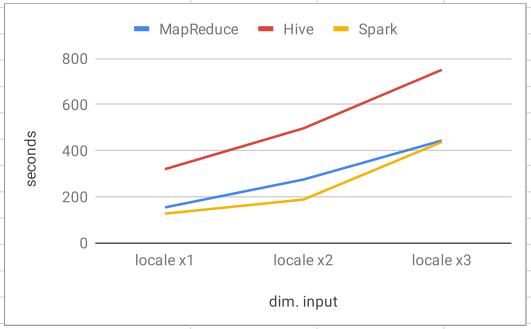
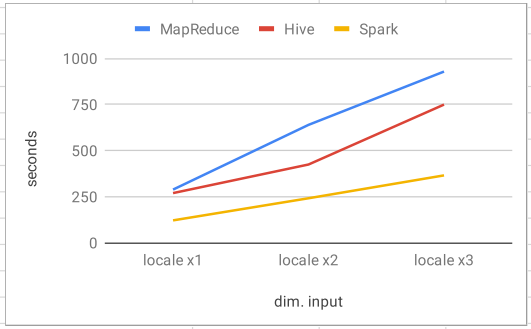
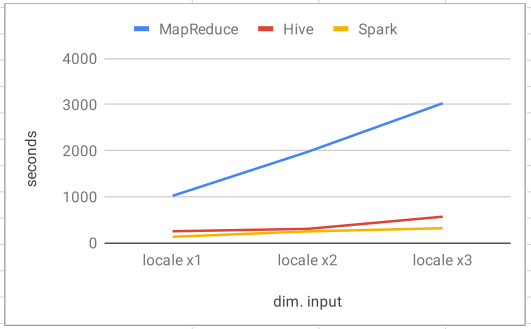
Tempi di esecuzione Job 1	MapReduce	Hive	Spark
locale x1	153	319	126
locale x2	274	496	187
locale x3	443	750	437



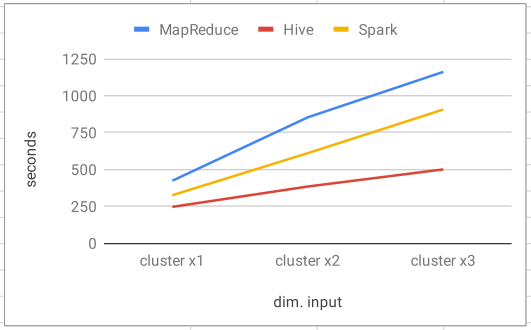
Tempi di esecuzione Job 2	MapReduce	Hive	Spark
locale x1	82+206	269	121
locale x2	252+387	424	241
locale x3	362+566	749	365



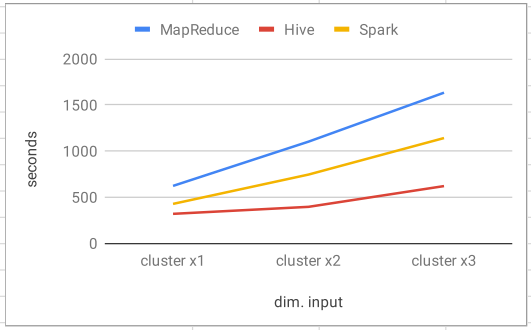
Tempi di esecuzione Job 3	MapReduce	Hive	Spark
locale x1	89+905+22	248	128
locale x2	194+1754+22	300	245
locale x3	327+2675+22	564	314



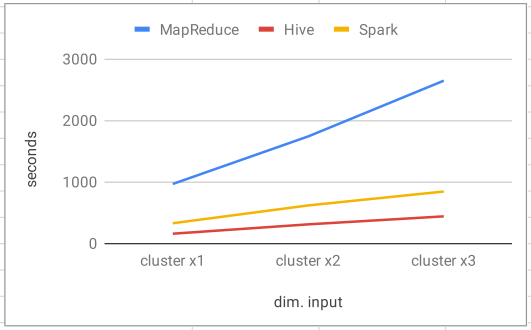
Tempi di esecuzione Job 1	MapReduce	Hive	Spark
cluster x1	423	246	325
cluster x2	853	384	610
cluster x3	1161	500	906



Tempi di esecuzione Job 2	MapReduce	Hive	Spark
cluster x1	188+433	318	426
cluster x2	321+780	394	744
cluster x3	488+1144	619	1140



Tempi di esecuzione Job 3	MapReduce	Hive	Spark
cluster x1	189+692+82	154	324
cluster x2	328+1329+80	306	614
cluster x3	451+2110+82	436	840



x1-2-3 : dimensioni input

dataset: historical_stock_prices.csv (2GB) e historical_stocks.csv (431.9KB)

AWS instances : m3.xlarge
8 vCore, 15 GiB di memoria, 80 SSD GB di storage

1 NameNode (master) and
2 DataNodes (slaves/workers)

Tutti i tempi di esecuzione comprendono anche le operazioni di I/O

Hive su AWS sfrutta Tez e non Map Reduce come in locale