

ALLUXIO ENABLES COMPUTE



DATA LOCALITY

Make your data as close as possible to compute.

Make your data local to compute workloads like Spark, Presto, and Tensorflow, across hybrid and multiple clouds.



DATA ACCESSIBILITY

Make your data accessible in many ways.

No matter if it sits on-prem or in the cloud, HDFS or S3, make your files and objects accessible in many different ways.



DATA ON-DEMAND

Make your data as elastic as compute.

Effortlessly orchestrate your data for compute in any cloud, even if data is spread across multiple clouds.

Bring your data to compute, on-demand

Data orchestration challenges for today's data engineer

Big data in the cloud too slow?

S3 performance for big data workloads is inconsistent and data egress is expensive.

Want to use object stores for big data workloads?

Object storage performance, particularly for metadata operations, is unpredictable.

Hybrid cloud for data too hard to implement?

Making HDFS or object store data accessible to any compute in any cloud is complex, expensive, and slow.

Multi-cloud data access too complex?

Orchestrating data from multiple public clouds for big data workloads is complex and expensive.

Alluxio can help

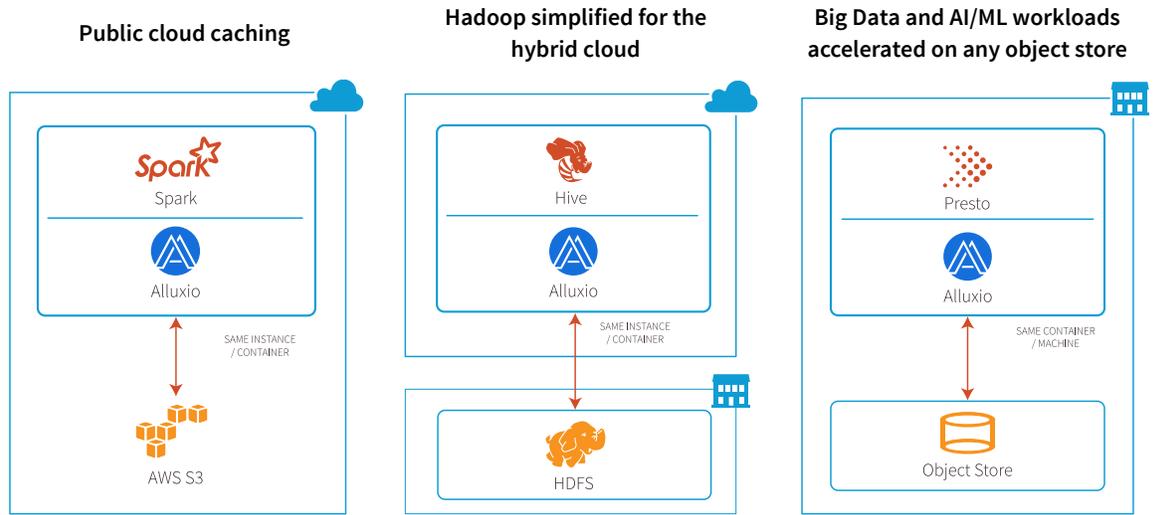
- Data is local to compute, giving you memory-speed access.** Alluxio makes your data local to frameworks like Spark, Presto, Tensorflow, and more, accelerating speed to insights with hot data made local to compute faster. Leverage Alluxio's intelligent tiering as a distributed caching layer on top of slow or remote storage to avoid reading data repeatedly from the cloud.
- Data is accessible through one unified namespace, regardless of where it resides.** No matter where your data sits, Alluxio makes your files and objects accessible with frameworks like Spark, Hive, Presto, and more to easily integrate with storage systems like HDFS and S3. Reduce costs associated with ETL and eliminate multiple persisted copies.
- Data is as elastic and flexible as compute.** Burst data elastically with compute anytime, in any cloud environment. With Alluxio's unified namespace capability, independently scale compute and storage to reduce costs and resources needed for maintenance, particularly for hybrid and multiple cloud environments.

Meet Alluxio

Alluxio brings your data to compute, on demand. It's an open source data orchestration layer for big data and AI/ML workloads in the cloud, enabling data locality, data accessibility, and data elasticity.



Use Case and Deployments



Get in-memory data access for Spark, Presto, and Hive on any of the public clouds – AWS S3, Google Cloud Platform, or Microsoft Azure.

Simplify Hadoop for the hybrid cloud by making on-prem HDFS accessible to any compute in the cloud.

Accelerate your Spark, Presto and Tensorflow workloads for both on-prem and cloud object stores.

bazaarvoice:

Bazaarvoice, a digital marketing SaaS company, faced challenges in performance and cost to build an efficient big data analytics platform on AWS. Specifically, they needed a way to accelerate data access for S3 in order to serve their retail customers at internet-scale. With Alluxio, they achieve 10x acceleration of Spark and Hive jobs on AWS S3 with Alluxio’s tiered storage.

BARCLAYS

Leading financial services company Barclays wanted to enable big data compute frameworks on their traditional data warehouse. They faced challenges around slow/not interactive queries, which was resulting in major operational inefficiencies. With Alluxio, they can execute Spark workloads with dramatically better performance on data extracted from Teradata and stored within Alluxio.

About Alluxio

www.alluxio.io

Proven at global web scale in production for modern data services, Alluxio is the developer of open source data orchestration software for the cloud. Alluxio moves data closer to big data and machine learning compute frameworks in any cloud across clusters, regions, clouds and countries, providing memory-speed data access to files and objects. Intelligent data tiering and data management deliver consistent high performance to customers in financial services, high tech, retail and telecommunications. Alluxio is in production use today at seven out of the top ten internet companies. Venture-backed by Andreessen Horowitz and Seven Seas Partners, Alluxio was founded at UC Berkeley’s AMPLab by the creators of the Tachyon open source project.