

Hadoop Results

Pokec

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
4185c7fc3b2d	resourcemanager	0.02%	1.062MiB / 7.664GiB	0.01%	9.01kB / 6.24kB	28.7kB / 65.5kB	2
2381d111b064	spark-worker-1	0.07%	167.6MiB / 7.664GiB	2.13%	45.5kB / 384kB	20.6MB / 2.84MB	64
6f15e4e31b8f	datanode	0.30%	590.5MiB / 7.664GiB	7.52%	304MB / 773MB	519MB / 307MB	109
82058742d73b	namenode	106.17%	1.336GiB / 7.664GiB	17.44%	773MB / 304MB	476MB / 14.8MB	126
0acf249be712	spark-master	0.17%	184.4MiB / 7.664GiB	2.35%	386kB / 43.8kB	50.5MB / 3.06MB	64
24806a469951	nodemanager	0.00%	3.195MiB / 7.664GiB	0.04%	5.55kB / 6.18kB	2.7MB / 115kB	2

```
real    0m53.964s
user    1m0.426s
sys     0m4.089s
```

Email

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
4185c7fc3b2d	resourcemanager	0.00%	1.062MiB / 7.664GiB	0.01%	13.5kB / 9.62kB	28.7kB / 65.5kB	2
2381d111b064	spark-worker-1	0.09%	167.6MiB / 7.664GiB	2.14%	48kB / 407kB	20.6MB / 2.97MB	64
6f15e4e31b8f	datanode	3.87%	682.6MiB / 7.664GiB	8.70%	320MB / 897MB	524MB / 322MB	108
82058742d73b	namenode	338.46%	737MiB / 7.664GiB	9.39%	897MB / 320MB	476MB / 15.7MB	124
0acf249be712	spark-master	0.08%	184.3MiB / 7.664GiB	2.35%	409kB / 46.3kB	50.5MB / 3.19MB	64
24806a469951	nodemanager	0.00%	1012KiB / 7.664GiB	0.01%	2.4kB / 2.05kB	0B / 152kB	2

```
real    0m4.934s
user    0m7.113s
sys     0m1.404s
```

Patents

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
4185c7fc3b2d	resourcemanager	0.00%	1004KiB / 7.664GiB	0.01%	18.8kB / 13.7kB	28.7kB / 65.5kB	2
2381d111b064	spark-worker-1	0.11%	167.6MiB / 7.664GiB	2.14%	51kB / 433kB	20.6MB / 3.13MB	64
6f15e4e31b8f	datanode	8.00%	686.1MiB / 7.664GiB	8.74%	321MB / 1.16GB	789MB / 323MB	109
82058742d73b	namenode	139.01%	1.3GIB / 7.664GiB	16.96%	1.16GB / 321MB	476MB / 16.1MB	126
0acf249be712	spark-master	0.10%	183.2MiB / 7.664GiB	2.33%	435kB / 49.3kB	50.5MB / 3.35MB	64
24806a469951	nodemanager	0.00%	1016KiB / 7.664GiB	0.01%	6.43kB / 7.29kB	0B / 152kB	2

```
real    0m36.636s
user    0m42.639s
sys     0m3.885s
```

LiveJournal

Reducer = 2

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.10%	169MiB / 7.664GiB	2.15%	81.5kB / 708kB	20.6MB / 4.6MB	64
6f15e4e31b8f	datanode	0.24%	287.2MiB / 7.664GiB	3.66%	1.45GB / 2.08GB	1.67GB / 1.45GB	109
82058742d73b	namenode	105.40%	1.166GiB / 7.664GiB	15.22%	1.36GB / 1.45GB	1.64GB / 19.1MB	121
0acf249be712	spark-master	0.09%	277.9MiB / 7.664GiB	3.54%	720MB / 1.17MB	129MB / 16.1MB	65
4185c7fc3b2d	resourcemanager	0.02%	1.07MiB / 7.664GiB	0.01%	5.89kB / 3.83kB	20.5kB / 77.8kB	2
24806a469951	nodemanager	0.00%	3.227MiB / 7.664GiB	0.04%	1.62kB / 1.03kB	2.95MB / 156kB	2

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.09%	169MiB / 7.664GiB	2.15%	81.7kB / 709kB	20.6MB / 4.62MB	64
6f15e4e31b8f	datanode	0.29%	294.3MiB / 7.664GiB	3.75%	1.45GB / 2.32GB	1.92GB / 1.45GB	111
82058742d73b	namenode	105.91%	1.382GiB / 7.664GiB	18.03%	1.6GB / 1.45GB	1.64GB / 19.2MB	126
0acf249be712	spark-master	0.09%	277.2MiB / 7.664GiB	3.53%	720MB / 1.17MB	129MB / 16.1MB	65
4185c7fc3b2d	resourcemanager	0.00%	1012KiB / 7.664GiB	0.01%	6.18kB / 4.05kB	20.5kB / 77.8kB	2
24806a469951	nodemanager	0.00%	3.223MiB / 7.664GiB	0.04%	1.84kB / 1.33kB	2.95MB / 156kB	2

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.10%	169MiB / 7.664GiB	2.15%	81.8kB / 710kB	20.6MB / 4.62MB	64
6f15e4e31b8f	datanode	2.33%	298.3MiB / 7.664GiB	3.80%	1.45GB / 2.55GB	2.14GB / 1.45GB	111
82058742d73b	namenode	134.64%	1.393GiB / 7.664GiB	18.17%	1.83GB / 1.45GB	1.64GB / 19.8MB	126
0acf249be712	spark-master	0.10%	277.2MiB / 7.664GiB	3.53%	720MB / 1.17MB	129MB / 16.1MB	65
4185c7fc3b2d	resourcemanager	0.00%	1.074MiB / 7.664GiB	0.01%	6.52kB / 4.31kB	20.5kB / 77.8kB	2
24806a469951	nodemanager	0.00%	3.223MiB / 7.664GiB	0.04%	2.1kB / 1.67kB	2.95MB / 156kB	2

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.09%	169MiB / 7.664GiB	2.15%	82.1kB / 713kB	20.6MB / 4.64MB	64
6f15e4e31b8f	datanode	0.36%	302.2MiB / 7.664GiB	3.85%	1.45GB / 2.96GB	2.55GB / 1.45GB	109
82058742d73b	namenode	105.45%	1.39GiB / 7.664GiB	18.14%	2.24GB / 1.45GB	1.64GB / 97.2MB	125
0acf249be712	spark-master	0.09%	277.2MiB / 7.664GiB	3.53%	720MB / 1.17MB	129MB / 16.1MB	65
4185c7fc3b2d	resourcemanager	0.00%	1016KiB / 7.664GiB	0.01%	7.12kB / 4.77kB	20.5kB / 77.8kB	2
24806a469951	nodemanager	0.45%	3.227MiB / 7.664GiB	0.04%	2.56kB / 2.27kB	2.95MB / 156kB	2

real	1m49.392s
user	2m1.416s
sys	0m7.636s

Reducer = 4

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.10%	169MiB / 7.664GiB	2.15%	94.4kB / 824kB	20.6MB / 5.19MB	64
6f15e4e31b8f	datanode	0.27%	403MiB / 7.664GiB	5.14%	1.54GB / 5.38GB	3.71GB / 1.54GB	109
82058742d73b	namenode	187.81%	1015MiB / 7.664GiB	12.94%	3.56GB / 1.54GB	1.68GB / 274MB	121
0acf249be712	spark-master	0.08%	277.2MiB / 7.664GiB	3.53%	1.82GB / 2.46MB	188MB / 17MB	65
24806a469951	nodemanager	0.00%	8.266MiB / 7.664GiB	0.11%	6.12kB / 6.75kB	10.1MB / 180kB	2
4185c7fc3b2d	resourcemanager	0.00%	1.094MiB / 7.664GiB	0.01%	6.41kB / 4.34kB	164kB / 90.1kB	2

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.09%	169MiB / 7.664GiB	2.15%	94.7kB / 827kB	20.6MB / 5.2MB	64
6f15e4e31b8f	datanode	3.32%	733.2MiB / 7.664GiB	9.34%	1.54GB / 5.73GB	3.71GB / 1.54GB	108
82058742d73b	namenode	105.81%	1.38GiB / 7.664GiB	18.01%	3.91GB / 1.54GB	1.68GB / 275MB	126
0acf249be712	spark-master	0.09%	277.2MiB / 7.664GiB	3.53%	1.82GB / 2.46MB	188MB / 17.1MB	65
24806a469951	nodemanager	0.00%	8.273MiB / 7.664GiB	0.11%	6.49kB / 7.23kB	10.1MB / 180kB	2
4185c7fc3b2d	resourcemanager	0.00%	1.016MiB / 7.664GiB	0.01%	6.89kB / 4.7kB	164kB / 90.1kB	2

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.13%	169MiB / 7.664GiB	2.15%	94.9kB / 828kB	20.6MB / 5.21MB	64
6f15e4e31b8f	datanode	2.77%	969.8MiB / 7.664GiB	12.36%	1.54GB / 5.97GB	3.71GB / 1.54GB	109
82058742d73b	namenode	136.33%	1.397GiB / 7.664GiB	18.22%	4.15GB / 1.54GB	1.68GB / 275MB	126
0acf249be712	spark-master	0.11%	277.2MiB / 7.664GiB	3.53%	1.82GB / 2.46MB	188MB / 17.1MB	65
24806a469951	nodemanager	0.00%	8.285MiB / 7.664GiB	0.11%	6.75kB / 7.57kB	10.1MB / 180kB	2
4185c7fc3b2d	resourcemanager	0.00%	1.098MiB / 7.664GiB	0.01%	7.23kB / 4.96kB	164kB / 90.1kB	2

real	1m54.264s
user	2m5.827s
sys	0m7.069s

Reducer = 8

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.10%	169MiB / 7.664GiB	2.15%	97.1kB / 848kB	20.6MB / 5.34MB	64
6f15e4e31b8f	datanode	0.27%	1.359GiB / 7.664GiB	17.73%	1.59GB / 6.57GB	3.71GB / 1.59GB	109
82058742d73b	namenode	105.96%	1.13GiB / 7.664GiB	14.74%	4.75GB / 1.59GB	1.68GB / 379MB	121
0acf249be712	spark-master	0.08%	277.2MiB / 7.664GiB	3.53%	1.82GB / 2.46MB	188MB / 17.2MB	65
4185c7fc3b2d	resourcemanager	0.00%	1.102MiB / 7.664GiB	0.01%	11.1kB / 7.84kB	164kB / 90.1kB	2
24806a469951	nodemanager	0.00%	3.141MiB / 7.664GiB	0.04%	2.97kB / 2.8kB	2.71MB / 180kB	2

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.09%	169MiB / 7.664GiB	2.15%	97.4kB / 850kB	20.6MB / 5.34MB	64
6f15e4e31b8f	datanode	0.37%	1.359GiB / 7.664GiB	17.74%	1.59GB / 6.87GB	3.71GB / 1.59GB	108
82058742d73b	namenode	106.85%	1.388GiB / 7.664GiB	18.10%	5.05GB / 1.59GB	1.68GB / 392MB	124
0acf249be712	spark-master	0.09%	277.2MiB / 7.664GiB	3.53%	1.82GB / 2.46MB	188MB / 17.2MB	65
4185c7fc3b2d	resourcemanager	0.00%	1.094MiB / 7.664GiB	0.01%	11.6kB / 8.25kB	164kB / 90.1kB	2
24806a469951	nodemanager	3.26%	3.156MiB / 7.664GiB	0.04%	3.37kB / 3.33kB	2.71MB / 180kB	2

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.09%	169MiB / 7.664GiB	2.15%	97.6kB / 852kB	20.6MB / 5.36MB	64
6f15e4e31b8f	datanode	1.13%	1.366GiB / 7.664GiB	17.82%	1.59GB / 7.15GB	3.71GB / 1.59GB	109
82058742d73b	namenode	143.81%	1.382GiB / 7.664GiB	18.03%	5.33GB / 1.59GB	1.68GB / 419MB	125
0acf249be712	spark-master	0.09%	277.2MiB / 7.664GiB	3.53%	1.82GB / 2.46MB	188MB / 17.2MB	65
4185c7fc3b2d	resourcemanager	0.00%	1.094MiB / 7.664GiB	0.01%	12kB / 8.52kB	164kB / 90.1kB	2
24806a469951	nodemanager	0.47%	3.125MiB / 7.664GiB	0.04%	3.64kB / 3.7kB	2.71MB / 180kB	2

```
real    2m1.826s
user    2m9.844s
sys     0m6.788s
```

Spark Results

Pokec

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.09%	166.4MiB / 7.664GiB	2.12%	68.9kB / 594kB	20.6MB / 4.02MB	64
6f15e4e31b8f	datanode	2.92%	275.8MiB / 7.664GiB	3.51%	354MB / 1.3GB	897MB / 357MB	112
82058742d73b	namenode	0.17%	316.1MiB / 7.664GiB	4.03%	1.21GB / 354MB	476MB / 17.8MB	77
0acf249be712	spark-master	462.91%	1.067GiB / 7.664GiB	13.93%	82.5MB / 175kB	125MB / 4.4MB	181
4185c7fc3b2d	resourcemanager	3.74%	1MiB / 7.664GiB	0.01%	5.78kB / 3.77kB	28.7kB / 73.7kB	2
24806a469951	nodemanager	0.00%	1000KiB / 7.664GiB	0.01%	4.2kB / 4.41kB	0B / 156kB	2

```
real    0m57.178s
user    0m25.554s
sys     0m4.864s
```

Email

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.07%	166.6MiB / 7.664GiB	2.12%	71kB / 614kB	20.6MB / 4.12MB	64
6f15e4e31b8f	datanode	0.33%	271.1MiB / 7.664GiB	3.45%	354MB / 1.64GB	1.23GB / 358MB	108
82058742d73b	namenode	0.19%	320.1MiB / 7.664GiB	4.08%	1.21GB / 354MB	476MB / 18.1MB	77
0acf249be712	spark-master	229.92%	551.3MiB / 7.664GiB	7.02%	430MB / 689kB	129MB / 15.2MB	157
4185c7fc3b2d	resourcemanager	0.00%	1016KiB / 7.664GiB	0.01%	9.76kB / 6.81kB	28.7kB / 73.7kB	2
24806a469951	nodemanager	0.00%	1012KiB / 7.664GiB	0.01%	7.24kB / 8.4kB	0B / 156kB	2

```
real    0m7.695s
user    0m13.032s
sys     0m2.621s
```

Patents

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.08%	168.9MiB / 7.664GiB	2.15%	72.5kB / 627kB	20.6MB / 4.19MB	64
6f15e4e31b8f	datanode	2.32%	280MiB / 7.664GiB	3.57%	355MB / 1.78GB	1.37GB / 358MB	110
82058742d73b	namenode	0.14%	320.4MiB / 7.664GiB	4.08%	1.21GB / 354MB	476MB / 18.2MB	77
0acf249be712	spark-master	228.01%	1.158GiB / 7.664GiB	15.11%	561MB / 907kB	129MB / 15.4MB	173
4185c7fc3b2d	resourcemanager	0.00%	1.082MiB / 7.664GiB	0.01%	12.4kB / 8.77kB	28.7kB / 73.7kB	2
24806a469951	nodemanager	0.00%	1012KiB / 7.664GiB	0.01%	2.62kB / 2.38kB	0B / 156kB	2

```
real    0m50.954s
user    0m20.974s
sys     0m4.147s
```

LiveJournal

No optimization

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.10%	169MiB / 7.664GiB	2.15%	85.2kB / 741kB	20.6MB / 4.78MB	64
6f15e4e31b8f	datanode	16.49%	387.2MiB / 7.664GiB	4.93%	1.49GB / 3.12GB	2.62GB / 1.49GB	116
82058742d73b	namenode	5.85%	347.3MiB / 7.664GiB	4.42%	2.36GB / 1.49GB	1.64GB / 146MB	78
0acf249be712	spark-master	973.84%	913.9MiB / 7.664GiB	11.65%	766MB / 1.23MB	185MB / 16.3MB	197
4185c7fc3b2d	resourcemanager	0.29%	1.074MiB / 7.664GiB	0.01%	12.5kB / 8.89kB	20.5kB / 77.8kB	2
24806a469951	nodemanager	0.71%	3.23MiB / 7.664GiB	0.04%	6.67kB / 7.65kB	2.95MB / 156kB	2

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.10%	169MiB / 7.664GiB	2.15%	85.4kB / 743kB	20.6MB / 4.79MB	64
6f15e4e31b8f	datanode	3.46%	891.4MiB / 7.664GiB	11.36%	1.49GB / 3.64GB	2.62GB / 1.49GB	116
82058742d73b	namenode	0.23%	347.3MiB / 7.664GiB	4.42%	2.36GB / 1.49GB	1.64GB / 146MB	78
0acf249be712	spark-master	885.11%	2.086GiB / 7.664GiB	27.22%	1.29GB / 1.81MB	185MB / 16.3MB	197
4185c7fc3b2d	resourcemanager	0.02%	1.066MiB / 7.664GiB	0.01%	12.8kB / 9.14kB	20.5kB / 77.8kB	2
24806a469951	nodemanager	0.00%	3.227MiB / 7.664GiB	0.04%	6.93kB / 7.98kB	2.95MB / 156kB	2

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.13%	169MiB / 7.664GiB	2.15%	85.6kB / 745kB	20.6MB / 4.79MB	64
6f15e4e31b8f	datanode	2.31%	1.318GiB / 7.664GiB	17.19%	1.49GB / 4.16GB	2.62GB / 1.49GB	111
82058742d73b	namenode	0.31%	347.3MiB / 7.664GiB	4.43%	2.36GB / 1.49GB	1.64GB / 146MB	78
0acf249be712	spark-master	969.68%	3.149GiB / 7.664GiB	41.09%	1.8GB / 2.39MB	188MB / 16.6MB	192
4185c7fc3b2d	resourcemanager	0.00%	1.074MiB / 7.664GiB	0.01%	13.2kB / 9.4kB	20.5kB / 77.8kB	2
24806a469951	nodemanager	0.00%	3.234MiB / 7.664GiB	0.04%	7.19kB / 8.32kB	2.95MB / 156kB	2

```
real    1m9.781s
user    0m51.756s
sys     0m8.776s
```

With cache optimization

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.07%	169MiB / 7.664GiB	2.15%	102kB / 893kB	20.6MB / 5.57MB	64
6f15e4e31b8f	datanode	0.25%	1.376GiB / 7.664GiB	17.96%	1.64GB / 7.6GB	3.71GB / 1.63GB	108
82058742d73b	namenode	0.74%	355.8MiB / 7.664GiB	4.53%	5.78GB / 1.63GB	1.68GB / 516MB	78
0acf249be712	spark-master	184.60%	598.8MiB / 7.664GiB	7.63%	1.82GB / 2.47MB	243MB / 17.5MB	138
4185c7fc3b2d	resourcemanager	0.00%	1.098MiB / 7.664GiB	0.01%	20kB / 14.5kB	164kB / 90.1kB	2
24806a469951	nodemanager	0.00%	1016KiB / 7.664GiB	0.01%	3.09kB / 2.98kB	0B / 180kB	2

```
real    0m5.535s
user    0m7.921s
sys     0m1.673s
```

With repartition tuning

CONTAINER ID	NAME	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O	PIDS
2381d111b064	spark-worker-1	0.07%	169MiB / 7.664GiB	2.15%	105kB / 917kB	20.6MB / 5.69MB	64
6f15e4e31b8f	datanode	0.35%	364.1MiB / 7.664GiB	4.64%	1.64GB / 7.6GB	3.71GB / 1.63GB	109
82058742d73b	namenode	0.18%	335.5MiB / 7.664GiB	4.28%	5.78GB / 1.63GB	1.68GB / 516MB	78
0acf249be712	spark-master	176.39%	435.7MiB / 7.664GiB	5.55%	1.82GB / 2.47MB	278MB / 17.6MB	104
4185c7fc3b2d	resourcemanager	0.00%	8.008MiB / 7.664GiB	0.10%	6.78kB / 7.67kB	10MB / 180kB	2
24806a469951	nodemanager	0.00%	1.016MiB / 7.664GiB	0.01%	2.84kB / 1.61kB	160kB / 90.1kB	2

```
real    0m5.701s
user    0m8.454s
sys     0m1.745s
```