Alluxio: Open Source Data Orchestration for Analytics and AI in the Cloud

Haoyuan (H.Y.) Li | Founder, Chairman & CTO | haoyuan@alluxio.com

2019-11-18 @ PDSW 2019

The Alluxio Story

amplab
2013

Originated as Tachyon project, at the UC Berkley's AMP Lab by then Ph.D. student & now Alluxio CTO, Haoyuan (H.Y.) Li.

ALLUXIO
2015

Open Source project established & company to commercialize Alluxio founded

ANDREESSEN HOROWITZ

Goal: **Orchestrate Data at Memory Speed for the Cloud** for data driven apps such as Big Data Analytics, ML and AI.
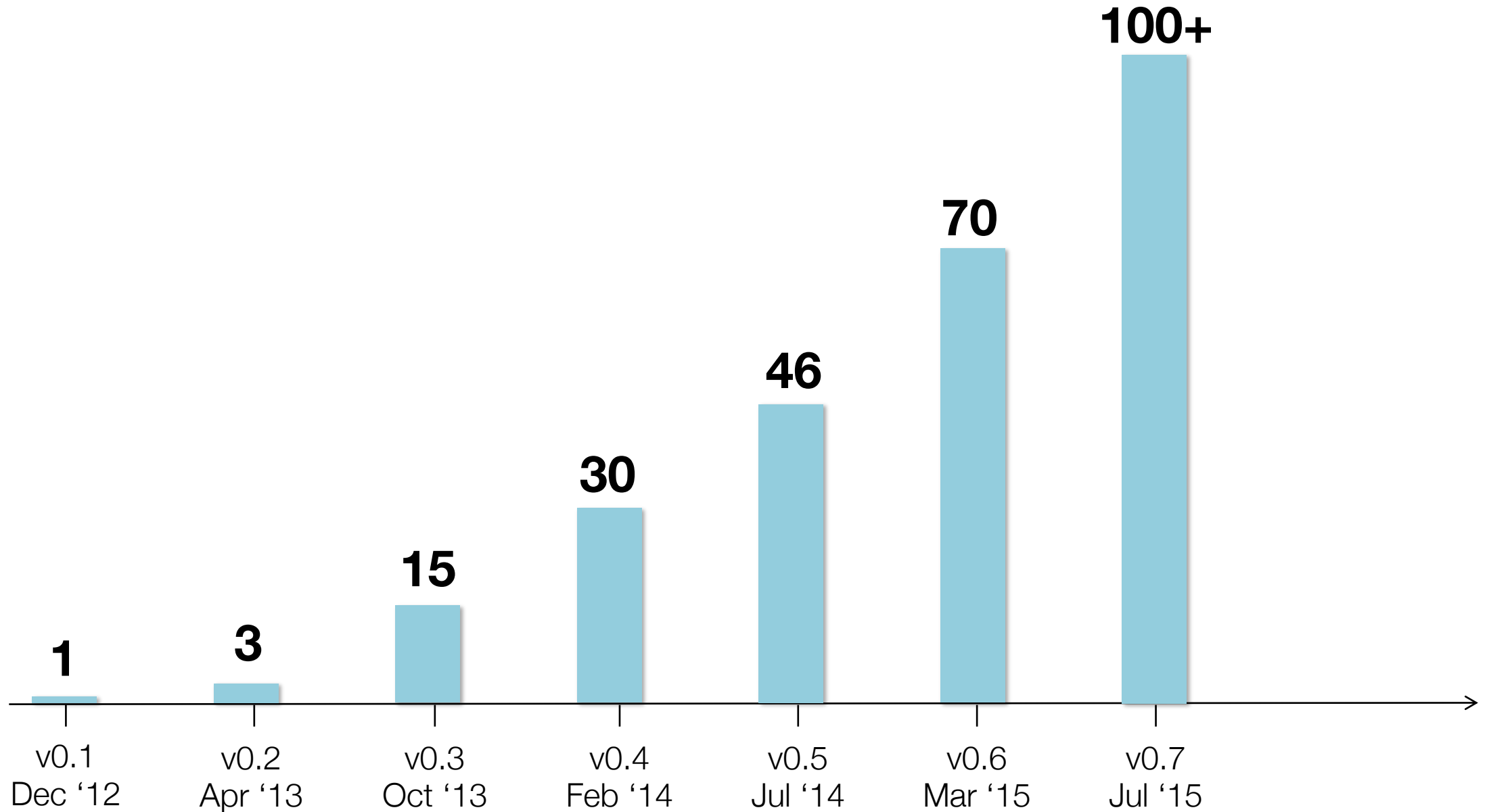
CRN
2018

NETWORKWORLD
FROM IDG
2018

IMPACT 50
THE INDUSTRY'S MOST IMPACTFUL COMPANIES
inside BIGDATA
2019

Open Source Started From UC Berkeley AMPLab

1000+ contributors & growing

4000+ Git Stars

Apache 2.0 Licensed

GitHub's Top 100 Most Valuable Repositories Out of 96 Million

Join the conversation on Slack

slackin.alluxio.io

Companies Running Alluxio (Learn More)

Financial Services

DBS · ING · BARCLAYS · WELLS FARGO · PayPal · HUATAI SECURITIES

Retail & Entertainment

Walmart · Shopee · ROBLOX · CAESARS ENTERTAINMENT · NetEase Games · Myntra · JD.COM · 苏宁易购 suning.com · VIP 唯品会 VIPSHOP.COM

Data & Analytics Services

Aunalytics rethink data · LigaData · SAMSUNG SAMSUNG SDS · TalkingData · RYTE

Technology

Tencent 腾讯 · PERCEPTIN · esri · Tencent Cloud · @WalmartLabs · eSENTIRE · Baidu 百度 · ORACLE · Alibaba Cloud · Baidu Cloud · Sogou 搜狗 · CUELOGIC · Lucidworks · QINIU 七牛云存储 · walkme · MOMO · Qunar.Com 去哪儿? 聪明你的旅行

Consumer

SAMSUNG · nielsen · YAHOO! · 链家 · Lenovo · bazaarvoice · GUARDANT HEALTH · 网易严选

Telco & Media

COMCAST · HUAWEI · 虎牙直播 Huya.com · swisscom · 中国移动 China Mobile · China unicom 中国联通

Travel & Transportation

waze · Ctrip · DiDi
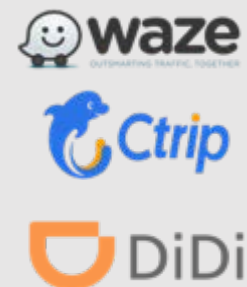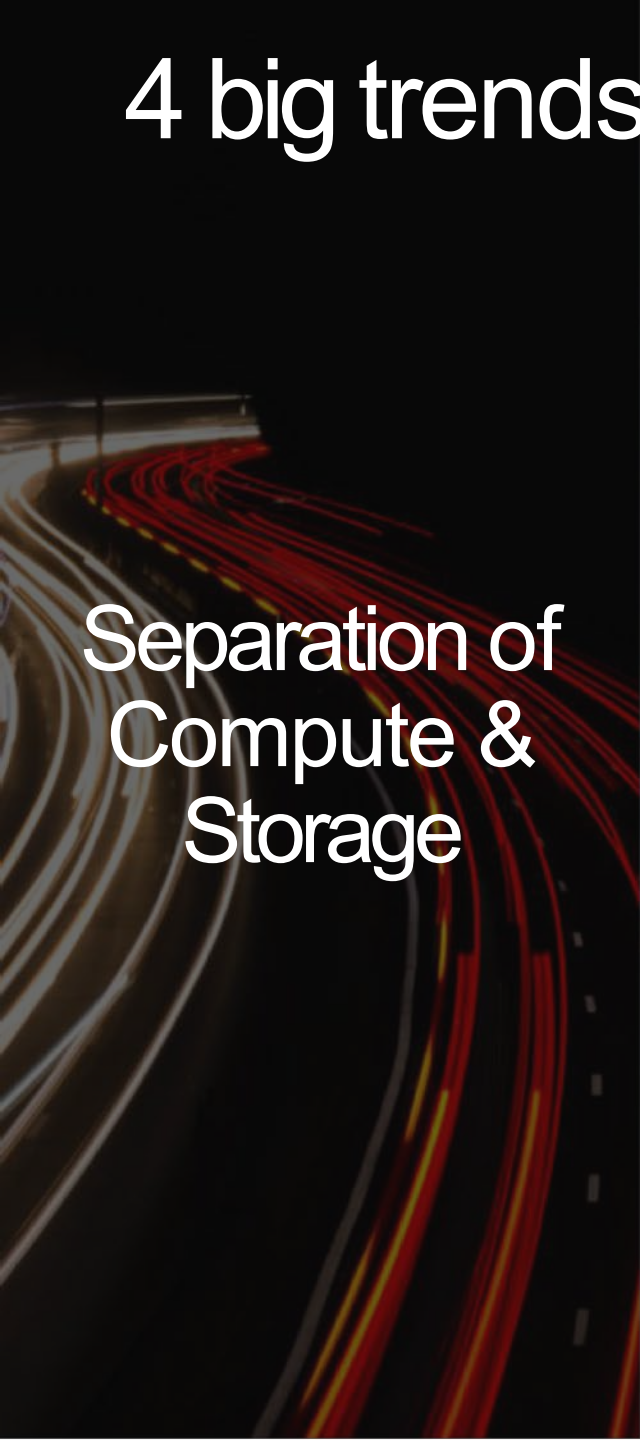
4 big trends driving the need for a new architecture

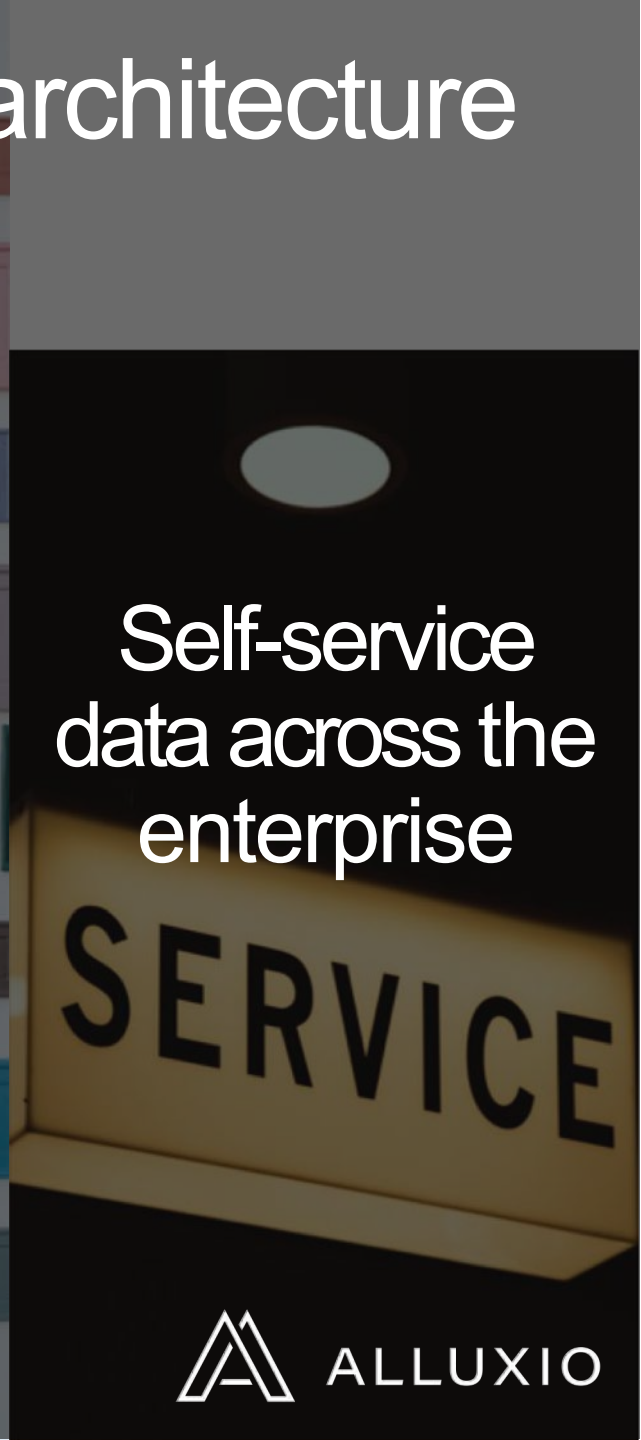**Separation of Compute & Storage**

Hybrid – Multi cloud environments

Rise of the object store

Self-service data across the enterprise

ALLUXIO

Data Ecosystem 1.0 – The Challenges

COMPUTE

Complex

presto   Spark   Flink   hadoop Map Reduce

HIVE   kafka   Caffe   APACHE HBASE   TensorFlow

Low performance

amazon webservices S3   Microsoft Azure   Google Cloud Storage   ceph   hadoop HDFS

IBM   NetApp   Hewlett Packard Enterprise   DELL EMC ECS   HITACHI Hitachi Data Systems

Expensive

STORAGE

ALLUXIO
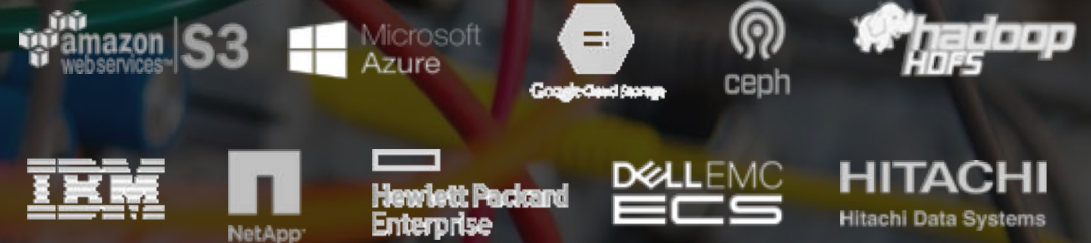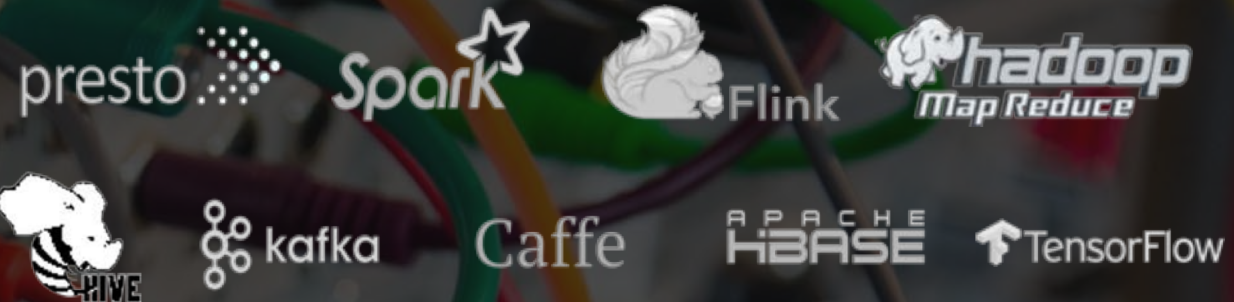
Data stack journey and innovation paths

Co-located

Disaggregated

Support more frameworks

Co-located compute & HDFS on the same cluster

Disaggregated compute & HDFS on the same cluster

Support Presto, Spark across DCs without app changes

HDFS for Hybrid Cloud

MR / Hive

HDFS

Hive

HDFS

Burst HDFS data in the cloud, public or private

Transition to Object store

Enable & accelerate big data on object stores
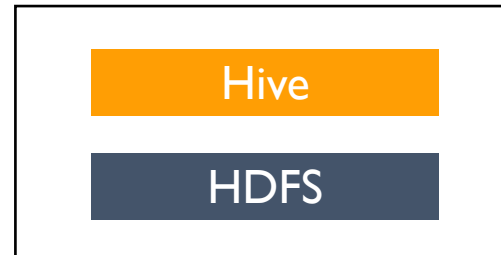
- Typically compute-bound clusters over 100% capacity
- Compute & I/O need to be scaled together even when not needed

- Compute & I/O can be scaled independently but I/O still needed on HDFS which is expensive

ALLUXIO

Independent scaling of compute & storage

Java File API	HDFS Interface	S3 Interface	POSIX Interface	REST API

ALLUXIO Data Orchestration for the Cloud

HDFS Driver	Swift Driver	S3 Driver	NFS Driver

APIs to Interact with data in Alluxio

Application have great flexibility to read / write data with many options

Spark

```
> rdd = sc.textFile("alluxio://localhost:19998/myInput")
```

Presto

```
CREATE SCHEMA hive.web
WITH (location = 'alluxio://master:port/my-table/')
```

**POSIX**

```
$ cat /mnt/alluxio/myInput
```

Java

```
FileSystem fs = FileSystem.Factory.get();
FileInStream in = fs.openFile(new AlluxioURI("/myInput"));
```

Challenges with supporting more frameworks across data centers

Support more frameworks

- Running new frameworks on existing an HDFS cluster can dramatically affect performance of existing workloads

- Orchestrating data to compute clusters in another data center is typically a manual effort and time consuming

- Storing and managing multiple copies of the data becomes expensive

**On-premise satellite compute clusters across data centers**

Data center A

Presto

Alluxio

Hive   MapReduce

hadoop HDFS

Data center B

ALLUXIO
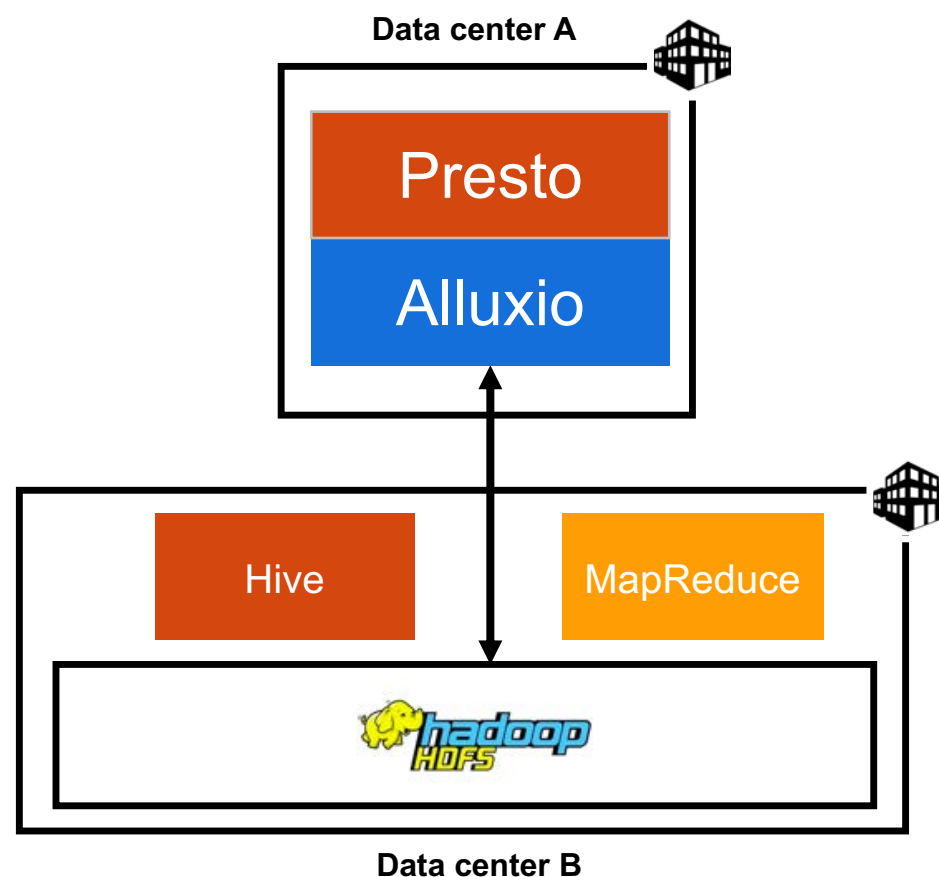
Compute caching for S3 / GCS

- S3 performance is variable and consistent query SLAs are hard to achieve

- S3 metadata operations are expensive making workloads run longer

- S3 egress costs add up making the solution expensive

- S3 is eventually consistent making it hard to predict query results

### Accelerate analytical frameworks on the public cloud
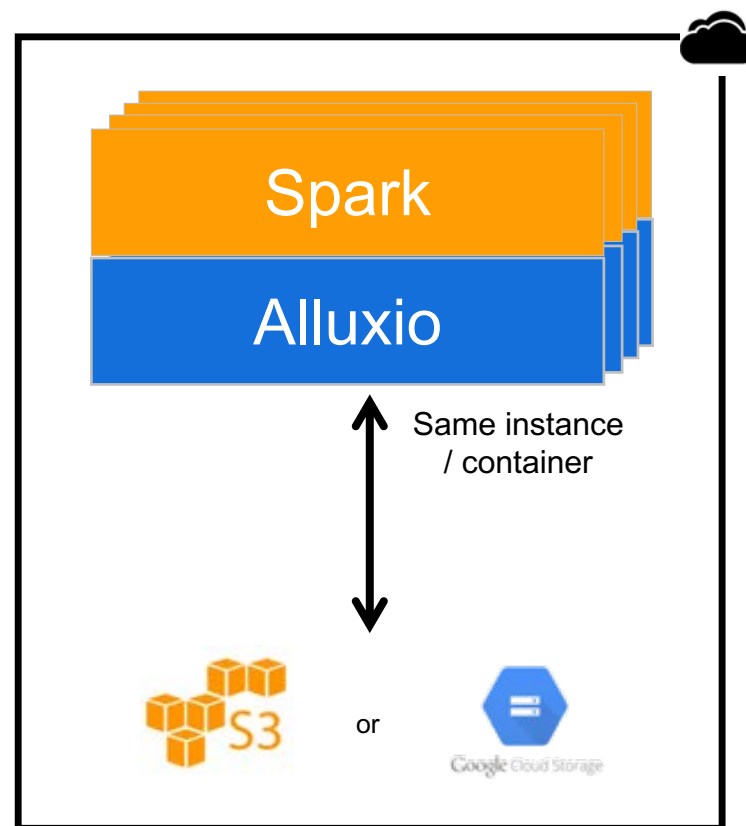
Challenges with Hybrid Cloud

HDFS for Hybrid Cloud

- Accessing data over WAN too slow

- Copying data to compute cloud time consuming and complex

- Using another storage system like S3 means expensive application changes

- Using S3 via HDFS connector leads to extremely low performance

Burst big data workloads in hybrid cloud environments

Presto

Alluxio

Same instance / container
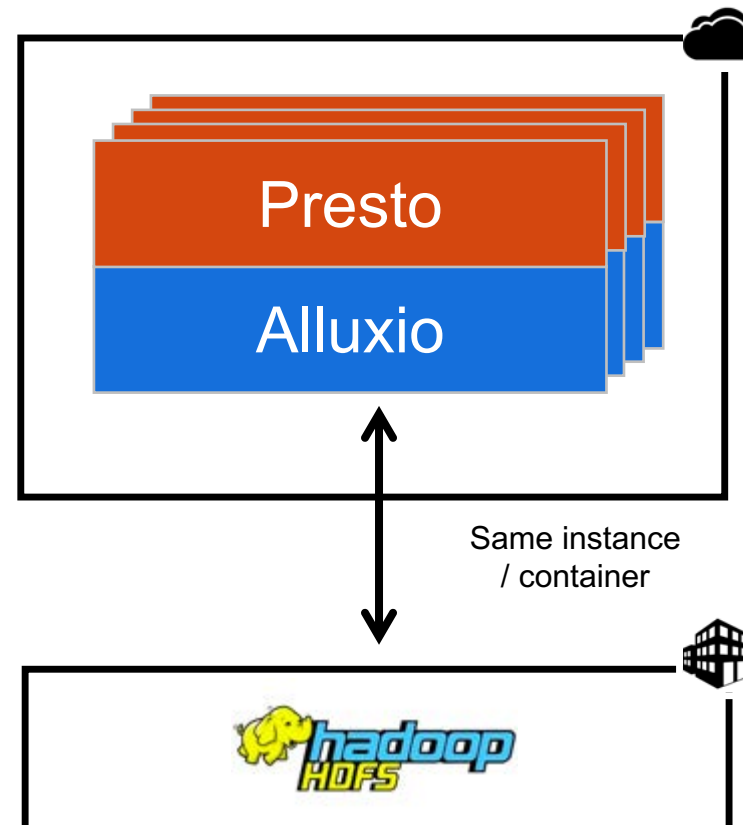
Solution Benefits
- Same performance as local
- Same end-user experience

- 100% of I/O is offloaded

ALLUXIO

Challenges running Big Data on Object Stores & Alluxio Solution

Transition to Object store

- Object stores performance for big data workloads can be very poor

- No native support for popular frameworks

- Expensive metadata operations reduce performance even more

- No support for hybrid environments directly

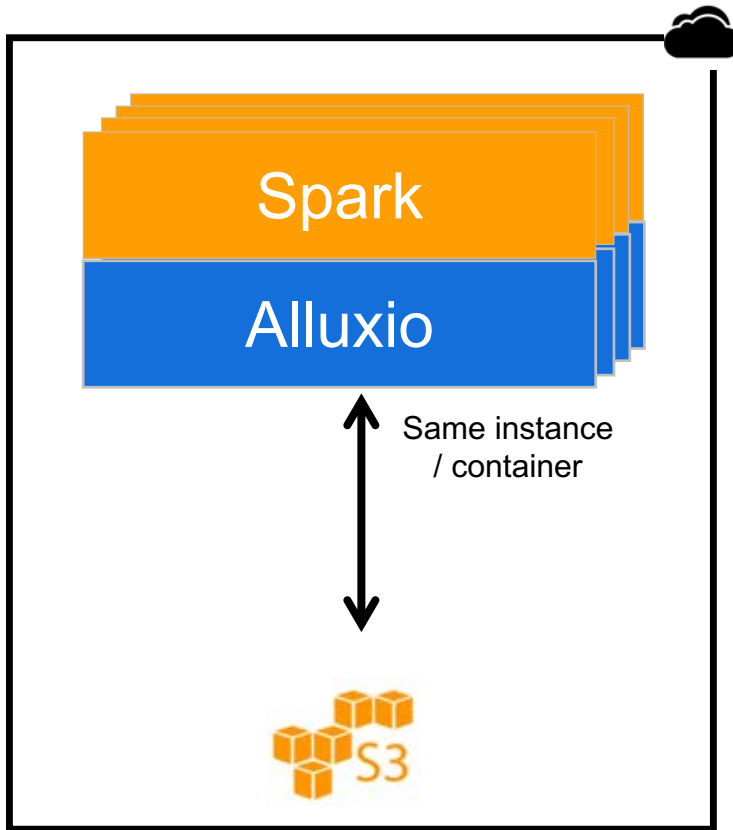**Dramatically speed-up big data on object stores on premise**

Presto

Alluxio

Same container / machine

ceph   or   cleversafe   or   DELLEMC ECS
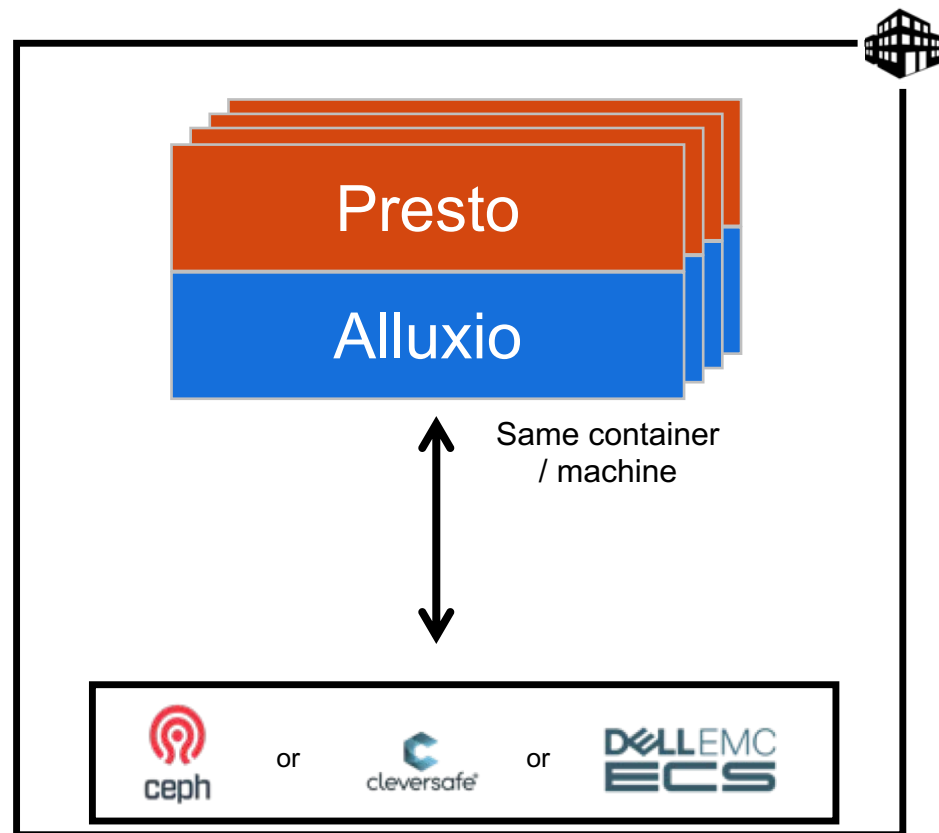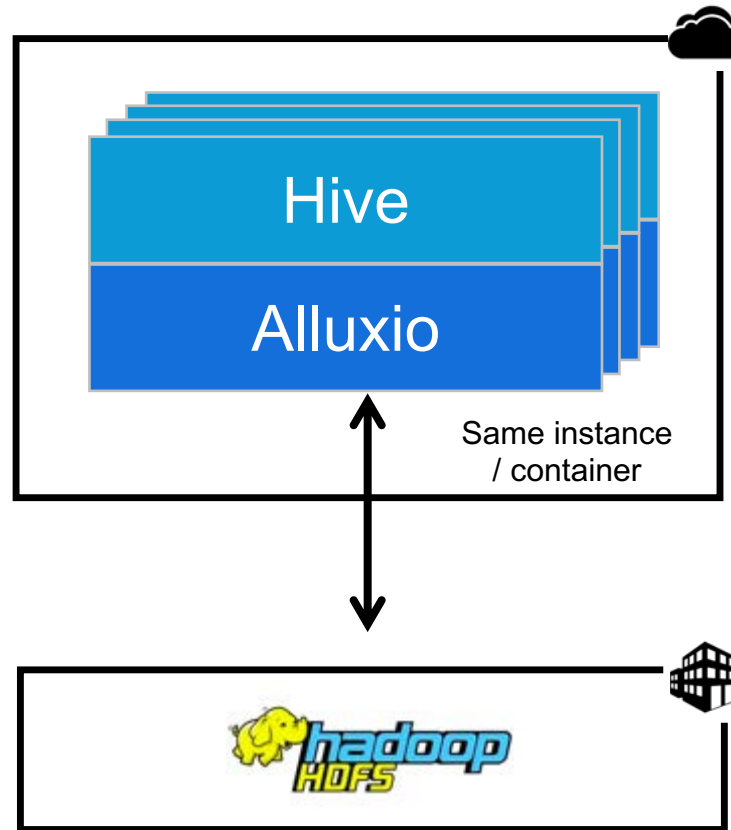
Solution Benefits

- Same performance as HDFS
- Uses HDFS APIs
- Same end-user experience

- Storage at fraction of the cost of HDFS

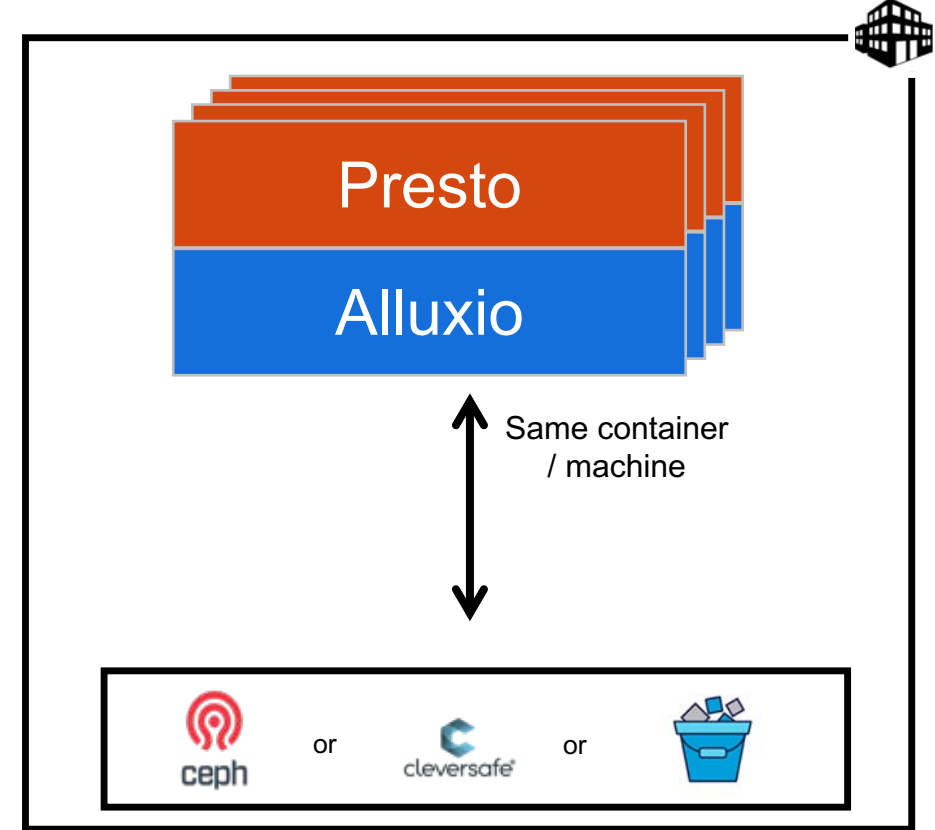ALLUXIO

Use Cases Alluxio Enables

Accelerate big data frameworks on the public cloud

Burst big data workloads in hybrid cloud environments

Dramatically speed-up big data on object stores on premise

Spark
Alluxio

Same instance / container

S3

Hive
Alluxio

Same instance / container

hadoop HDFS

Presto
Alluxio

Same container / machine

ceph   or   cleversafe   or
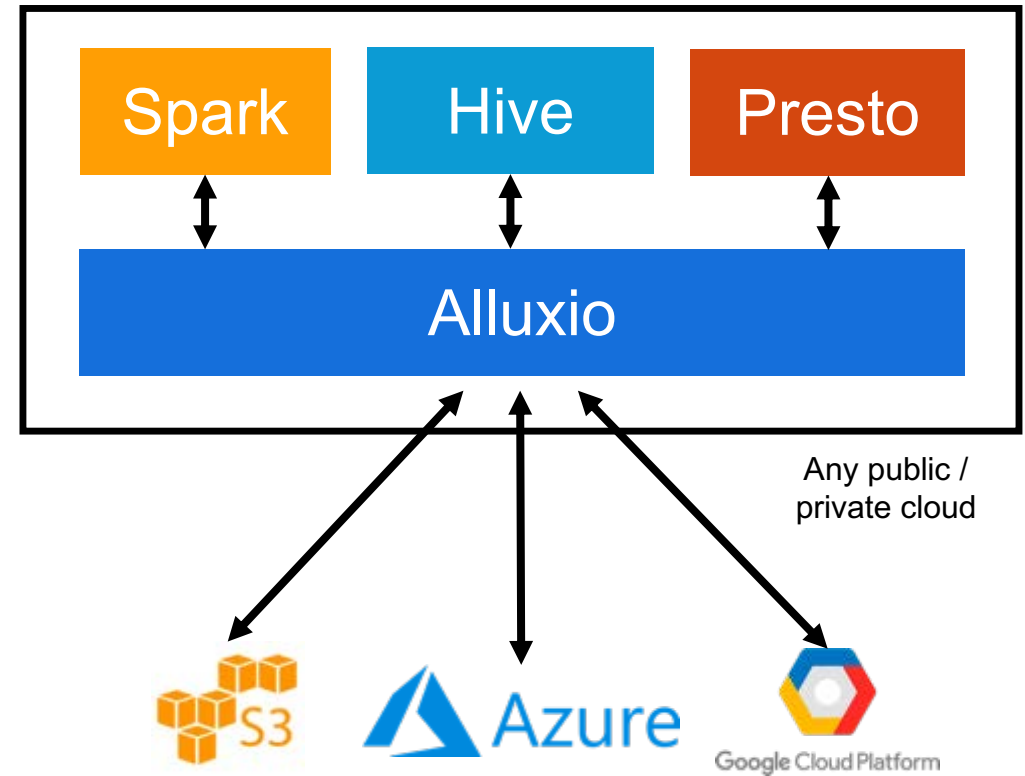
ALLUXIO

Advanced Use Cases
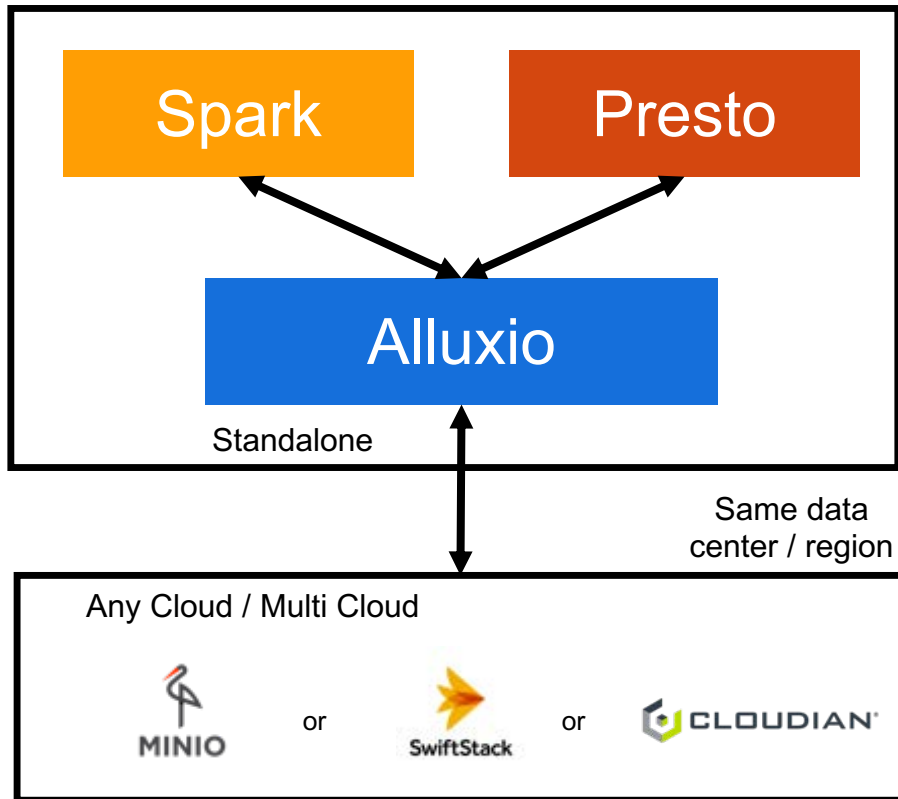
Enable big data on object stores across single or multiple clouds

Orchestrate data frameworks on the public cloud

Alluxio – Key innovations

Data Locality
with Intelligent Multi-tiering

Accelerate big data workloads with transparent tiered local data

Data Accessibility
for popular APIs & API translation

Run Spark, Hive, Presto, ML workloads on your data located anywhere

Data Elasticity
with a unified namespace

Abstract data silos & storage systems to independently scale data on-demand with compute

Data Locality with Intelligent Multi-tiering

Local performance from remote data using multi-tier storage
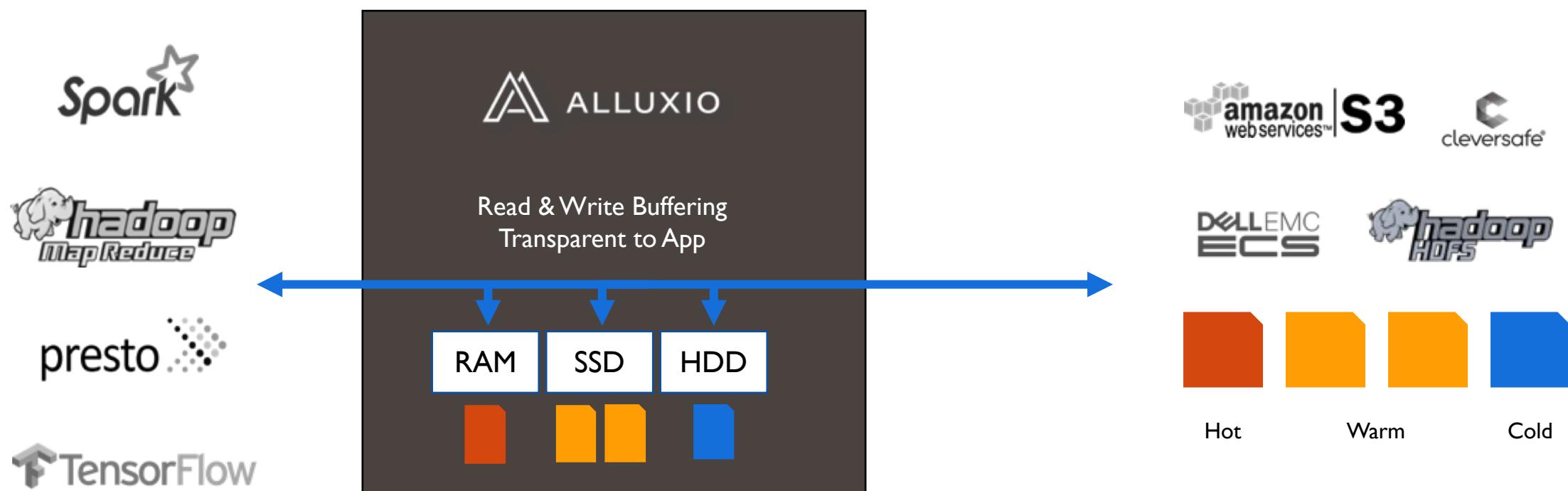
Data Accessibility via popular APIs and API Translation

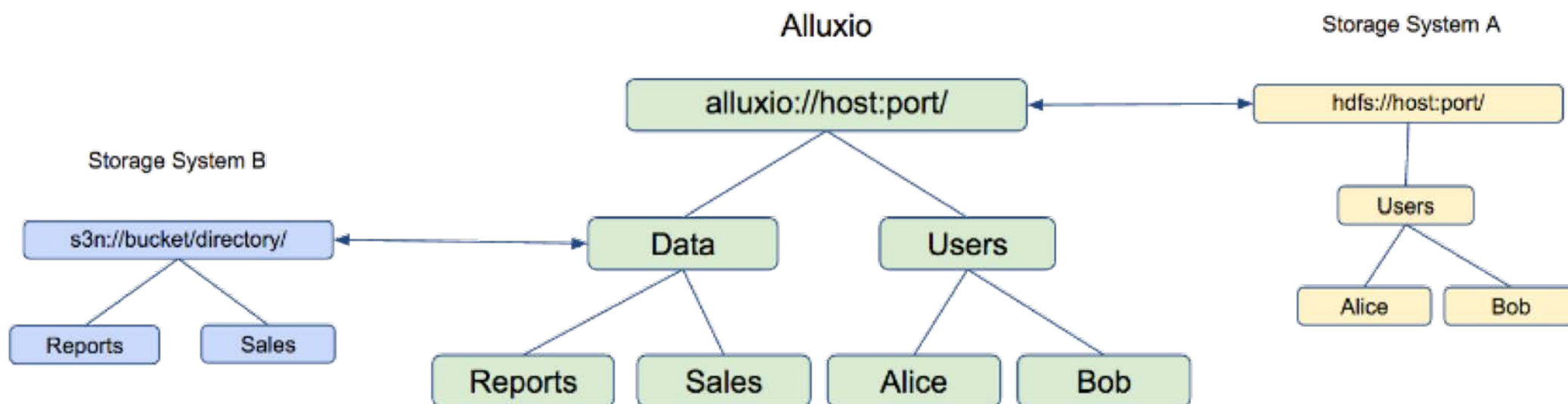**Convert from Client-side Interface to native Storage Interface**
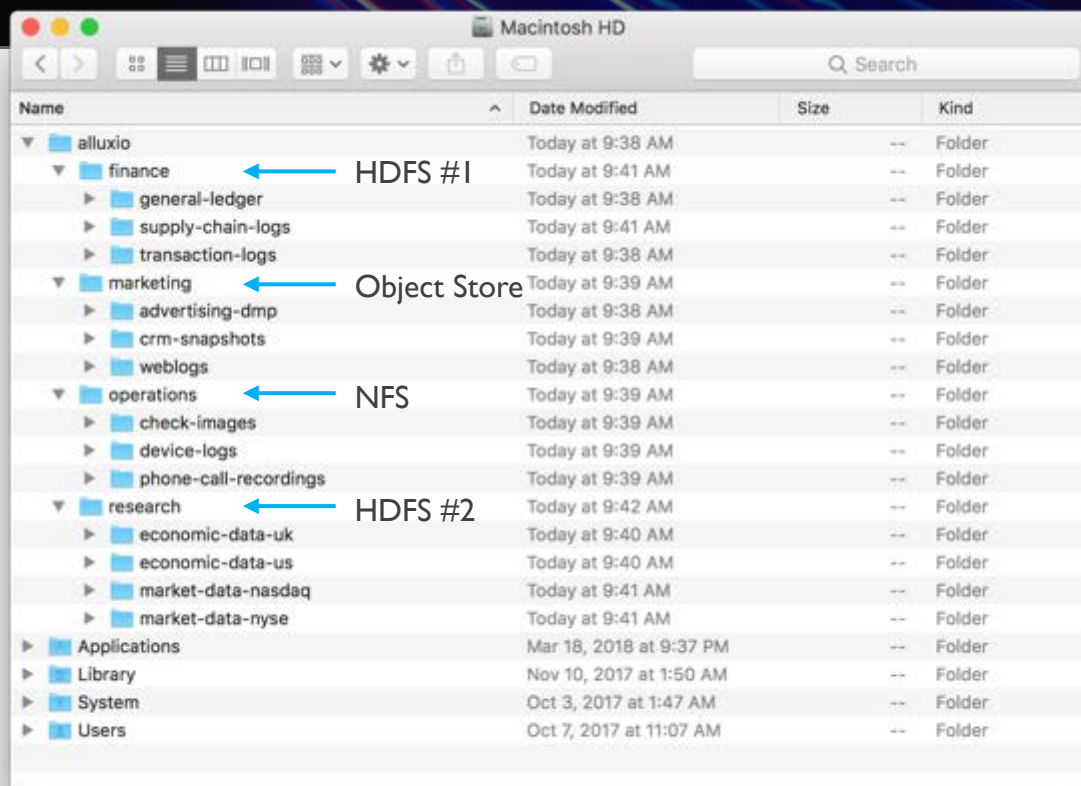
Data Elasticity via Unified Namespace

Enables effective data management across different Under Store

- **Uses Mounting with Transparent Naming**

Unified Namespace: Global Data Accessibility

Transparent access to understorage makes all enterprise data available locally

| Name		Date Modified	Size	Kind
▼ 📁 alluxio		Today at 9:38 AM	--	Folder
▼ 📁 finance	← HDFS #1	Today at 9:41 AM	--	Folder
▶ 📁 general-ledger		Today at 9:38 AM	--	Folder
▶ 📁 supply-chain-logs		Today at 9:41 AM	--	Folder
▶ 📁 transaction-logs		Today at 9:38 AM	--	Folder
▼ 📁 marketing	← Object Store	Today at 9:39 AM	--	Folder
▶ 📁 advertising-dmp		Today at 9:38 AM	--	Folder
▶ 📁 crm-snapshots		Today at 9:39 AM	--	Folder
▶ 📁 weblogs		Today at 9:38 AM	--	Folder
▼ 📁 operations	← NFS	Today at 9:39 AM	--	Folder
▶ 📁 check-images		Today at 9:39 AM	--	Folder
▶ 📁 device-logs		Today at 9:39 AM	--	Folder
▶ 📁 phone-call-recordings		Today at 9:39 AM	--	Folder
▼ 📁 research	← HDFS #2	Today at 9:42 AM	--	Folder
▶ 📁 economic-data-uk		Today at 9:40 AM	--	Folder
▶ 📁 economic-data-us		Today at 9:40 AM	--	Folder
▶ 📁 market-data-nasdaq		Today at 9:41 AM	--	Folder
▶ 📁 market-data-nyse		Today at 9:41 AM	--	Folder
▶ 📁 Applications		Mar 18, 2018 at 9:37 PM	--	Folder
▶ 📁 Library		Nov 10, 2017 at 1:50 AM	--	Folder
▶ 📁 System		Oct 3, 2017 at 1:47 AM	--	Folder
▶ 📁 Users		Oct 7, 2017 at 11:07 AM	--	Folder

SUPPORTS

- HDFS
- NFS
- OpenStack
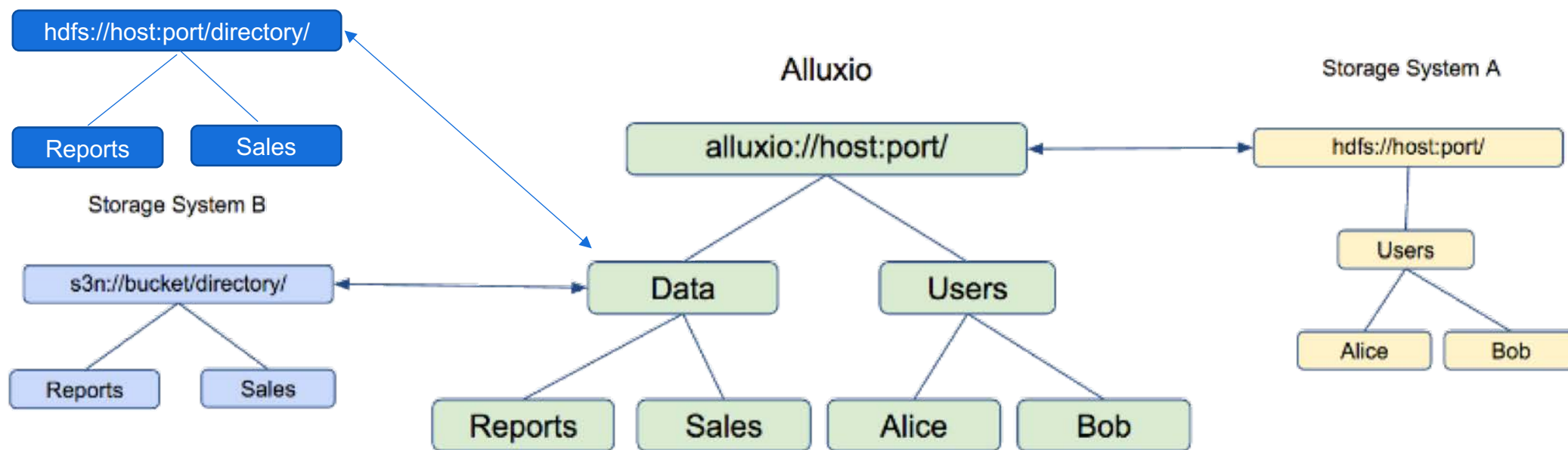- Ceph
- Amazon S3
- Azure
- Google Cloud

IT OPS FRIENDLY

- Storage mounted into Alluxio by central IT
- Security in Alluxio mirrors source data
- Authentication through LDAP/AD
- Wireline encryption

ALLUXIO

Alluxio Reference Architecture
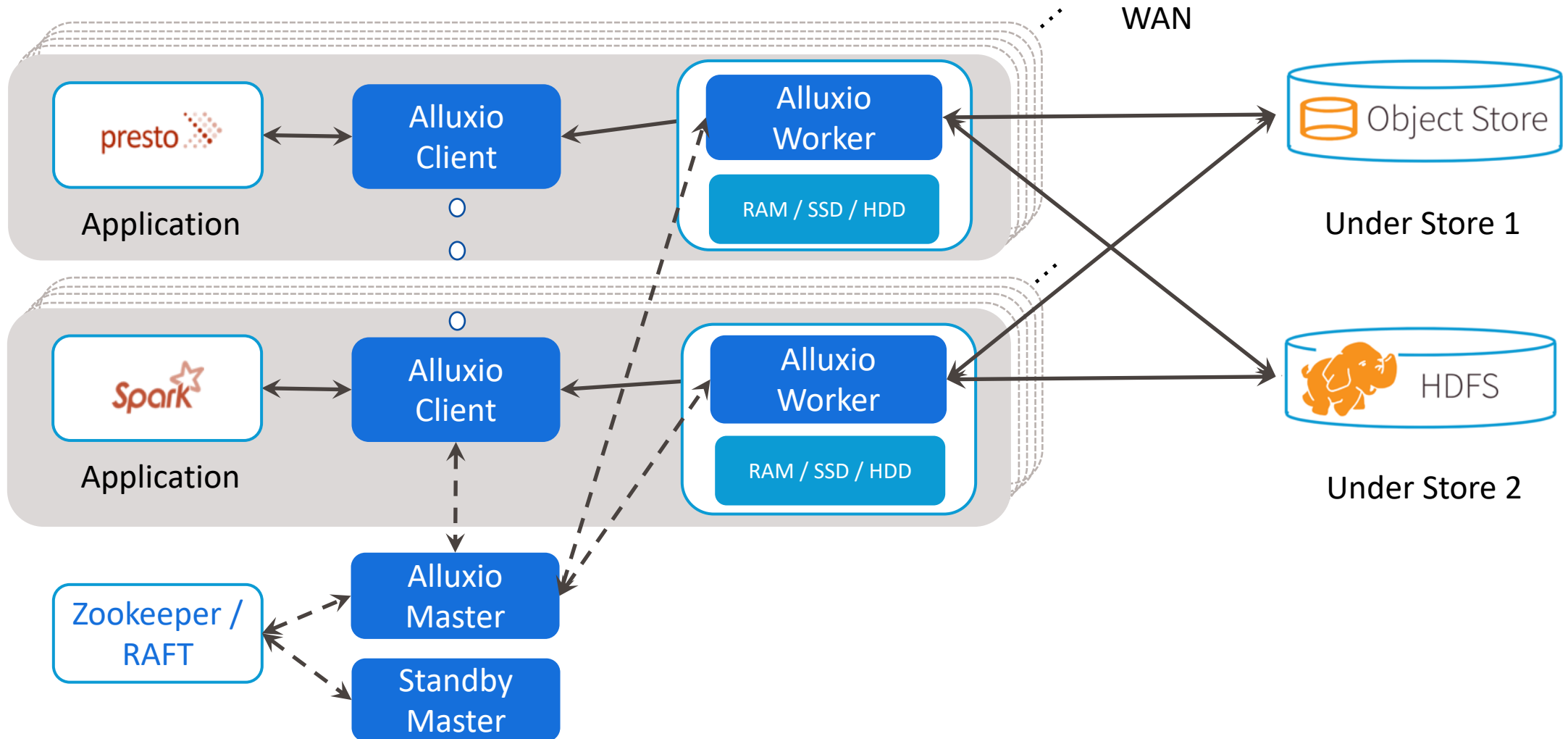
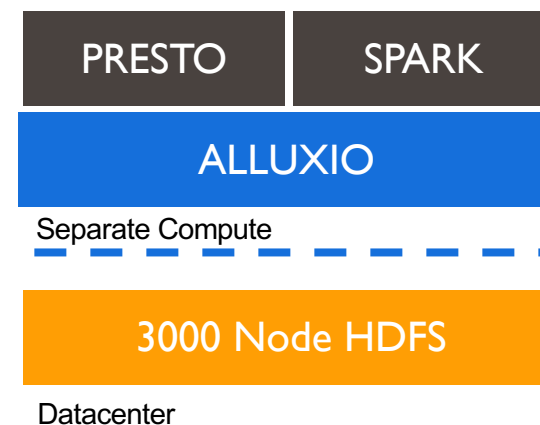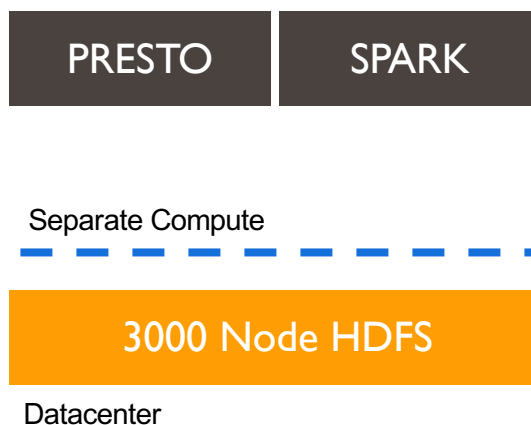# Policy Driven under File System Migration

Research Directions

- ▲ Machine-learning based Data Orchestration Policies
- ▲ Scalable and High-performance File System Metadata service
- ▲ Optimization for in-memory data partition / format
- ▲ Cross-layer optimization for distributed compute and storage systems

$70B e-commerce retailer

PRESTO	SPARK

Separate Compute

3000 Node HDFS

Datacenter

➡️

PRESTO	SPARK
ALLUXIO	

Separate Compute

3000 Node HDFS

Datacenter

Project:
- Offload HDFS with separate clusters of Presto and Spark

Problem:
- HDFS cluster is compute and network bound
- Performance is inconsistent

Alluxio solution:
- Alluxio offloads the network I/O as well as the compute

Result:
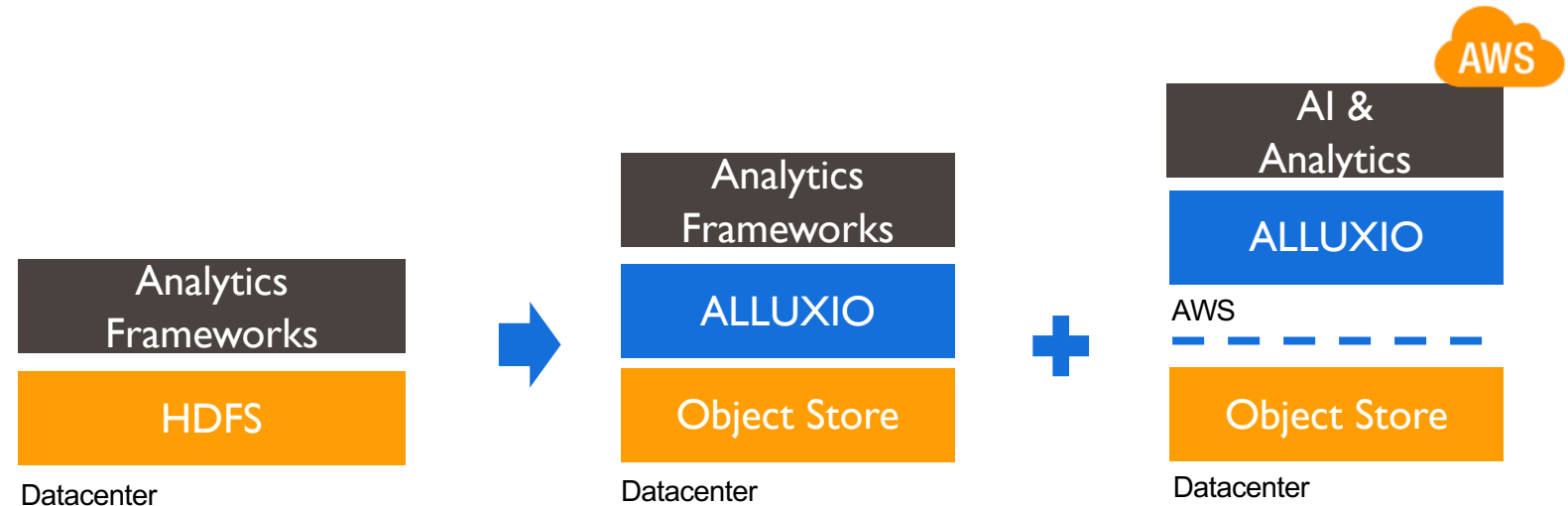- Teams can run additional workloads without taxing the existing HDFS cluster

ALLUXIO

DBS Bank | Performance & Hybrid

Largest bank in Southeast Asia

Initial Project:
- Digital Bank Initiative
- Solve scaling challenges by separating compute and using object storage

Problem:
- Coupled systems were not flexible to scale
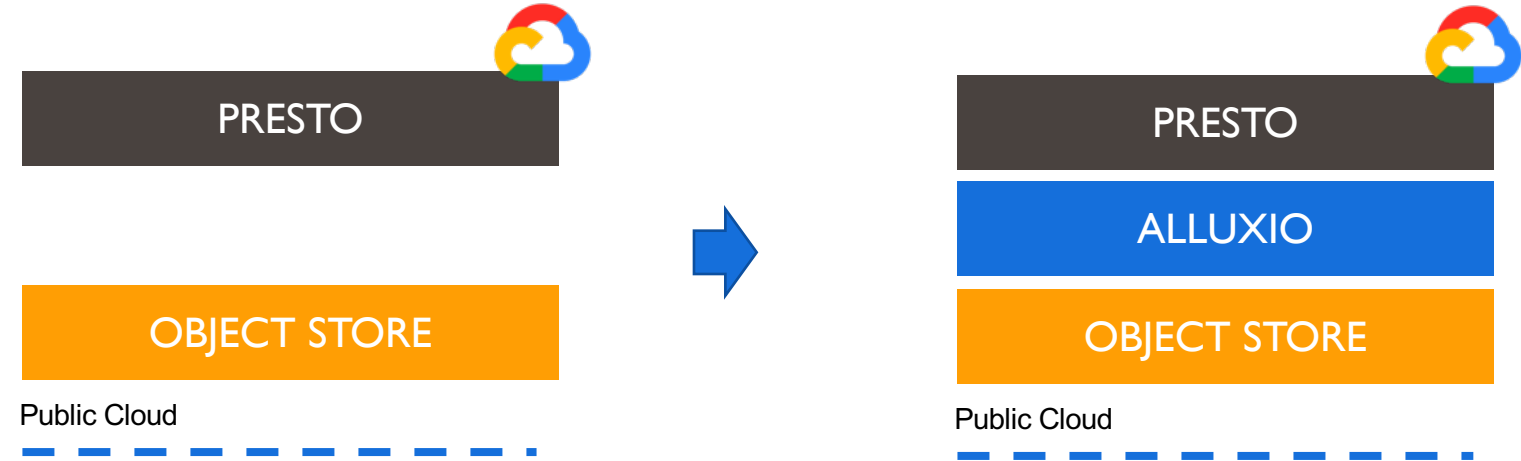
Alluxio solution:
1. Alluxio provides intelligent caching layer for object storage
2. Burst workloads to hybrid cloud

Result:
- Enables data on-demand, Alluxio now considered mature layer in stack

ALLUXIO

Walmart | Performance Use Case in Cloud

PRESTO

OBJECT STORE

Public Cloud

→

PRESTO
ALLUXIO
OBJECT STORE

Public Cloud

Project:
- Utilize Presto for interactive queries on cloud object store compute

Problem:
- Low performance of queries too slow to be usable
- Inconsistent performance of queries

Alluxio solution:
- Alluxio provides intelligent distributed caching layer for object storage

Result:
- High performance queries
- Consistent performance
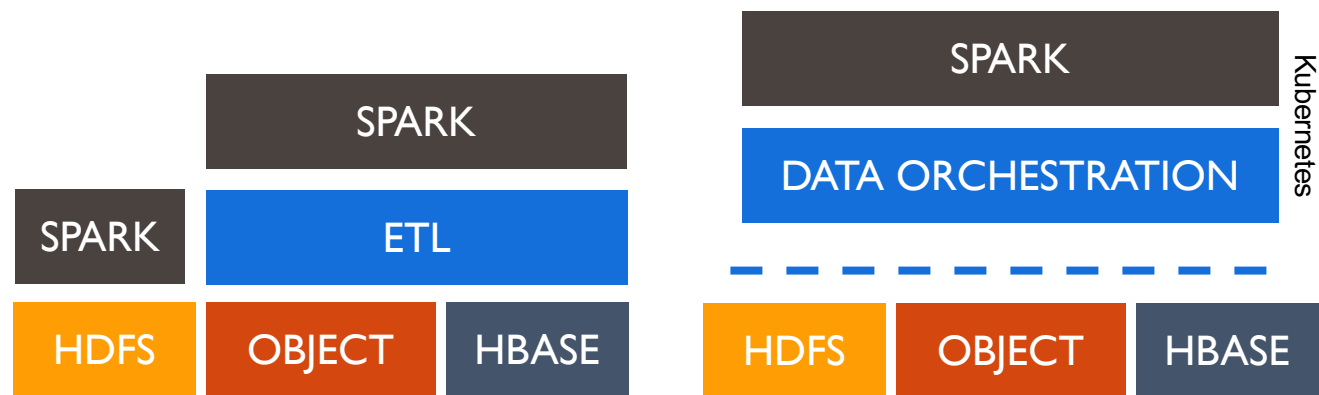- Interactive query performance for analysts

ALLUXIO

China Unicom
Leading Chinese Telco serving 320 million subscribers

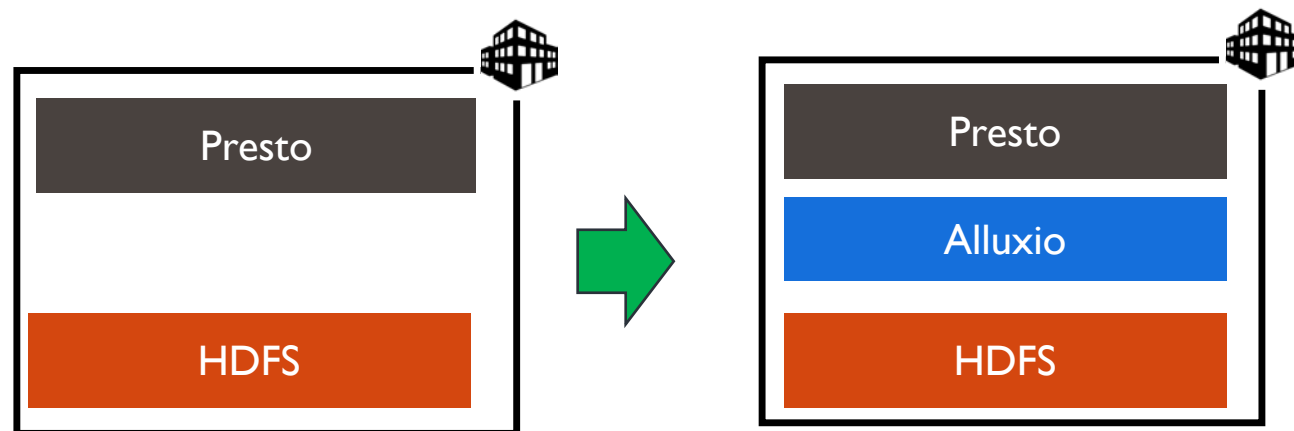## Use case | Data orchestration for agility

- Single namespace to access & address all data
- Data local to compute accelerates workloads

ALLUXIO

NetEase Games
Leading Online Game Company in China

Use Case | On-premise Caching for Presto

- Large query variance during peak hours before
- Alluxio brings data local to Presto to reduce the latency during peak hours

https://www.alluxio.io/blog/presto-on-alluxio-how-netease-games-leveraged-alluxio-to-boost-ad-hoc-sql-on-hdfs/

Next steps - Try it out!

- [Getting Started](#)
- [Try 10 Minutes Alluxio & Presto Tutorial on Laptop](#)
- [Try 10 Minutes Alluxio & Presto Tutorial on AWS](#)
- [Spark and Alluxio in 5 minutes](#)

Questions or Suggestions? [Engage with our Community in Slack!](#)

Questions?

Welcome to join the Alluxio Open Source Community!

www.alluxio.io | slackin.alluxio.io | @alluxio