Morticia: Visualize and Debug Complex Spark Workflows

Jacob Perkins
Stitchfix

SPARK SUMMIT 2016
DATA SCIENCE AND ENGINEERING AT SCALE
JUNE 6-8, 2016 SAN FRANCISCO

Who am I?

What do we do?

- Developer enablement platform

Morticia?

- Self-service debugging of spark workflows

Why bother?

- Data scientists are not going to become spark experts

Current state of the universe...

Here's a simple query…

```
select count (distinct source)
  from test.marvel_social_graph
where target != 'CAPTAIN AMERICA'
```

Current state of the universe...

Questions:

How many input records?

What is the parallelism throughout?

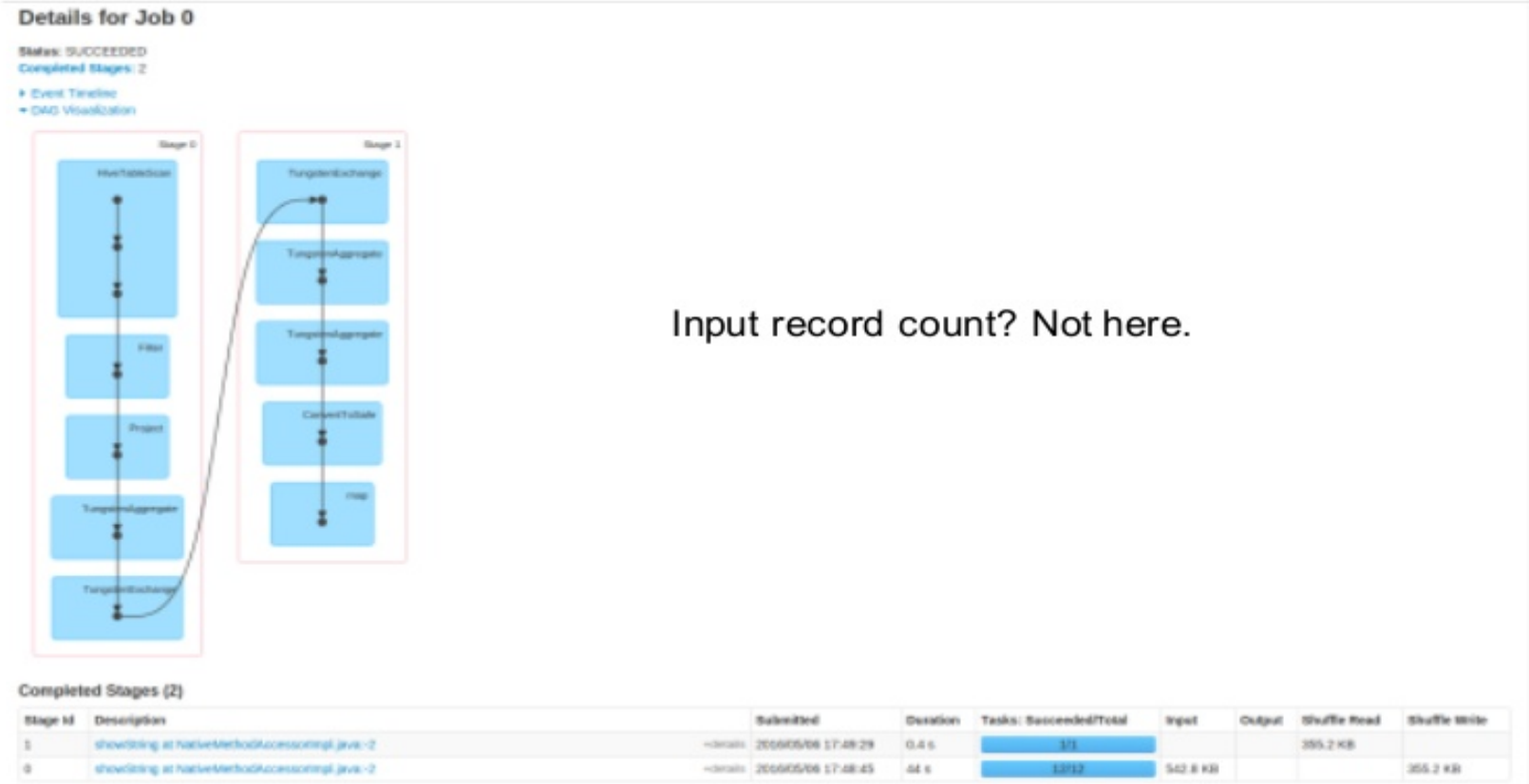
How did my query get mapped to actual work?

Logs?

Current state of the universe...

Input record count?

Current state of the universe...



Input record count? Not here.

Current state of the universe...

Details for Stage 0 (Attempt 0)

Total Time Across All Tasks: 4.1 min
Input Size / Records: 542.8 KB / 1144470 ← Input record count!
Shuffle Write: 355.2 KB / 20751

- DAG Visualization
- Show Additional Metrics
- Event Timeline

Summary Metrics for 12 Completed Tasks

Metric	Min	25th percentile	Median	75th percentile	Max
Duration	0.1 s	20 s	25 s	28 s	28 s
GC Time	0 ms	0.3 s	0.4 s	0.5 s	0.7 s
Input Size / Records	0.0 B / 44470	0.0 B / 100000	0.0 B / 100000	0.0 B / 100000	379.1 KB / 100000
Shuffle Write Size / Records	17.2 KB / 962	29.1 KB / 1702	29.5 KB / 1722	32.4 KB / 1902	34.8 KB / 2045

Aggregated Metrics by Executor

Executor ID	Address	Task Time	Total Tasks	Failed Tasks	Succeeded Tasks	Input Size / Records	Shuffle Write Size / Records
1	CANNOT FIND ADDRESS	32 s	1	0	1	0.0 B / 100000	28.1 KB / 1641
10	CANNOT FIND ADDRESS	22 s	3	0	3	542.8 KB / 244470	84.3 KB / 4909
2	CANNOT FIND ADDRESS	29 s	1	0	1	0.0 B / 100000	29.5 KB / 1714
3	CANNOT FIND ADDRESS	32 s	1	0	1	0.0 B / 100000	30.1 KB / 1760

Current state of the universe...

Number of tasks? Count!

Tasks

Index ▲	ID	Attempt	Status	Locality Level	Executor ID / Host	Launch Time	Duration	GC Time	Input Size / Records	Write Time	Shuffl
0	0	0	SUCCESS	RACK_LOCAL	5 / ip-10-0-75-34.ec2.internal	2016/05/06 17:48:53	23 s	0.3 s	0.0 B (hadoop) / 100000	12 ms	29.1 K
1	1	0	SUCCESS	RACK_LOCAL	2 / ip-10-0-75-32.ec2.internal	2016/05/06 17:48:54	26 s	0.4 s	0.0 B (hadoop) / 100000	14 ms	29.5 K
2	2	0	SUCCESS	RACK_LOCAL	9 / ip-10-0-83-53.ec2.internal	2016/05/06 17:48:54	23 s	0.3 s	0.0 B (hadoop) / 100000	10 ms	29.4 K
3	3	0	SUCCESS	RACK_LOCAL	10 / ip-10-0-75-36.ec2.internal	2016/05/06 17:48:54	19 s	0.3 s	0.0 B (hadoop) / 100000	11 ms	32.4 K
4	4	0	SUCCESS	RACK_LOCAL	4 / ip-10-0-75-30.ec2.internal	2016/05/06 17:48:54	25 s	0.4 s	0.0 B (hadoop) / 100000	10 ms	34.8 K
5	5	0	SUCCESS	RACK_LOCAL	1 / ip-10-0-86-202.ec2.internal	2016/05/06 17:48:55	28 s	0.6 s	0.0 B (hadoop) / 100000	25 ms	28.1 K
6	6	0	SUCCESS	RACK_LOCAL	6 / ip-10-0-75-32.ec2.internal	2016/05/06 17:48:55	26 s	0.4 s	0.0 B (hadoop) / 100000	10 ms	28.8 K
7	7	0	SUCCESS	RACK_LOCAL	8 / ip-10-0-75-30.ec2.internal	2016/05/06 17:48:56	20 s	0.5 s	0.0 B (hadoop) / 100000	10 ms	29.2 K
8	8	0	SUCCESS	RACK_LOCAL	3 / ip-10-0-86-202.ec2.internal	2016/05/06 17:48:56	28 s	0.7 s	0.0 B (hadoop) / 100000	33 ms	30.1 K
9	9	0	SUCCESS	RACK_LOCAL	7 / ip-10-0-86-202.ec2.internal	2016/05/06 17:48:58	28 s	0.5 s	0.0 B (hadoop) / 100000	10 ms	32.1 K
10	10	0	SUCCESS	RACK_LOCAL	10 / ip-10-0-75-36.ec2.internal	2016/05/06 17:49:15	0.2 s	27 ms	379.1 KB (hadoop) / 100000		34.7 K
11	11	0	SUCCESS	RACK_LOCAL	10 / ip-10-0-75-36.ec2.internal	2016/05/06 17:49:16	0.1 s		163.6 KB (hadoop) / 44470		17.2 K

Current state of the universe...

Number of tasks? Count!

Tasks

Index ▲	ID	Attempt	Status	Locality Level	Executor ID / Host	Laur
0	12	0	SUCCESS	PROCESS_LOCAL	1 / ip-10-0-86-202.ec2.internal	2016

Current state of the universe...

Details for Query 0

Submitted Time: 2016/05/06 17:48:43
Duration: 46 s
Succeeded Jobs: 0
Detail:

```
== Parsed Logical Plan ==
Limit 21
 Aggregate [COUNT(DISTINCT source#0) AS _c0#2L]
  Filter NOT (target#1 = CAPTAIN AMERICA)
   MetastoreRelation test, marvel_social_graph, None

== Analyzed Logical Plan ==
_c0: bigint
Limit 21
 Aggregate [COUNT(DISTINCT source#0) AS _c0#2L]
  Filter NOT (target#1 = CAPTAIN AMERICA)
   MetastoreRelation test, marvel_social_graph, None

== Optimized Logical Plan ==
Limit 21
 Aggregate [COUNT(DISTINCT source#0) AS _c0#2L]
  Project [source#0]
   Filter NOT (target#1 = CAPTAIN AMERICA)
    MetastoreRelation test, marvel_social_graph, None

== Physical Plan ==
Limit 21
 ConvertToSafe
  TungstenAggregate(key=[], functions=[(count(source#0),mode=Complete,isDistinct=true)], output=[_c0#2L])
   TungstenAggregate(key=[source#0], functions=[], output=[source#0])
    TungstenExchange SinglePartition
     TungstenAggregate(key=[source#0], functions=[], output=[source#0])
      Project [source#0]
       Filter NOT (target#1 = CAPTAIN AMERICA)
        HiveTableScan [source#0,target#1], (MetastoreRelation test, marvel_social_graph, None)

Code Generation: true
```

How did my query map to actual work? Uhhh

Current state of the universe...

Logs?

You're on your own

Enter Morticia...

- Interactive, coherent, unified view
- Logical information
- Status
- Archival

Morticia

Search

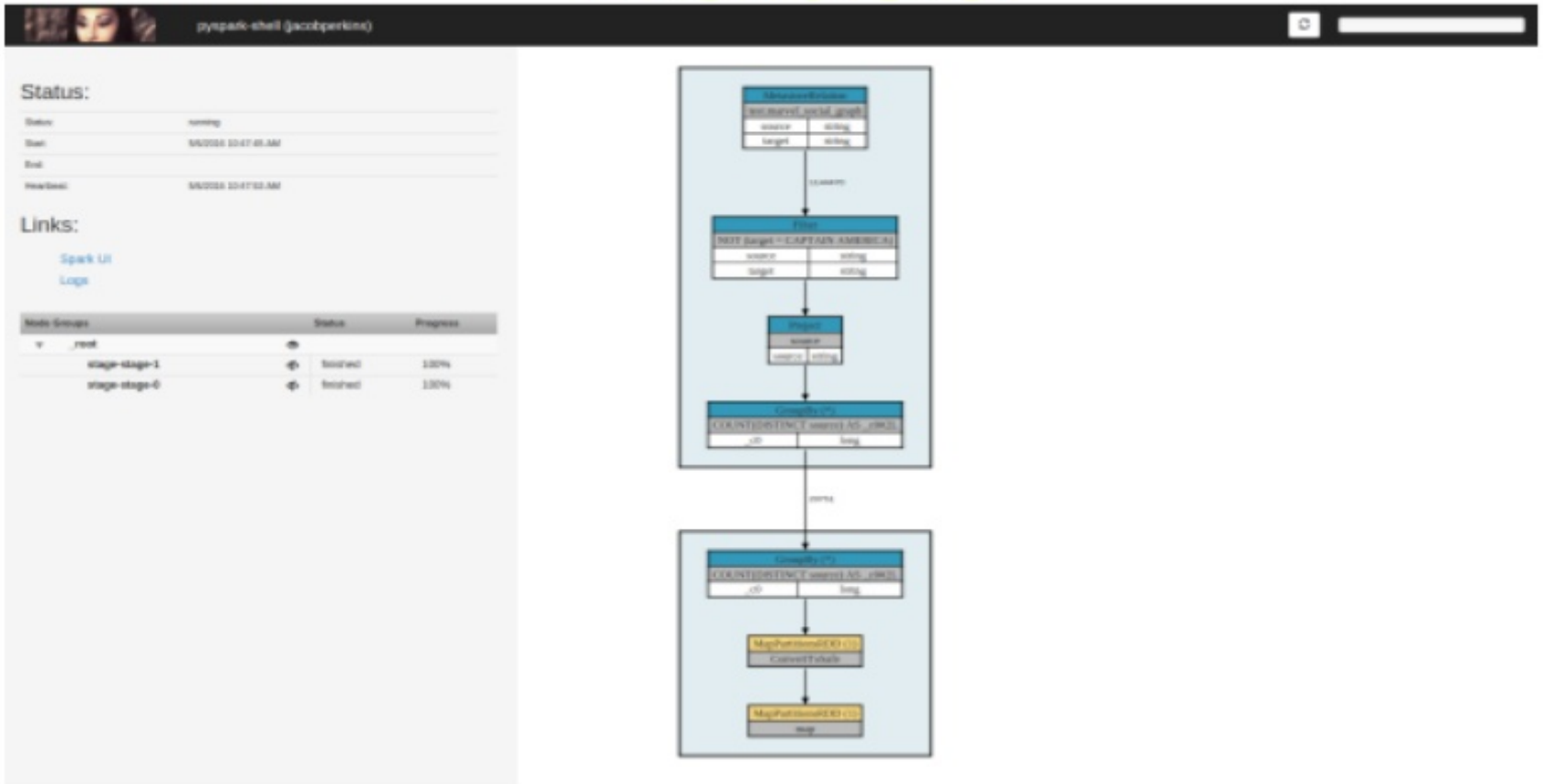
All

User	Name	Created At	Updated At	Progress
brad	pyspark-shell	5/6/2016, 11:11:56 AM	5/6/2016, 11:13:30 AM	
jacobperkins	pyspark-shell	5/6/2016, 10:48:46 AM	5/6/2016, 10:49:30 AM	
jacobperkins	pyspark-shell	5/6/2016, 7:15:50 AM	5/6/2016, 10:47:42 AM	100%
tomcat	res_logs_20160131_20160201	5/6/2016, 9:53:13 AM	5/6/2016, 9:55:43 AM	100%
tomcat	res_logs_20160116_20160130	5/6/2016, 9:49:03 AM	5/6/2016, 9:52:11 AM	100%
tomcat	res_logs_20160101_20160115	5/6/2016, 9:45:00 AM	5/6/2016, 9:48:06 AM	100%
tomcat	res_logs_20160101_20160107	5/6/2016, 9:18:20 AM	5/6/2016, 9:21:42 AM	100%
tomcat	inventory_by_hizzy	5/6/2016, 8:38:40 AM	5/6/2016, 8:38:41 AM	100%
tomcat	box:sbox:anaplan_on_order_upload_date	5/6/2016, 8:37:40 AM	5/6/2016, 8:37:50 AM	100%
tomcat	box:sbox:anaplan_sales_upload_date	5/6/2016, 8:25:35 AM	5/6/2016, 8:35:55 AM	100%

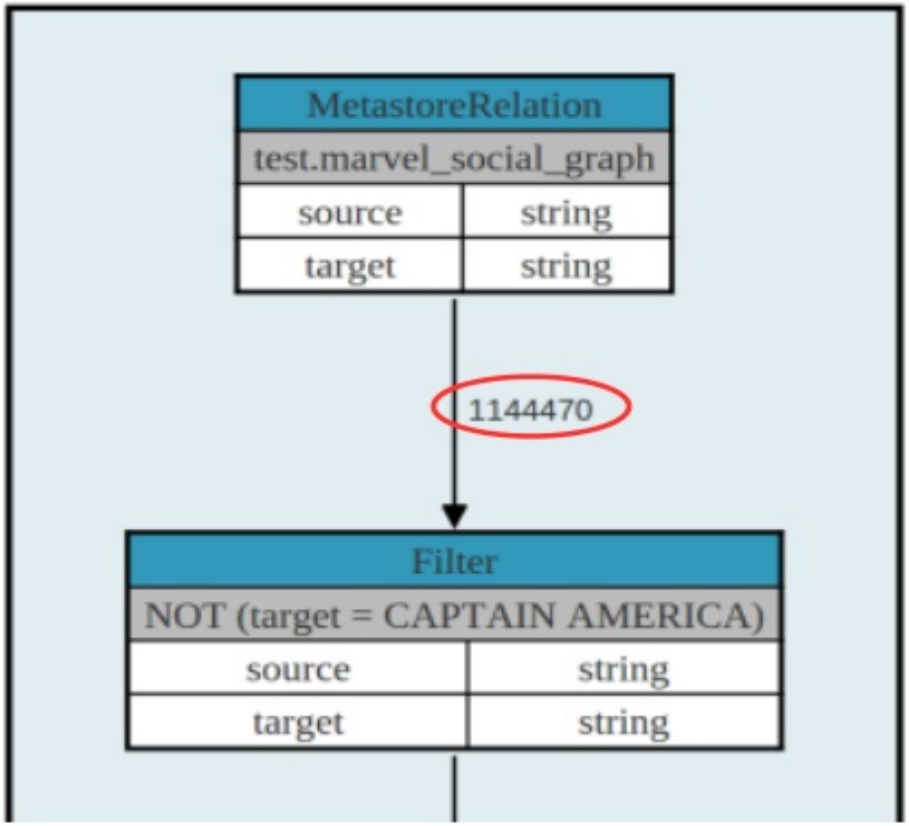
1 - 10 of 2091 items 10 Per Page

Page 1 of 210

Morticia


pyspark-shell (jacobperkins)
Status:
Links:
Spark UI
Logs
Node Groups
Status
Progress
_root
stage-stage-1
stage-stage-0
finished
100%
finished
100%

Morticia


MetastoreRelation
test.marvel_social_graph
source
string
target
string
1144470
Filter
NOT (target = CAPTAIN AMERICA)
source
string
target
string

Morticia

stage-stage-0

Progress:

100%

Properties:

shuffle_output_records	20751
shuffle_input_bytes	
shuffle_output_bytes	363706
output_bytes	
input_bytes	555798
memory_bytes_spilled	
shuffle_input_records	
peak_execution_memory	405823488
disk_bytes_spilled	
input_records	1144470
output_records	
internal_name	showString at NativeMethodAccessorImpl.java:-2
num_tasks	12

Morticia

stage-stage-1

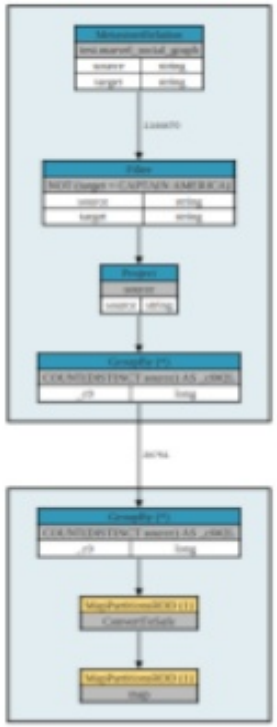
Progress:

100%

Properties:

shuffle_output_records	
shuffle_input_bytes	363706
shuffle_output_bytes	
output_bytes	
input_bytes	
memory_bytes_spilled	
shuffle_input_records	20751
peak_execution_memory	67637248
disk_bytes_spilled	
input_records	
output_records	
internal_name	showString at NativeMethodAccessorImpl.java:-2
num_tasks	1

Morticia



morticia.vertigo.stitchfix.com/workflows.html

Apps gmail git jnkns aws nexus spark keyboardt ipython crashes pyspark

Search

All

User	Name	Created At	Updated At	Progress
jacobperkins	pyspark-shell	5/9/2016, 12:36:13 PM	5/9/2016, 12:37:01 PM	
tomcat	client_size_qual	5/9/2016, 11:54:47 AM	5/9/2016, 12:04:14 PM	100%
rawdek	pyspark-shell	5/9/2016, 11:35:07 AM	5/9/2016, 11:36:21 AM	
tomcat	jas	5/9/2016, 11:05:18 AM	5/9/2016, 11:06:46 AM	100%
uraza	pyspark-shell	5/9/2016, 11:03:43 AM	5/9/2016, 11:03:43 AM	
tomcat	client_size_qual	5/9/2016, 10:03:06 AM	5/9/2016, 10:18:03 AM	100%
crinaldi	pyspark-shell	5/9/2016, 10:12:57 AM	5/9/2016, 10:15:41 AM	
crinaldi	pyspark-shell	5/9/2016, 9:57:28 AM	5/9/2016, 10:11:13 AM	100%
tomcat	client_states_20160508	5/9/2016, 10:04:53 AM	5/9/2016, 10:06:35 AM	100%
tomcat	jas	5/9/2016, 10:01:37 AM	5/9/2016, 10:03:55 AM	100%

1 - 10 of 2000 items 10 Per Page

Page 1 of 200

box:sbox:simulation_by_week (tomcat)

100%

Status:

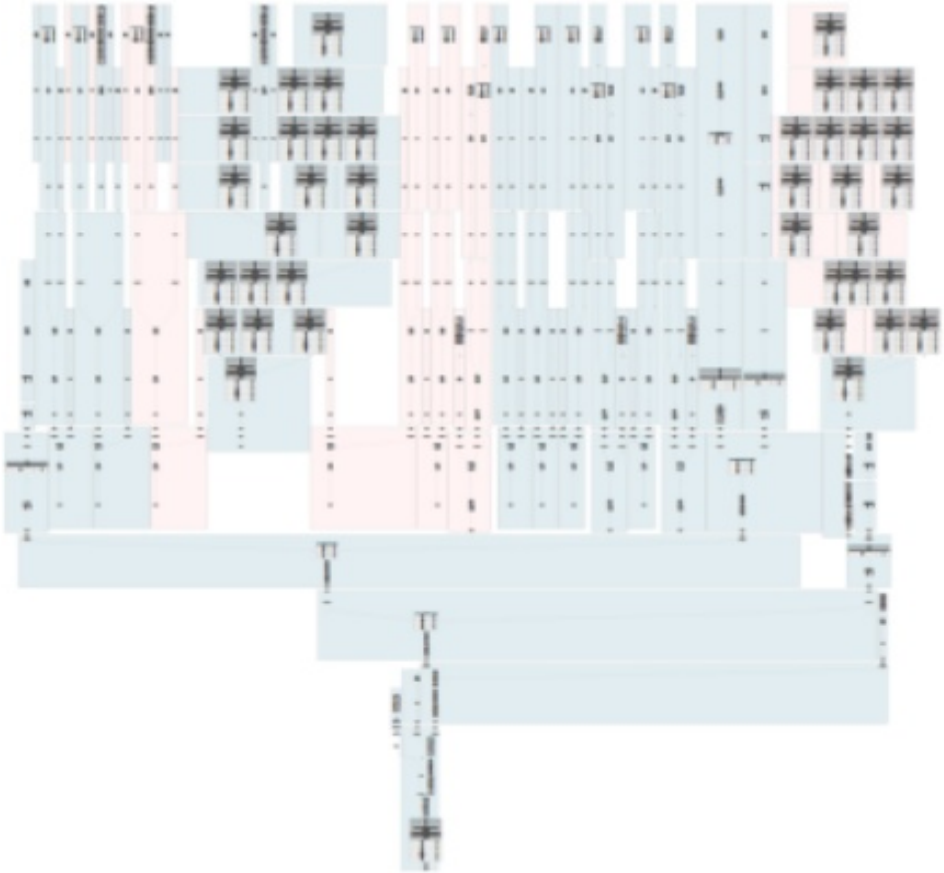
Status:	finished
Start:	5/3/2016 6:19:43 AM
End:	5/3/2016 6:23:05 AM
Heartbeat:	5/3/2016 6:19:52 AM

Links:

Spark UI

Logs

Node Groups	Status	Progress
_root		
stage-stage-58	finished	100%
stage-stage-59	finished	100%
stage-stage-56	finished	100%
stage-stage-57	finished	100%
stage-stage-50	finished	100%
stage-stage-51	finished	100%
stage-stage-54	finished	100%
stage-stage-55	finished	100%
stage-stage-52	finished	100%
stage-stage-53	finished	100%
stage-stage-69	finished	100%
stage-stage-67	finished	100%
stage-stage-68	finished	100%
stage-stage-61	finished	100%
stage-stage-62	finished	100%
stage-stage-60	finished	100%
stage-stage-65	finished	100%
stage-stage-66	finished	100%
stage-stage-63	finished	100%
stage-stage-64	finished	100%
stage-stage-36	finished	100%



How?

- Public SparkListener interface + AspectJ pointcuts to *access internal state*

Please help!

- Public interface for logical and physical planning events

Btw, why Morticia?

- Morticia Addams is inspiring and powerful
- Initially a tool for post-mortem analysis
- AspectJ pointcuts == basically witchcraft; Morticia is a witch
- Amidst chaos and complexity, Morticia remains calm and incisive

THANK YOU.

SPARK SUMMIT 2016
DATA SCIENCE AND ENGINEERING AT SCALE
JUNE 6-8, 2016 SAN FRANCISCO