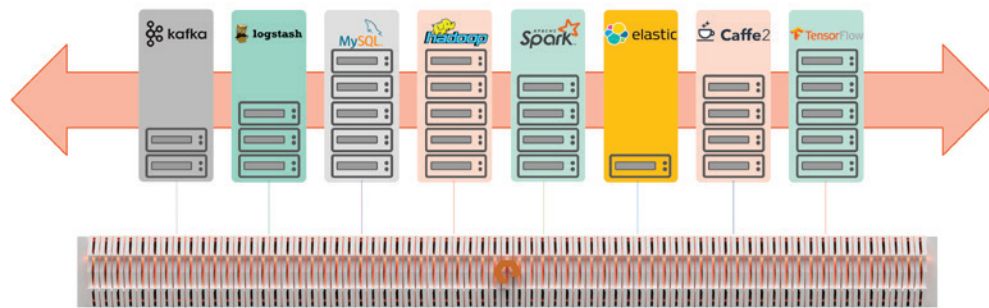
With roughly 50 quants relying on IT tools, they decided to build their infrastructure on FlashBlade.



"Many of our researchers have found that the introduction of FlashBlade has made it easier to use Spark for performing multiple simulations. One of them experienced a 10 to 20x improvement in throughput for his Spark workloads compared to the previous storage system," said James Blackburn, Head of Data Engineering. It is a game-changer that creates time-to-market advantage for the company.

FLASHBLADE IS ARCHITECTED FOR THE UNKNOWN IN MODERN DATA

FlashBlade is the industry's first dynamic data hub purpose-built for modern analytics, offering real-time response for any unstructured data, with cloud-like agility. FlashBlade is massively parallel at its core, and delivers unprecedented performance and simplicity for data scientists and engineers. From small, metadata-heavy workloads to random, large file accesses, it is architected for the unknown – whatever your next big data initiative may be.



FLASHBLADE: DYNAMIC DATA HUB BUILT FOR TODAY'S HADOOP AND SPARK WORKLOADS

While DDAS architecture was appropriate for the old world of analytics, modern data requires a new class of scale-out storage – a data hub built for unstructured workloads with cloud service-like agility and simplicity. FlashBlade is purpose-built for modern analytics, with six essential qualities required to power the next-generation of big data.

TUNED FOR EVERYTHING



7.5M IOPS, 75 GB/s, Small to Large,
Random to Sequential, Structured to Unstructured

REAL-TIME



Millisecond Latency for
Responsive Workloads

ALL-FLASH



Optimized for Flash
From Software to Hardware

PARALLEL



Massively Parallel Purity SW
75 Blades w/ Scale-Out Fabric

ELASTIC



Linear Performance
Growth w/ Each Blade

SIMPLE



Evergreen – No TB Re-Buys
No Manual Tuning Required



VISIT PURESTORAGE.COM/ANALYTICS TO LEARN MORE

© 2017 Pure Storage, Inc. All rights reserved. Pure Storage, FlashBlade, and the P logo are trademarks or registered trademarks of Pure Storage, Inc.
ps_sb2p_data-analytics_itr_01