

No.	Workload	Input Data Size	#Retired Instructions (Billions)	Source
1	Sort	150 GB documents	4578	Hadoop example
2	WordCount	154 GB documents	3533	Hadoop example
3	Grep	154 GB documents	1499	Hadoop example
4	Naïve Bayes	147 GB text	68131	Mahout
5	SVM	148 GB html file	2051	our implementation
6	K-means	150 GB vector	3227	Mahout
7	Fuzzy K-means	150 GB vector	15470	Mahout
8	IBCF	147 GB ratings data	32340	Mahout
9	HMM	147 GB html file	1841	our implementation
10	PageRank	187 GB web page	18470	Mahout
11	Hive-bench	156 GB DBtable	3659	Hivebench