

PFLOCK Report

Andres Calderon

University of California, Riverside

January 10, 2020

Task analysis...

- ▶ Apache Spark divides the work into a number of Stages, each one is also divide into a number of Tasks.
- ▶ Each Task evaluates the data from a particular Partition and it is sent to an Executor for evaluation.
- ▶ Tracking the Tasks will give us a notion of the Partitions and Executors performance...

Task analysis...

- ▶ SparkListener class allows us to monitor Task metrics. We can capture info about:
 - ▶ executor: The id of the executor where the Task was evaluated.
 - ▶ duration: Execution time of the task.
 - ▶ recordsRead/Written: Total number of records read or written.
 - ▶ bytesRead/Written: Total number of bytes read or written.
 - ▶ shuffleRead/Written: Total number of records read or written from the shuffle by this task.

Task analysis...

- ▶ Running experiments with Epsilon=30, Mu=3 and Delta=3.
- ▶ Collecting Task metrics. Focus on Time by time implementation...

timestamp	executor	duration	recordsRead	bytesRead	recordsWritten	bytesWritten	shuffleRead	shuffleWritten	taskId	index	stageId	jobId	phase	phaseTime	timeInstant	appId
2020-01-08 12:27:04.377	10.10.1.7:0	378.99	0	0	0	0	14037	0	97742	10	73588	736	3.Flocks reported	392.73	19	0146
2020-01-08 12:19:36.086	10.10.1.9:9	237.699	0	0	0	0	12127	0	92689	9	66336	699	3.Flocks reported	248.25	18	0146
2020-01-08 13:42:55.240	10.10.1.10:5	235.786	0	0	0	0	10223	0	244083	9	448326	1809	3.Flocks reported	251.44	48	0146
2020-01-08 13:52:17.960	10.10.1.2:0	177.572	0	0	0	0	13210	0	254095	10	485888	1883	3.Flocks reported	195.21	58	0146
2020-01-08 13:47:35.776	10.10.1.1:2	171.391	0	0	0	0	8826	0	249124	4	466918	1846	3.Flocks reported	191.37	49	0146
2020-01-08 12:18:29.040	10.10.1.5:7	170.653	0	0	0	0	12360	0	92692	12	66336	699	3.Flocks reported	248.25	18	0146
2020-01-08 12:23:35.851	10.10.1.6:4	170.463	0	0	0	0	12582	0	97740	8	73588	736	3.Flocks reported	392.73	19	0146
2020-01-08 12:23:33.030	10.10.1.4:3	167.643	0	0	0	0	11792	0	97743	11	73588	736	3.Flocks reported	392.73	19	0146
2020-01-08 12:18:15.858	10.10.1.10:5	157.472	0	0	0	0	10643	0	92691	11	66336	699	3.Flocks reported	248.25	18	0146
2020-01-08 12:14:47.144	10.10.1.5:7	154.757	0	0	0	0	10398	0	87630	10	59462	662	3.Flocks reported	166.08	17	0146
2020-01-08 12:30:29.388	10.10.1.3:1	141.801	0	0	0	0	11811	0	102773	9	81218	773	3.Flocks reported	154.8	20	0146
2020-01-08 12:59:24.048	10.10.1.2:0	140.66	0	0	0	0	6457	0	158051	11	190096	1180	3.Flocks reported	152.02	31	0146
2020-01-08 13:51:33.706	10.10.1.7:0	133.318	0	0	0	0	12423	0	254203	8	485888	1883	3.Flocks reported	195.21	58	0146
2020-01-08 12:48:35.329	10.10.1.6:4	132.943	0	0	0	0	9087	0	138093	9	145212	1032	3.Flocks reported	149.56	27	0146
2020-01-08 12:30:28.020	10.10.1.2:0	131.642	0	0	0	0	10302	0	102775	11	81218	773	3.Flocks reported	154.8	20	0146
2020-01-08 12:39:16.284	10.10.1.5:7	127.817	0	0	0	0	9266	0	102774	10	81218	773	3.Flocks reported	154.8	20	0146
2020-01-08 12:30:07.990	10.10.1.7:0	126.585	0	0	0	0	9495	0	102771	7	81218	773	3.Flocks reported	154.8	20	0146
2020-01-08 13:46:49.124	10.10.1.5:7	124.738	0	0	0	0	10508	0	249129	9	466918	1846	3.Flocks reported	191.37	49	0146
2020-01-08 12:33:36.213	10.10.1.6:4	119.625	0	0	0	0	8159	0	107825	11	89226	810	3.Flocks reported	135.32	21	0146
2020-01-08 12:14:11.183	10.10.1.8:6	118.796	0	0	0	0	8959	0	87631	11	59462	662	3.Flocks reported	166.08	17	0146
2020-01-08 13:46:39.349	10.10.1.6:4	118.423	0	0	0	0	8351	0	249123	3	466918	1846	3.Flocks reported	191.37	49	0146
2020-01-08 13:48:51.948	10.10.1.6:4	112.496	0	0	0	0	8718	0	244082	8	448326	1809	3.Flocks reported	251.44	48	0146
2020-01-08 13:48:51.083	10.10.1.3:1	112.428	0	0	0	0	9981	0	244085	11	448326	1809	3.Flocks reported	251.44	48	0146
2020-01-08 13:46:33.402	10.10.1.10:5	109.914	0	0	0	0	12698	0	249130	10	466918	1846	3.Flocks reported	191.37	49	0146
2020-01-08 12:13:59.913	10.10.1.10:5	107.525	0	0	0	0	10710	0	87633	13	59462	662	3.Flocks reported	166.08	17	0146
2020-01-08 12:09:04.546	10.10.1.3:1	107.16	0	0	0	0	8475	0	77485	8	46848	588	3.Flocks reported	117.55	15	0146
2020-01-08 12:29:54.481	10.10.1.8:6	106.013	0	0	0	0	11466	0	102776	12	81218	773	3.Flocks reported	154.8	20	0146
2020-01-08 12:22:29.870	10.10.1.9:9	104.483	0	0	0	0	15153	0	97739	7	73588	736	3.Flocks reported	392.73	19	0146
2020-01-08 13:51:01.278	10.10.1.10:5	100.891	0	0	0	0	10685	0	254204	9	485888	1883	3.Flocks reported	195.21	58	0146
2020-01-08 12:22:25.645	10.10.1.3:1	100.258	0	0	0	0	10954	0	97741	9	73588	736	3.Flocks reported	392.73	19	0146

Stage analysis...

- ▶ SparkListener also provide info about the whole stage. Some interesting metrics are:
 - ▶ numTasks: The number of tasks in which the stage is divided.
 - ▶ name: The Spark function which invoke the stage and its line of code.
 - ▶ details: The stack trace when the stage was called.

Timestamp	executor	duration	recordsRead	bytesRead	recordsWritten	bytesWritten	shuffleRead	shuffleWritten	taskId	index	stageId	jobId	phase	phaseTime	timeInstant	appId
2020-01-08 12:27:04.377	10.10.1.7:0	370.99	0	0	0	0	14037	0	97742	10	73588	736	[3.Flocks reported]	392.73	19	0146
2020-01-08 12:19:36.086	10.10.1.9:9	237.699	0	0	0	0	12127	0	92689	9	66336	699	[3.Flocks reported]	248.25	18	0146
2020-01-08 12:12:55.105	10.10.1.10:5	10223	0	0	0	0	244083	9	448326	1089	3.Flocks reported	251.44	48	0146		
2020-01-08 13:52:17.900	10.10.1.2:8	177.572	0	0	0	0	13210	0	254205	10	485888	1883	[3.Flocks reported]	195.21	50	0146
2020-01-08 13:47:35.776	10.10.1.1:2	171.391	0	0	0	0	8826	0	249124	4	466918	1846	[3.Flocks reported]	191.37	49	0146
2020-01-08 12:18:29.040	10.10.1.5:7	170.653	0	0	0	0	12360	0	92692	12	66336	699	[3.Flocks reported]	248.25	18	0146
2020-01-08 12:23:35.851	10.10.1.6:4	170.463	0	0	0	0	12582	0	97740	8	73588	736	[3.Flocks reported]	392.73	19	0146
2020-01-08 12:23:33.030	10.10.1.4:3	167.643	0	0	0	0	11702	0	97743	11	73588	736	[3.Flocks reported]	392.73	19	0146
2020-01-08 12:18:15.058	10.10.1.10:5	157.472	0	0	0	0	10643	0	92691	11	66336	699	[3.Flocks reported]	248.25	18	0146
2020-01-08 12:14:47.144	10.10.1.5:7	154.257	0	0	0	0	10396	0	87039	10	59462	662	[3.Flocks reported]	166.08	17	0146
2020-01-08 12:30:29.388	10.10.1.3:1	141.801	0	0	0	0	11011	0	102773	9	81218	773	[3.Flocks reported]	154.8	20	0146
2020-01-08 12:59:24.048	10.10.1.2:8	140.66	0	0	0	0	8457	0	150651	11	190096	1180	[3.Flocks reported]	152.02	31	0146
2020-01-08 13:51:33.706	10.10.1.7:0	133.318	0	0	0	0	12423	0	254203	8	485888	1883	[3.Flocks reported]	195.21	50	0146
2020-01-08 12:48:35.329	10.10.1.6:4	132.943	0	0	0	0	9087	0	138063	9	145212	1032	[3.Flocks reported]	149.56	27	0146
2020-01-08 12:30:20.028	10.10.1.2:8	131.642	0	0	0	0	10302	0	102775	11	81218	773	[3.Flocks reported]	154.8	20	0146
2020-01-08 12:30:16.204	10.10.1.5:7	127.817	0	0	0	0	9268	0	102774	10	81218	773	[3.Flocks reported]	154.8	20	0146
2020-01-08 12:30:07.990	10.10.1.7:0	126.505	0	0	0	0	9268	0	102771	7	81218	773	[3.Flocks reported]	154.8	20	0146
2020-01-08 13:46:49.124	10.10.1.5:7	124.738	0	0	0	0	10508	0	249120	9	466918	1846	[3.Flocks reported]	191.37	49	0146
2020-01-08 12:33:36.213	10.10.1.6:4	119.825	0	0	0	0	8159	0	107825	11	89226	810	[3.Flocks reported]	135.32	21	0146
2020-01-08 12:14:11.183	10.10.1.8:6	118.796	0	0	0	0	8959	0	87631	11	59462	662	[3.Flocks reported]	166.08	17	0146
2020-01-08 13:46:39.349	10.10.1.6:4	118.423	0	0	0	0	8351	0	249123	3	466918	1846	[3.Flocks reported]	191.37	49	0146
2020-01-08 13:40:51.948	10.10.1.6:4	112.496	0	0	0	0	8718	0	244082	8	448326	1089	[3.Flocks reported]	251.44	48	0146
2020-01-08 13:40:51.083	10.10.1.3:1	112.428	0	0	0	0	9981	0	244085	11	448326	1089	[3.Flocks reported]	251.44	48	0146
2020-01-08 13:46:33.402	10.10.1.10:5	109.014	0	0	0	0	12698	0	249130	10	466918	1846	[3.Flocks reported]	191.37	49	0146
2020-01-08 12:13:59.913	10.10.1.10:5	107.525	0	0	0	0	10710	0	87633	13	59462	662	[3.Flocks reported]	166.08	17	0146
2020-01-08 12:09:04.546	10.10.1.3:1	107.16	0	0	0	0	8475	0	77485	8	46848	588	[3.Flocks reported]	117.55	15	0146
2020-01-08 12:29:54.481	10.10.1.8:6	106.013	0	0	0	0	11466	0	102776	12	81218	773	[3.Flocks reported]	154.8	20	0146
2020-01-08 12:22:29.870	10.10.1.9:9	104.483	0	0	0	0	15153	0	97739	7	73588	736	[3.Flocks reported]	392.73	19	0146
2020-01-08 13:51:01.278	10.10.1.10:5	100.891	0	0	0	0	10685	0	254204	9	485888	1883	[3.Flocks reported]	195.21	50	0146
2020-01-08 12:22:25.645	10.10.1.3:1	100.258	0	0	0	0	10954	0	97741	9	73588	736	[3.Flocks reported]	392.73	19	0146