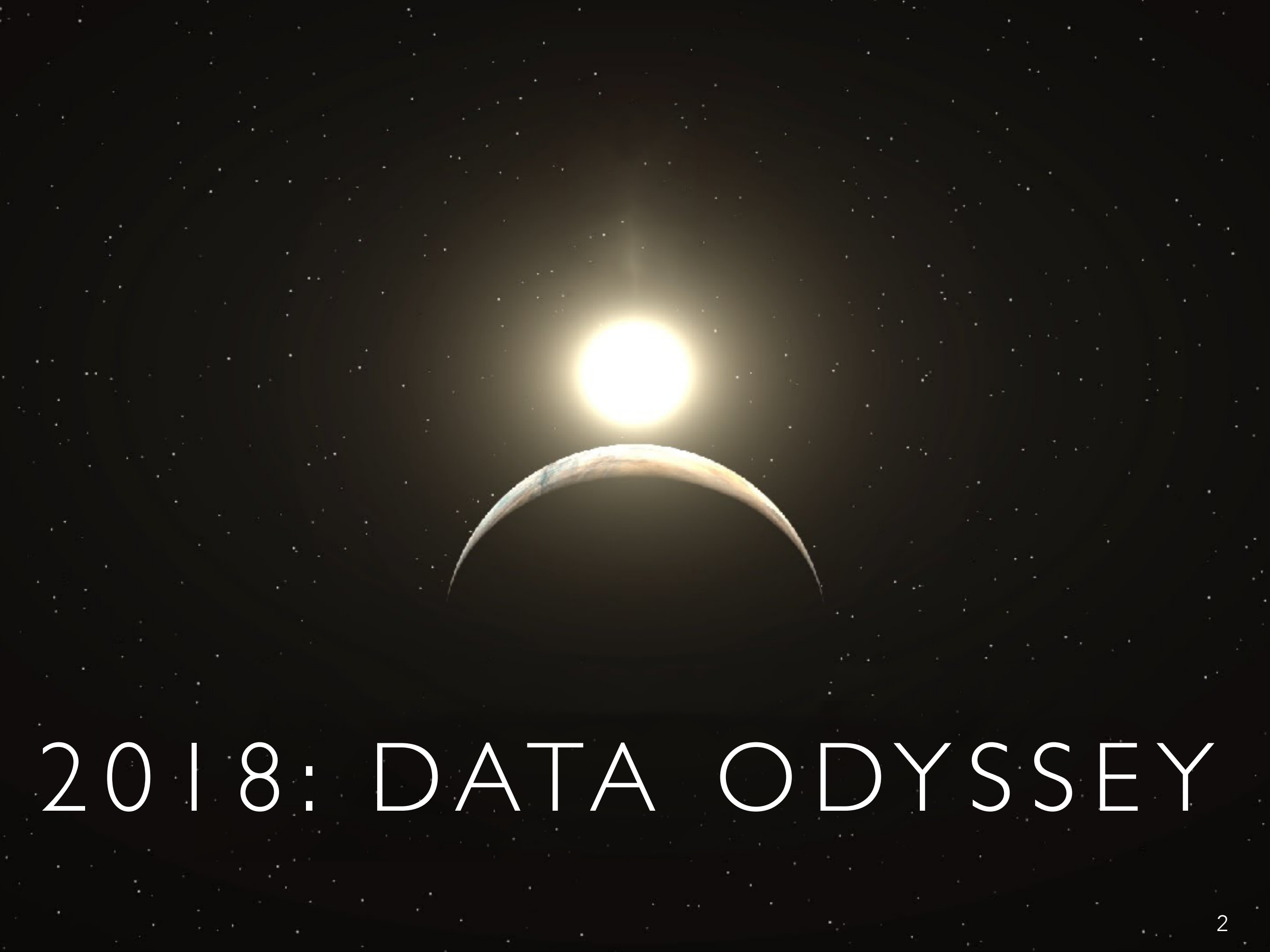




Event-Driven Data Architecture at Letgo

Ricardo Fanjul
Data Engineer



2018: DATA ODYSSEY



Ricardo Fanjul
Data Engineer



🚩 Founded in 2015

⬆ 100MM+ downloads

⬆ 400MM+ listings

Post in a Snap
in just seconds



LET DATA PLATFORM IN NUMBERS

1 billion

Events Processed Daily

50K

Peaks of events per Second

500GB

Data daily

600+

Event Types

200TB

Storage (S3)

< 1 sec

NRT Processing Time



THE DAWN OF LETGO

THE DAWN OF LETGO

CLASSICAL BI PLATFORM



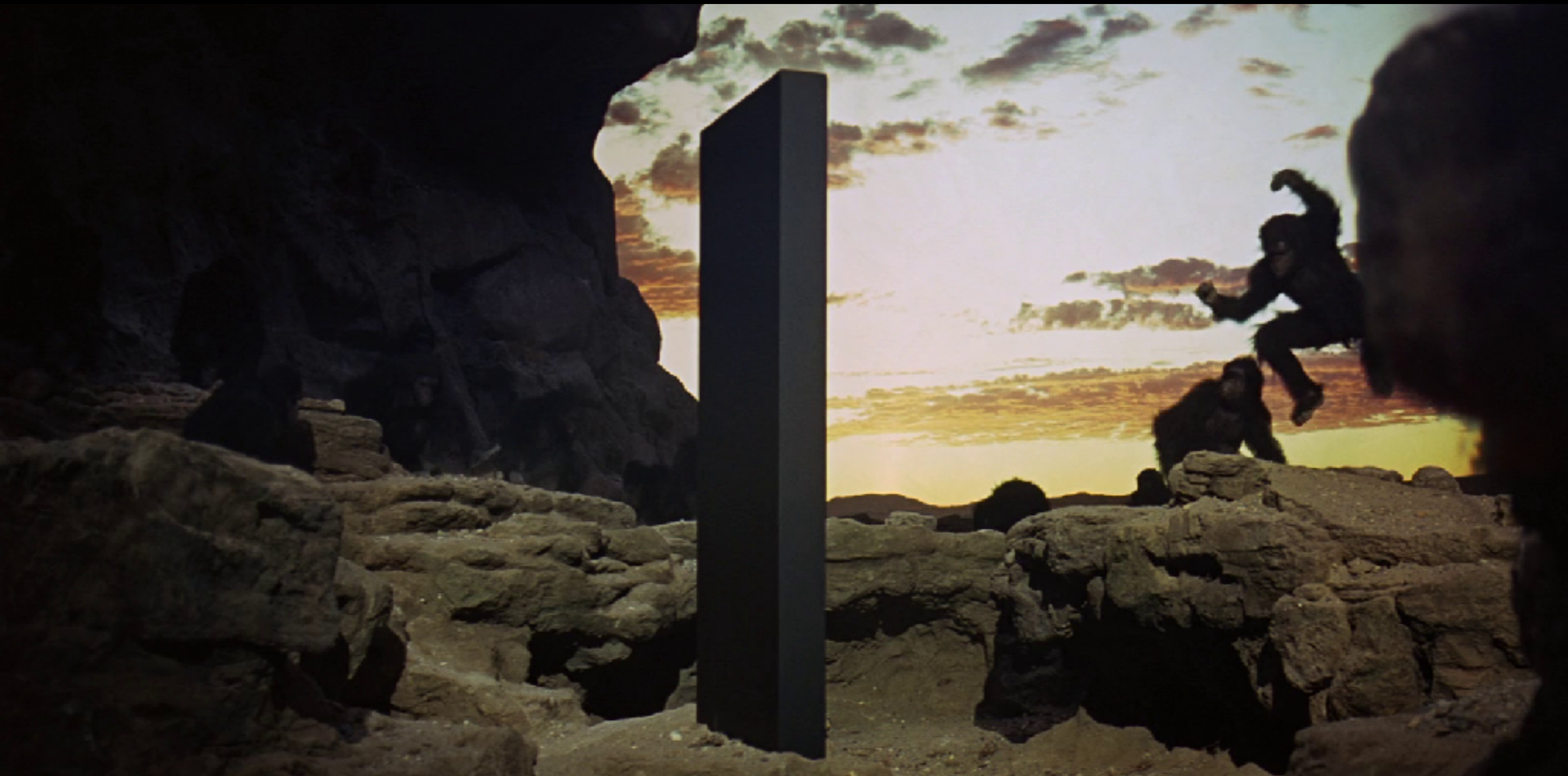
THE DAWN OF LETGO

CLASSICAL BI PLATFORM



THE DAWN OF LETGO

MOVING TO μ -SERVICES AND EVENTS



THE DAWN OF LETGO

MOVING TO μ -SERVICES AND EVENTS

Tracking Events

Domain Events

INGEST

Data Ingestion

Storage

PROCESS

Stream

Batch

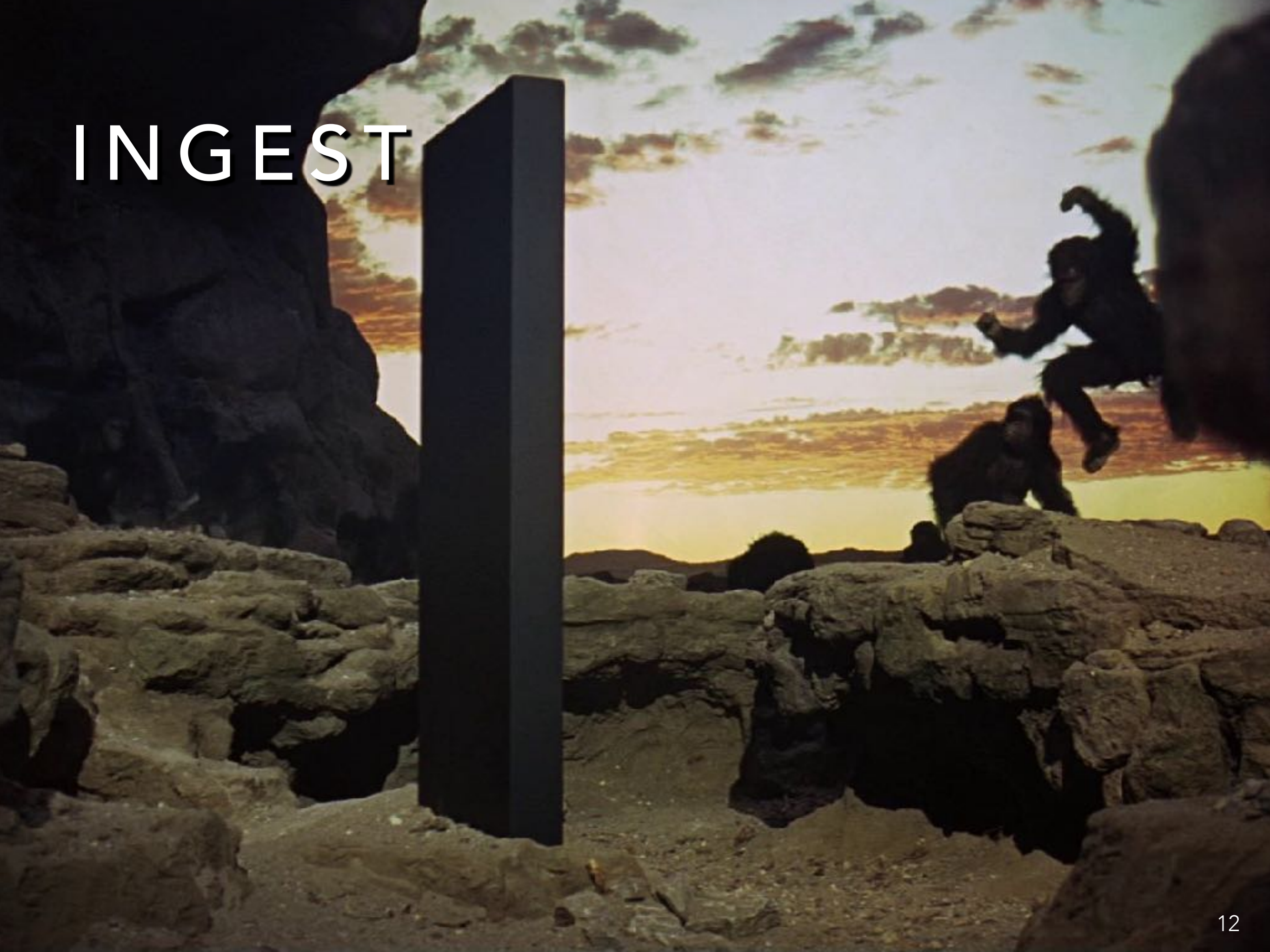
DISCOVER

Query

Data exploitation

Orchestration

INGEST



INGEST

Data Ingestion

Storage

PROCESS

Stream

Batch

DISCOVER

Query

Data exploitation

Orchestration

OUR GOAL



amazon
SQS



kafka

DATA INGESTION

THE DISCOVERY



KAFKA CONNECT







THE JOURNEY BEGINS

INGEST

Data Ingestion

Storage

PROCESS

Stream

Batch

DISCOVER

Query

Data exploitation

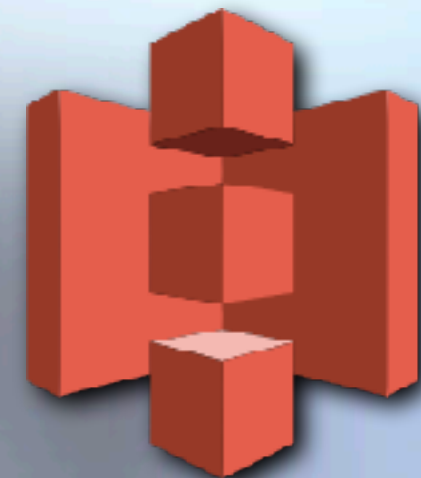
Orchestration

STORAGE

BUILDING THE DATA LAKE



BUILDING THE DATA LAKE

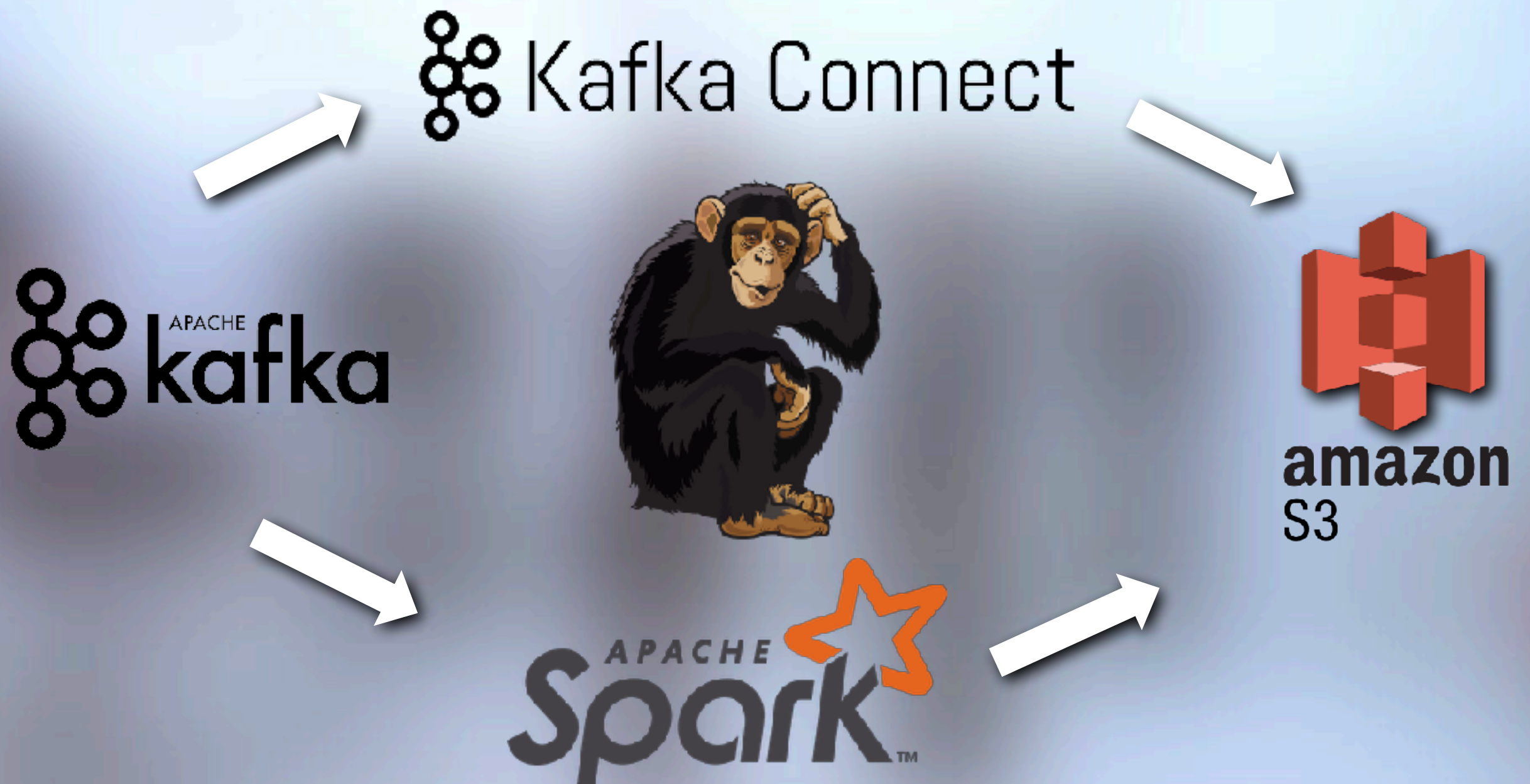


amazon
S3

We want to store all events coming from Kafka to S3.

STORAGE

BUILDING THE DATA LAKE



STORAGE

SOMETIMES... SHIT HAPPENS



ZERO GRAVITY TOILET

PASSENGERS ARE ADVISED TO
READ INSTRUCTIONS BEFORE USE

- 1 The toilet is of the standard zero-gravity type, depending on requirements, either a water system or a vacuum system. It is designed to be used in the zero-gravity environment of space. When operating, please follow the instructions and do not use the toilet until you have received the necessary training. The toilet is designed to be used in the zero-gravity environment of space. When operating, please follow the instructions and do not use the toilet until you have received the necessary training.
- 2 The toilet is now ready for use. The vacuum chamber is activated by the control panel on the left. When using, please follow the instructions and do not use the toilet until you have received the necessary training. The toilet is designed to be used in the zero-gravity environment of space. When operating, please follow the instructions and do not use the toilet until you have received the necessary training.
- 3 The controls for operation are located on the control panel. The red button is used to activate the vacuum system. The green button is used to activate the water system. The yellow button is used to activate the fan. The blue button is used to activate the light. The red button is used to activate the vacuum system. The green button is used to activate the water system. The yellow button is used to activate the fan. The blue button is used to activate the light.
- 4 The toilet is now ready for use. The vacuum chamber is activated by the control panel on the left. When using, please follow the instructions and do not use the toilet until you have received the necessary training. The toilet is designed to be used in the zero-gravity environment of space. When operating, please follow the instructions and do not use the toilet until you have received the necessary training.
- 5 To use the toilet, first activate the vacuum system. Then, activate the water system. The toilet is now ready for use. The vacuum chamber is activated by the control panel on the left. When using, please follow the instructions and do not use the toilet until you have received the necessary training. The toilet is designed to be used in the zero-gravity environment of space. When operating, please follow the instructions and do not use the toilet until you have received the necessary training.
- 6 The toilet is now ready for use. The vacuum chamber is activated by the control panel on the left. When using, please follow the instructions and do not use the toilet until you have received the necessary training. The toilet is designed to be used in the zero-gravity environment of space. When operating, please follow the instructions and do not use the toilet until you have received the necessary training.
- 7 The toilet is now ready for use. The vacuum chamber is activated by the control panel on the left. When using, please follow the instructions and do not use the toilet until you have received the necessary training. The toilet is designed to be used in the zero-gravity environment of space. When operating, please follow the instructions and do not use the toilet until you have received the necessary training.



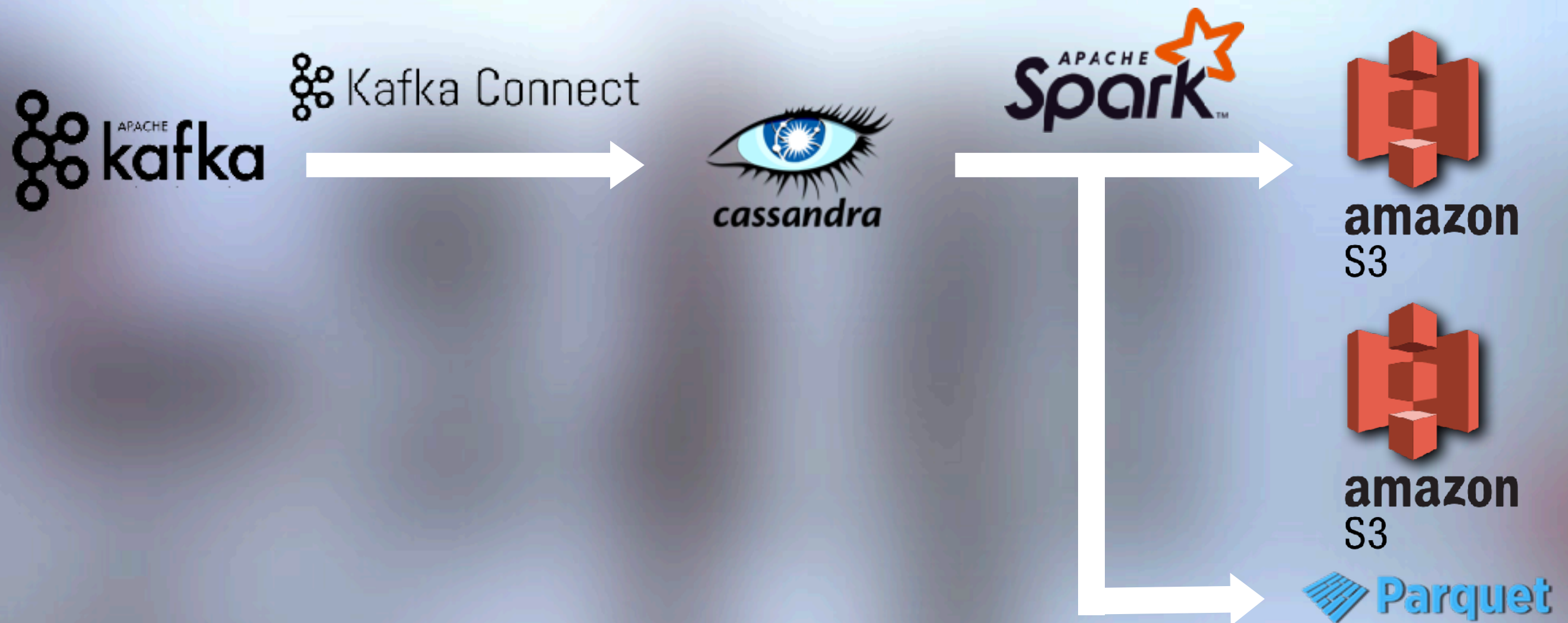
STORAGE

DUPLICATED EVENTS



STORAGE

DUPLICATED EVENTS

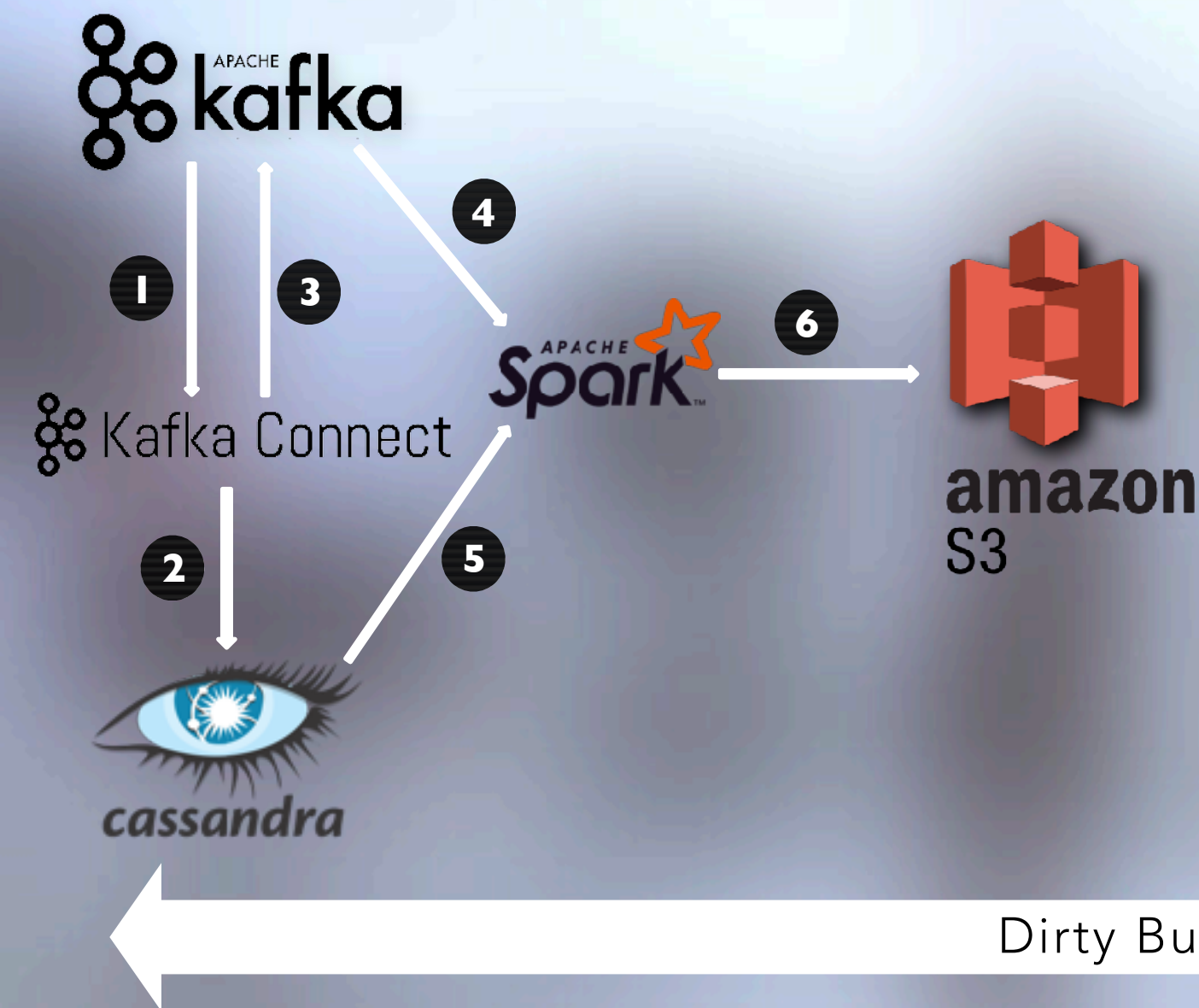


STORAGE

(VERY) LATE EVENTS



(VERY) LATE EVENTS



- 1 Read batch of events from Kafka.
- 2 Write each event to Cassandra.
- 3 Write "dirty hours" to compact topic: Key=(event_type, hour).
- 4 Read dirty "hours" topic.
- 5 Read all events with dirty hours.
- 6 Store in S3

STORAGE

S3 PROBLEMS



S3 PROBLEMS



VS



amazon
S3

SOME S3 BIG DATA PROBLEMS:

1. Eventual consistency
2. Very slow renames

STORAGE

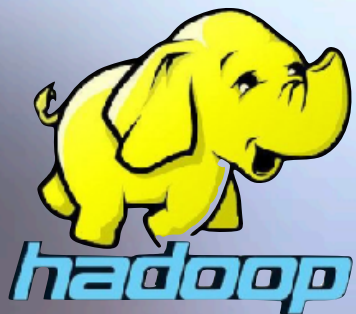
S3 PROBLEMS: EVENTUAL CONSISTENCY



S3 PROBLEMS: EVENTUAL CONSISTENCY



S3GUARD



FileSystem
Operations

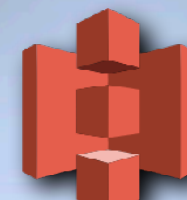


DynamoDB Client

S3 Client



amazon
DynamoDB



amazon
S3

Write Path 1 2

Read Path 1 2 3

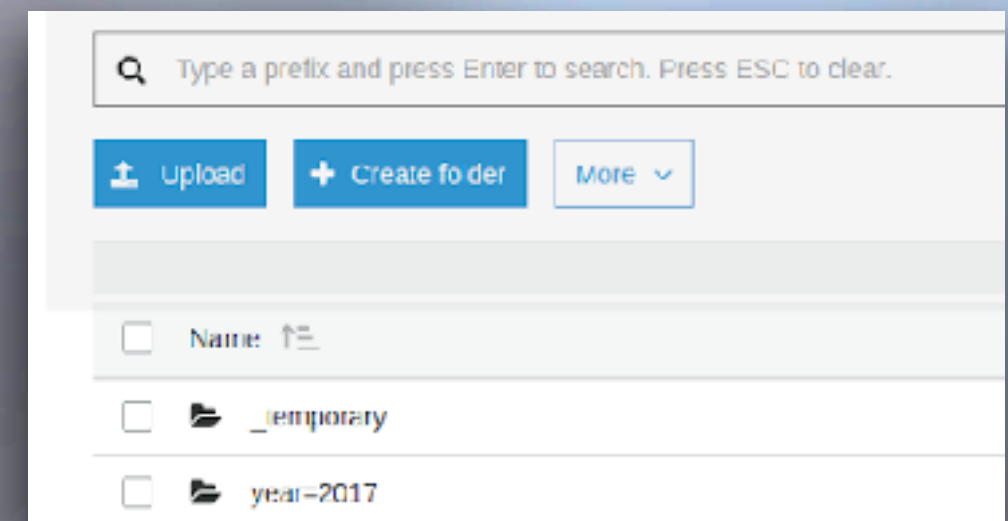
S3 PROBLEMS: SLOW RENAMES



Summary

	RDD Blocks	Storage Memory	Disk Used	Cores	Active Tasks	Failed Tasks	Complete Tasks	Total Tasks
Active(2)	4	107.1 KB / 23.1 GB	0.0 B	8	0	0	8	8
Dead(10)	4	128.3 KB / 95.5 GB	0.0 B	80	0	0	1436	1436
Total(12)	8	235.4 KB / 118.6 GB	0.0 B	88	0	0	1444	1444

Executors



¿Job freeze?

S3 PROBLEMS: SLOW RENAMES



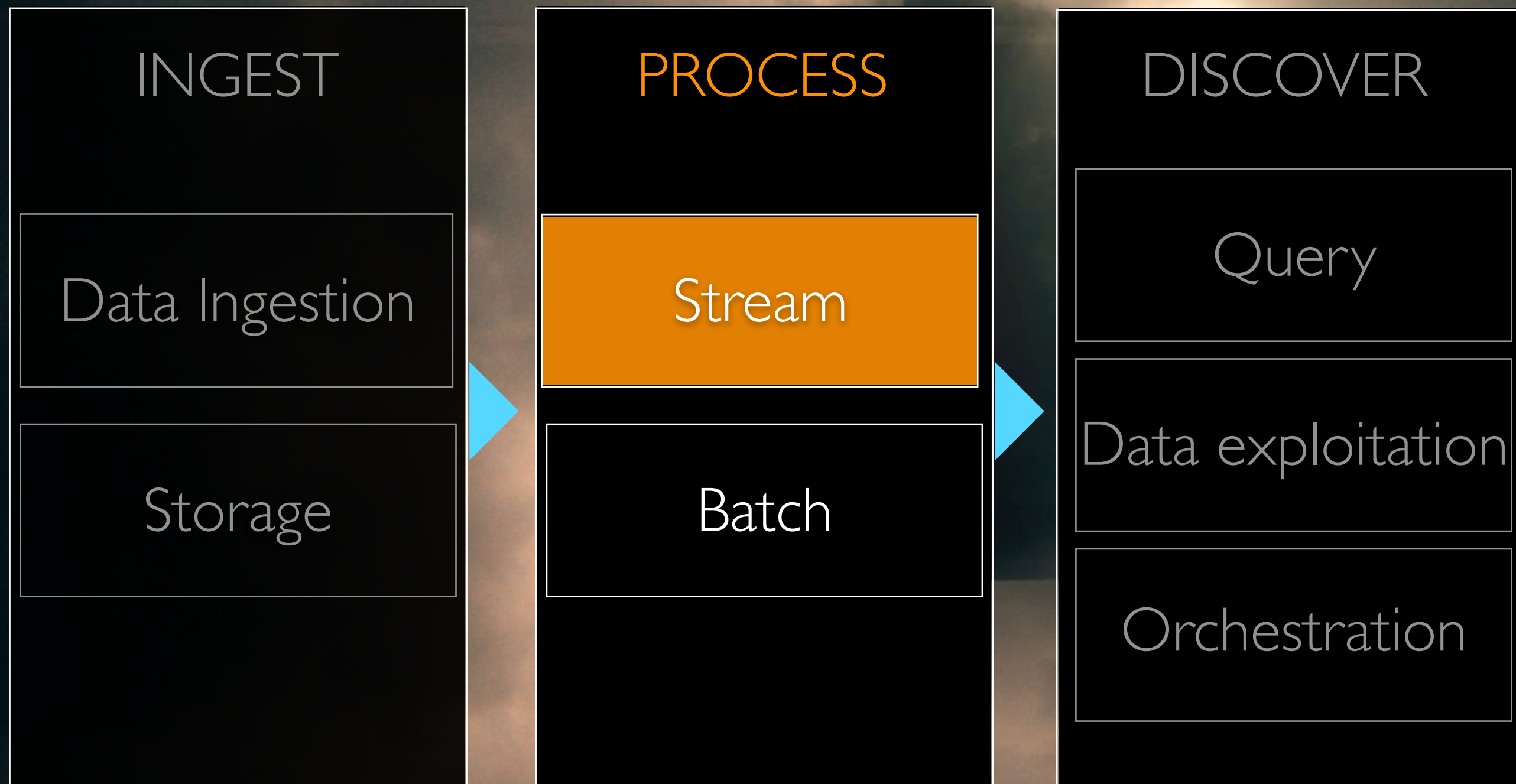
New **Hadoop 3.1** S3A committers:

- Staging
- Partitioned
- Magic





PROCESS



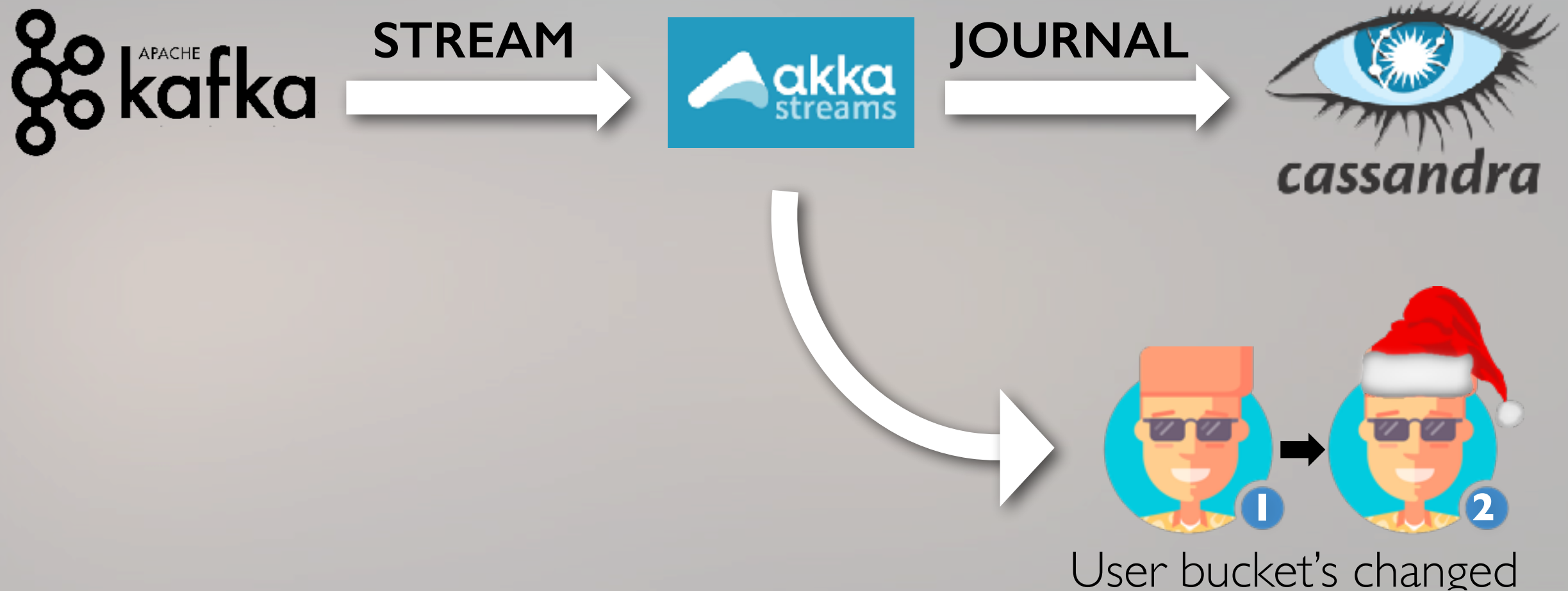
STREAM

REAL TIME USER SEGMENTATION



STREAM

REAL TIME USER SEGMENTATION



STREAM

REAL TIME PATTERN DETECTION



Is it still available?

Is the price negotiable?

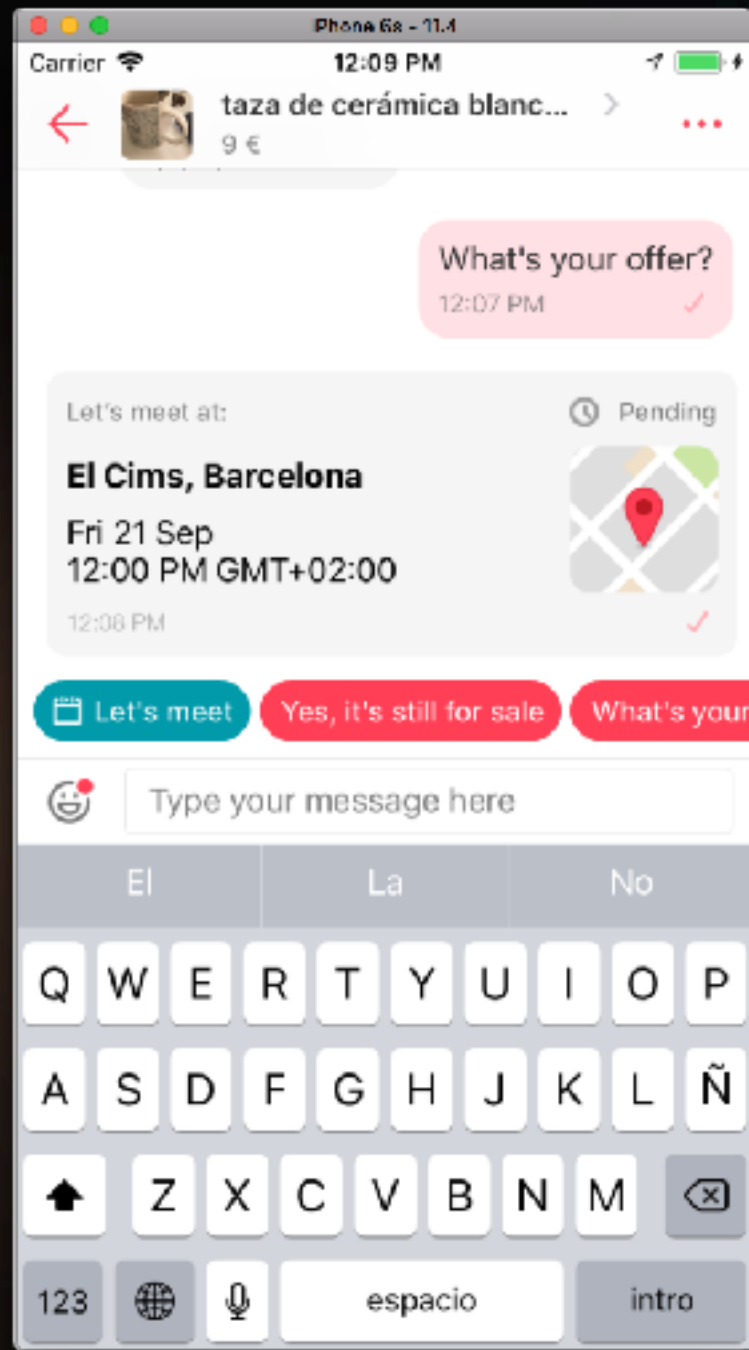
What condition is it in?

I offer you....\$

Could we meet at.....?

STREAM

REAL TIME PATTERN DETECTION

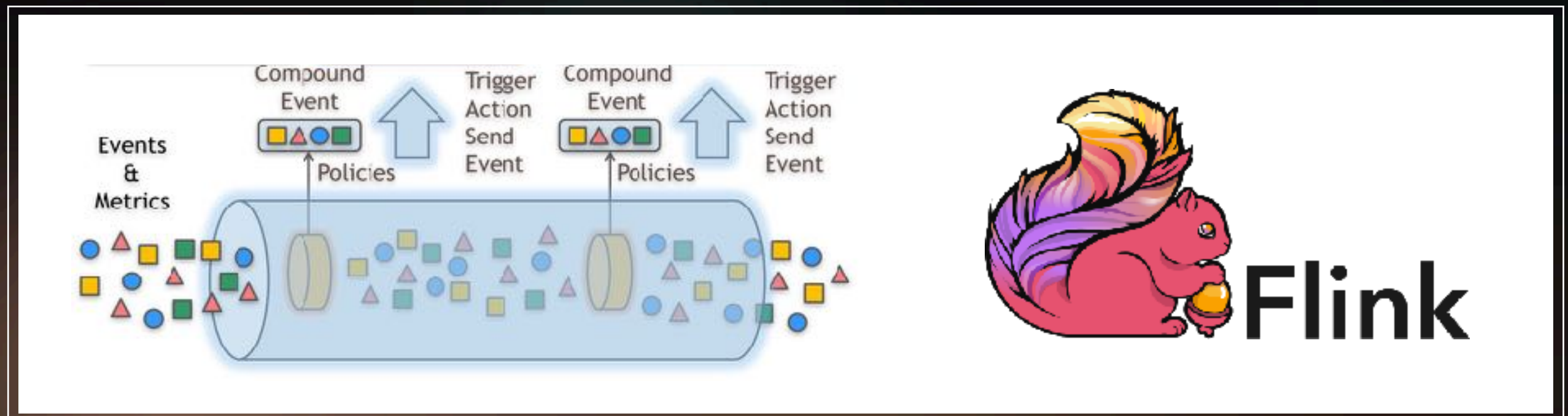


Structured data

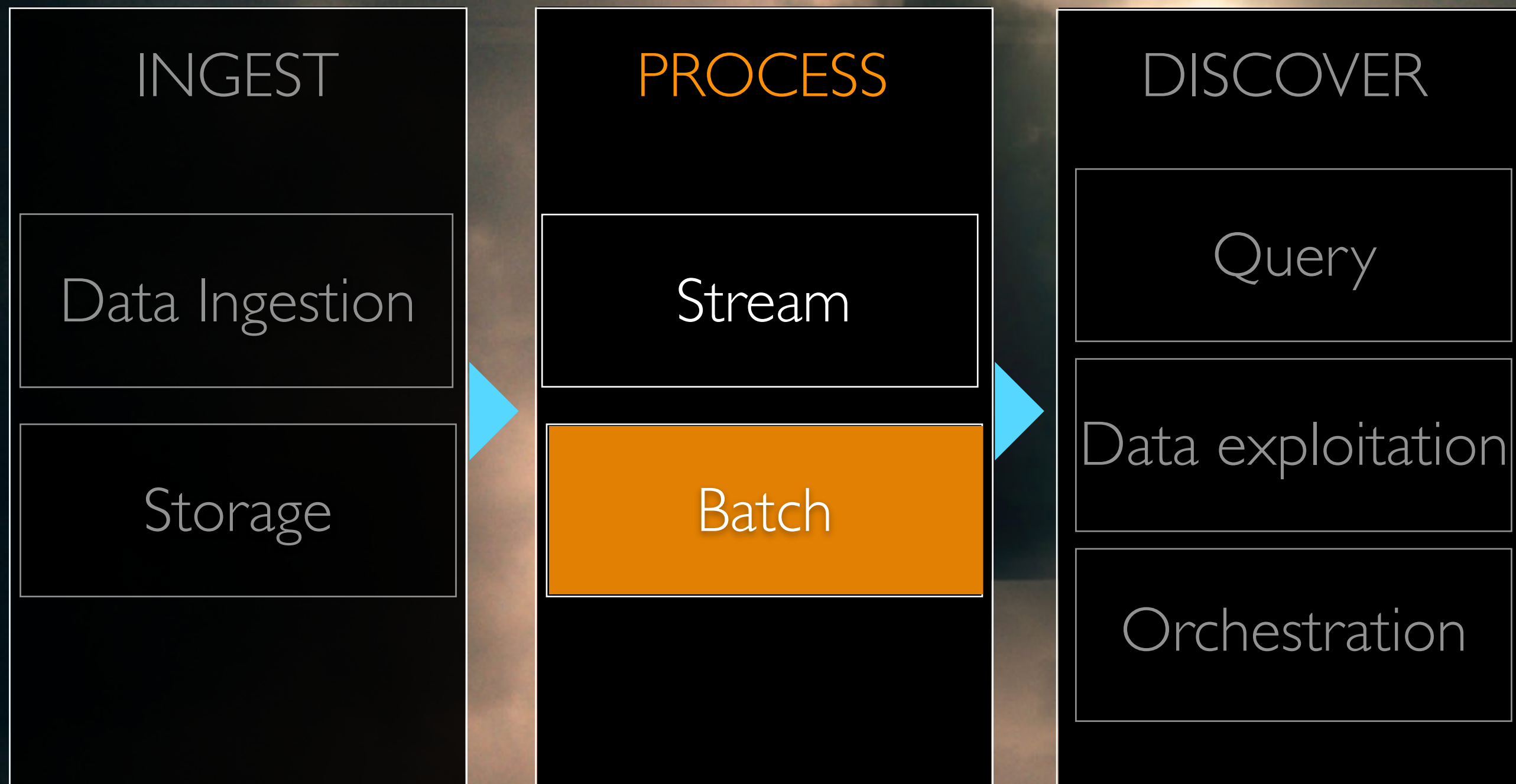
```
{  
  "type": "meeting_proposal",  
  "properties": {  
    "location_name": "Letgo HQ",  
    "geo": {  
      "lat": "41.390205",  
      "lon": "2.154007"  
    },  
    "date": "1511193820350",  
    "meeting_id": "23213213213"  
  }  
}
```


STREAM

REAL TIME PATTERN DETECTION

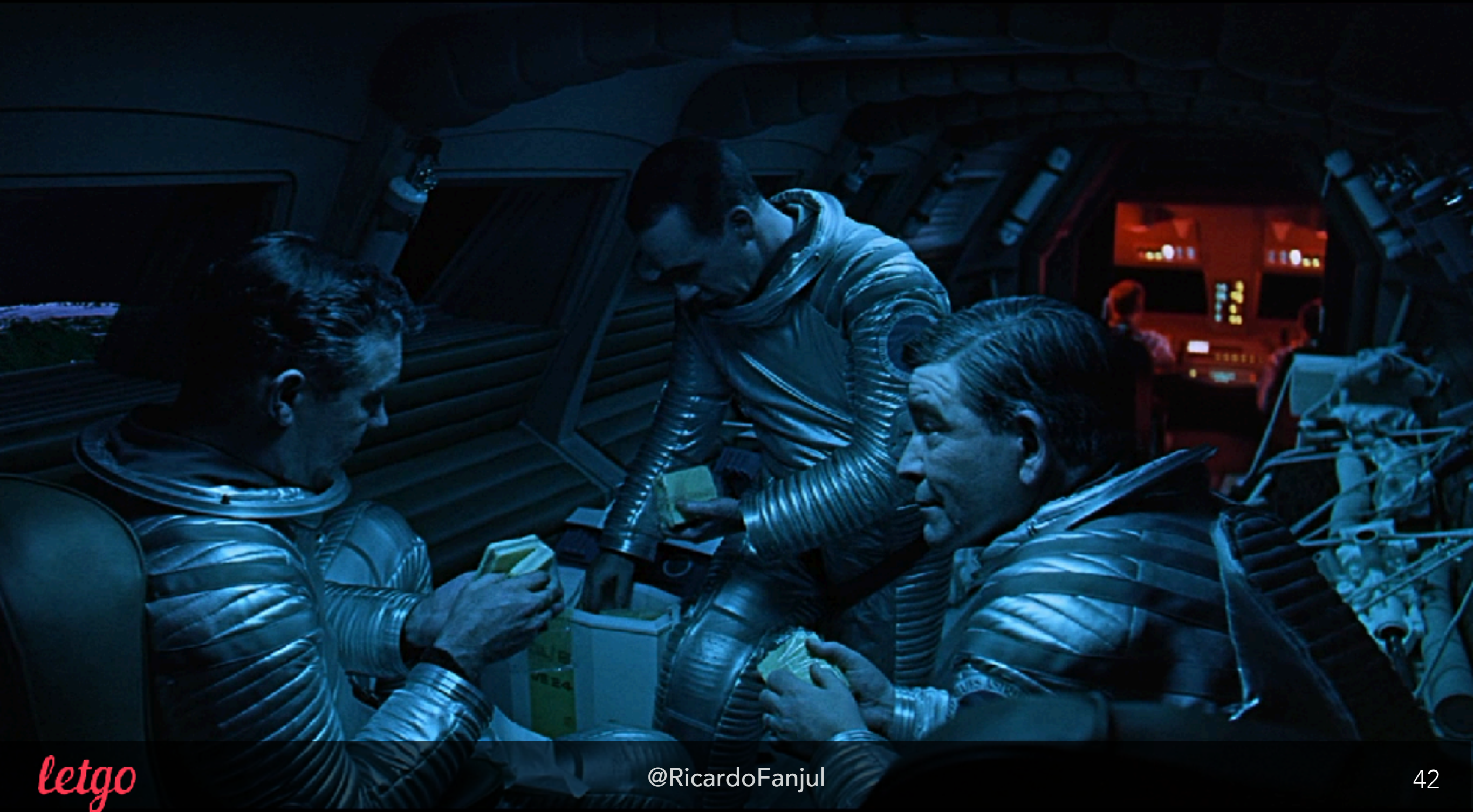


Meeting proposed + meeting accepted = emit accepted-meeting event
Meeting proposed + nothing in X time = "You have a proposal to meet"



BATCH

GEODATA ENRICHMENT



BATCH

GEO DATA ENRICHMENT

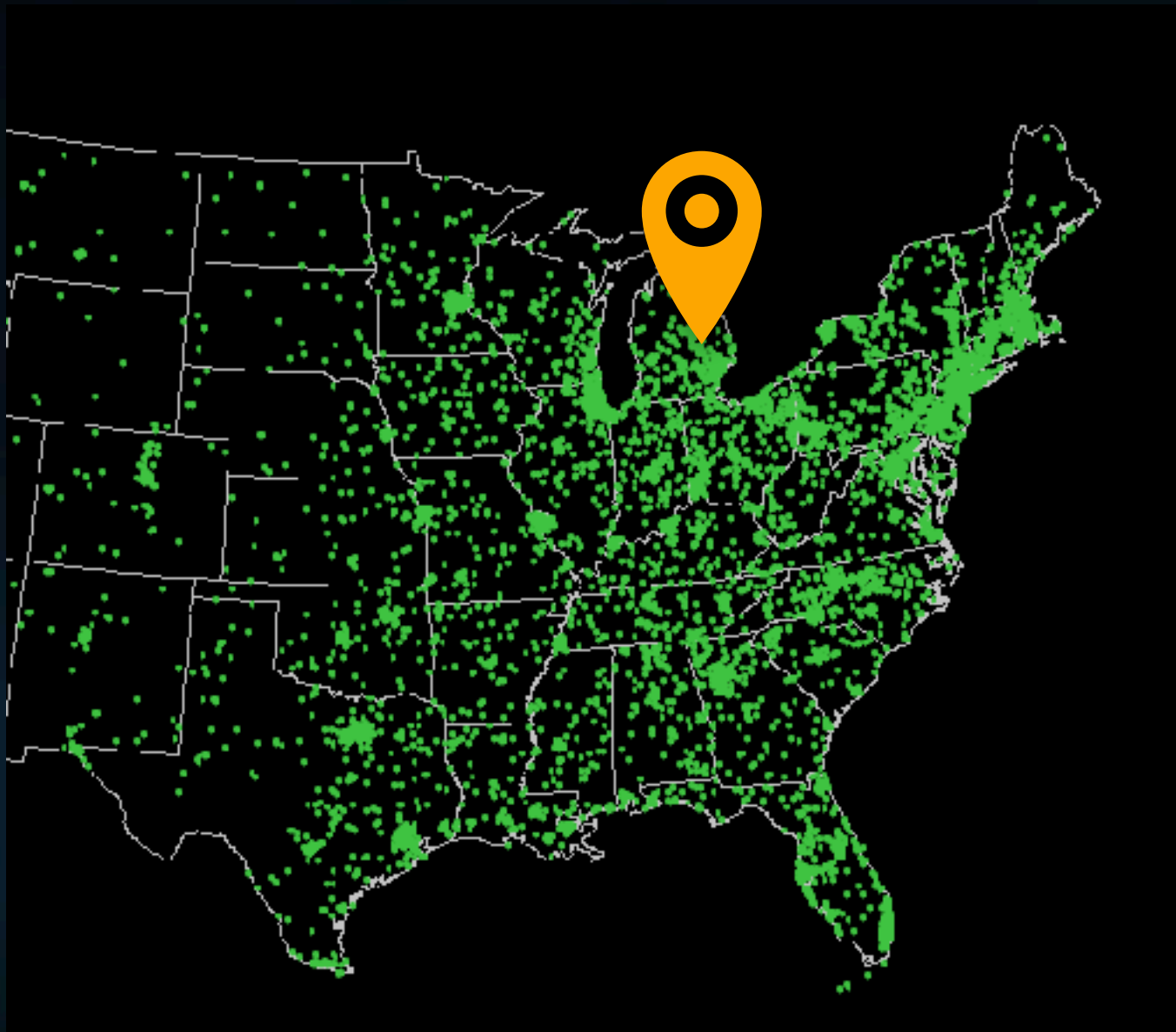
Technically correct but
not very actionable

```
{  
  "data": {  
    "id": "105dg3272-8e5f-426f-  
bca0-704e98552961",  
    "type": "some_event"  
    "attributes": {  
      "latitude": 42.3677203,  
      "longitude": -83.1186093  
    }  
  },  
  "meta": {  
    "created_at": 1522886400036  
  }  
}
```


BATCH

GEO DATA ENRICHMENT

What we know:



(42.3677203, -83.1186093)

- City: Detroit
- Postal code: 48206
- State: Michigan
- DMA: Detroit
- Country: US

BATCH

GEODATA ENRICHMENT

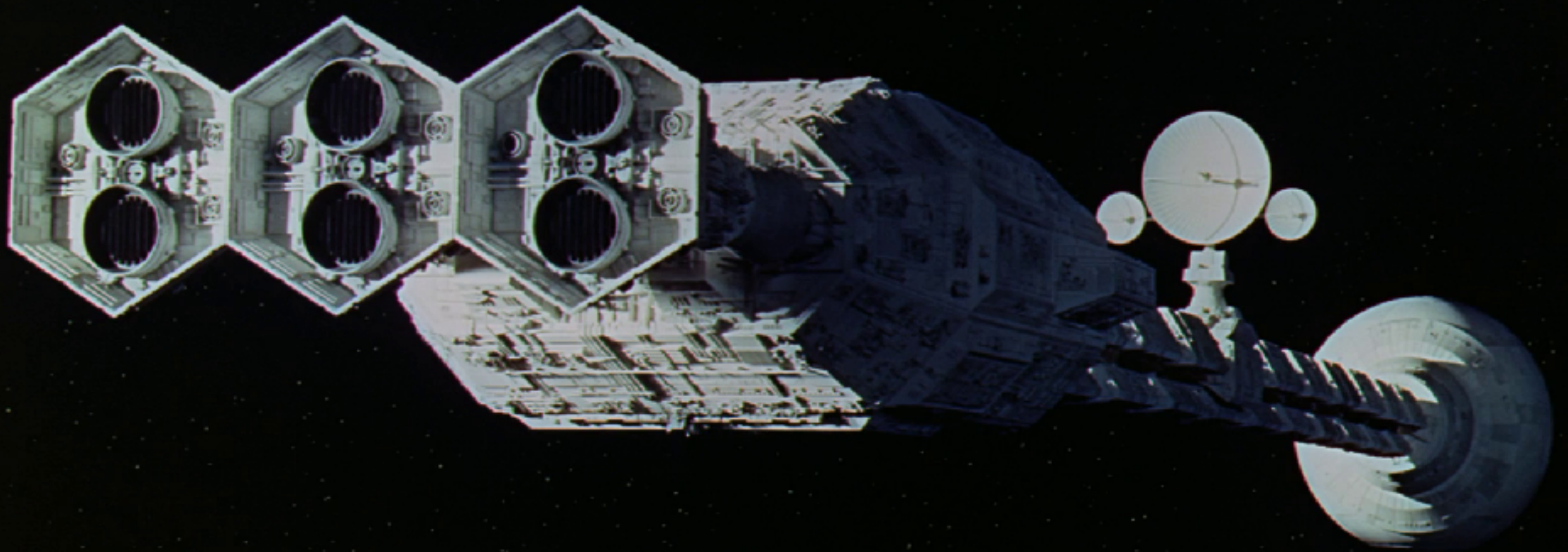
How we do it:

- Populating JTS indices from WKT polygon data
- Custom Spark SQL UDF



```
SELECT geodata.dma_name  
geodata.dma_number AS dma_number  
geodata.city      AS city  
geodata.state     AS state  
geodata.zip_code  AS zip_code  
FROM (  
  SELECT  
    geodata(longitude, latitude) AS geodata  
  FROM ....  
)
```


DISCOVER



INGEST

Data Ingestion

Storage

PROCESS

Stream

Batch

DISCOVER

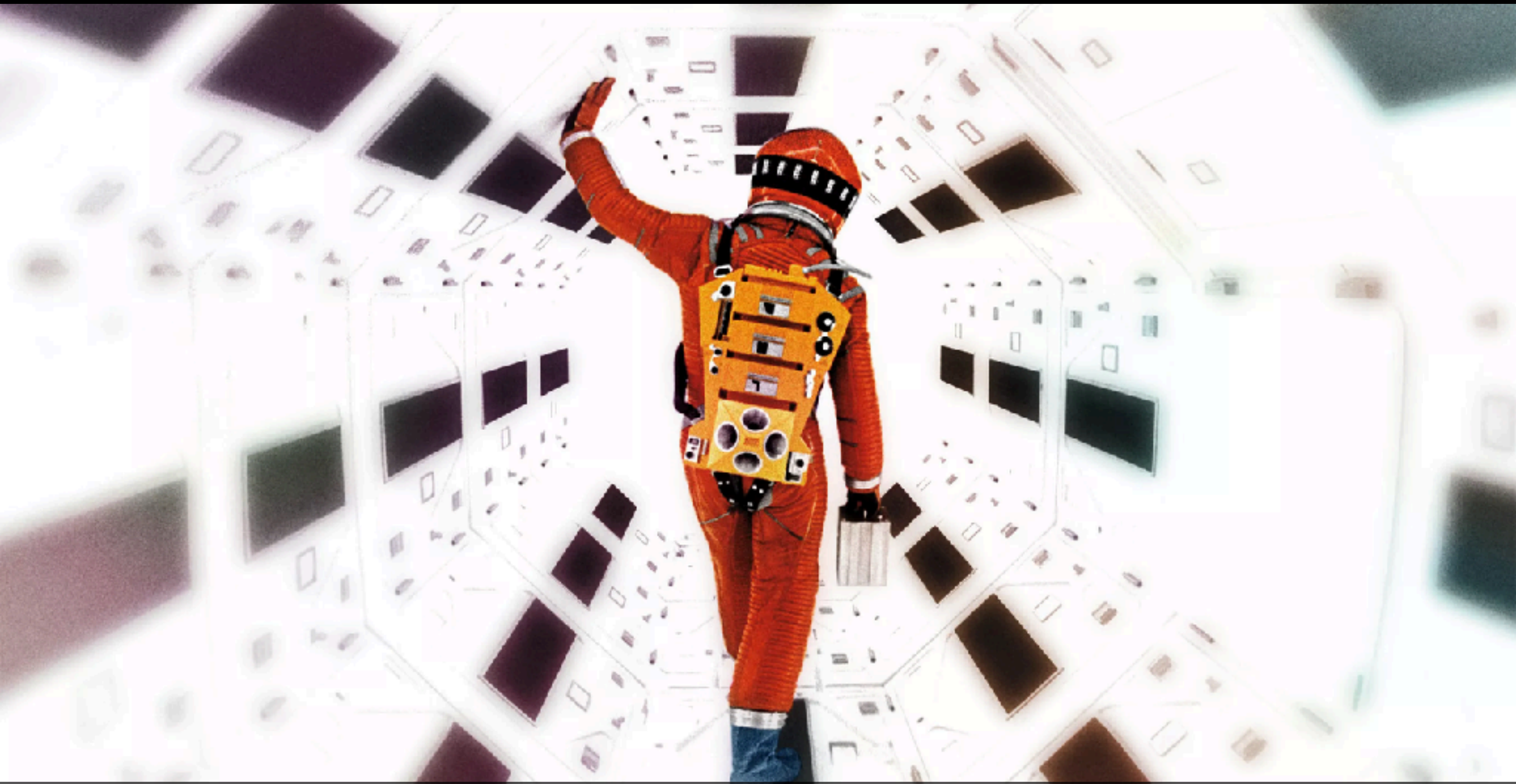
Query

Data exploitation

Orchestration

QUERY

QUERYING DATA



QUERY

QUERYING DATA



QUERY

QUERYING DATA



QUERY

QUERYING DATA



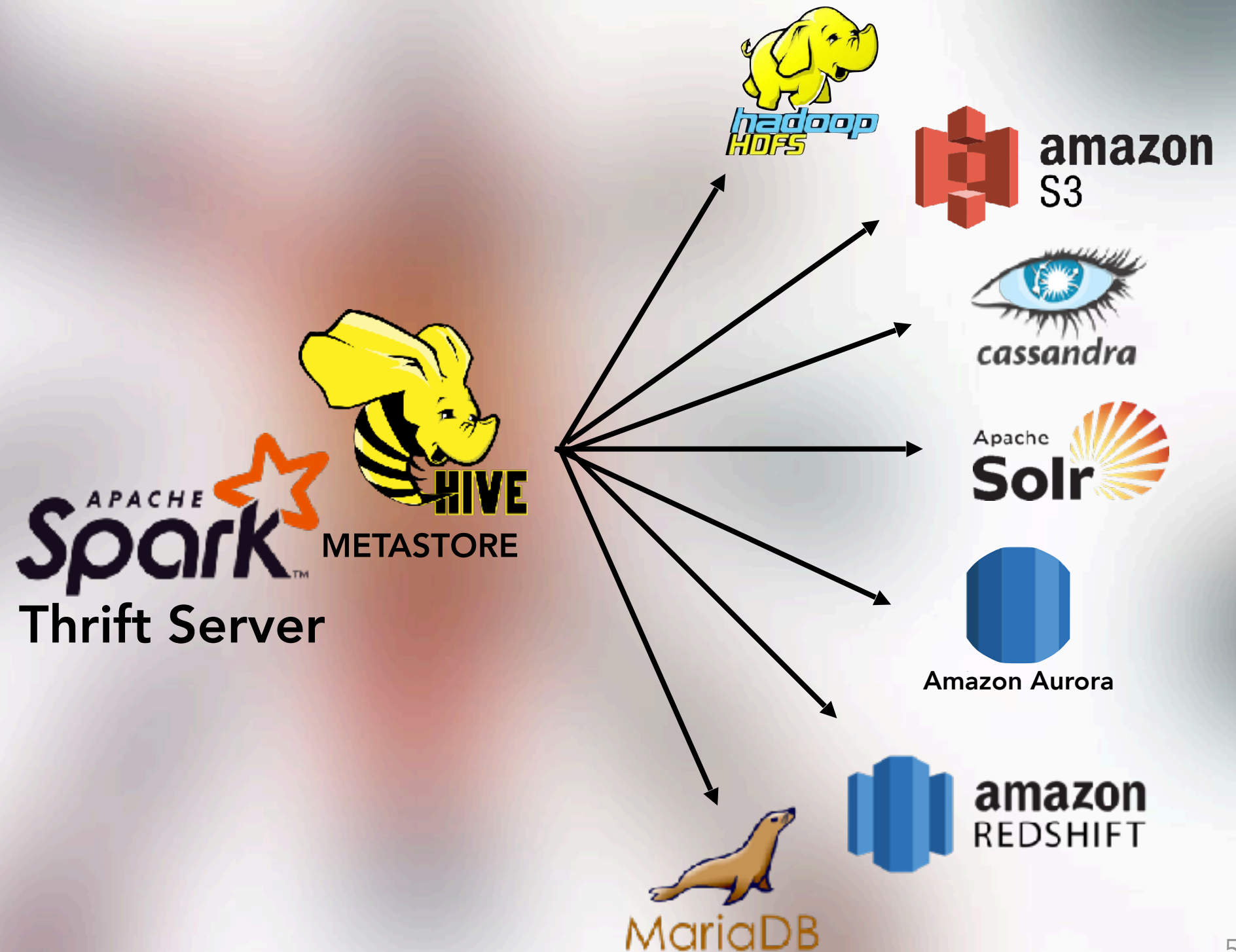
QUERY

QUERYING DATA



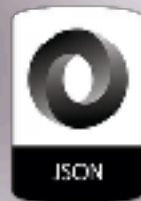
QUERY

QUERYING DATA



QUERY

QUERYING DATA



CREATE TABLE IF NOT EXISTS

```
database_name.table_name(  
  some_column STRING,  
  ...  
  dt DATE  
)  
USING json  
PARTITIONED BY (`dt`)
```



CREATE TEMPORARY VIEW table_name

```
USING org.apache.spark.sql.cassandra  
OPTIONS (  
  table "table_name",  
  keyspace "keyspace_name")
```



amazon
S3



Parquet

CREATE EXTERNAL TABLE IF NOT EXISTS

```
database_name.table_name(  
  some_column STRING...,  
  dt DATE  
)  
PARTITIONED BY (`dt`)  
USING PARQUET  
LOCATION 's3a://bucket-name/database_name/table_name'
```



amazon
REDSHIFT

CREATE TABLE IF NOT EXISTS database_name.table_name

```
using com.databricks.spark.redshift  
options (  
  dbtable 'schema.redshift_table_name',  
  tempdir 's3a://redshift-temp',  
  url 'jdbc:redshift://xxxx.redshift.amazonaws.com:5439/letgo?  
    user=xxx&password=xxx',  
  forward_spark_s3_credentials 'true')
```

QUERY

QUERYING DATA

CREATE TABLE ...
STORED AS...

VS

CREATE TABLE ...
USING [parquet,json,csv...]



70%

Higher performance!

QUERY

QUERYING DATA: BATCHES WITH SQL

1

Creating the
table

2

Inserting data

QUERYING DATA: BATCHES WITH SQL

1

Creating the
table

```
CREATE EXTERNAL TABLE IF NOT EXISTS database.some_name(  
  user_id STRING,  
  column_b STRING,  
  ...  
)  
USING PARQUET  
PARTITIONED BY (`dt` STRING)  
LOCATION 's3a://example/some_table'
```


QUERYING DATA: BATCHES WITH SQL

2

Inserting data

```
INSERT OVERWRITE TABLE database.some_name PARTITION(dt)
SELECT
user_id,
column_b,
dt
FROM other_table
...
```

QUERY

QUERYING DATA: BATCHES WITH SQL

Viewing 1 to 200

Name ↑	Last modified ↑	Size ↑	Storage class ↑
part-00000-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:38 AM GMT+0100	6.5 MB	Standard
part-00001-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:39 AM GMT+0100	10.6 MB	Standard
part-00002-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:39 AM GMT+0100	10.5 MB	Standard
part-00003-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:40 AM GMT+0100	8.9 MB	Standard
part-00004-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:41 AM GMT+0100	9.3 MB	Standard
part-00005-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:42 AM GMT+0100	7.5 MB	Standard
part-00006-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:42 AM GMT+0100	10.0 MB	Standard
part-00007-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:43 AM GMT+0100	7.2 MB	Standard
part-00008-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:44 AM GMT+0100	9.6 MB	Standard
part-00009-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:45 AM GMT+0100	9.6 MB	Standard
part-00010-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:46 AM GMT+0100	8.7 MB	Standard
part-00011-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000			
part-00012-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000			

Problem?



QUERY

QUERYING DATA: BATCHES WITH SQL

Viewing 1 to 200

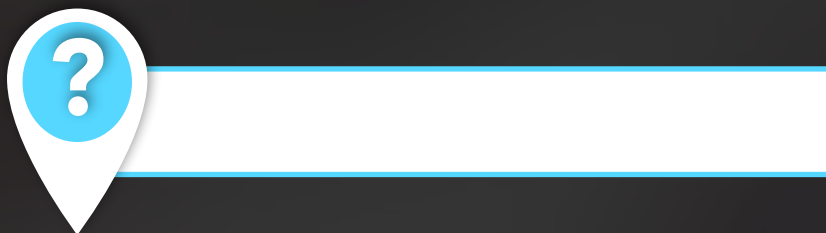
Name ↑	Last modified ↑	Size ↑	Storage class ↑
part-00000-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:38 AM GMT+0100	6.5 MB	Standard
part-00001-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:39 AM GMT+0100	10.6 MB	Standard
part-00002-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:39 AM GMT+0100	10.5 MB	Standard
part-00003-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:40 AM GMT+0100	8.9 MB	Standard
part-00004-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:41 AM GMT+0100	9.3 MB	Standard
part-00005-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:42 AM GMT+0100	7.5 MB	Standard
part-00006-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:42 AM GMT+0100	10.0 MB	Standard
part-00007-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:43 AM GMT+0100	7.2 MB	Standard
part-00008-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:44 AM GMT+0100	9.6 MB	Standard
part-00009-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:45 AM GMT+0100	9.6 MB	Standard
part-00010-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:46 AM GMT+0100	8.7 MB	Standard
part-00011-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:47 AM GMT+0100	9.2 MB	Standard
part-00012-d0f97fca-b705-49e1-b3d1-b8ab2af5b174-c000	Mar 5, 2018 10:58:48 AM GMT+0100	12.9 MB	Standard

200 files because default value of
"spark.sql.shuffle.partition"

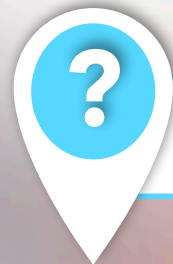
QUERYING DATA: BATCHES WITH SQL



```
INSERT OVERWRITE TABLE database.some_name PARTITION(dt)
SELECT
user_id,
column_b,
dt
FROM other_table
...
```



QUERYING DATA: BATCHES WITH SQL



DISTRIBUTE BY (dt):

Only one file not Sorted



CLUSTERED BY (dt, user_id, column_b):

Multiple files



DISTRIBUTE BY (dt) SORT BY (user_id, column_b):

Only one file sorted by user_id, column_b.


Good for joins using this properties.

QUERYING DATA: BATCHES WITH SQL

```
INSERT OVERWRITE TABLE database.some_name  
PARTITION(dt)  
SELECT  
user_id,  
column_b,  
dt  
FROM other_table  
...  
DISTRIBUTE BY (dt) SORT BY (user_id)
```


QUERY

QUERYING DATA: BATCHES WITH SQL

				Viewing 1 to 1
Name	Last modified	Size	Storage class	
 part-00020-d8690/red-dbc5-4c5f-a399-51a72852f980.c000	Mar 5, 2018 12:35:38 PM GMT+0100	2.5 GB	Standard	
				Viewing 1 to 1

INGEST

Data Ingestion

Storage

PROCESS

Stream

Batch

DISCOVER

Query

Data exploitation

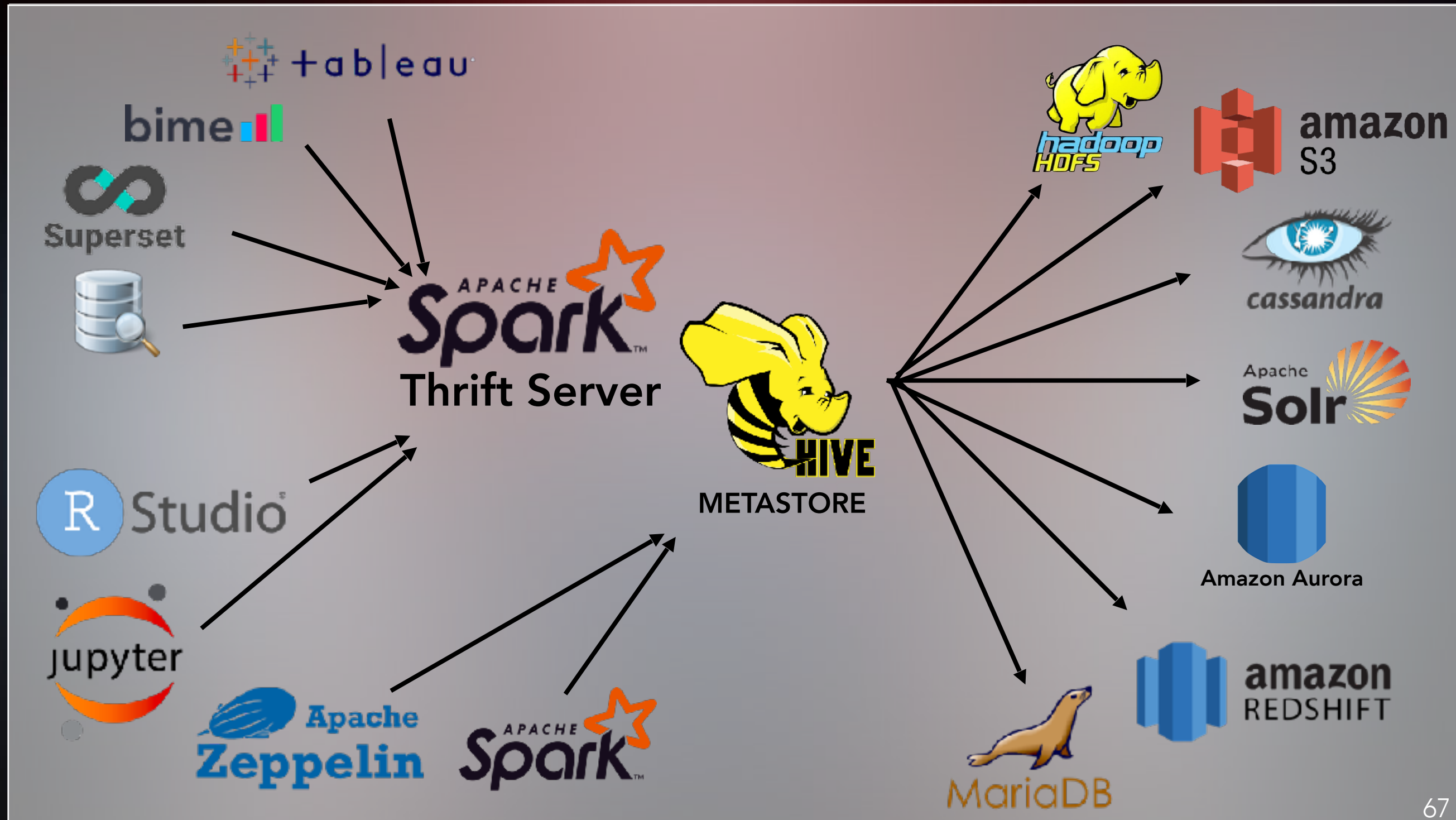
Orchestration

DATA EXPLOITATION

OUR ANALYTICAL STACK



OUR ANALYTICAL STACK



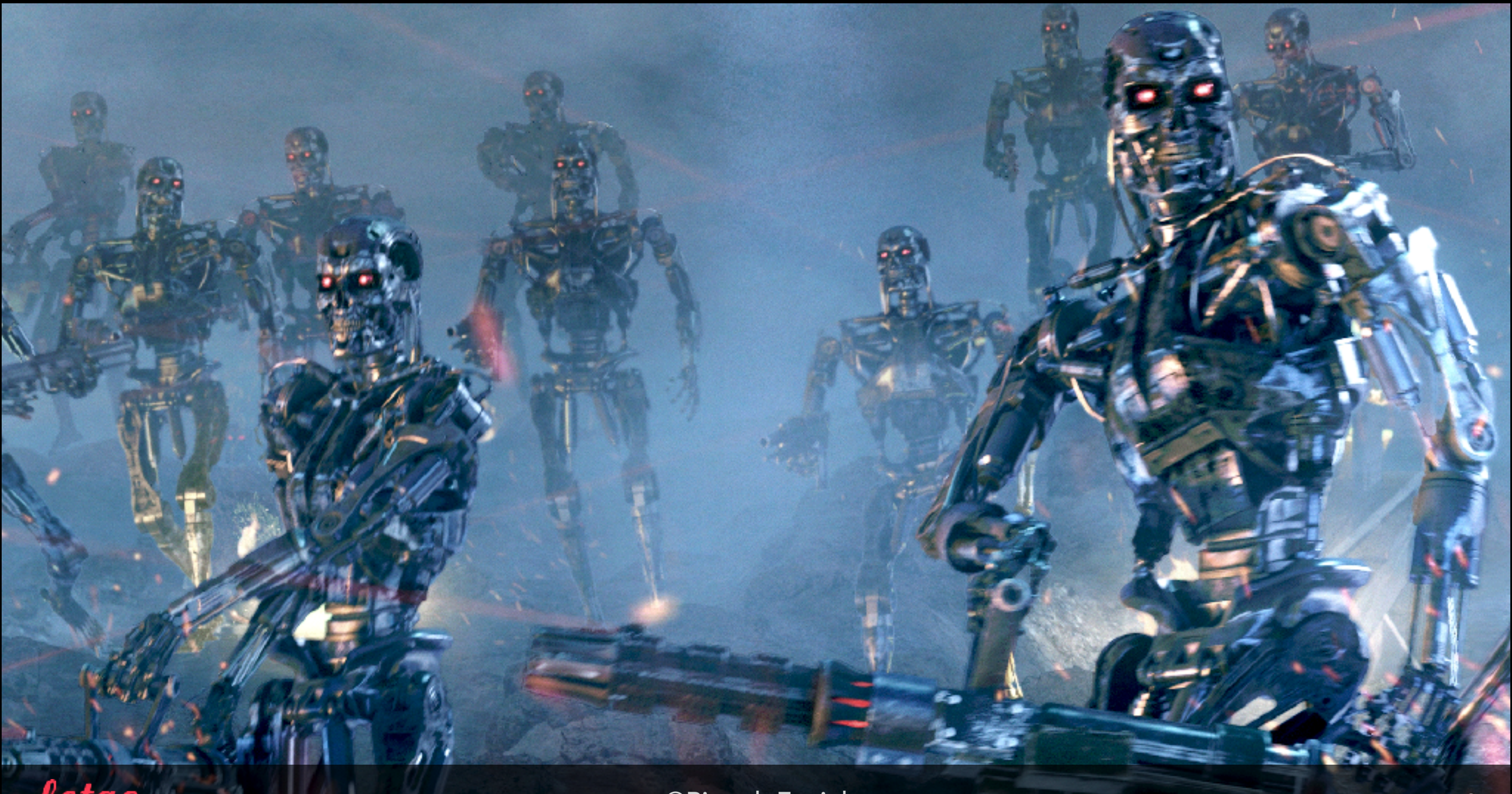
DATA EXPLOITATION

DATA SCIENTISTS TEAM?



DATA EXPLOITATION

DATA SCIENTISTS AS I SEE THEM



DATA EXPLOITATION

DATA SCIENTISTS SINS

**COMPUTER
MALFUNCTION**

DATA SCIENTISTS SINS

The screenshot shows the Hadoop Browse Directory interface. The top navigation bar includes links for Hadoop, Overview, Datanodes, Snapshot, Startup Progress, and Utilities. A red box highlights the top right corner, showing '128 MB' and '1 de 3.701'. Below the navigation bar, the 'Browse Directory' section shows a path '/user/hive/warehouse/dbtmp/user/retention' and a 'Go!' button. A table lists files with columns: Permission, Owner, Group, Size, Last Modified, Replication, Block Size, and Name. The 'Block Size' column shows '128 MB' for several files, which is highlighted by a speech bubble saying 'Too many small files!'. A red warning light icon is also present.

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	admin	hive	0 B	Wed Mar 14 12:56:56 +0100 2018	3	128 MB	_SUCCESS
-rw-r--r--	root	hive	1.8 MB	Tue Mar 13 17:34:13 +0100 2018	3	128 MB	part-00000-2143d4d-8faa-432c-8e00-c94a9e6a9e01-c000.snappy.parquet
-rw-r--r--	root	hive	2.71 MB	Wed Mar 14 11:15:29 +0100 2018	3	128 MB	part-00000-4de6c2b1-a21c-4774-a674-354b95b036ec-c000.snappy.parquet
-rw-r--r--	root	hive	2.55 MB	Tue Mar 13 20:20:36 +0100 2018	3	128 MB	part-00000-4de6c2b1-a21c-4774-a674-354b95b036ec-c000.snappy.parquet
-rw-r--r--	root	hive	1.99 MB	Tue Mar 13 18:16:14 +0100 2018	3	128 MB	part-00000-4de6c2b1-a21c-4774-a674-354b95b036ec-c000.snappy.parquet
-rw-r--r--	root	hive	2.89 MB	Tue Mar 13 22:55:18 +0100 2018	3	128 MB	part-00000-4de6c2b1-a21c-4774-a674-354b95b036ec-c000.snappy.parquet

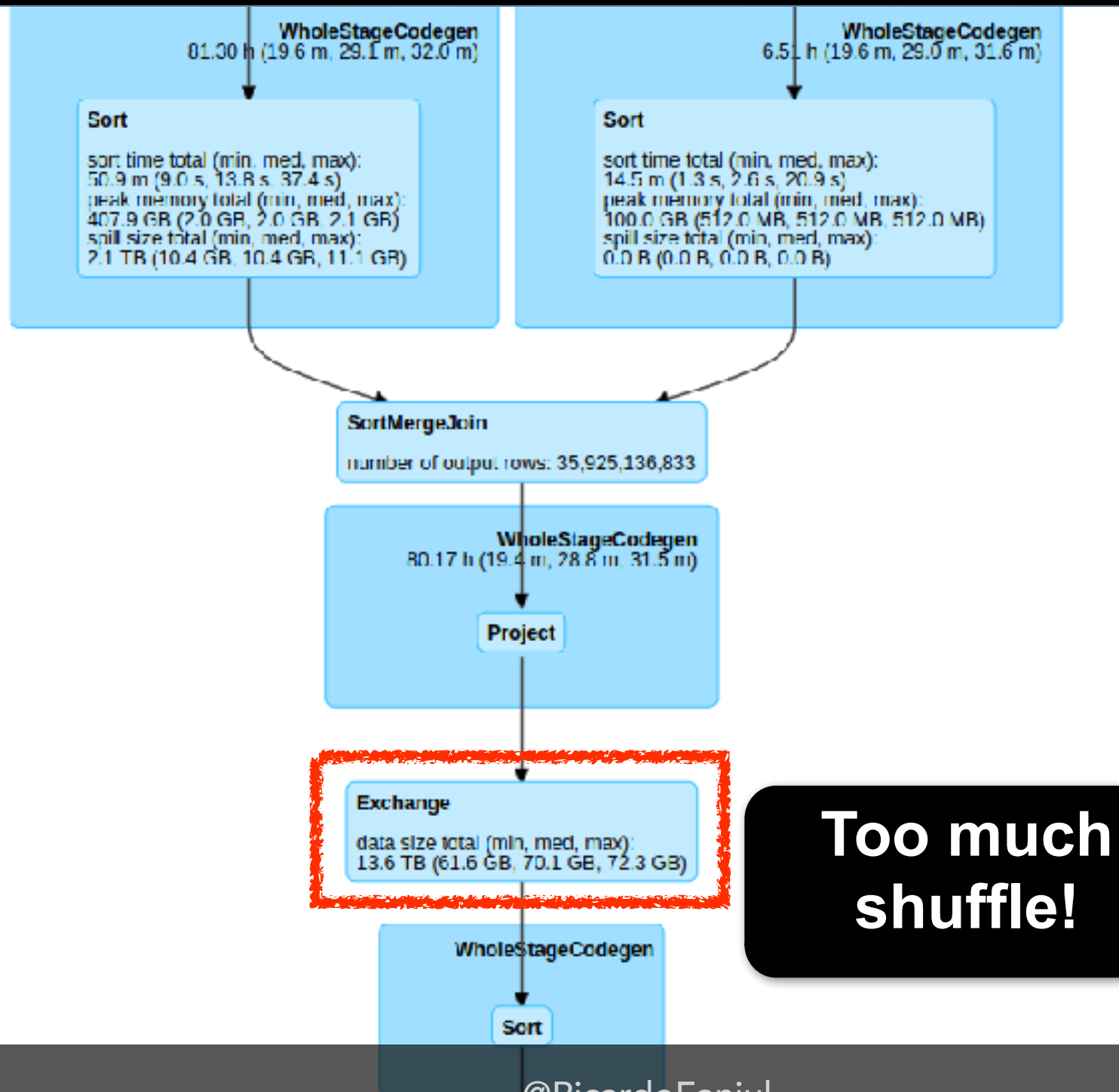
DATA SCIENTISTS SINS



Huge Query!



DATA SCIENTISTS SINS



Too much shuffle!



INGEST

Data Ingestion

Storage

PROCESS

Stream

Batch

DISCOVER

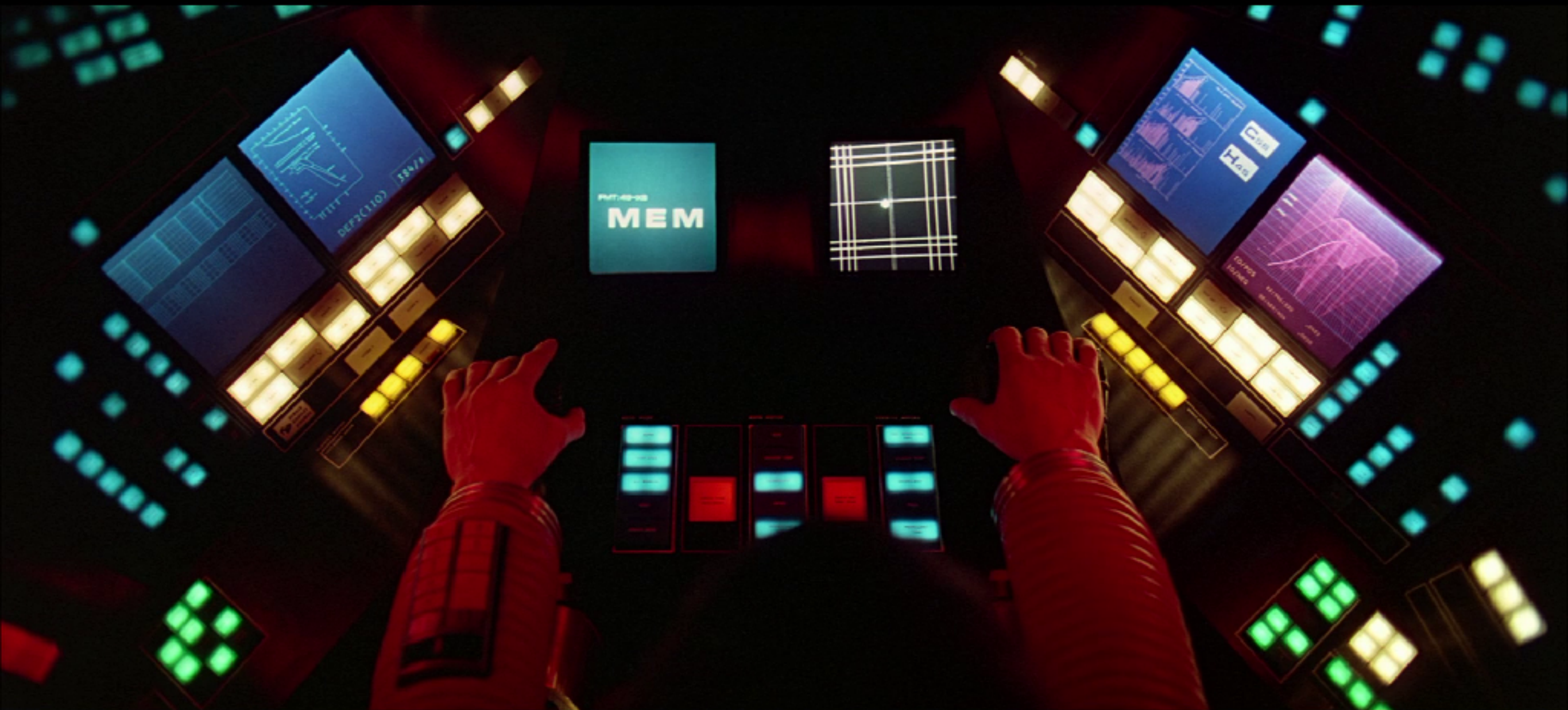
Query

Data exploitation

Orchestration

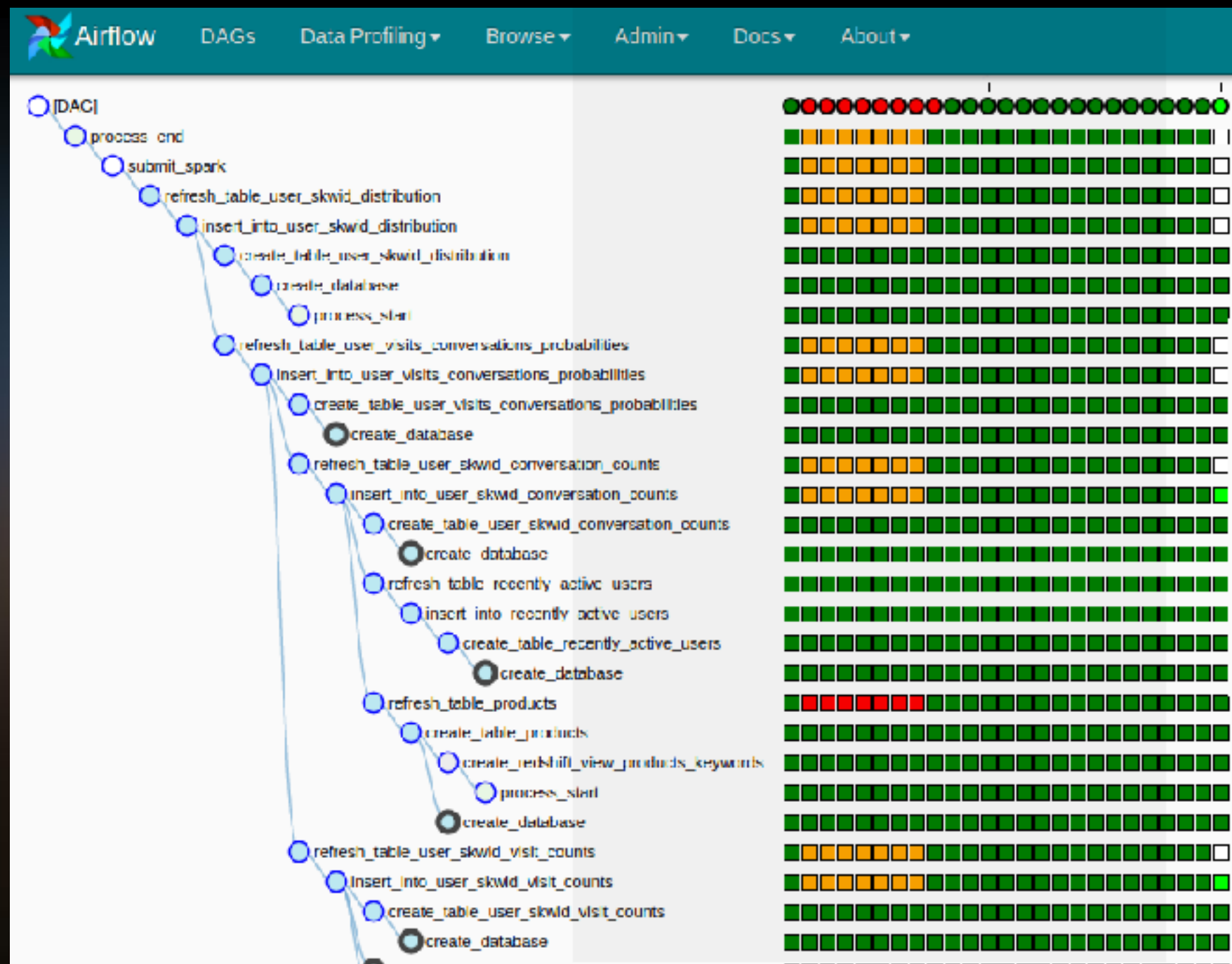
ORCHESTRATION

AIRFLOW

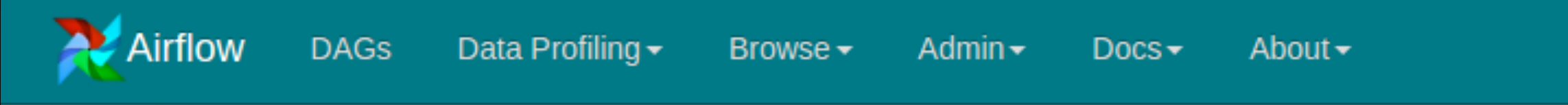


ORCHESTRATION

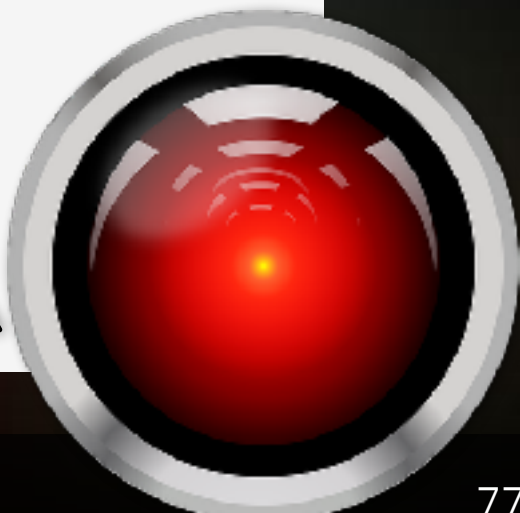
AIRFLOW



AIRFLOW



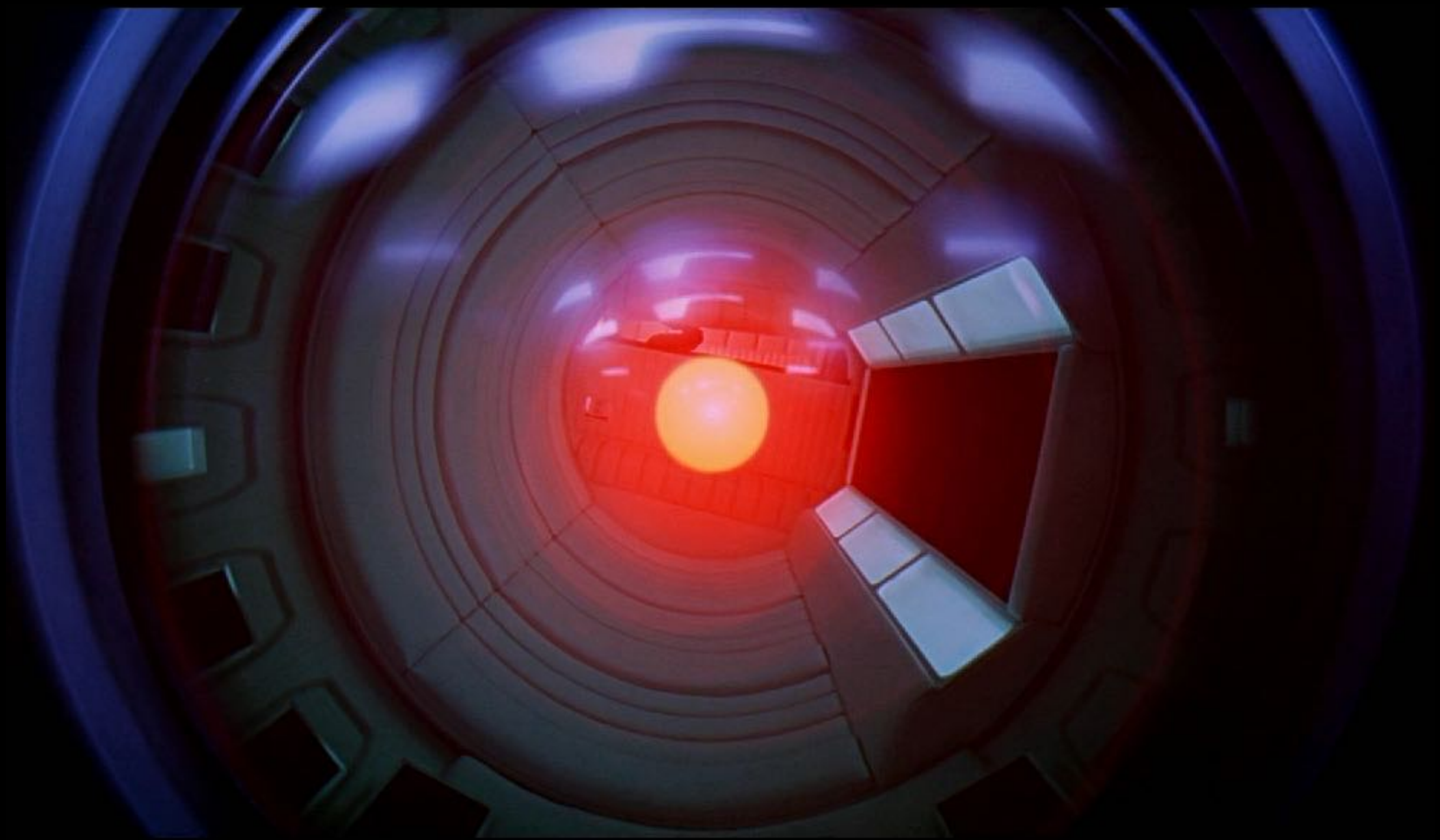
```
127 )
128
129 create_table_user_skwid_distribution = HiveOperatorLetgo(
130     task_id='create_table_user_skwid_distribution',
131     dag=dag,
132     sql="personalization/create_table_user_skwid_distribution.sql",
133     parameters={
134         's3_location': path.join(location, 'user_skwid_distribution')
135     }
136 )
137
138 insert_into_recently_active_users = HiveOperatorLetgo(
139     task_id='insert_into_recently_active_users',
140     dag=dag,
141     sql="personalization/insert_into_recently_active_users.sql",
142     parameters={
143         'current_ts': "{{ ts_nodash }}",
144         'lookback': lookback
145     }
146 )
```



I'm happy!!



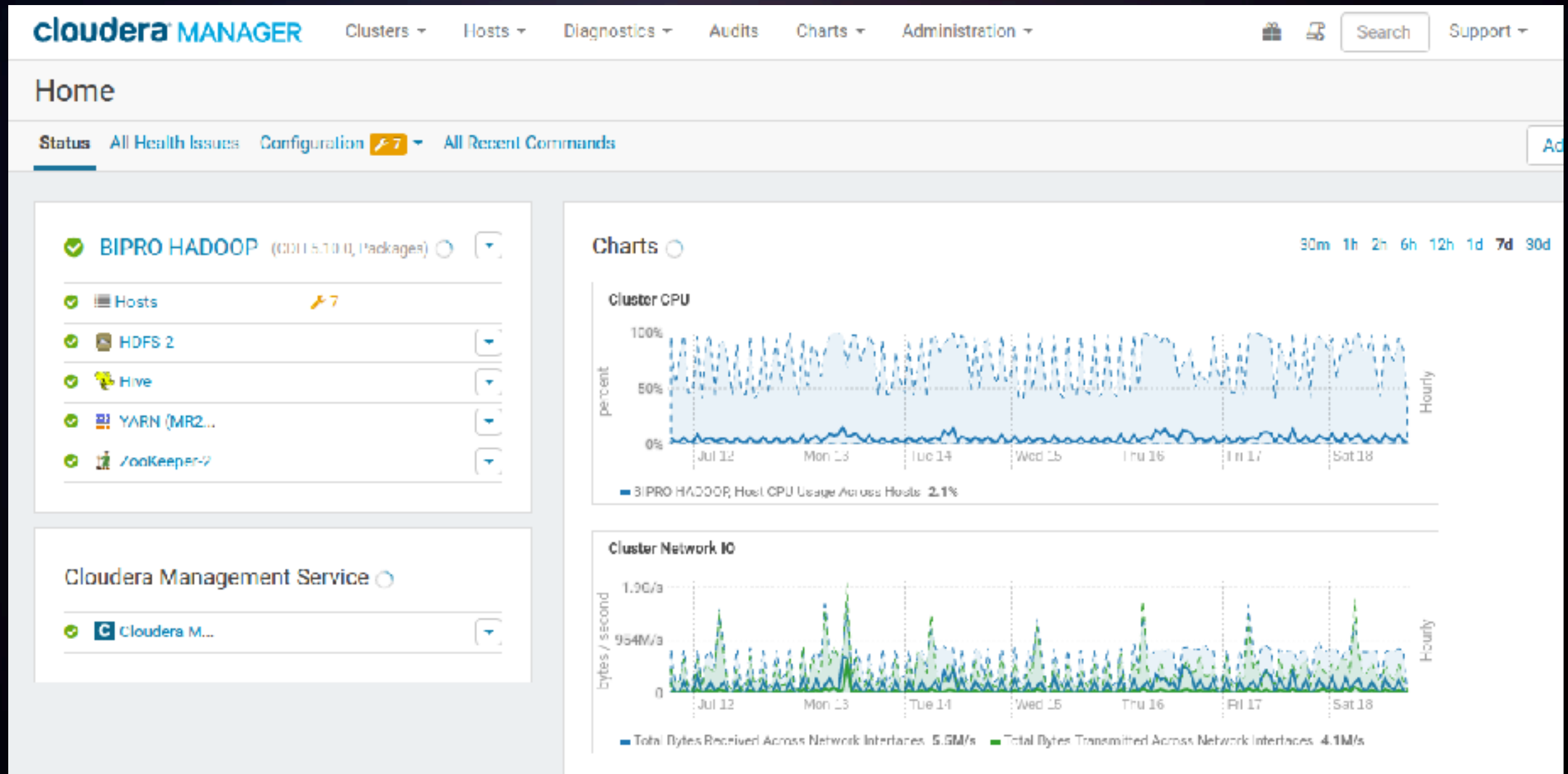
MOVING TO STATELESS CLUSTER



I'M SORRY RICARDO. I'M
AFRAID I CAN'T DO THAT.

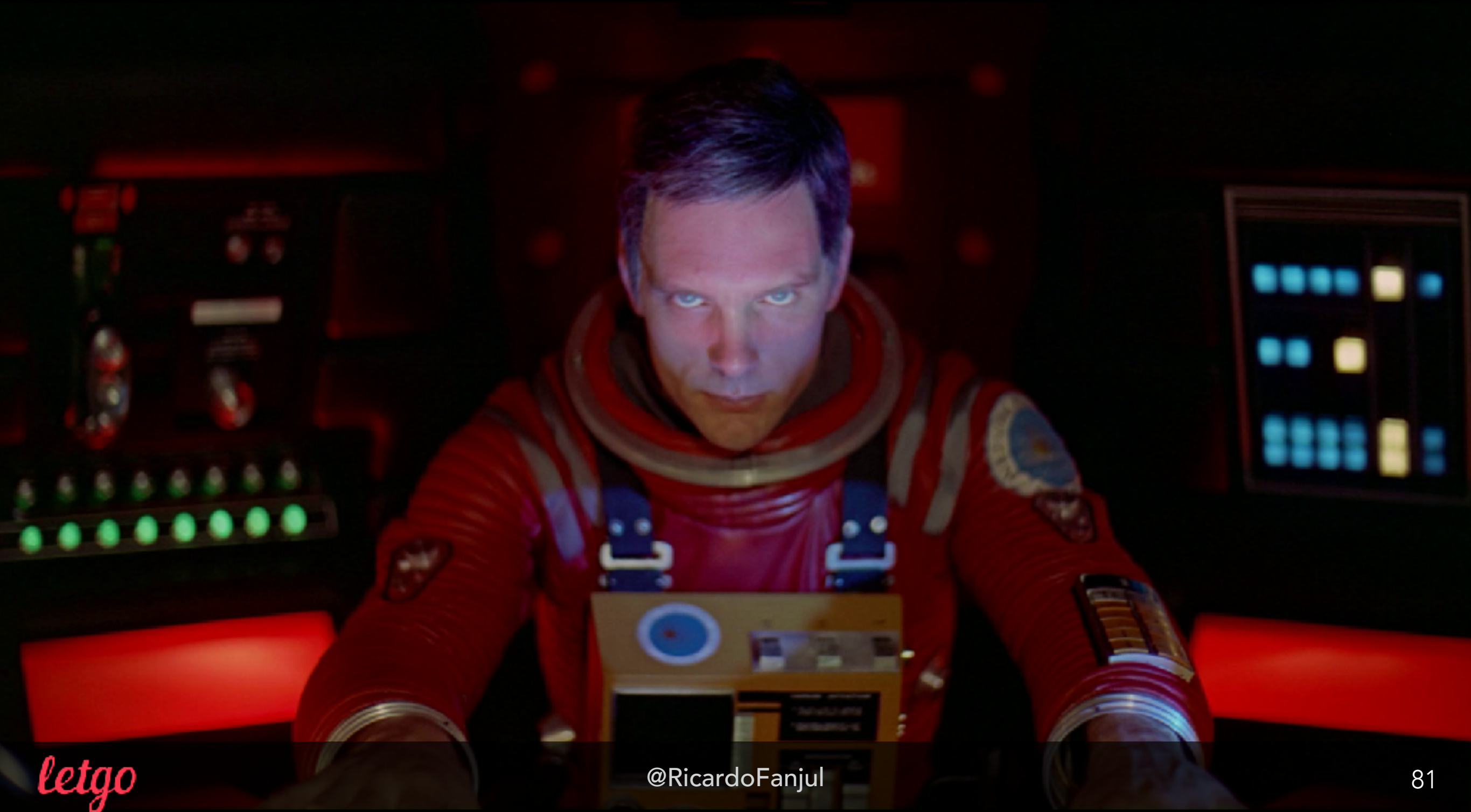
MOVING TO STATELESS CLUSTER

PLATFORM LIMITATIONS



MOVING TO STATELESS CLUSTER

PLANNING THE SOLUTION



MOVING TO STATELESS CLUSTER

PLANNING THE SOLUTION



amazon
S3

MOVING TO STATELESS CLUSTER

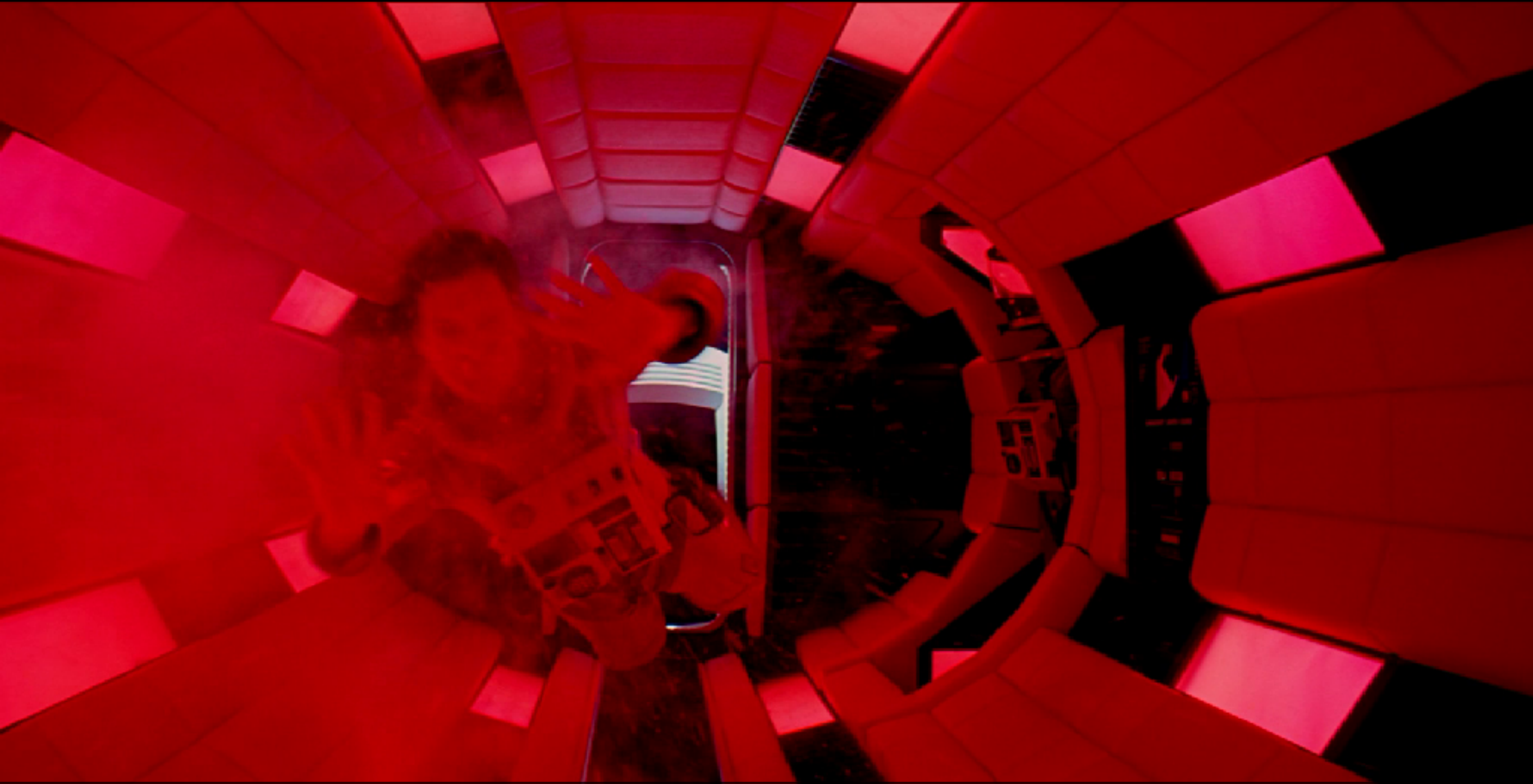
PLANNING THE SOLUTION



Cloudbreak

MOVING TO STATELESS CLUSTER

A LONG AND WINDING PROCESS



MOVING TO STATELESS CLUSTER

A LONG AND WINDING PROCESS

DATA Software project

Data Scrum Board

Backlog

Active sprints

Reports

Releases

Issues and filters →

Pages

Components

Add item

Project settings

71 Issues in this epic

DATA-331	Research Hive metastore deployment options	✓	DONE
DATA-301	List of errors made in Cloudera cluster	✓	DONE
DATA-520	Test node labels configuration in HDP3 cluster	✓	DONE
DATA-385	Create DNSs for Hadoop components	✓	DONE
DATA-39	Create a IAM role for production environment	✓	DONE
DATA-431	Migrate Kafka offsets from HDFS to S3	✓	DONE
DATA-153	Execute Spark Streaming Job in DEV environment	✓	DONE
DATA-23	HDP 3.0 / Hadoop 3.0	✓	DONE
DATA-140	Add spark.yarn.archive for spark-defaults.conf in puppet manifest	✓	DONE
DATA-456	Spark job history and logs are not working properly	✗	DONE
DATA-455	Configure inbound rules for Security groups in DEV environment	✓	DONE
DATA-148	Run some Airflow/STS Job In DEV	✓	DONE

There are no tickets in this sprint

Development

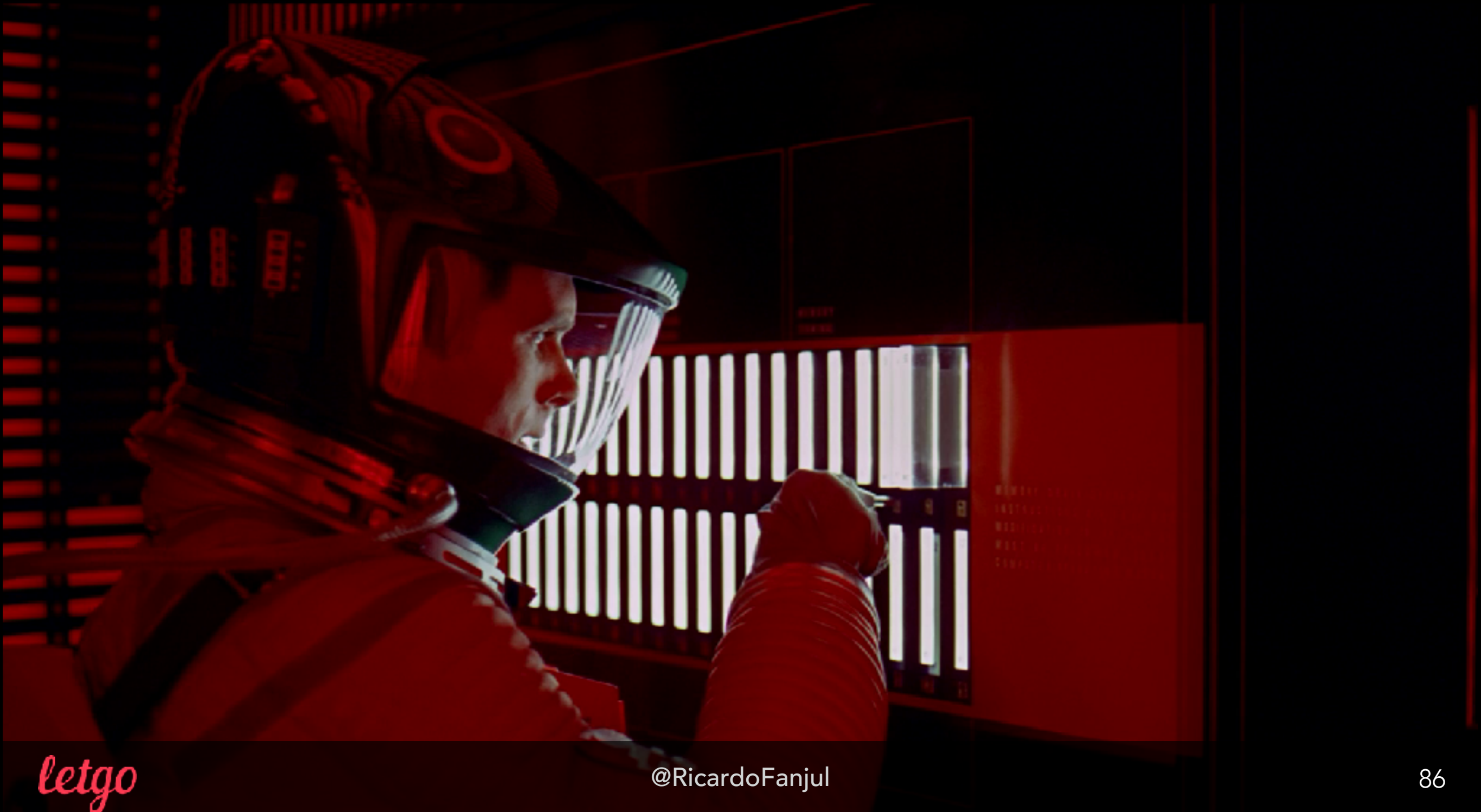
Create branch

Agile

View on Board

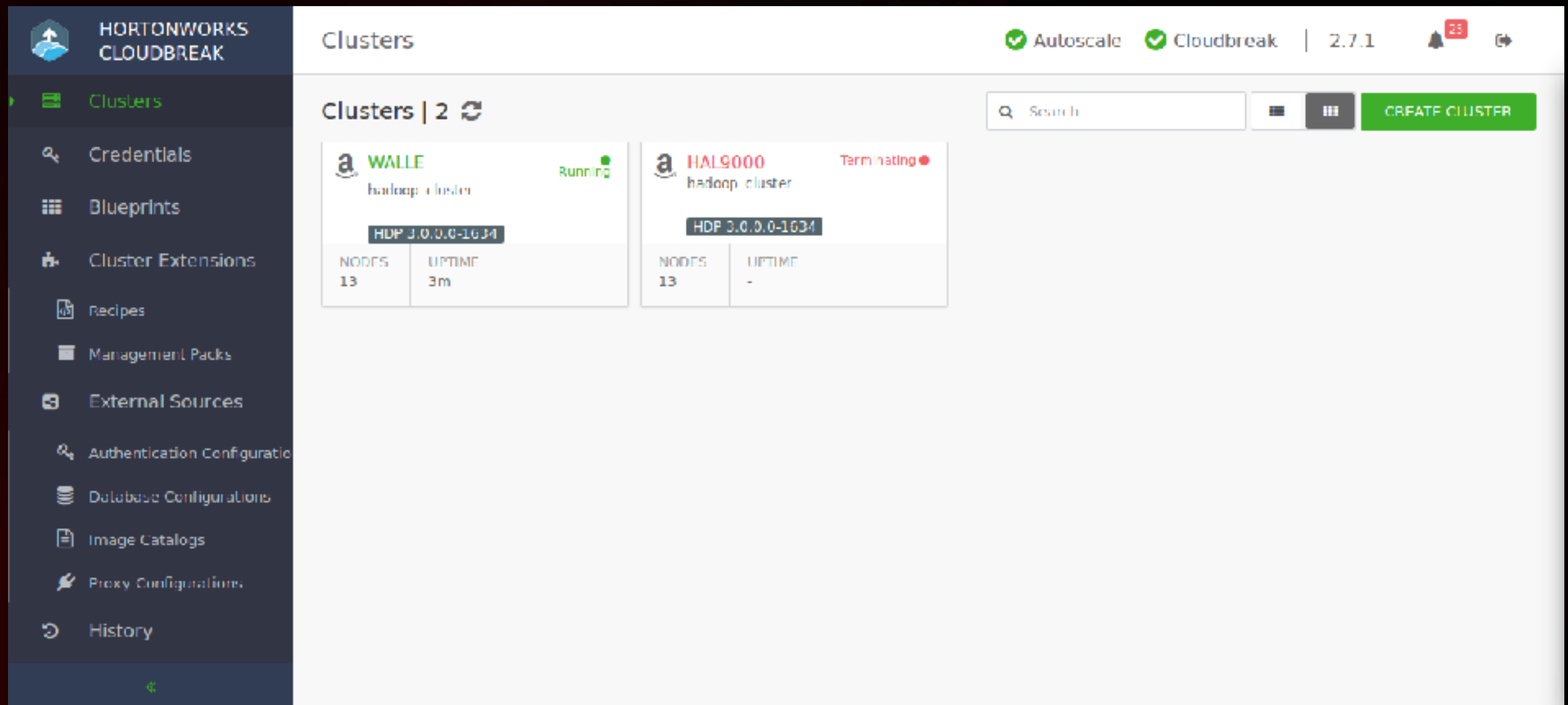
MOVING TO STATELESS CLUSTER

NEW CAPABILITIES OF THE PLATFORM



MOVING TO STATELESS CLUSTER

NEW CAPABILITIES OF THE PLATFORM



The screenshot displays the Hortonworks Cloudbreak interface. The left sidebar contains navigation links: Clusters, Credentials, Blueprints, Cluster Extensions, Recipes, Management Packs, External Sources, Authentication Configuration, Database Configurations, Image Catalogs, Proxy Configurations, and History. The main content area is titled 'Clusters' and shows a list of two Hadoop clusters. The first cluster, 'WALLE', is in a 'Running' state with 13 nodes and an uptime of 3m. The second cluster, 'HAL9000', is in a 'Terminating' state with 13 nodes and no uptime. Both clusters are running HDP 3.0.0.0-1634. The top right of the interface shows 'Autoscale' and 'Cloudbreak' status, the version '2.7.1', and a 'CREATE CLUSTER' button.

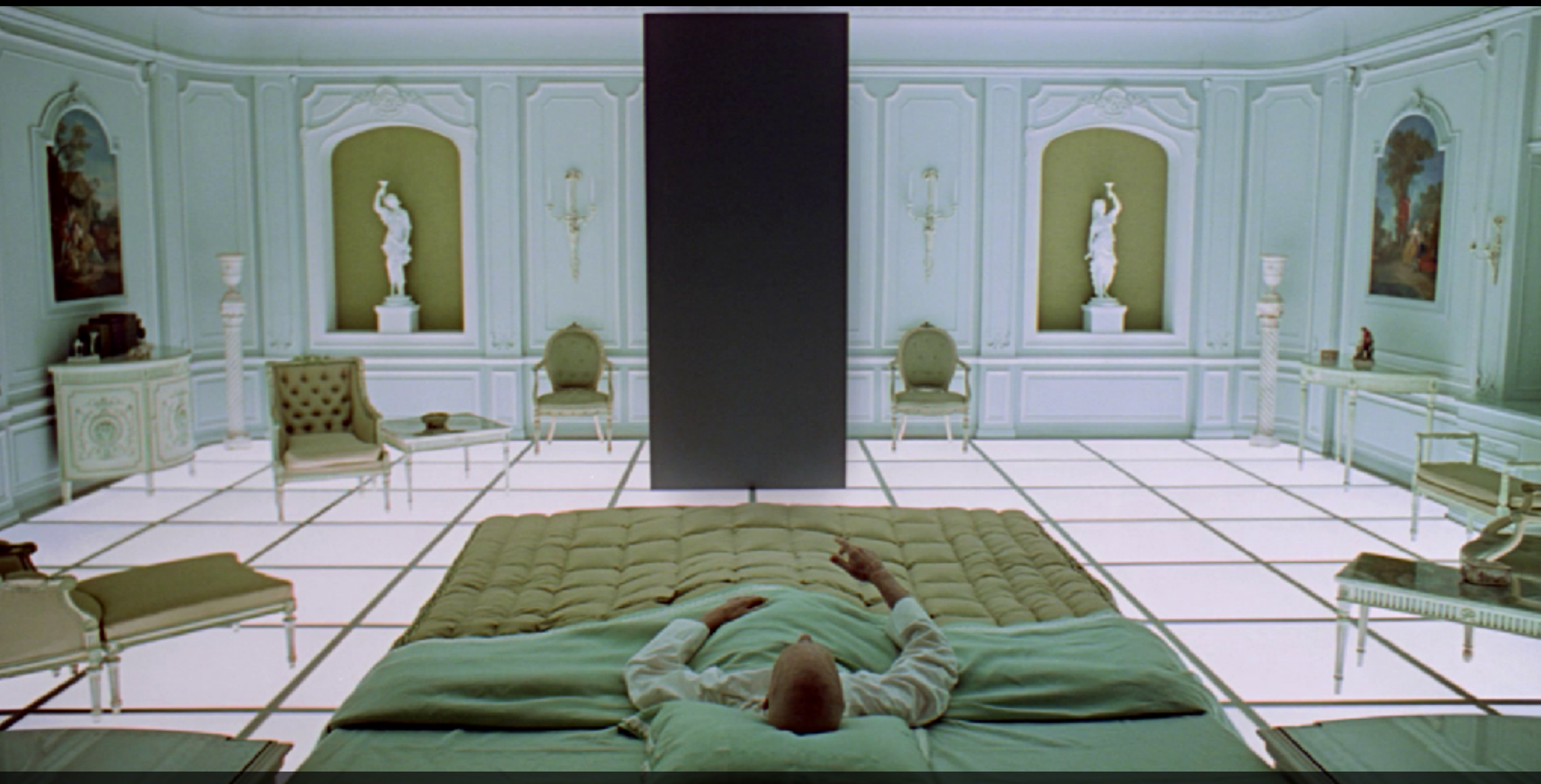
Cluster Name	Status	Nodes	Uptime	HDP Version
WALLE	Running	13	3m	HDP 3.0.0.0-1634
HAL9000	Terminating	13	-	HDP 3.0.0.0-1634



JUPITER AND BEYOND THE INFINITE

JUPITER AND BEYOND THE INFINITE

SOMETHING WONDERFUL WILL
HAPPEN





<https://youtu.be/CInMDMuSFwc>

“The only way to discover the limits of the possible
is to go beyond them into the impossible.”

–ARTHUR C. CLARKE



THE FUTURE...?

DO YOU WANT TO JOIN US?

