

**Only average time have been recorded here**

*All times are in ms*

Patch	Description	Is Write Impacted?	Is Read Impacted?
BASE:	branch-3.0	N/A	N/A
A:	Use GenericData instead of ReflectData when writing avro data	Yes	No
B:	AvroDeserializer improve struct and array performance by doing some calculation during construction	No	Yes
C:	Reuse objects when reading/writing avro	Yes	Yes
D:	Native avro datum reader and writers, reuseObj in SparkAvroDatumReader commented out	Yes	Yes
E:	SparkAvroDatumReader: reuse objects	No	Yes
F:	SparkAvroDatumReader: always reuse struct, from pool if needed	No	Yes
G:	Expand ArrayData instead of using ArrayBuffer	No	Yes
K:	Reader: skip data not needed, move inner classes outside, remove reusing from pool, and union always	No	Yes

### AvroReadBenchmark

Benchmark	Patch=>	BASE	A	B	C	D	E	F	G	K
SQL Single TINYINT Column Scan:		3578	3788	3695	3169	2698	2731	2744	2835	2821
SQL Single SMALLINT Column Scan:		3691	3765	3757	3319	2849	2877	2918	2927	2824
SQL Single INT Column Scan:		3657	3904	3823	3349	2864	2907	2929	3042	2824
SQL Single BIGINT Column Scan:		4246	4400	4284	3805	3522	3459	3503	3678	3425
SQL Single FLOAT Column Scan:		3688	3659	3762	3164	2920	2977	2937	2961	2752
SQL Single DOUBLE Column Scan:		3733	3938	3951	3326	3125	3177	3099	3300	3010
Int and String Scan:		5821	6774	7172	6058	5525	5411	5412	5566	5449
Partitioned Table:	Data column	4207	4387	4448	5792	3930	4029	3906	4155	3860
	Partition column	3903	4075	3924	3795	3789	3627	3636	3591	3225
	Both columns	4253	4833	4466	3894	3863	3652	3716	3855	3635
Repeated String:		4265	4496	4446	3734	3868	3806	3701	3963	3571
String with Nulls Scan		7091	7173	7262	6199	6289	6505	6212	6439	6208
String with Nulls Scan (50.0%):		5597	5833	5751	5399	5199	5578	5173	5359	4989
String with Nulls Scan (95.0%):		3901	4309	4093	3886	3591	3624	3580	3594	3426
Single Column Scan from 100 columns:		6844	7281	7160	6825	6829	6526	6702	6776	4715
Single Column Scan from 200 columns:		14466	14056	14254	13050	13299	12877	12823	13353	9646
Single Column Scan from 300 columns:		20198	21253	20951	19660	19710	19641	19744	19991	14147
Nested Struct Scan:		75306	110186	53598	49112	48366	46881	50596	47448	46232
Array Scan:		6251	6498	5798	6113	7714	7164	8935	6115	5563
Array of Struct Scan:		47392	36769	20211	20323	18911	18557	18918	17921	16932

### AvroWriteBenchmark

<u>Benchmark</u>	<u>Patch=&gt;</u>	<u>BASE</u>	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>	<u>F</u>	<u>G</u>	<u>K</u>
Output Single Int Column		4096	3044	3299	2650	2205	2296	2274	3369	2227
Output Single Double Column		4373	3846	3412	3008	2519	2437	3220	2423	2366
Output Int and String Column		9211	7388	8553	7514	6651	6515	6848	6589	6699
Output Partitions		8545	8120	6742	7181	6301	6301	6341	6513	6439
Output Buckets		9867	8446	8454	7860	7768	7925	8574	8245	7750
Nested Struct		147142	88118	98717	93397	34484	36648	33181	36907	35946
Array		33355	19635	22280	19855	7183	6926	7932	9434	9303
Array of Struct		139380	83614	95421	87900	31805	33861	29230	33763	33992

**AvroNestedSchemaPruningBenchmark**

<b>Benchmark</b>	<b>Patch=&gt;</b>	<b>BASE</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>K</b>
Selection:	Top-level column	4474	4525	4376	4044	4117	4364	4104	4082	3323
	Nested column	6151	5974	4959	4568	5044	4961	4958	4619	4014
	Nested column in array	8903	7979	5794	5526	5329	5818	5133	5218	5156
Limiting:	Top-level column	5037	4829	4994	4412	4591	4599	4558	4336	3598
	Nested column	6863	6085	5389	5278	5414	5640	5420	5274	4611
	Nested column in array	10109	8920	6654	6227	6519	6572	5944	5900	6140
Repartitioning:	Top-level column	4892	4914	4756	4335	4404	4701	4490	4331	3637
	Nested column	6780	5988	5418	5085	5404	5437	5325	5009	4565
	Nested column in array	10077	8827	6442	6243	6051	6643	5902	5845	5807
Repartitioning by exprs:	Top-level column	5054	4869	4919	4336	4468	4476	4463	4196	3615
	Nested column	8687	7634	7054	6968	7010	7405	7032	7084	6110
	Nested column in array	11210	9865	7737	7261	7457	7469	6878	6832	7045
Sample:	Top-level column	4541	4349	4388	3918	4025	4129	4042	3957	3149
	Nested column	6102	5447	4939	4664	5023	4912	4850	4685	4084
	Nested column in array	9036	8183	5838	5449	5431	5673	5192	5170	5178
Sorting:	Top-level column	5385	5184	5280	4753	4812	4865	4782	4535	3917
	Nested column	9741	8743	8126	7977	8138	8339	8172	8117	7226
	Nested column in array	12795	11355	9061	8741	8686	8915	8444	8110	8622