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

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

Open Source project established & company to commercialize Alluxio founded

Goal: **Orchestrate Data at Memory Speed for the Cloud** for data driven apps such as Big Data Analytics, ML and AI.
Early Days Contributors Growth

- **v0.1**: 1 contributor (Dec '12)
- **v0.2**: 3 contributors (Apr '13)
- **v0.3**: 15 contributors (Oct '13)
- **v0.4**: 30 contributors (Feb '14)
- **v0.5**: 46 contributors (Jul '14)
- **v0.6**: 70 contributors (Mar '15)
- **v0.7**: 100+ contributors (Jul '15)
Companies Running Alluxio (Learn More)

<table>
<thead>
<tr>
<th>Financial Services</th>
<th>Retail &amp; Entertainment</th>
<th>Data &amp; Analytics Services</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBS</td>
<td>Walmart</td>
<td>Aunalytics</td>
<td>Tencent</td>
</tr>
<tr>
<td>ING</td>
<td>Roblox</td>
<td>LigaData</td>
<td>esri</td>
</tr>
<tr>
<td>Wells Fargo</td>
<td>NetEase Games</td>
<td>Samsung SDS</td>
<td>Perceptin</td>
</tr>
<tr>
<td>Barclays</td>
<td>Caesars Entertainment</td>
<td>TalkingData</td>
<td>Tencent Cloud</td>
</tr>
<tr>
<td>PayPal</td>
<td>Mynta</td>
<td>TalkingData</td>
<td>Baidu</td>
</tr>
<tr>
<td>Huatai Securities</td>
<td>JD.com</td>
<td>TalkingData</td>
<td>Alibaba Cloud</td>
</tr>
<tr>
<td>Suning.com</td>
<td>Shopee</td>
<td>RYTE</td>
<td>Sogou</td>
</tr>
<tr>
<td></td>
<td>VIP Shopaholic</td>
<td></td>
<td>Baidu Cloud</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lucidworks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CUELOGIC</td>
</tr>
</tbody>
</table>
|                             |                        |                           | Travel & Transportation
|                             |                        |                           | Travel & Transportation
|                             |                        |                           | Waze                |
|                             |                        |                           | Ctrip               |
|                             |                        |                           | DiDi                |
|                             |                        |                           | MOMO                |
|                             |                        |                           | Guaner.Com          |

Consumer

<table>
<thead>
<tr>
<th>Telco &amp; Media</th>
<th>Travel &amp; Transportation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samsung</td>
<td>Waze</td>
</tr>
<tr>
<td>Nielsen</td>
<td>Ctrip</td>
</tr>
<tr>
<td>Yahoo!</td>
<td>DiDi</td>
</tr>
<tr>
<td>Baidu</td>
<td>Momo</td>
</tr>
<tr>
<td>Lenovo</td>
<td>walkme</td>
</tr>
<tr>
<td>Chainhome</td>
<td>去哪儿</td>
</tr>
</tbody>
</table>

Data & Analytics Services

<table>
<thead>
<tr>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tencent</td>
</tr>
<tr>
<td>esri</td>
</tr>
<tr>
<td>Tencent Cloud</td>
</tr>
<tr>
<td>Baidu Cloud</td>
</tr>
<tr>
<td>Alibaba Cloud</td>
</tr>
<tr>
<td>Sogou</td>
</tr>
<tr>
<td>Lucidworks</td>
</tr>
<tr>
<td>CUELOGIC</td>
</tr>
</tbody>
</table>
| Travel & Transportation
| Guaner.Com |

Retail & Entertainment

<table>
<thead>
<tr>
<th>Data &amp; Analytics Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>TalkingData</td>
</tr>
<tr>
<td>RYTE</td>
</tr>
<tr>
<td>Didi</td>
</tr>
</tbody>
</table>

Financial Services

<table>
<thead>
<tr>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tencent</td>
</tr>
<tr>
<td>Tencent Cloud</td>
</tr>
<tr>
<td>Baidu Cloud</td>
</tr>
<tr>
<td>Alibaba Cloud</td>
</tr>
<tr>
<td>Sogou</td>
</tr>
<tr>
<td>Lucidworks</td>
</tr>
<tr>
<td>CUELOGIC</td>
</tr>
</tbody>
</table>

Retail & Entertainment

<table>
<thead>
<tr>
<th>Data &amp; Analytics Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>TalkingData</td>
</tr>
<tr>
<td>RYTE</td>
</tr>
<tr>
<td>Didi</td>
</tr>
</tbody>
</table>

Consumer

<table>
<thead>
<tr>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tencent</td>
</tr>
<tr>
<td>Tencent Cloud</td>
</tr>
<tr>
<td>Baidu Cloud</td>
</tr>
<tr>
<td>Alibaba Cloud</td>
</tr>
<tr>
<td>Sogou</td>
</tr>
<tr>
<td>Lucidworks</td>
</tr>
<tr>
<td>CUELOGIC</td>
</tr>
</tbody>
</table>

Travel & Transportation

<table>
<thead>
<tr>
<th>Data &amp; Analytics Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>TalkingData</td>
</tr>
<tr>
<td>RYTE</td>
</tr>
<tr>
<td>Didi</td>
</tr>
</tbody>
</table>

Financial Services

<table>
<thead>
<tr>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tencent</td>
</tr>
<tr>
<td>Tencent Cloud</td>
</tr>
<tr>
<td>Baidu Cloud</td>
</tr>
<tr>
<td>Alibaba Cloud</td>
</tr>
<tr>
<td>Sogou</td>
</tr>
<tr>
<td>Lucidworks</td>
</tr>
<tr>
<td>CUELOGIC</td>
</tr>
</tbody>
</table>

Telco & Media

<table>
<thead>
<tr>
<th>Data &amp; Analytics Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>TalkingData</td>
</tr>
<tr>
<td>RYTE</td>
</tr>
<tr>
<td>Didi</td>
</tr>
</tbody>
</table>

Retail & Entertainment

<table>
<thead>
<tr>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tencent</td>
</tr>
<tr>
<td>Tencent Cloud</td>
</tr>
<tr>
<td>Baidu Cloud</td>
</tr>
<tr>
<td>Alibaba Cloud</td>
</tr>
<tr>
<td>Sogou</td>
</tr>
<tr>
<td>Lucidworks</td>
</tr>
<tr>
<td>CUELOGIC</td>
</tr>
</tbody>
</table>

Data & Analytics Services

<table>
<thead>
<tr>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tencent</td>
</tr>
<tr>
<td>Tencent Cloud</td>
</tr>
<tr>
<td>Baidu Cloud</td>
</tr>
<tr>
<td>Alibaba Cloud</td>
</tr>
<tr>
<td>Sogou</td>
</tr>
<tr>
<td>Lucidworks</td>
</tr>
<tr>
<td>CUELOGIC</td>
</tr>
</tbody>
</table>
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
Data Ecosystem 1.0 – The Challenges

- Complex
- Low performance
- Expensive

STORAGE

COMPUTE

- presto
- Spark
- Flink
- Hadoop
  Map Reduce
- kafa
- Caffe
- HBase
- TensorFlow

STORAGE

COMPUTE

STORAGE

COMPUTE
Co-located compute & HDFS on the same cluster

Disaggregated compute & HDFS on the same cluster

Support more frameworks
- Support Presto, Spark across DCs without app changes
- HDFS for Hybrid Cloud
  - 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
Independent scaling of compute & storage

Java File API  HDFS Interface  S3 Interface  POSIX Interface  REST API

HDFS Driver  Swift Driver  S3 Driver  NFS Driver

presto  Spark  Flink  Hadoop MapReduce  Hive  HBase  Caffe  TensorFlow

ALLUXIO  Data Orchestration for the Cloud

Amazon S3  Microsoft Azure  Ceph  MINIO  HDFS  IBM NetApp  SwiftStack  Dell EMC  ECS  Hitachi Data Systems  Cloudian
APIs to Interact with data in Alluxio

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

**Spark**

```scala
> 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**

```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

Data center B

- Hive
- MapReduce
Challenges with running workloads on cloud storage

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

Spark

Alluxio

Same instance / container

or
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

Solution Benefits
- Same performance as local
- Same end-user experience
- 100% of I/O is offloaded

Burst big data workloads in hybrid cloud environments
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

Solution Benefits

- Same performance as HDFS
- Uses HDFS APIs
- Same end-user experience
- Storage at fraction of the cost of HDFS
**Use Cases Alluxio Enables**

**Accelerate big data frameworks on the public cloud**

- Spark
- Alluxio
- Same instance/container

**Burst big data workloads in hybrid cloud environments**

- Hive
- Alluxio
- Same instance/container

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

- Presto
- Alluxio
- Same container/machine

- S3
- HDFS
- or
- ceph
- or
- cleversafe
- or
- garbage can
Enable big data on object stores across single or multiple clouds

Orchestrate data frameworks on the public cloud
Data Elasticity
with a unified namespace

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

Data Accessibility
for popular APIs & API translation

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

Data Locality
with Intelligent Multi-tiering

Accelerate big data workloads with transparent tiered local data
Data Locality with Intelligent Multi-tiering

Local performance from remote data using multi-tier storage

- **RAM**
- **SSD**
- **HDD**

Read & Write Buffering
Transparent to App

Hot
Warm
Cold
Data Accessibility via popular APIs and API Translation

Convert from Client-side Interface to native Storage Interface

Java File API       HDFS Interface       S3 Interface       POSIX Interface       REST API

HDFS Driver       S3 Driver       Swift Driver       NFS Driver

Amazon S3       Microsoft Azure       ceph       MINIO       Hadoop       IBM       NetApp       SwiftStack       Dell EMC       ECS       Hitachi Data Systems       Cloudian
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

**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 Reference Architecture
Policy Driven under File System Migration

Storage System B
- hdfs://host:port/directory/
  - Reports
  - Sales

Alluxio
- alluxio://host:port/
  - Data
  - Users
    - Reports
    - Sales
    - Alice
    - Bob

Storage System A
- hdfs://host:port/
  - Users
    - Alice
    - Bob
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
### 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
**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
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
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
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