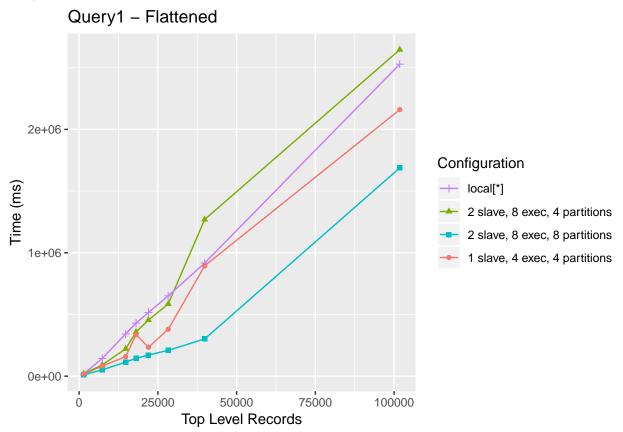
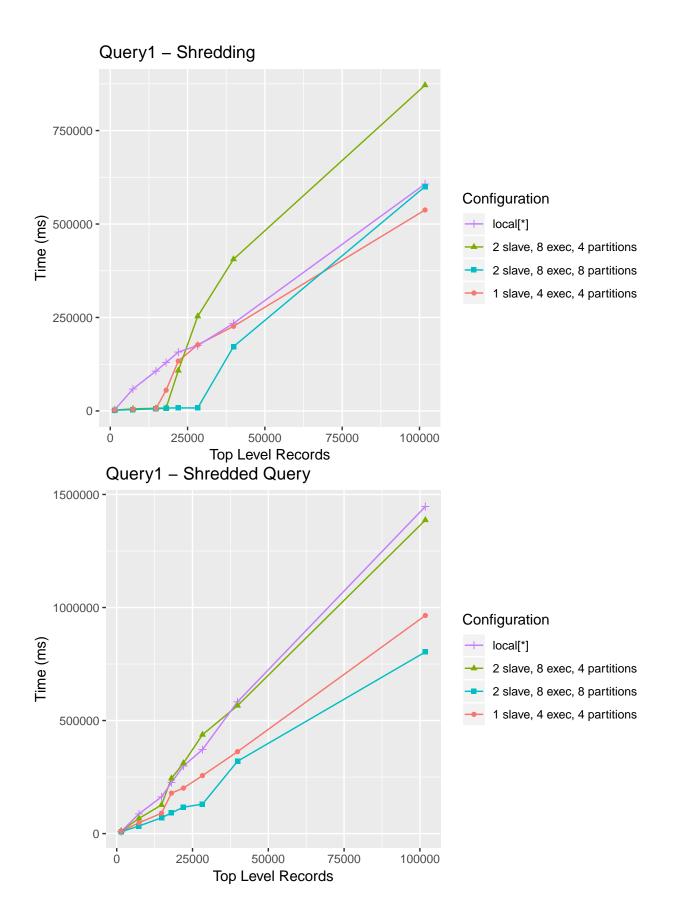
# cluster configurations

These tests were run on the Converge data (across 10640 samples) on the ODA cluster.

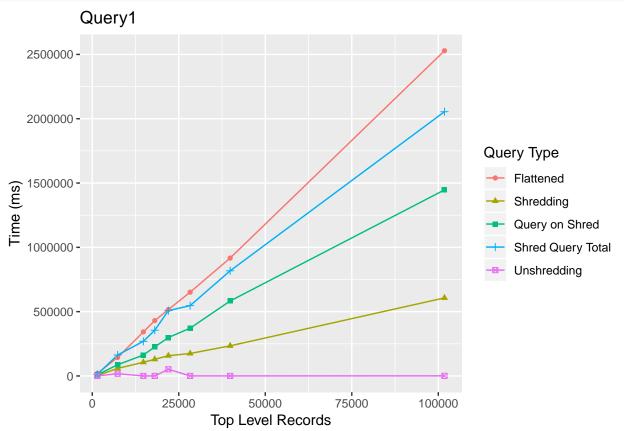
Query 1 Results





Local mode, Spark 2.1.1, Scala 2.11.8, 48 cores, 32G Driver, materializing shredded query stages

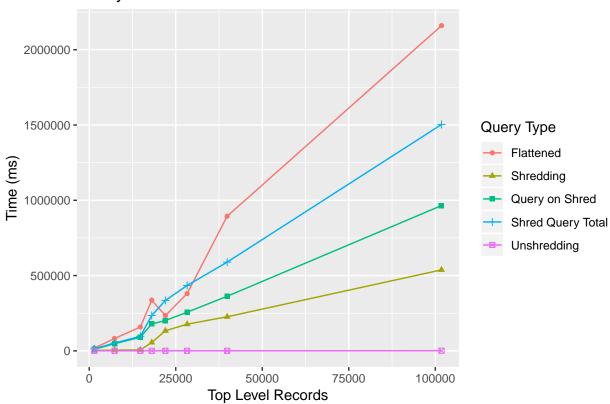
```
--master local[*] \
--executor-memory 24G \
--driver-memory 32G \
```



Converge, 1 slave, Spark 2.1.1, Scala 2.11.8, 4 partitions, materializing shredded query stages

```
--master spark://192.168.11.235:7077 \
--executor-cores 4 \
--num-executors 4 \
--executor-memory 24G \
--driver-memory 32G \
```

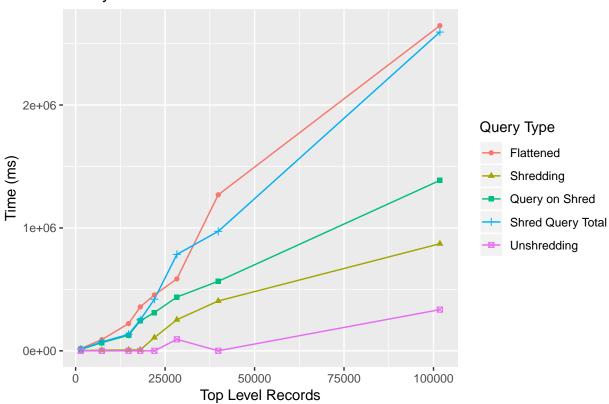
## Query1



 $Converge,\ 2\ slaves,\ Spark\ 2.1.1,\ Scala\ 2.11.8,\ 4\ partitions,\ materializing\ shredded\ query\ stages$ 

```
--master spark://192.168.11.235:7077 \
--executor-cores 4 \
--num-executors 8 \
--executor-memory 24G \
--driver-memory 32G \
```

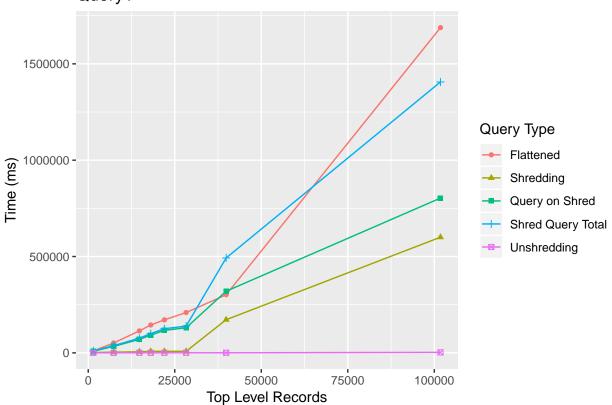




Converge, 2 slaves, Spark 2.1.1, Scala 2.11.8, 8 partitions, materializing shredded query stages

```
--master spark://192.168.11.235:7077 \
--executor-cores 4 \
--num-executors 8 \
--executor-memory 24G \
--driver-memory 32G \
```

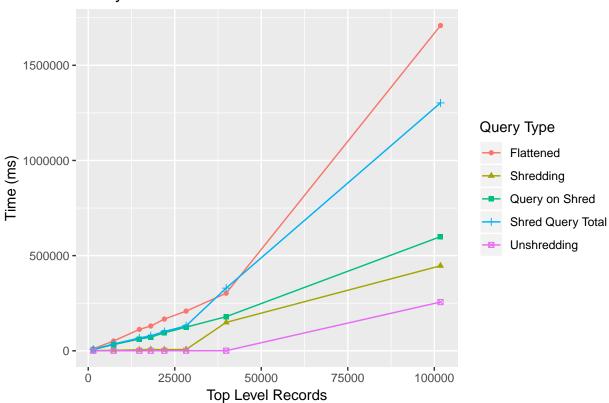
## Query1



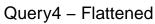
Optimizations to query 1: Converge, 2 slaves, Spark 2.1.1, Scala 2.11.8, 8 partitions, materializing shredded query stages

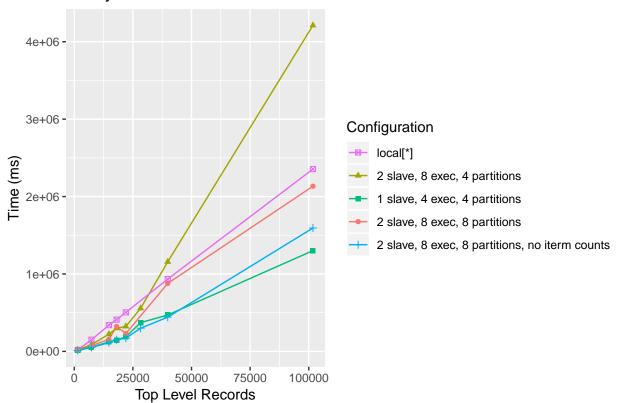
```
--master spark://192.168.11.235:7077 \
--executor-cores 4 \
--num-executors 8 \
--executor-memory 24G \
--driver-memory 32G \
```

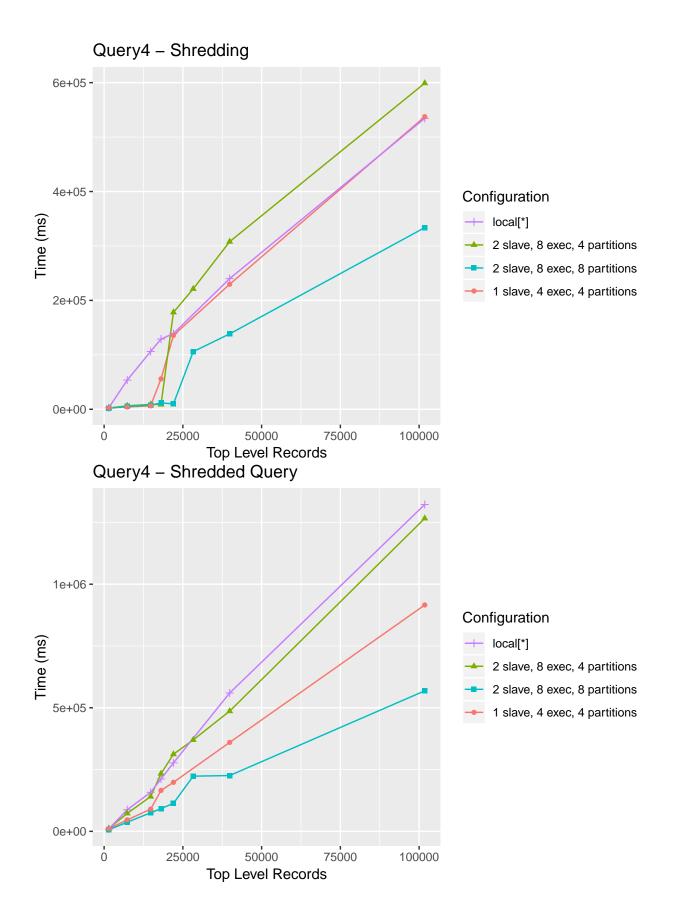


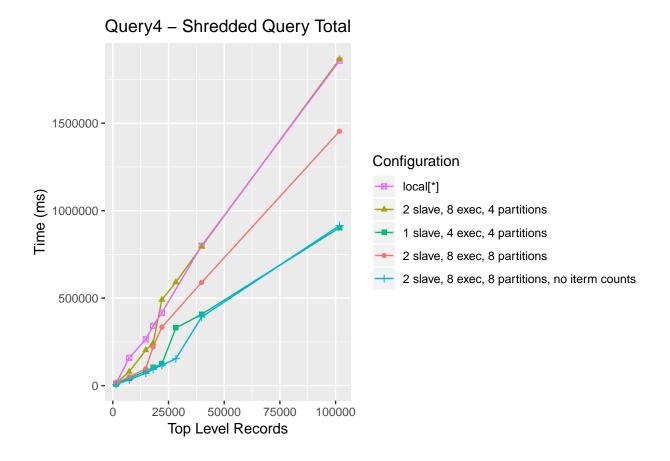


Query 4 Results





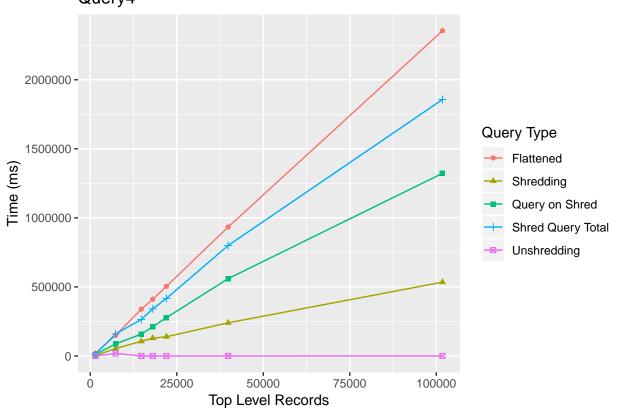




Local mode, Spark 2.1.1, Scala 2.11.8, 48 cores, 32G Driver, materializing shredded query stages

```
--master local[*] \
--executor-memory 24G \
--driver-memory 32G \
```

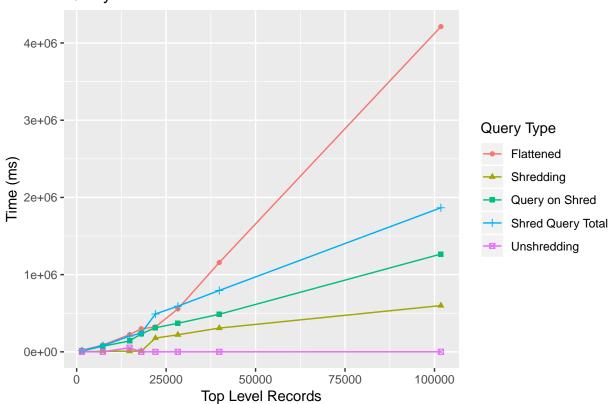




 $Converge,\ 2\ slaves,\ Spark\ 2.1.1,\ Scala\ 2.11.8,\ 4\ partitions,\ materializing\ shredded\ query\ stages$ 

```
--master spark://192.168.11.235:7077 \
--executor-cores 4 \
--num-executors 8 \
--executor-memory 24G \
--driver-memory 32G \
```

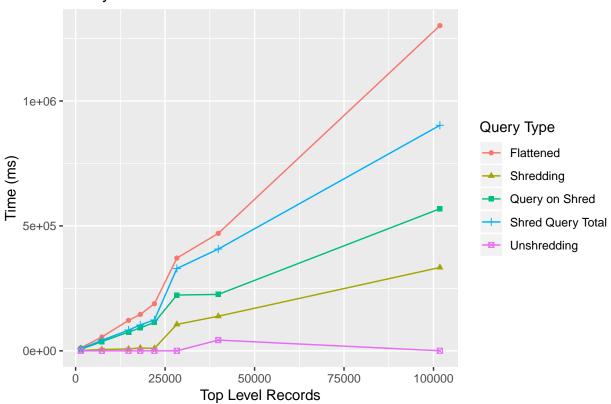




Converge, 2 slaves, Spark 2.1.1, Scala 2.11.8, 8 partitions, materializing shredded query stages

```
--master spark://192.168.11.235:7077 \
--executor-cores 4 \
--num-executors 8 \
--executor-memory 24G \
--driver-memory 32G \
```

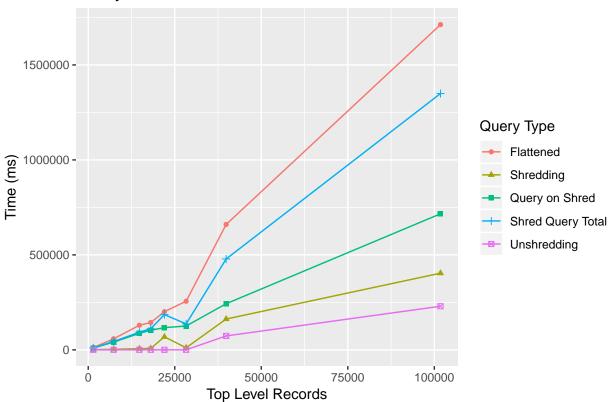




Converge, 2 slaves, Spark 2.1.1, Scala 2.11.8, 8 partitions, materializing shredded query stages

```
--master spark://192.168.11.235:7077 \
--executor-cores 4 \
--num-executors 8 \
--executor-memory 24G \
--driver-memory 32G \
```

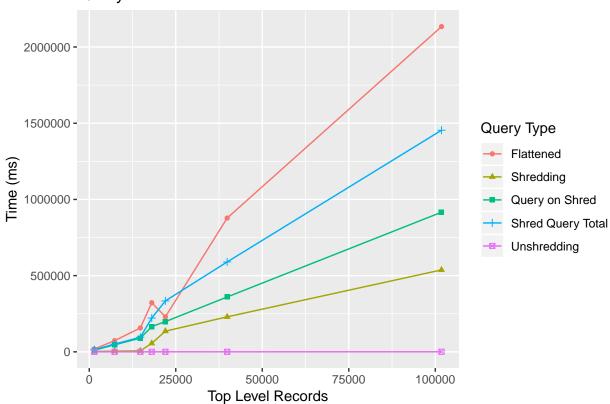
## Query4new



Converge, 1 slave, Spark 2.1.1, Scala 2.11.8, 4 partitions, materializing shredded query stages

```
--master spark://192.168.11.235:7077 \
--executor-cores 4 \
--num-executors 4 \
--executor-memory 24G \
--driver-memory 32G \
```

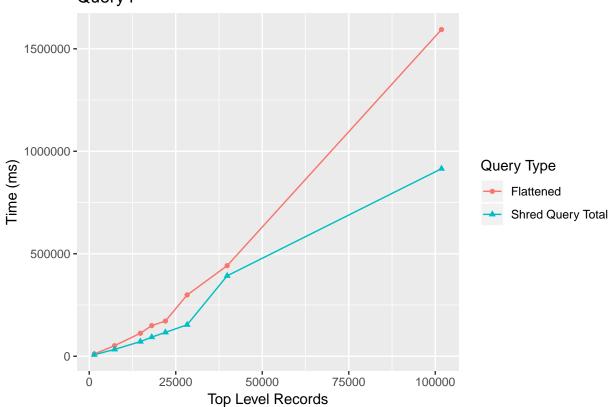




Converge, 2 slaves, Spark 2.1.1, Scala 2.11.8, 8 partitions, no intermediate materialization in the shredding query

```
--master spark://192.168.11.235:7077 \
--executor-cores 4 \
--num-executors 8 \
--executor-memory 24G \
--driver-memory 32G \
```





Query 5  ${\it Converge, 2 slaves, Spark 2.1.1, Scala 2.11.8, 8 partitions, materialization }$ 

10000

0

20000

Top Level Records

```
--master spark://192.168.11.235:7077 \
--executor-cores 4 \
--num-executors 8 \
--executor-memory 24G \
--driver-memory 32G \
```

# Query Type Flattened Shredding Query on Shred Shred Query Total Unshredding

30000

40000