Cerro Porteño Final Report

Load Balancer 2; Indexer 2

Amdahl's Experiment

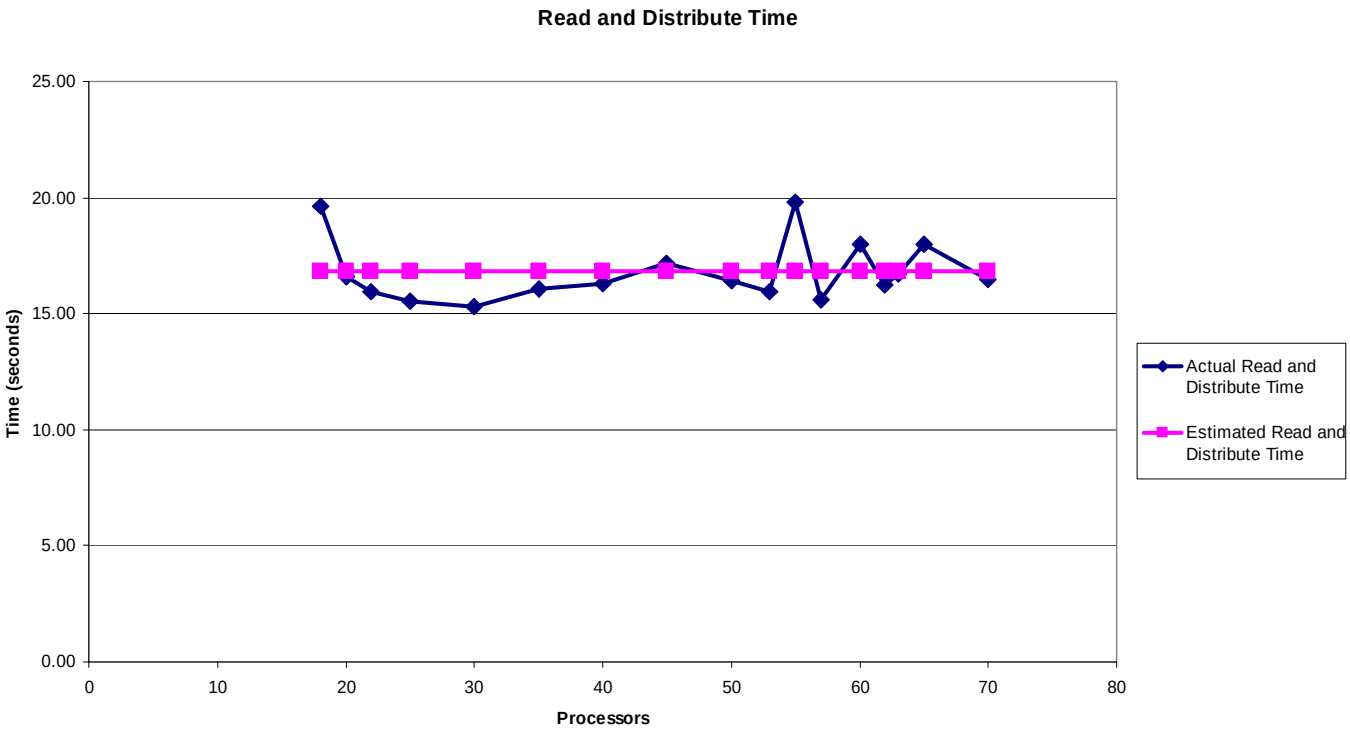
Read and Distribute Time

Read and distribute time is the time since the first byte is read from the disk to the time when indexing of first byte begins. This time varies for our runs as multiple root threads (in load-balancer configuration 2) are initiated to read the files. So there are factors like number of processors on the node the program is running on, and context switching of the threads that contribute to this variation. The estimated read and distribute time is taken as the average of times for different runs.

For *Amdahl's experiment*, the time varies between 15.28s and 19.82s, while the average read and distribute time is 16.81s.

For *Gustafson's experiment*, this time varies between 6.29s and 11.93s, averaging at 9.3s.

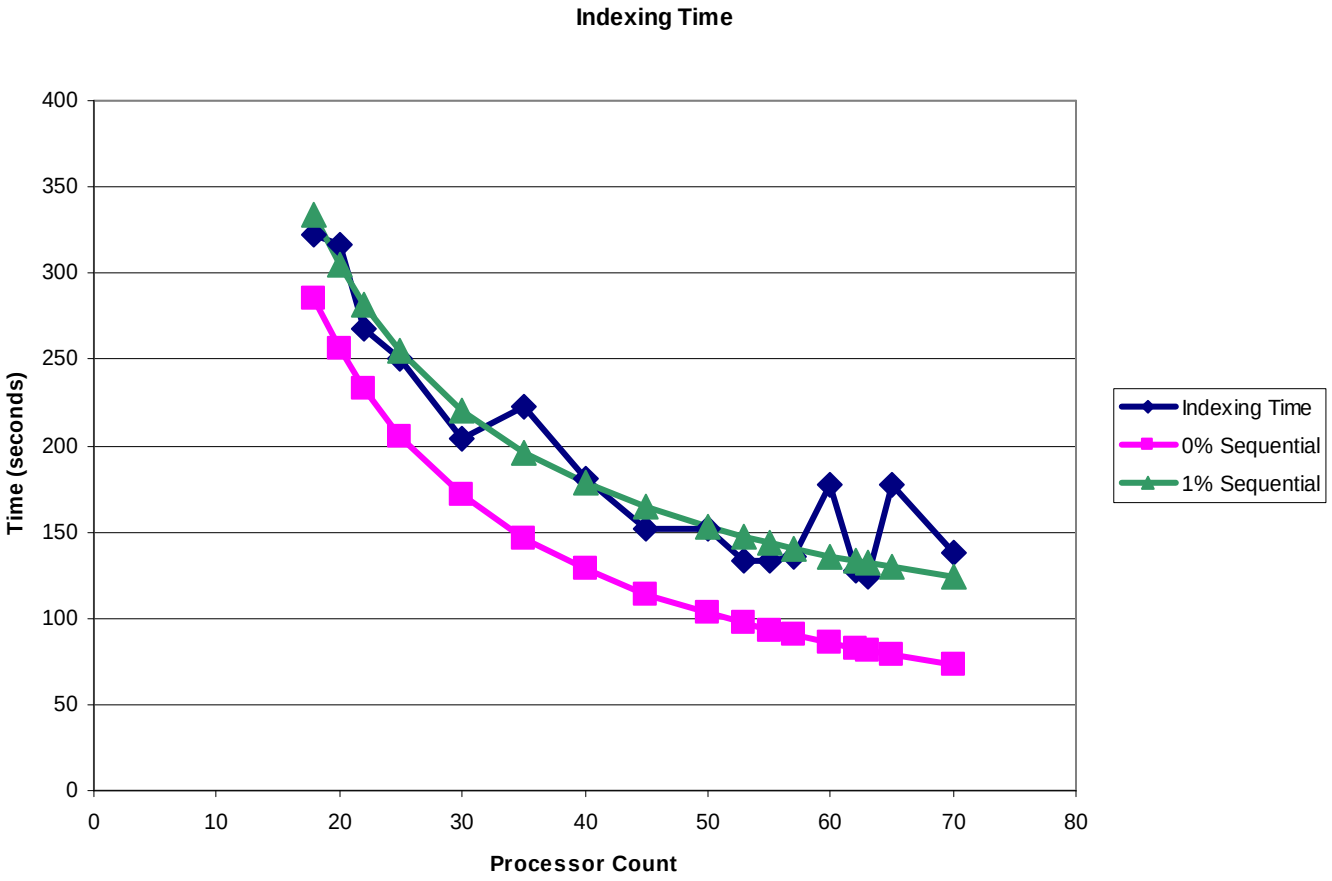
Processor Count	Actual Read and Distribute Time	Estimated Read and Distribute Time
18	19.64	16.81
20	16.60	16.81
22	15.93	16.81
25	15.53	16.81
30	15.28	16.81
35	16.09	16.81
40	16.28	16.81
45	17.20	16.81
50	16.44	16.81
53	15.96	16.81
55	19.82	16.81
57	15.59	16.81
60	17.97	16.81
62	16.24	16.81
63	16.70	16.81
65	17.97	16.81
70	16.45	16.81


Read and Distribute Time
25.00
20.00
15.00
10.00
5.00
0.00
Time (seconds)
0
10
20
30
40
50
60
70
80
Processors
Actual Read and Distribute Time
Estimated Read and Distribute Time

Indexing Time

The indexing itself has no serial component, so the graph of indexing time should look like Amdahl's Law when the serial percentage is 0. Our graph does not look like this. It instead looks very much like an Amdahl's Law graph with 1% serial. Reading and distributing the files must be somewhat serial. This is not unexpected as disk drives and network adapters can only do one thing at a time.

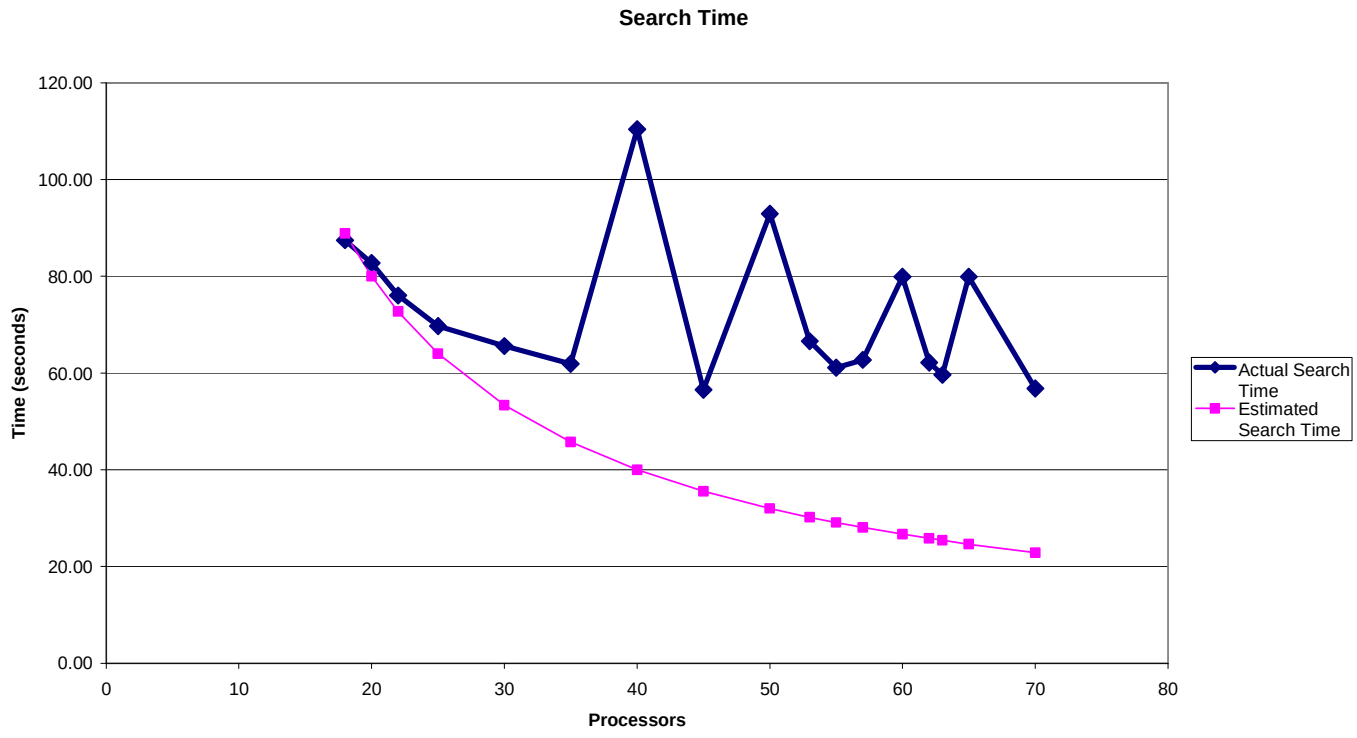
Processor Count	Indexing Time	0% Sequential	1% Sequential
18	322.38	285.16	333.63
20	316.84	256.64	305.40
22	268.13	233.31	282.30
25	250.38	205.31	254.59
30	203.86	171.09	220.71
35	222.53	146.65	196.51
40	180.66	128.32	178.36
45	152.25	114.06	164.25
50	152.23	102.66	152.96
53	132.99	96.85	147.20
55	133.19	93.32	143.72
57	136.16	90.05	140.48
60	176.83	85.55	136.02
62	127.10	82.79	133.29
63	123.82	81.47	131.99
65	176.83	78.97	129.50
70	137.83	73.33	123.92



Search Time

While we would like to see a nice Amdahl's Law type graph for our search times, the graph instead seems to show a very large serial component and seems to include a lot of noise. The serial component makes some sense, since the root process sequentially gathers results from the other nodes. The noise is less easy to explain.

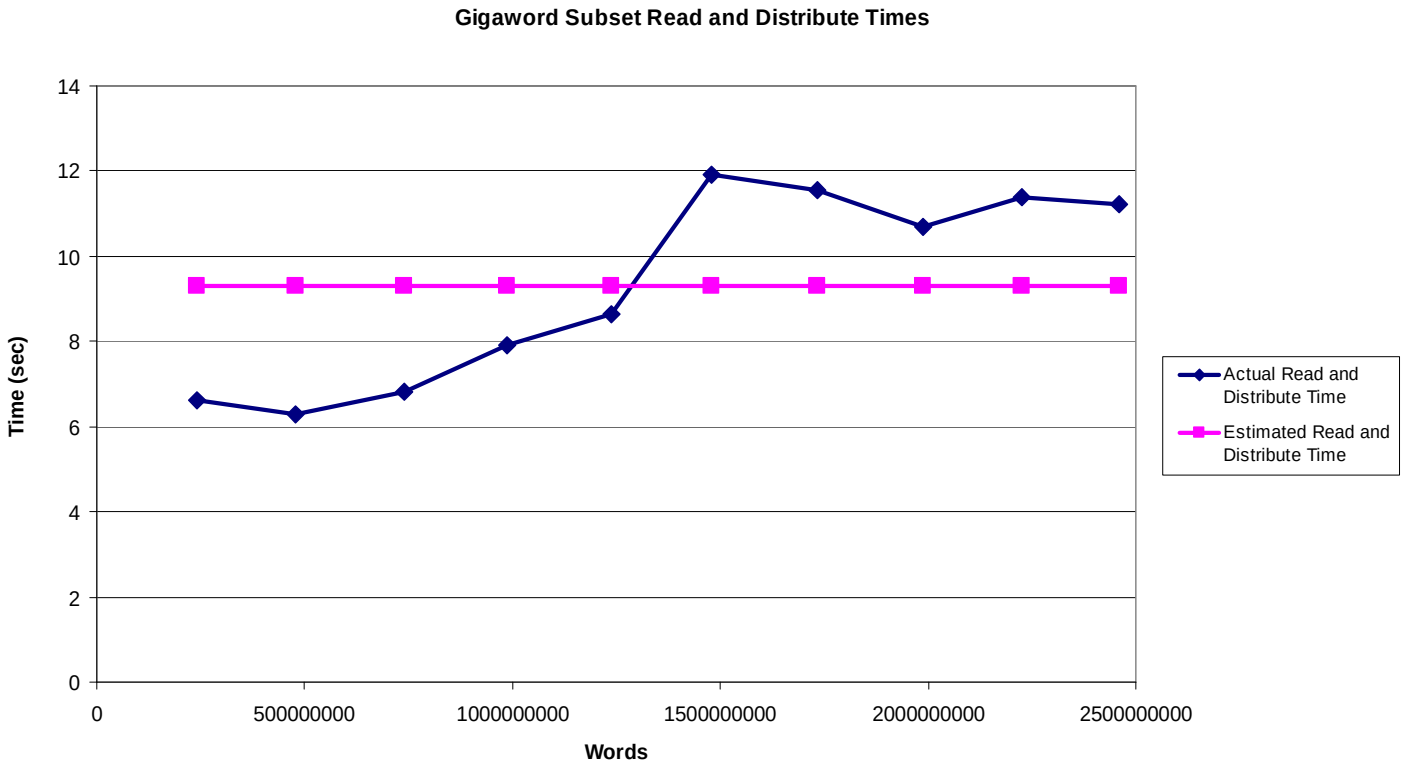
Processor Count	Actual Search Time	Estimated Search Time
18	87.46	88.89
20	82.76	80.00
22	76.02	72.73
25	69.72	64.00
30	65.57	53.33
35	61.92	45.71
40	110.40	40.00
45	56.52	35.56
50	92.94	32.00
53	66.58	30.19
55	61.08	29.09
57	62.72	28.07
60	79.93	26.67
62	62.17	25.81
63	59.64	25.40
65	79.93	24.62
70	56.81	22.86



Gustafson's Experiment

Read and Distribute Time

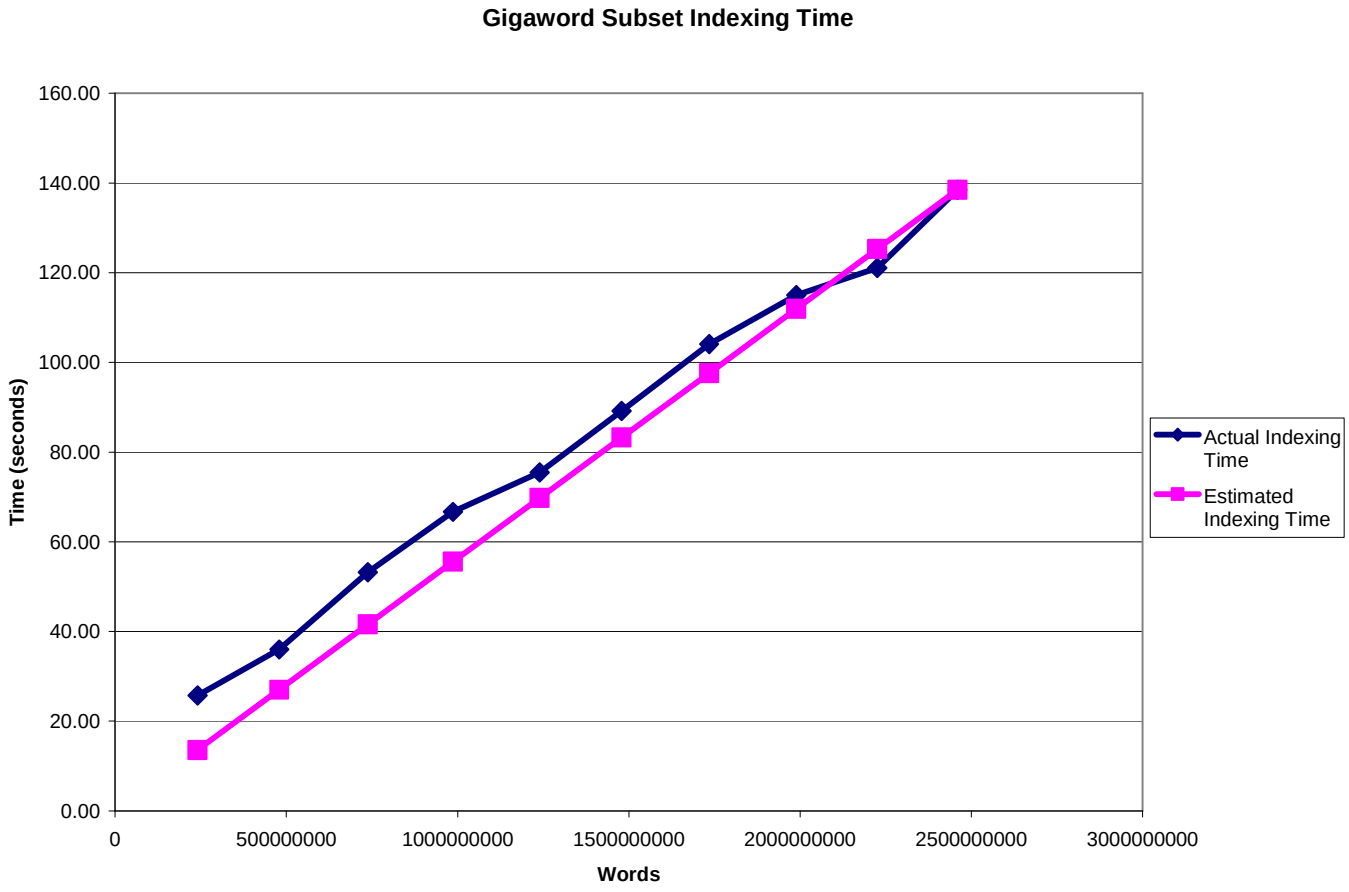
Word Count	Actual Read and Distribute Time	Estimated Read and Distribute Time
241347300	6.62	9.30
479529709	6.29	9.30
737909745	6.81	9.30
986912776	7.90	9.30
1239692536	8.65	9.30
1479025058	11.93	9.30
1734164783	11.55	9.30
1989126550	10.69	9.30
2225591317	11.39	9.30
2459788219	11.22	9.30



Indexing Time

As expected, indexing time is very nearly proportional to the number of words indexed.

Word Count	Actual Indexing Time	Estimated Indexing Time
241347300	25.78	13.58
479529709	35.99	26.99
737909745	53.16	41.53
986912776	66.64	55.54
1239692536	75.46	69.77
1479025058	89.18	83.24
1734164783	104.08	97.60
1989126550	114.99	111.95
2225591317	121.02	125.26
2459788219	138.44	138.44



Search Time (with Initial 25 queries, only)
Search times here are nearly exactly what one would expect. Search time is mainly related to the number of matches. One would expect that the number of matches in general would be proportional to the number of words indexed.

Word Count	Actual Search Time	Estimated Search Time
241347300	3.01	2.39
479529709	5.38	4.74
737909745	7.82	7.29
986912776	10.09	9.75
1239692536	14.33	12.25
1479025058	16.60	14.62
1734164783	17.57	17.14
1989126550	18.76	19.66
2225591317	20.80	22.00
2459788219	24.31	24.31