ALLUXIO

EFFECTIVE SPARK WITH ALLUXIO

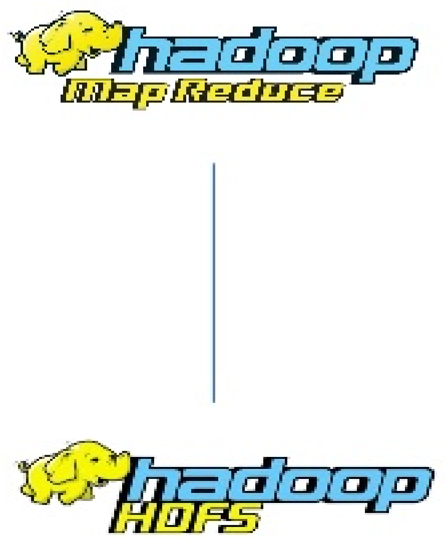
Jiri Simsa, Alluxio, Inc.

Spark Summit EU

OUTLINE

- Alluxio Overview
- Alluxio + Spark Use Cases
- Using Alluxio with Spark
- Performance Evaluation

BIG DATA ECOSYSTEM YESTERDAY

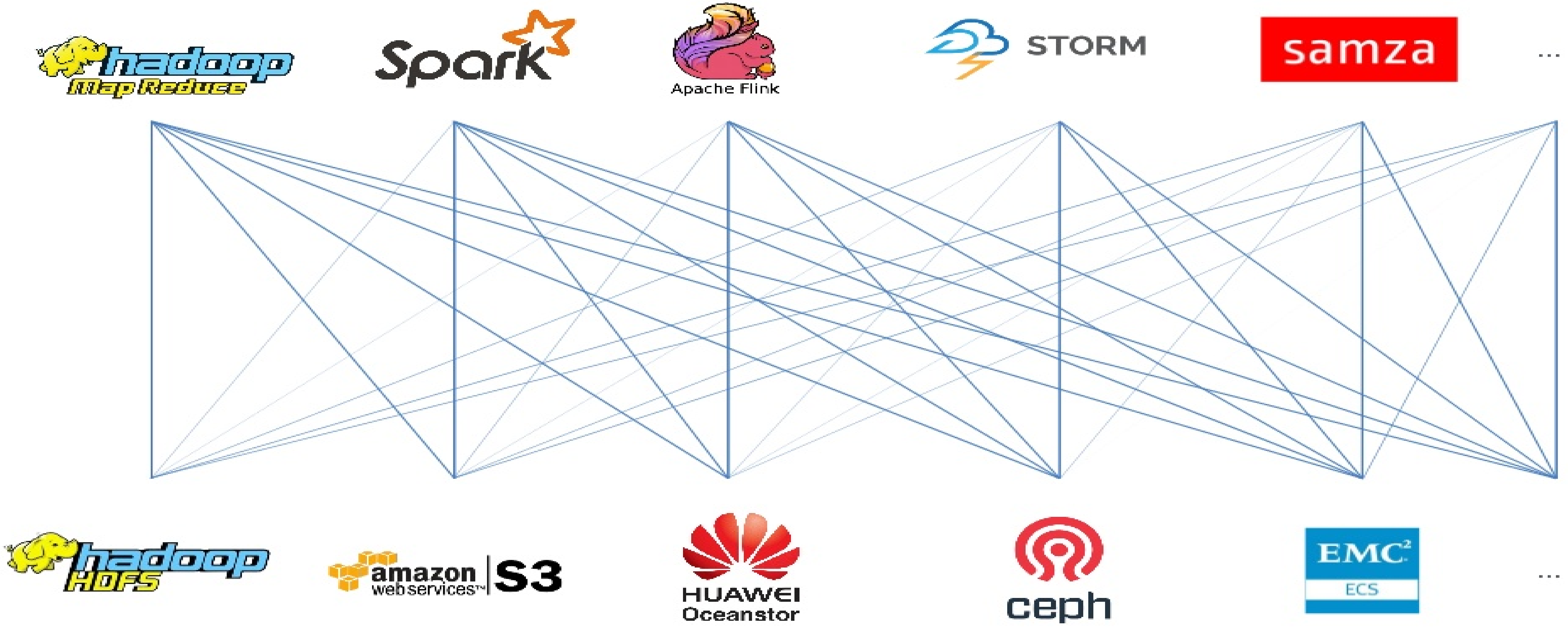


BIG DATA ECOSYSTEM TODAY

hadoop Map Reduce   Spark   Apache Flink   STORM   samza   ...

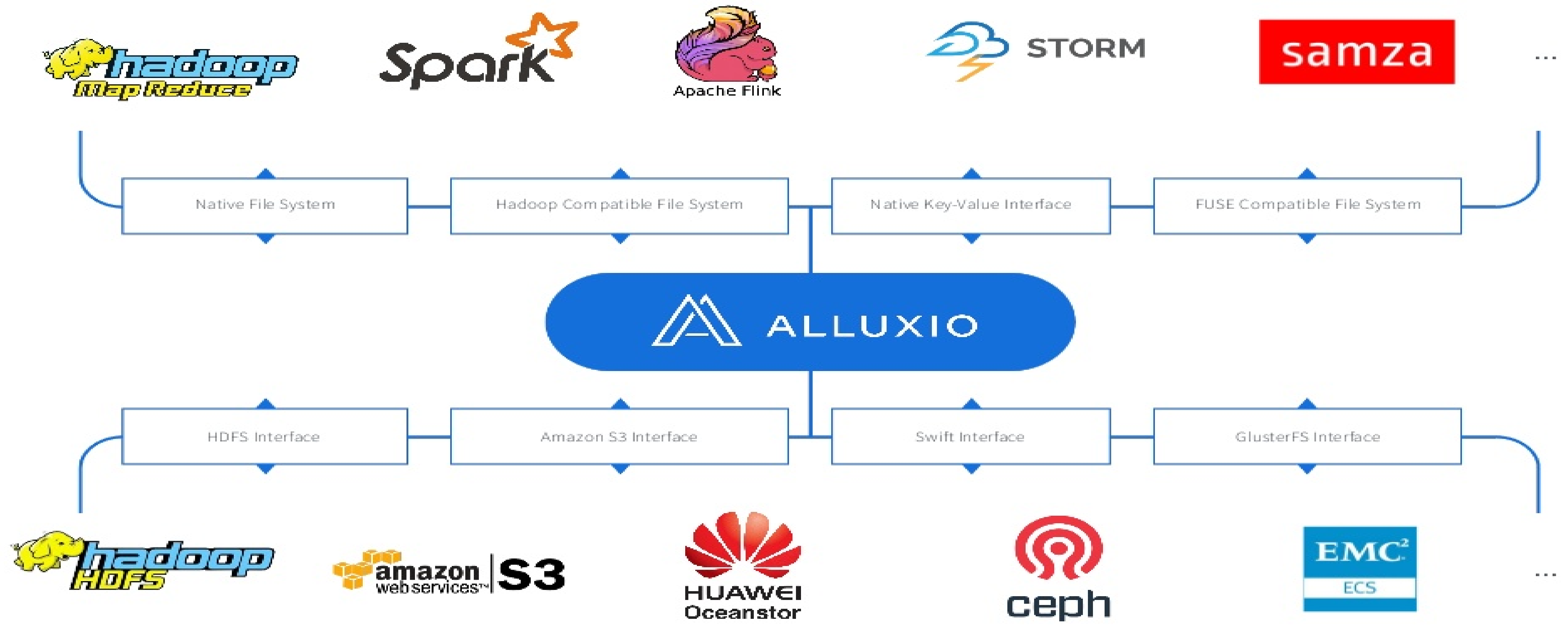
hadoop HDFS   amazon web services™ S3   HUAWEI Oceanstor   ceph   EMC² ECS   ...

BIG DATA ECOSYSTEM ISSUES

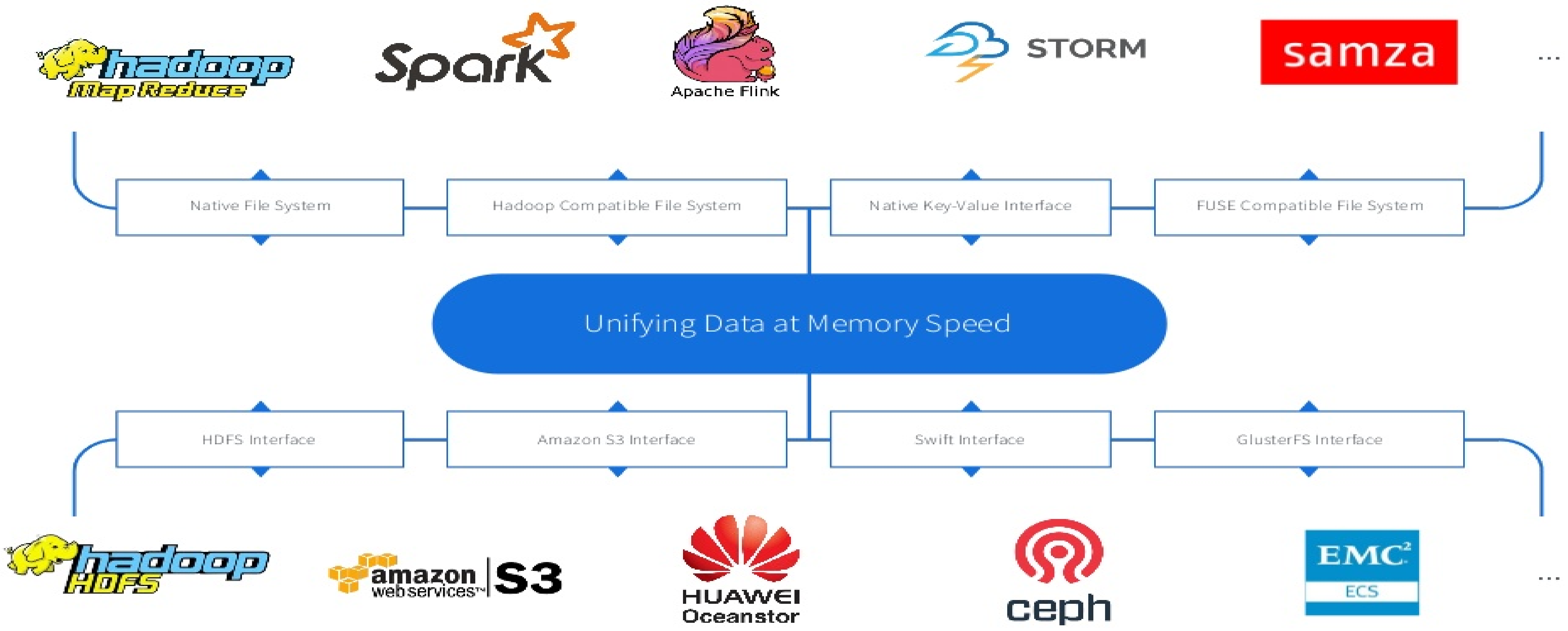


ALLUXIO

BIG DATA ECOSYSTEM WITH ALLUXIO



BIG DATA ECOSYSTEM WITH ALLUXIO



WHY ALLUXIO

- ✓ *Co-located compute and data with memory-speed access to data*
- ✓ *Virtualized across different storage systems under a unified namespace*
- ✓ *Scale-out architecture*
- ✓ *File system API, software only*

BENEFITS

Unification

New workflows across any data in any storage system

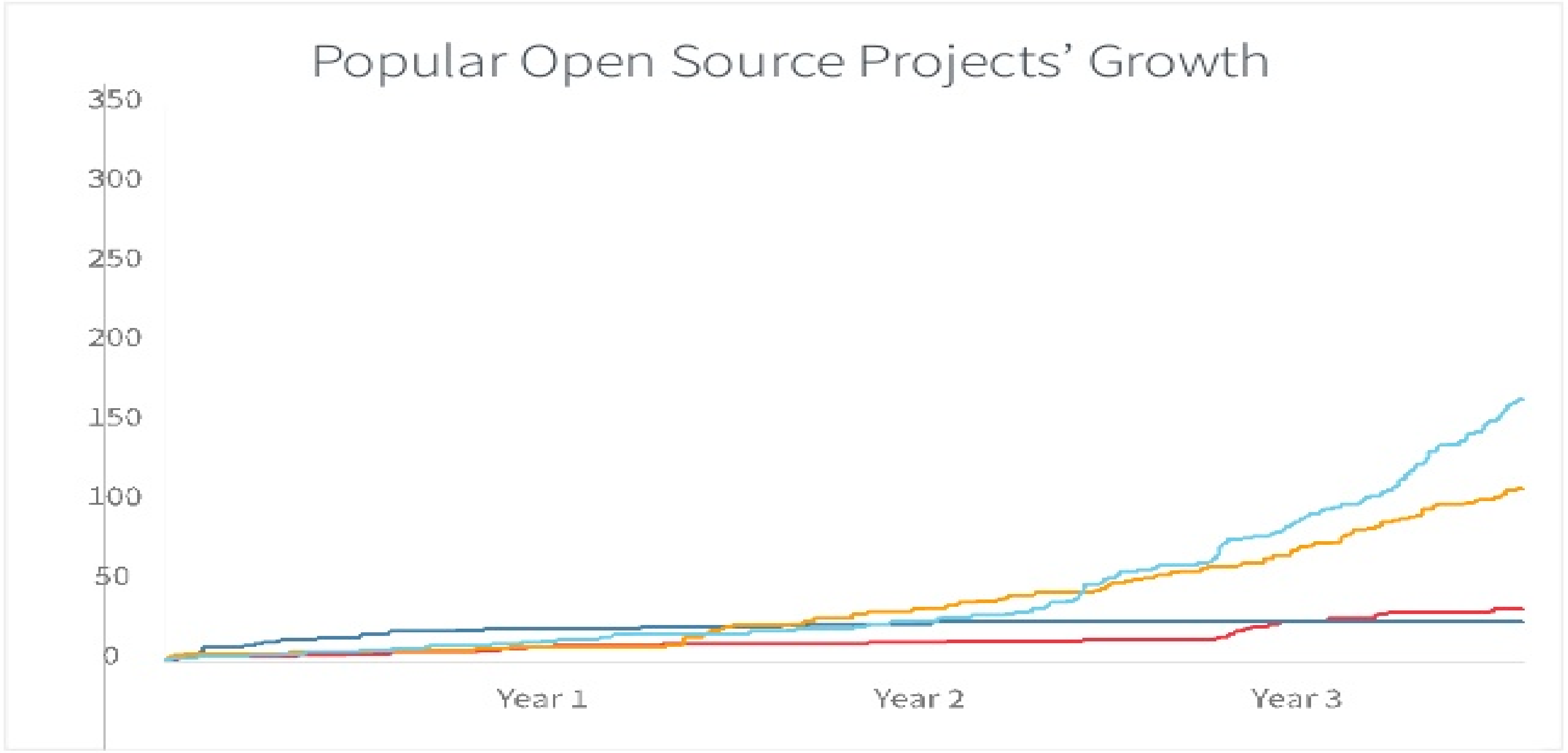
Performance

Orders of magnitude improvement in run time

Flexibility

Choice in compute and storage – grow each independently, buy only what is needed

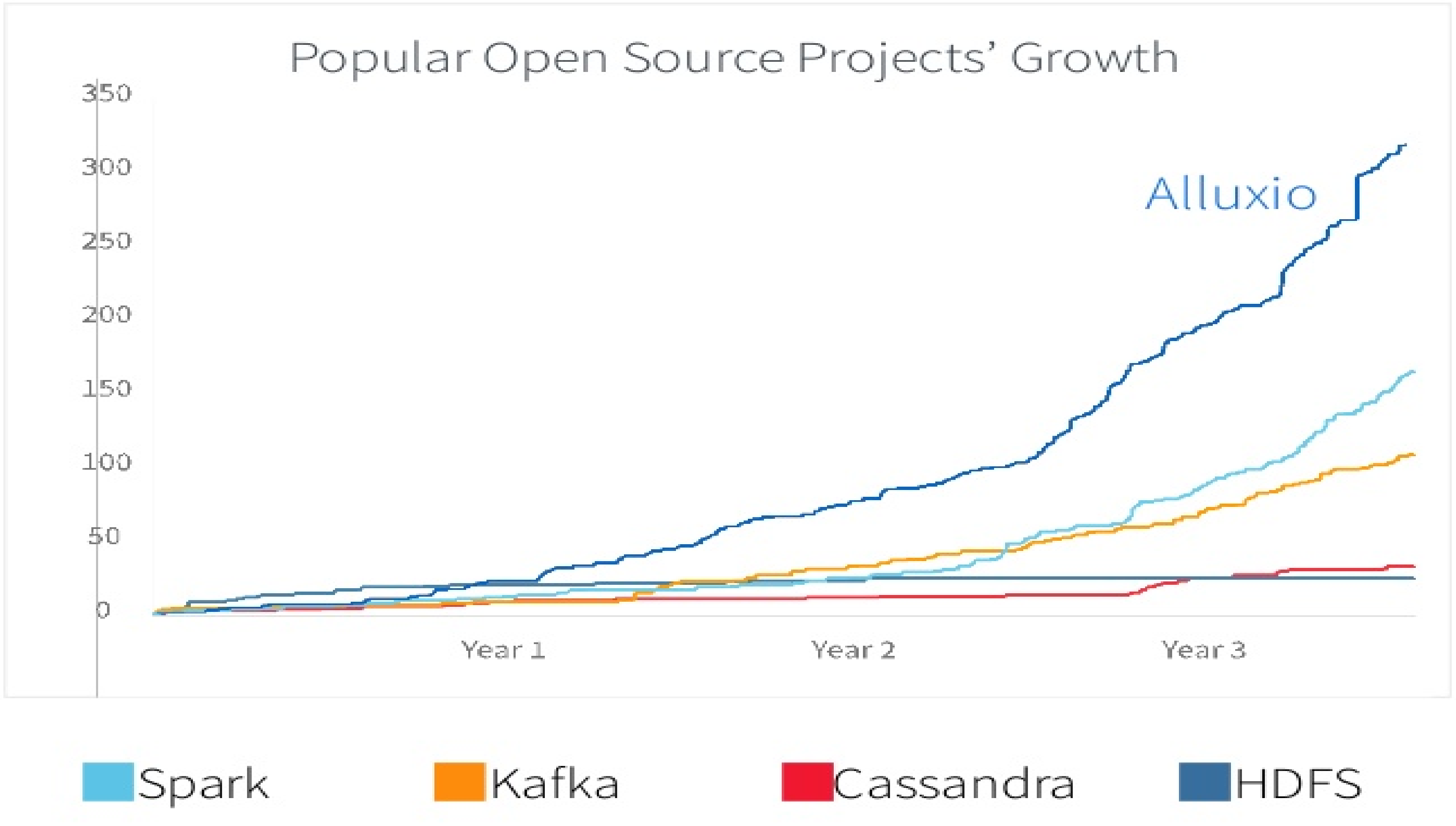
FASTEST GROWING BIG DATA PROJECTS



🟦 Spark 🟧 Kafka 🟥 Cassandra 🟦 HDFS

ALLUXIO

FASTEST GROWING BIG DATA PROJECTS



- Fastest growing open-source project in the big data ecosystem
- 300+ contributors from 100+ organizations
- Running in large production clusters

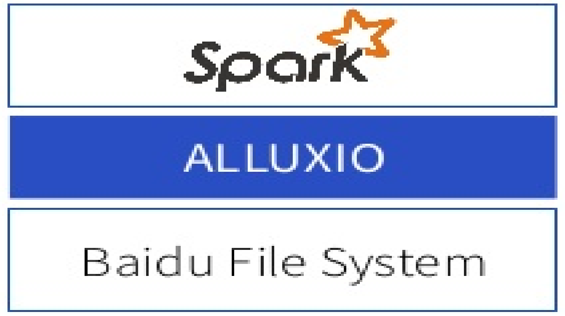
ALLUXIO

OUTLINE

- Alluxio Overview
- Alluxio + Spark Use Cases
- Using Alluxio with Spark
- Performance Evaluation

ACCELERATE I/O TO/FROM REMOTE STORAGE

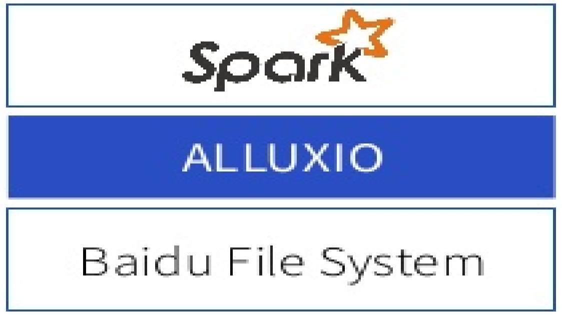
Baidu's PMs and analysts run interactive queries to gain insights into their products and business



- 200+ nodes deployment
- 2+ petabytes of storage
- Mix of memory + HDD

ACCELERATE I/O TO/FROM REMOTE STORAGE

Baidu's PMs and analysts run interactive queries to gain insights into their products and business



- 200+ nodes deployment
- 2+ petabytes of storage
- Mix of memory + HDD

The performance was amazing. With Spark SQL alone, it took 100-150 seconds to finish a query; using Alluxio, where data may hit local or remote Alluxio nodes, it took 10-15 seconds.

- Baidu

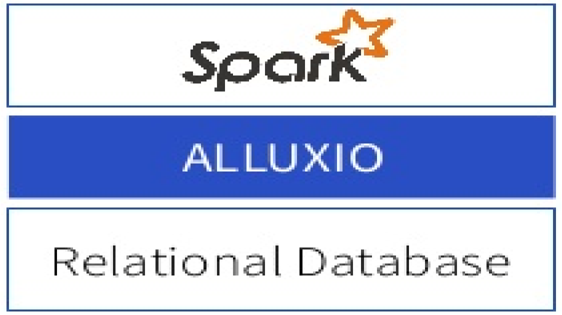
RESULTS

- Data queries are now 30x faster with Alluxio
- Alluxio cluster runs stably, providing over 50TB of RAM space
- By using Alluxio, batch queries usually lasting over 15 minutes were transformed into an interactive query taking less than 30 seconds

SHARE DATA ACROSS JOBS AT MEMORY SPEED

BARCLAYS

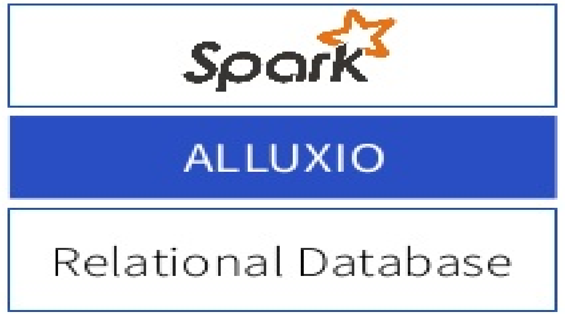
Barclays uses query and machine learning to train models for risk management



- 6 node deployment
- 1TB of storage
- Memory only

SHARE DATA ACROSS JOBS AT MEMORY SPEED

Barclays uses query and machine learning to train models for risk management



- 6 node deployment
- 1TB of storage
- Memory only

BARCLAYS

Thanks to Alluxio, we now have the raw data immediately available at every iteration and we can skip the costs of loading in terms of time waiting, network traffic, and RDBMS activity.

- Barclays

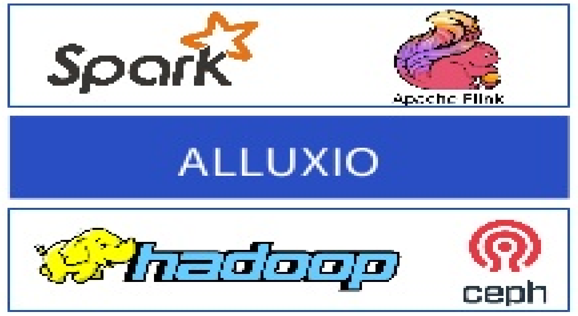
RESULTS

- Barclays workflow iteration time decreased from hours to seconds
- Alluxio enabled workflows that were impossible before
- By keeping data only in memory, the I/O cost of loading and storing in Alluxio is now on the order of seconds

ALLUXIO

MANAGE DATA ACROSS STORAGE SYSTEMS

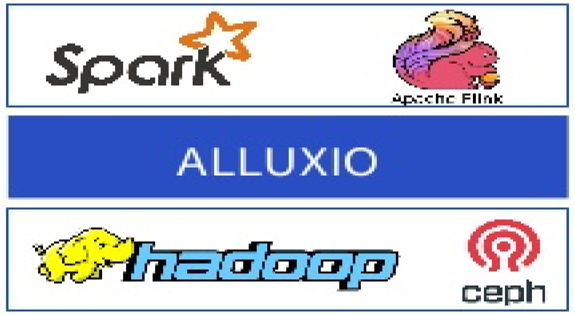
Qunar uses real-time machine learning for their website ads.



- 200+ nodes deployment
- 6 billion logs (4.5 TB) daily
- Mix of Memory + HDD

MANAGE DATA ACROSS STORAGE SYSTEMS

Qunar uses real-time machine learning for their website ads.



- 200+ nodes deployment
- 6 billion logs (4.5 TB) daily
- Mix of Memory + HDD

去哪儿? Qunar.Com 聪 明 你 的 旅 行

We've been running Alluxio in production for over 9 months, Alluxio's unified namespace enable different applications and frameworks to easily interact with data from different storage systems.

- Qunar

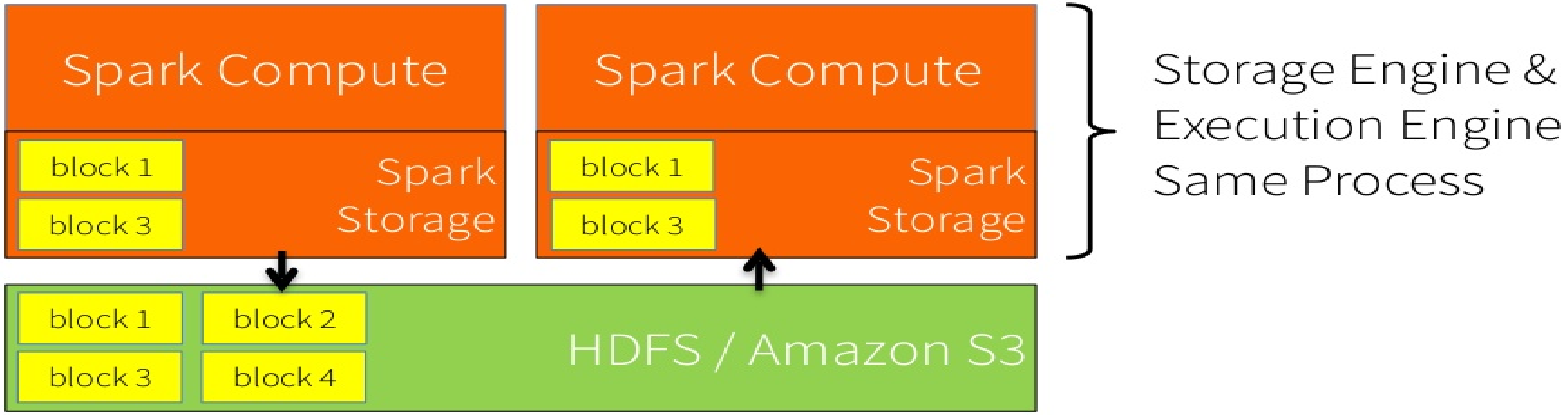
RESULTS

- Data sharing among Spark Streaming, Spark batch and Flink jobs provide efficient data sharing
- Improved the performance of their system with 15x – 300x speedups
- Tiered storage feature manages storage resources including memory and HDD

OUTLINE

- Alluxio Overview
- Alluxio + Spark Use Cases
- Using Alluxio with Spark
- Performance Evaluation

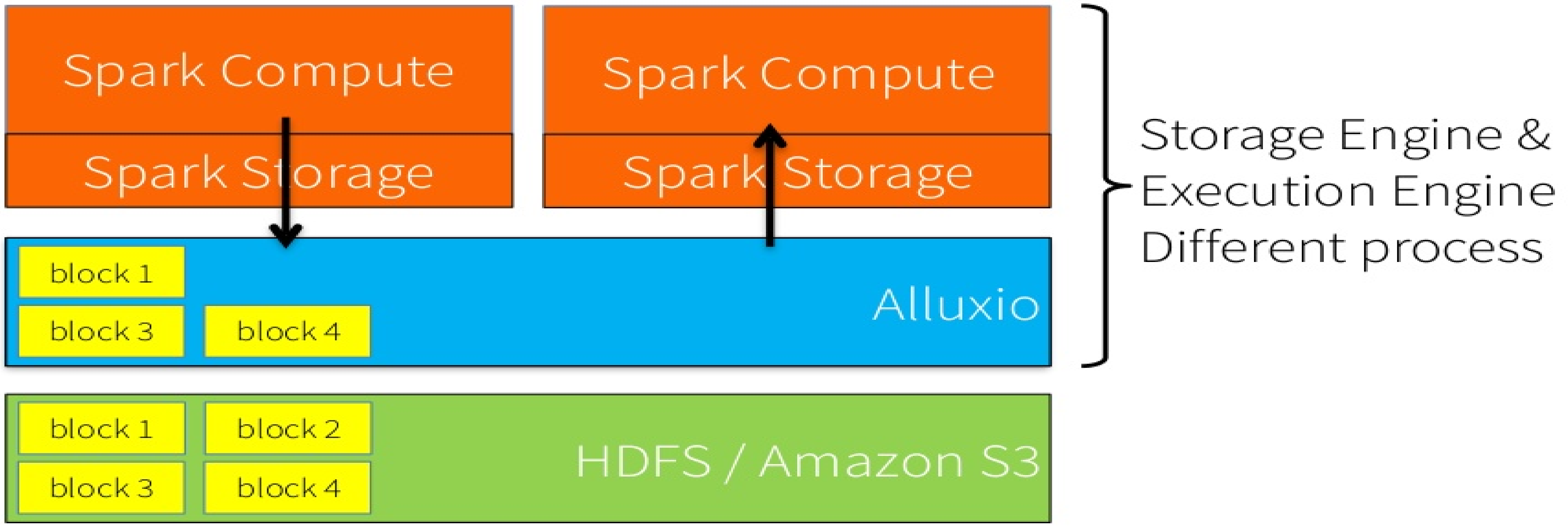
CONSOLIDATING MEMORY



- Two copies of data in memory – double the memory used
- Inter-process Sharing Slowed Down by Network / Disk I/O

ALLUXIO

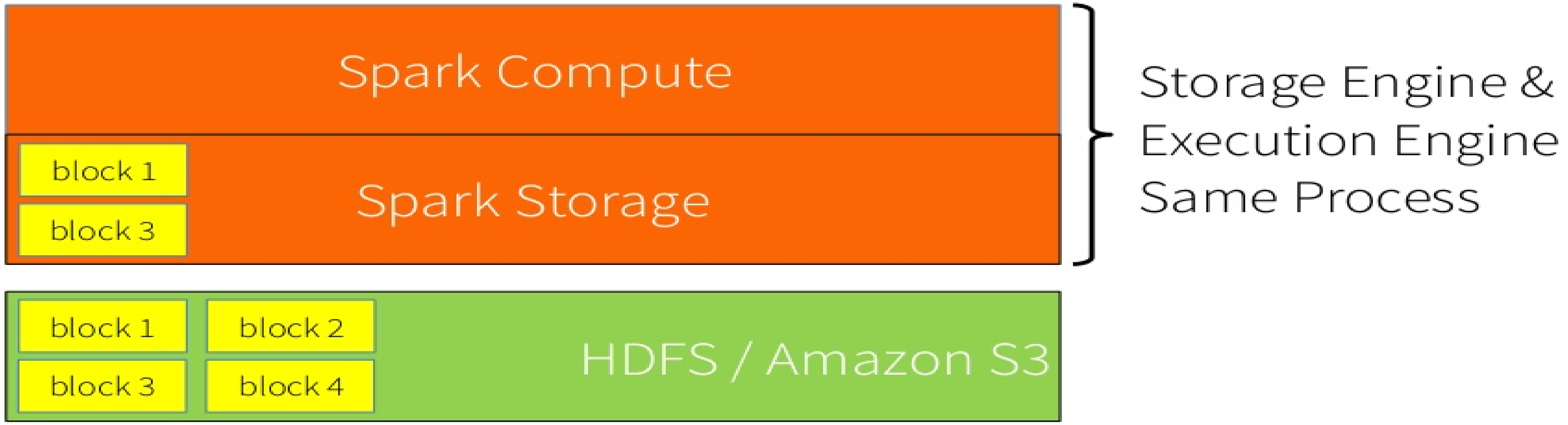
CONSOLIDATING MEMORY



- Half the memory used
- Inter-process Sharing Happens at Memory Speed

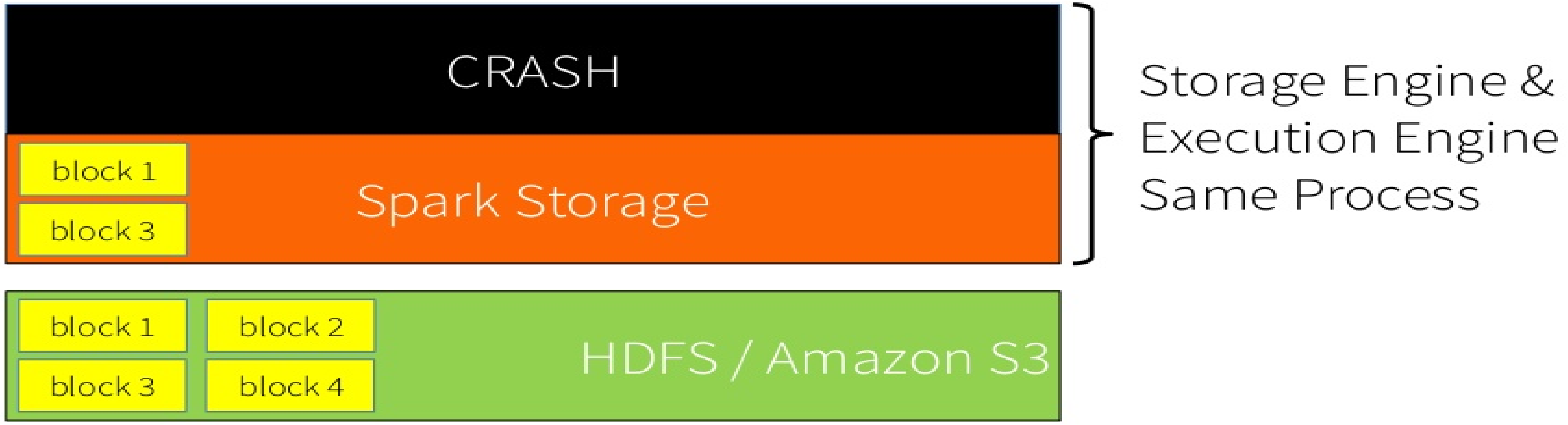
ALLUXIO

DATA RESILIENCE DURING CRASH



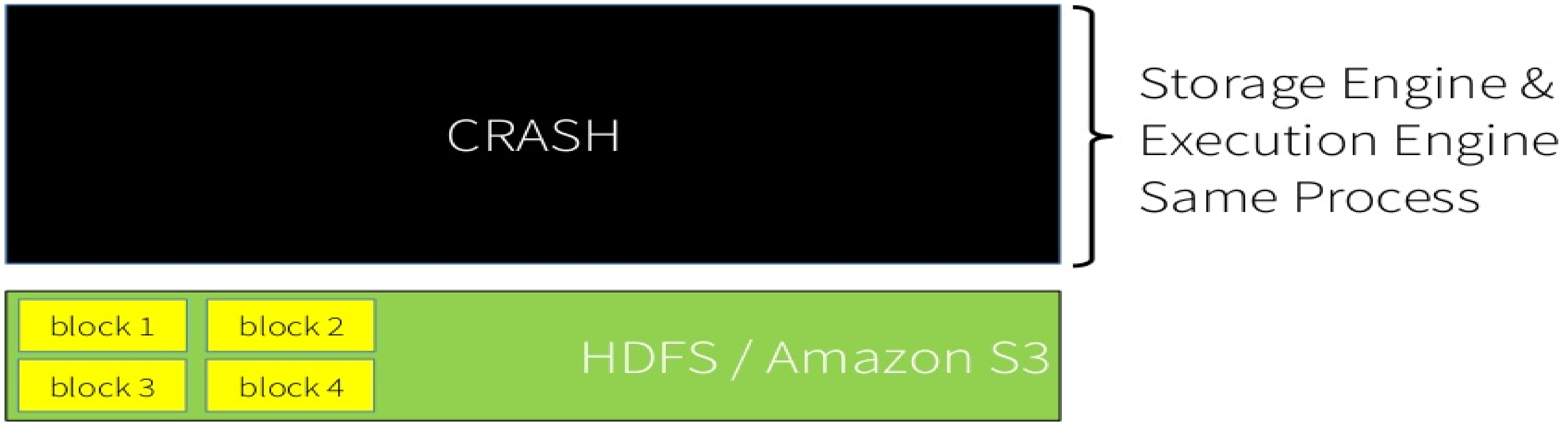
ALLUXIO

DATA RESILIENCE DURING CRASH



- Process Crash Requires Network and/or Disk I/O to Re-read Data

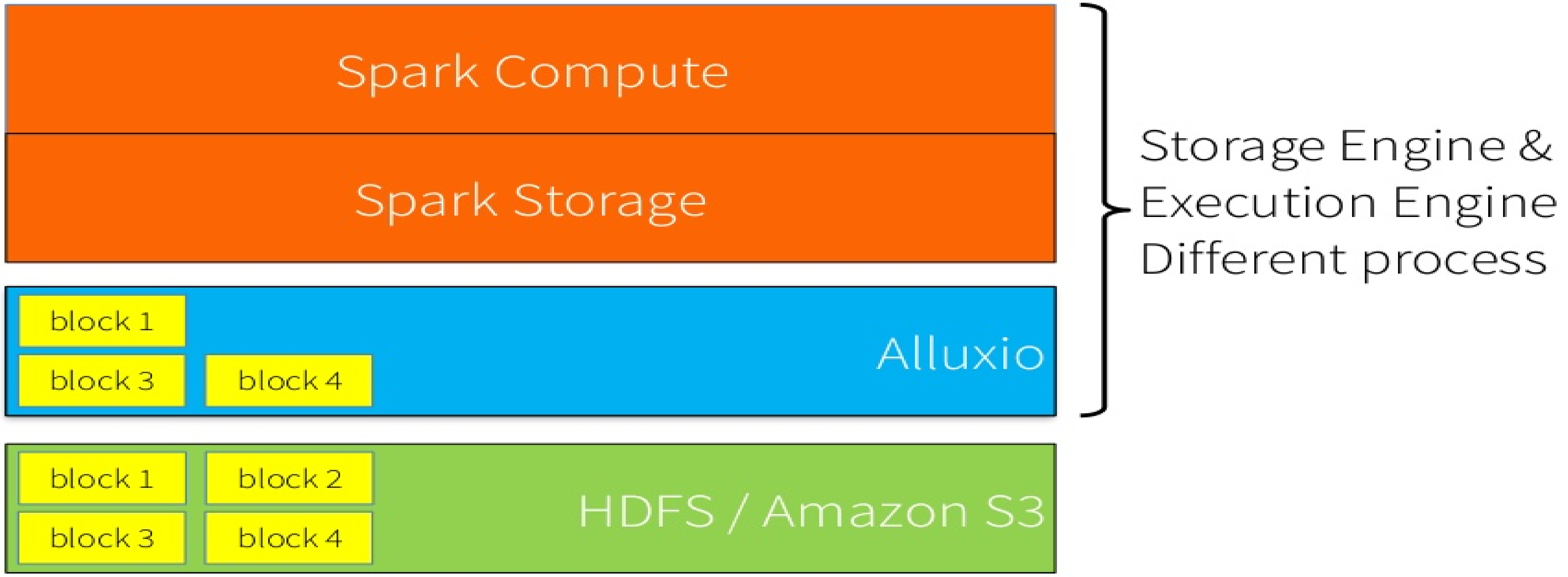
DATA RESILIENCE DURING CRASH



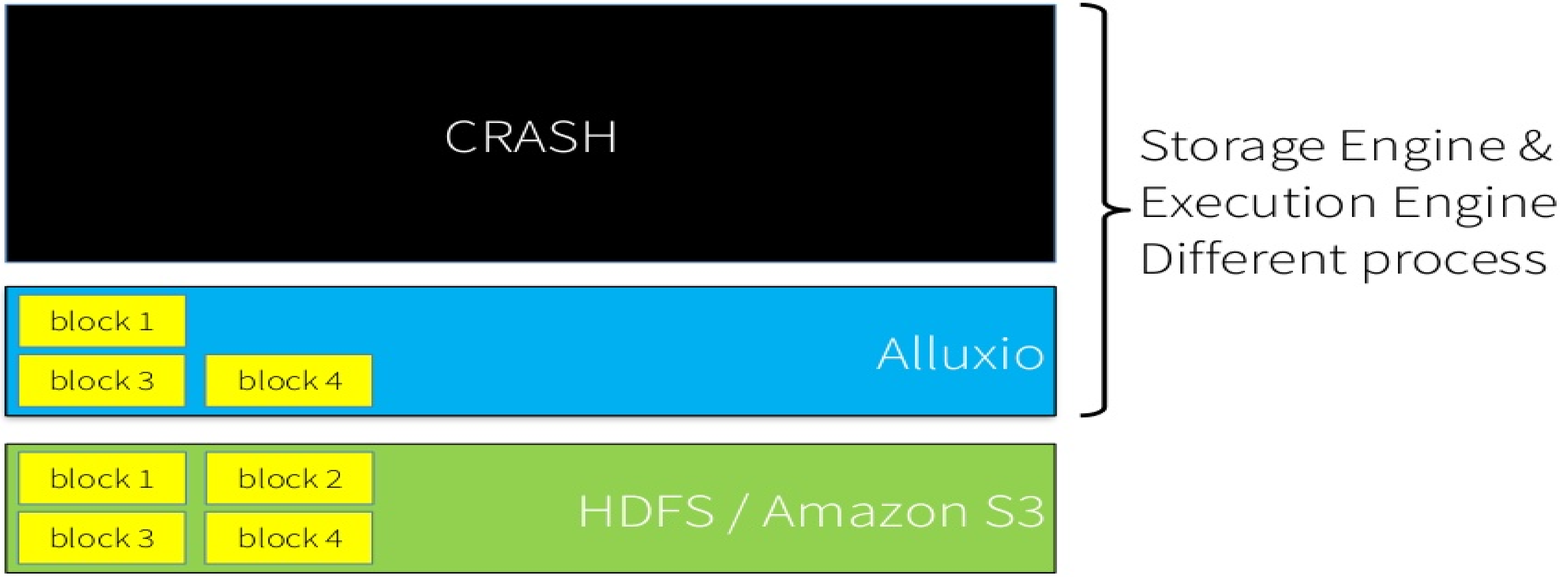
- Process Crash Requires Network and/or Disk I/O to Re-read Data

ALLUXIO

DATA RESILIENCE DURING CRASH



DATA RESILIENCE DURING CRASH



- Process Crash – Data is Re-read at Memory Speed

ACCESSING ALLUXIO DATA FROM SPARK

Writing Data	Write to an Alluxio file
Reading Data	Read from an Alluxio file

CODE EXAMPLE FOR SPARK RDDS

Writing RDD to Alluxio

```
rdd.saveAsTextFile(alluxioPath)
rdd.saveAsObjectFile(alluxioPath)
```

Reading RDD from Alluxio

```
rdd = sc.textFile(alluxioPath)
rdd = sc.objectFile(alluxioPath)
```

CODE EXAMPLE FOR SPARK DATAFRAMES

Writing DataFrame to Alluxio

```
df.write.parquet(alluxioPath)
```

Reading DataFrame from Alluxio

```
df = sc.read.parquet(alluxioPath)
```

OUTLINE

- Alluxio Overview
- Alluxio + Spark Use Cases
- Using Alluxio with Spark
- Performance Evaluation

ENVIRONMENT

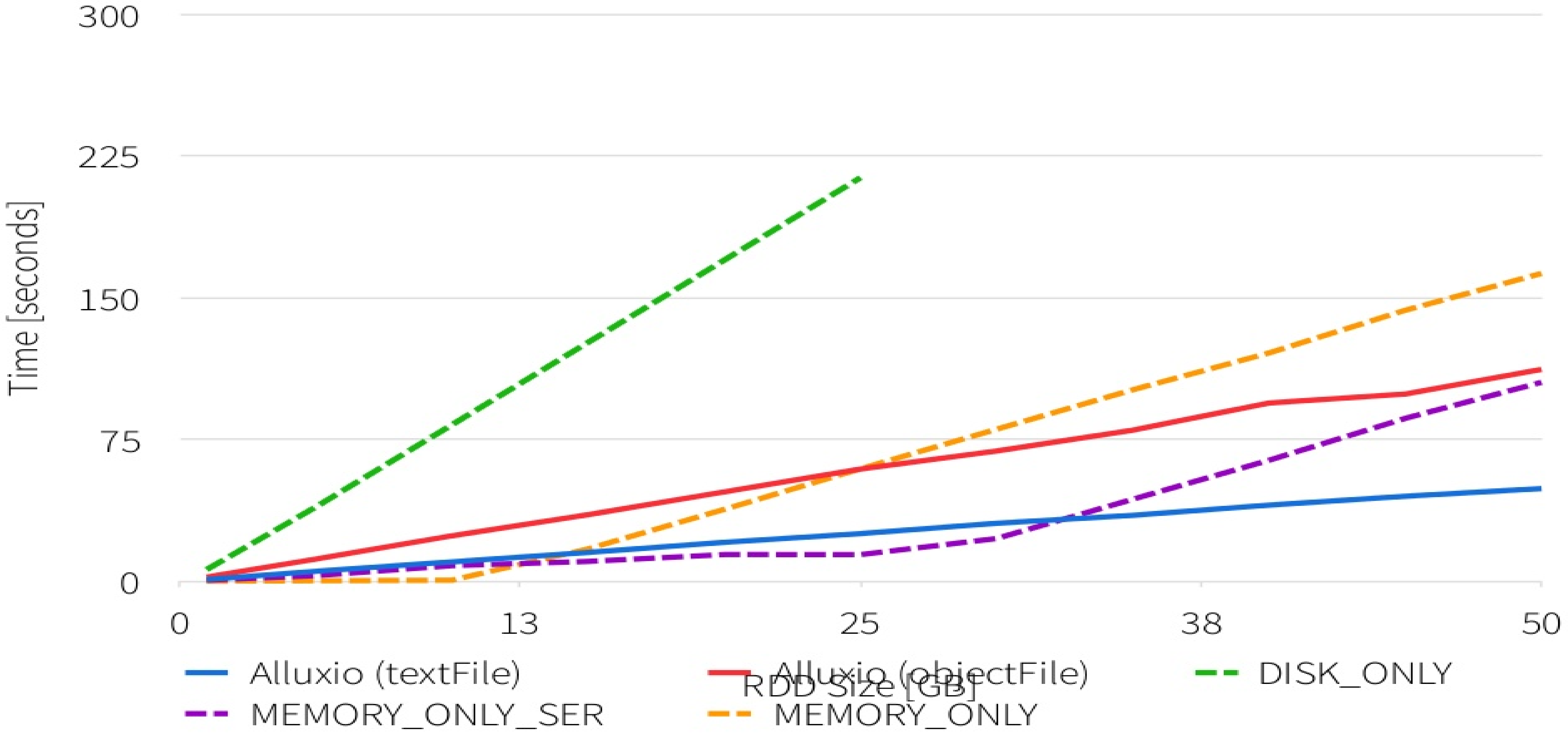
Spark 2.0.0 + Alluxio 1.2.0

Single r3.2xlarge instance (61GB RAM)

Comparisons:

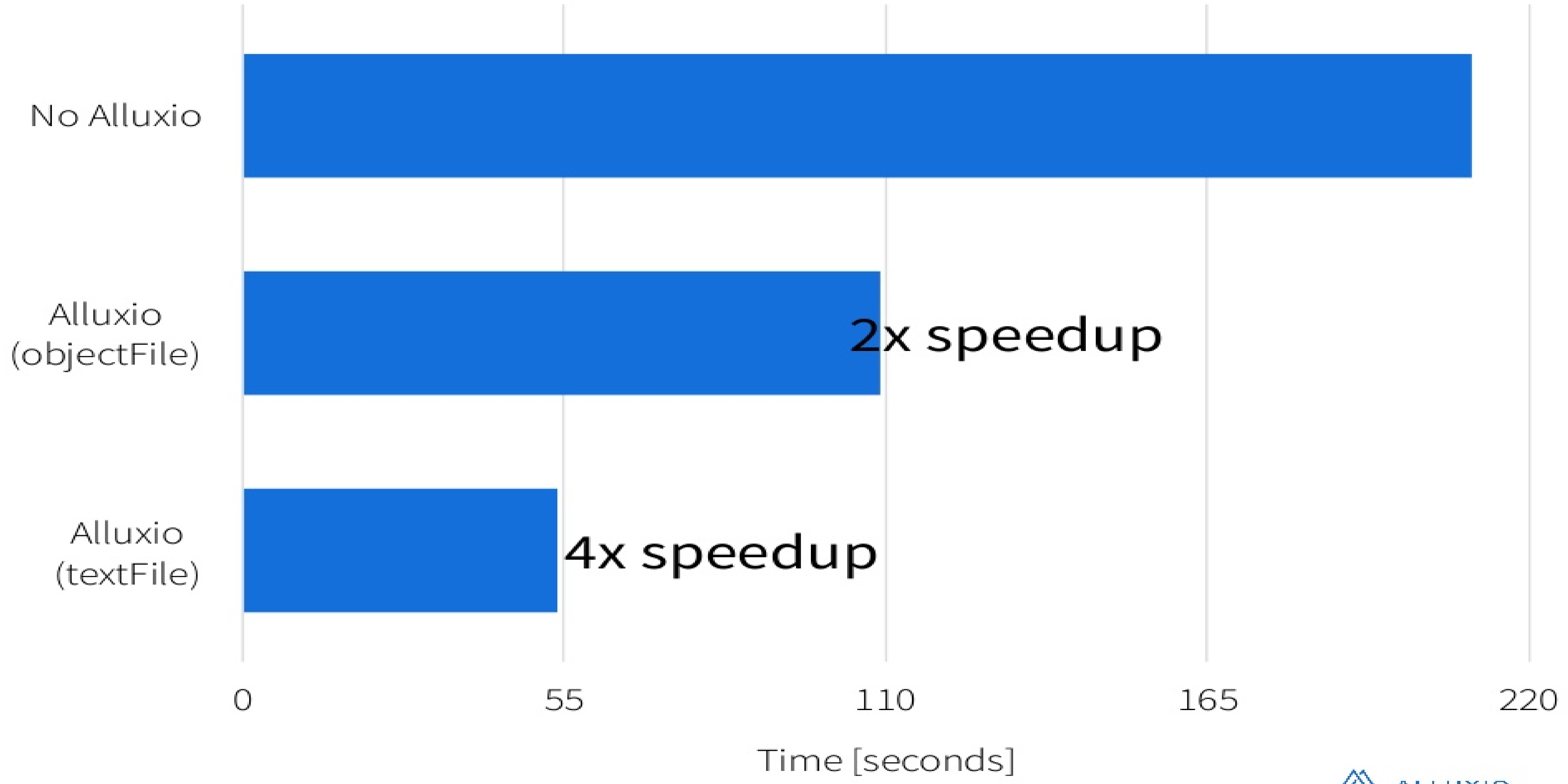
- Alluxio
- Spark Storage Level: MEMORY_ONLY
- Spark Storage Level: MEMORY_ONLY_SER
- Spark Storage Level: DISK_ONLY

Reading Cached RDD

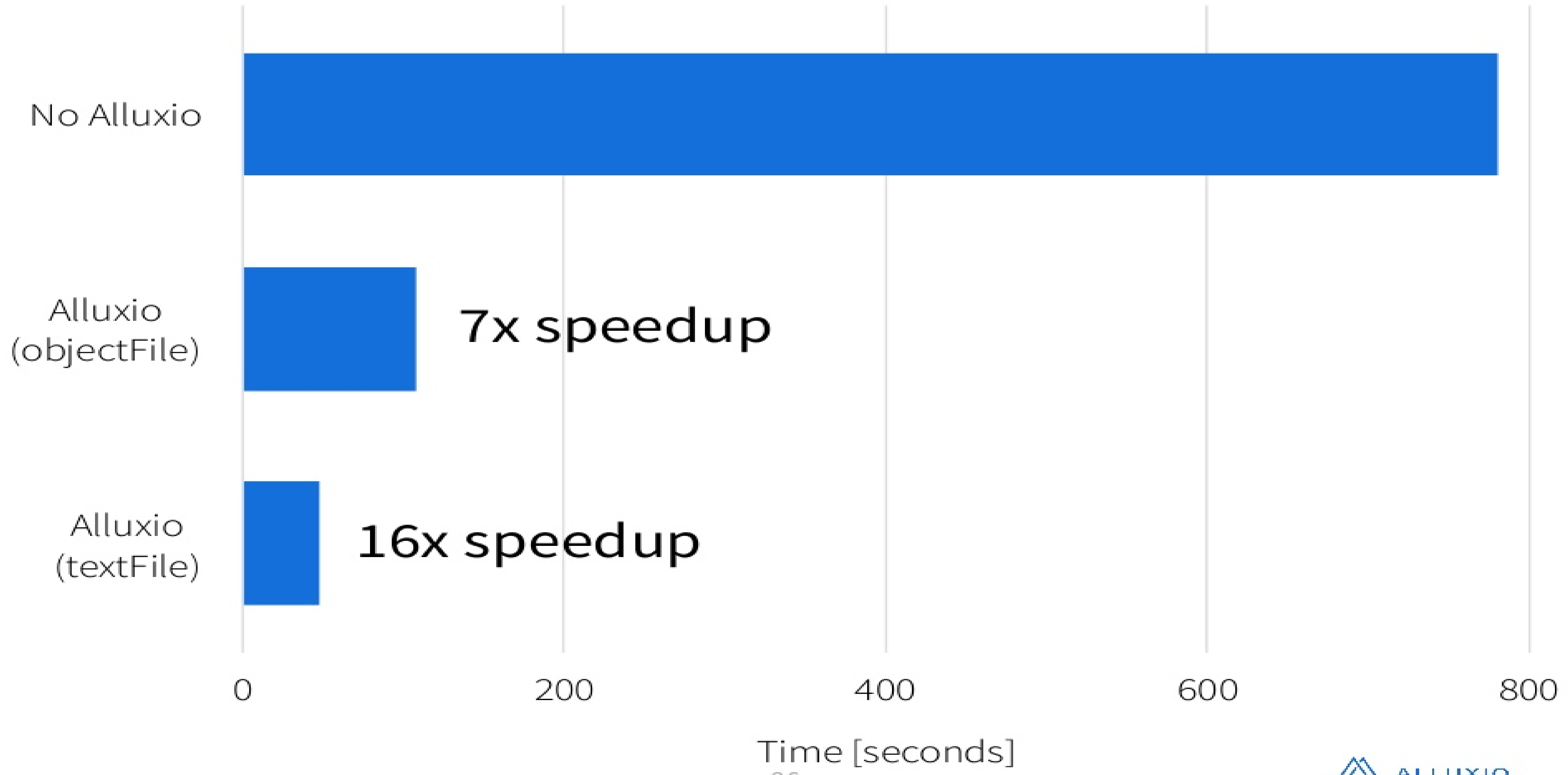


ALLUXIO

New Context: Read 50 GB RDD (SSD)

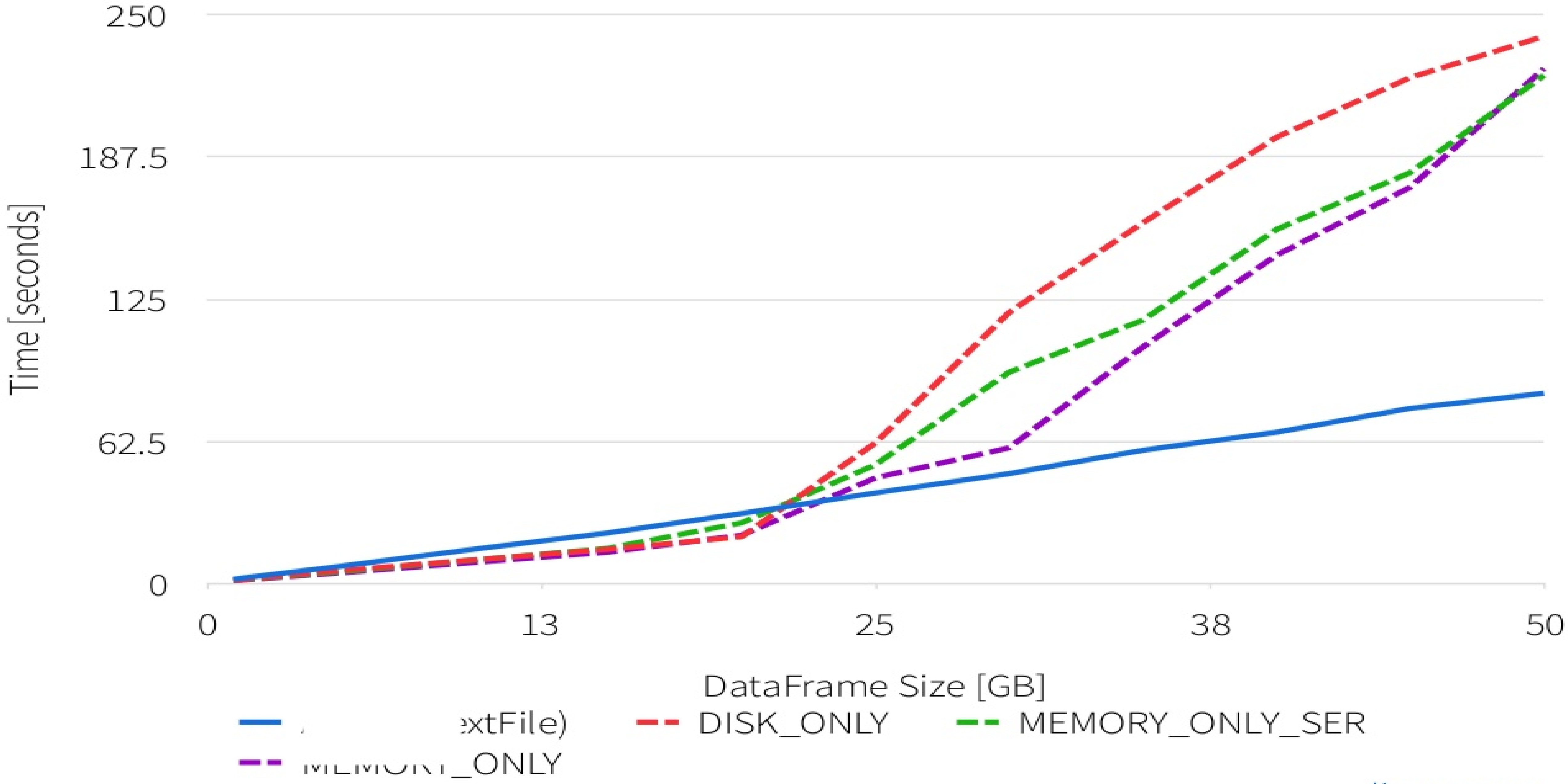


New Context: Read 50 GB RDD (S3)



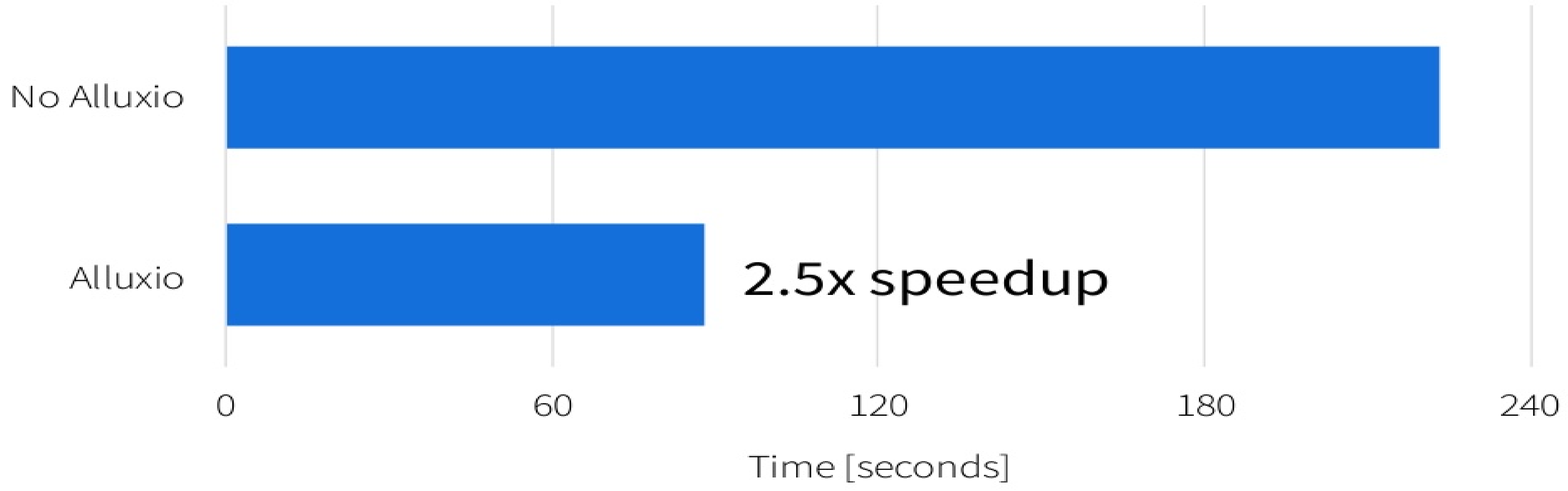
ALLUXIO

Reading CACHED DATAFRAME (parquet)

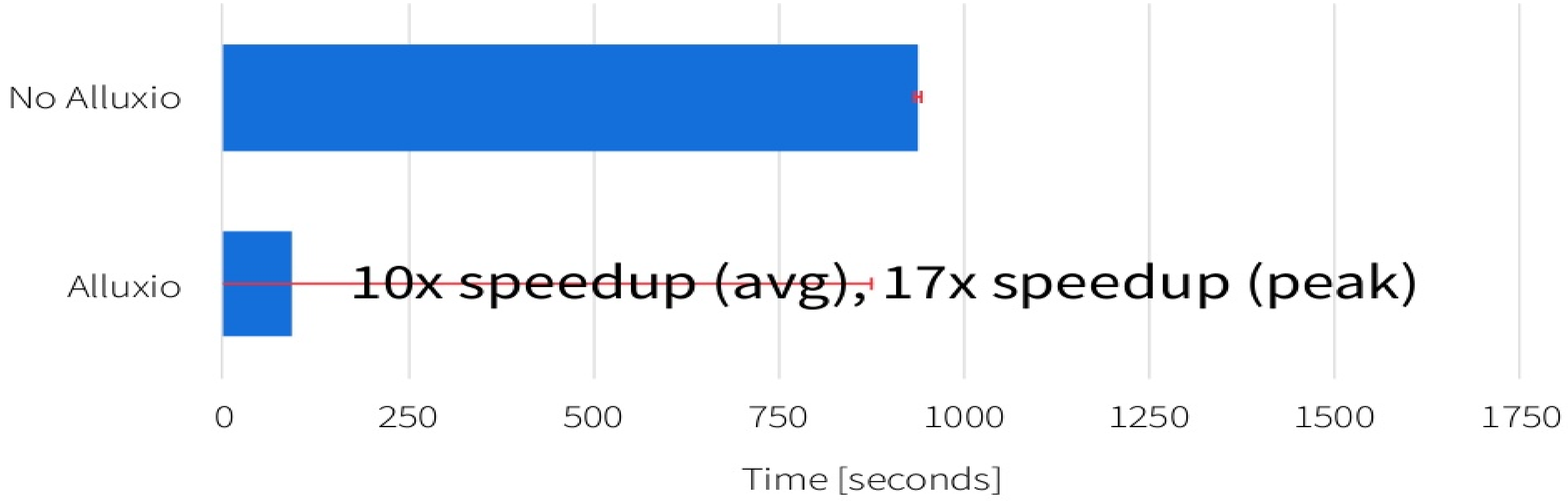


ALLUXIO

New Context: Read 50 GB DATAFRAME (SSD)



New CONTEXT: Read 50 GB DataFrame (S3)



CONCLUSION

- Easy to use with Spark
- Better Performance
- Predictable Performance

Thank you!

Contact: jiri@alluxio.com

Twitter: @jsimsa

Websites: www.alluxio.com and www.alluxio.org