

Table 1: Raw result (Mean and Standard Error) for MARIADB under different number of measurements within the search budget of 400 (for measurement-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	100th		200th		300th		400th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
IRACE	O1: average query runtime (ms)	60399	255	58375	38	57936	29	57716	132
GA		59224	145	57700	39	56957	4	56769	78
RS		62580	200	58473	15	57845	12	57321	95
MMO		59322	197	57184	8	56848	5	56661	78
PMO		60719	372	58059	141	56957	13	56680	82
AdMMO		59191	240	56860	17	56444	9	56432	65
IRACE	O2: CPU load (%)	127.16	0.84	116.32	0.33	110.75	0.3	104.98	1.05
GA		117.92	0.26	106.61	0.08	102.52	0.02	100.89	0.52
RS		127.07	0.25	114.34	0.14	111.01	0.12	106.05	1.0
MMO		119.6	0.34	105.76	0.09	101.89	0.03	100.73	0.56
PMO		124.76	0.4	116.13	0.19	113.1	0.16	106.87	1.13
AdMMO		120.35	0.5	108.14	0.21	104.37	0.14	101.62	0.45

Table 2: Raw result (Mean and Standard Error) for STORM under different number of measurements within the search budget of 400 (for measurement-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	100th		200th		300th		400th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
IRACE	O1: throughput (msgs/min)	30065	136	31678	60	32573	52	33556	170
GA		33049	46	34442	2	34501	1	34579	22
RS		28304	72	32345	27	33269	18	33678	135
MMO		32658	86	34508	3	34567	1	34615	21
PMO		29705	168	33444	41	34230	13	34558	16
AdMMO		32375	107	34528	7	34656	2	34718	10
IRACE	O2: average latency (ms)	3877	22	3649	9	3532	6	3439	9
GA		3459	8	3323	0	3317	0	3317	0
RS		5456	294	3599	4	3508	2	3448	13
MMO		3915	98	3331	1	3318	0	3317	0
PMO		4330	55	3578	17	3394	4	3338	9
AdMMO		3417	59	3316	1	3316	0	3317	0

Table 3: Raw result (Mean and Standard Error) for VP9 under different number of measurements within the search budget of 400 (for measurement-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	100th		200th		300th		400th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
IRACE	O1: total runtime (s)	55809	2093	36123	779	32579	261	31168	790
GA		54835	960	35359	200	31502	102	29048	489
RS		54085	396	38489	150	32492	95	30662	662
MMO		55901	1239	35882	371	31012	138	28542	8
PMO		55495	1205	41018	557	37356	475	33168	1394
AdMMO		46031	910	33671	303	30158	176	28509	7
IRACE	O2: CPU load (%)	383.67	4.75	353.26	0.8	349.04	0.15	348.06	0.46
GA		371.55	1.73	352.22	0.26	345.38	0.09	343.41	0.17
RS		392.09	1.19	354.92	0.13	350.97	0.08	348.83	0.34
MMO		377.1	2.25	354.5	0.28	349.15	0.16	353.52	8.52
PMO		394.5	2.94	364.27	0.64	360.05	0.5	353.52	8.52
AdMMO		385.48	3.02	351.97	0.38	347.35	0.22	345.57	0.95

Table 4: Raw result (Mean and Standard Error) for KERAS under different number of measurements within the search budget of 400 (for measurement-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	100th		200th		300th		400th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
IRACE	O1: AUC	0.5831	0.0041	0.6215	0.0032	0.6493	0.0032	0.6921	0.014
GA		0.5768	0.0028	0.6281	0.0017	0.6315	0.0011	0.6329	0.0127
RS		0.5621	0.0010	0.6024	0.0013	0.6342	0.0014	0.665	0.015
MMO		0.5638	0.0025	0.5978	0.0023	0.6129	0.0018	0.6357	0.0124
PMO		0.6513	0.0041	0.7625	0.0016	0.7857	0.0	0.7857	0.0
AdMMO		0.5803	0.0032	0.6459	0.0030	0.6726	0.0022	0.6836	0.0147
IRACE	O2: average inference runtime (ms)	80	1	70	0	69	0	69	0
GA		71	1	67	0	67	0	67	0
RS		149	7	70	0	69	0	68	0
MMO		82	3	67	0	67	0	67	0
PMO		82	3	67	0	67	0	67	0
AdMMO		80	4	62	0	62	0	62	0

Table 5: Raw result (Mean and Standard Error) for MONGODB under different number of measurements within the search budget of 400 (for measurement-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	100th		200th		300th		400th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
IRACE	O1: average query runtime (ms)	265415	3780	259488	3699	245590	2091	236088	3314
GA		260235	2164	242953	793	231667	430	236400	4967
RS		280734	794	260576	561	248677	470	239899	3938
MMO		275277	2612	242898	1016	233779	628	232420	4266
PMO		259463	1932	240135	877	230483	510	233744	4411
AdMMO		254644	2155	235855	908	228067	605	224759	3765
IRACE	O2: CPU load (%)	3561.16	48.6	3485.47	39.93	3309.25	24.73	3171.39	48.31
GA		3448.56	21.93	3250.71	7.32	3180.95	4.74	3122.61	69.63
RS		3538.63	10.9	3535.46	6.69	3444.99	5.71	3143.25	47.34
MMO		3432.81	24.55	3232.21	12.17	3145.78	7.14	3234.47	91.54
PMO		3465.58	22.26	3283.43	9.26	3235.05	6.03	3136.93	68.63
AdMMO		3431.17	26.34	3225.0	12.44	3122.51	9.34	3035.55	62.6

Table 6: Raw result (Mean and Standard Error) for x264 under different number of measurements within the search budget of 400 (for measurement-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	100th		200th		300th		400th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
IRACE	O1: PNSR	51.84	0.0	51.84	0.0	51.84	0.0	51.84	0.0
GA		52.41	0.05	53.5	0.04	53.92	0.03	54.09	0.11
RS		47.12	0.0	47.12	0.0	47.12	0.0	47.12	0.0
MMO		52.28	0.04	53.31	0.05	53.87	0.04	54.22	0.14
PMO		52.11	0.03	53.03	0.06	53.66	0.06	54.19	0.19
AdMMO		52.51	0.05	53.45	0.06	53.87	0.05	54.19	0.14
IRACE	O2: Energy (watt)	404.44	0.0	404.44	0.0	404.44	0.0	404.44	0.0
GA		385.84	1.06	369.6	0.79	359.2	0.64	357.59	2.18
RS		1064.83	0.0	1064.83	0.0	1064.83	0.0	1064.83	0.0
MMO		392.41	0.83	367.25	0.79	357.59	0.5	355.0	1.31
PMO		398.36	0.67	372.93	1.09	358.11	0.96	352.28	2.65
AdMMO		389.39	0.91	362.51	0.94	350.48	0.81	346.68	2.55

Table 7: Raw result (Mean and Standard Error) for LLVM under different number of measurements within the search budget of 400 (for measurement-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	100th		200th		300th		400th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
IRACE	O1: average runtime (ms)	59572	94	57601	55	56910	51	56299	205
GA		56730	121	52567	8	52285	0	52285	0
RS		61264	55	58227	24	57418	19	56786	176
MMO		57033	164	52514	10	52285	0	52285	0
PMO		58655	165	53802	63	52543	18	52463	97
AdMMO		56517	155	52508	20	52285	1	52285	0
IRACE	O2: CPU load (%)	2561.43	0.57	2544.03	0.36	2538.99	0.29	2536.01	1.18
GA		2558.85	0.5	2539.07	0.16	2535.98	0.1	2534.81	1.02
RS		2559.95	0.2	2547.2	0.14	2541.77	0.12	2537.13	1.16
MMO		2553.34	0.57	2538.78	0.23	2535.93	0.16	2533.72	0.78
PMO		2553.41	0.56	2540.84	0.21	2538.73	0.09	2536.04	0.6
AdMMO		2552.02	0.63	2540.26	0.36	2534.48	0.25	2530.88	0.67

Table 8: Raw result (Mean and Standard Error) for MARIADB under different number of measurements within the search budget of 50 (for model-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	35th		40th		45th		50th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
FLASH	O1: average query runtime (ms)	60742	337	60741	337	60478	138	60313	289
BOCA		61005	412	60826	400	60070	164	60026	167
AdMMO		58769	277	58231	80	58059	64	57983	160
FLASH	O2: CPU load (%)	130.16	1.18	122.57	0.63	122.57	0.33	122.57	0.88
BOCA		128.72	1.76	128.55	1.4	126.03	0.95	124.38	1.72
AdMMO		124.73	1.01	121.46	0.94	118.11	0.72	117.38	1.77

Table 9: Raw result (Mean and Standard Error) for STORM under different number of measurements within the search budget of 50 (for model-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	35th		40th		45th		50th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
FLASH	O1: throughput (msgs/min)	27095	242	27095	242	27095	171	27095	541
BOCA		28000	270	28031	266	28672	184	30171	538
AdMMO		30746	293	31560	264	32419	215	32433	506
FLASH	O2: average latency (ms)	4346	38	4346	38	4323	25	4320	80
BOCA		4240	37	4235	37	4175	25	4158	80
AdMMO		3886	50	3641	26	3517	18	3485	38

Table 10: Raw result (Mean and Standard Error) for VP9 under different number of measurements within the search budget of 50 (for model-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	35th		40th		45th		50th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
FLASH	O1: total time (s)	54245	2517	54245	2517	54245	2517	54245	2517
BOCA		54669	2676	54669	2676	54669	2676	54669	2676
AdMMO		57311	2498	51915	1946	51475	1909	51408	3577
FLASH	O2: CPU load (%)	367.84	1.77	367.84	1.77	367.84	1.77	367.84	1.77
BOCA		364.48	1.88	364.48	1.88	364.48	1.88	364.48	1.88
AdMMO		384.10	5.82	381.69	5.27	371.02	2.72	371.00	8.83

Table 11: Raw result (Mean and Standard Error) for KERAS under different number of measurements within the search budget of 50 (for model-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	35th		40th		45th		50th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
FLASH	O1: AUC	0.5514	0.0032	0.5514	0.0032	0.5514	0.0023	0.5514	0.0072
BOCA		0.5671	0.0051	0.5671	0.0051	0.5671	0.0036	0.5671	0.0115
AdMMO		0.5797	0.0082	0.5843	0.0086	0.5894	0.0082	0.5936	0.0128
FLASH	O2: average inference runtime (ms)	103	2	103	2	103	2	103	5
BOCA		98	2	97	2	86	2	84	5
AdMMO		81	4	75	3	74	3	74	7

Table 12: Raw result (Mean and Standard Error) for MONGODB under different number of measurements within the search budget of 50 (for model-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	35th		40th		45th		50th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
FLASH	O1: average query runtime (ms)	259811	767	259137	4561	259137	4561	259137	4561
BOCA		255311	933	245010	3902	245010	3902	245010	3902
AdMMO		272798	7562	256067	6007	261723	6831	261148	9143
FLASH	O2: CPU load (%)	3465.42	53.83	3361.41	48.81	3361.41	48.81	3361.41	48.81
BOCA		3654.32	31.34	3333.99	34.88	3333.99	34.88	3333.99	34.88
AdMMO		3636.02	93.84	3415.91	73.48	3361.68	55.14	3344.92	93.31

Table 13: Raw result (Mean and Standard Error) for x264 under different number of measurements within the search budget of 50 (for model-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	35th		40th		45th		50th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
FLASH	O1: PNSR	51.15	0.11	51.74	0.07	52.0	0.04	52.0	0.14
BOCA		49.33	0.73	49.63	0.83	49.63	0.83	49.63	0.83
AdMMO		51.83	0.04	51.88	0.05	52.93	0.07	52.98	0.09
FLASH	O2: Energy (watt)	512.2	12.82	505.11	4.46	498.32	2.99	498.32	9.44
BOCA		755.48	59.25	742.18	59.25	742.18	59.25	742.18	59.25
AdMMO		411.52	1.27	410.41	1.38	411.47	1.22	410.14	1.67

Table 14: Raw result (Mean and Standard Error) for LLVM under different number of measurements within the search budget of 50 (for model-based optimizers). The best is highlighted in green.

Optimizer	Performance Objective	35th		40th		45th		50th	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
FLASH	O1: average runtime (ms)	60843	180	60777	173	60777	122	60777	386
BOCA		61083	152	61068	150	60972	104	60810	321
AdMMO		56572	215	55954	191	55051	182	54402	264
FLASH	O2: CPU load (%)	2561.32	0.81	2560.48	0.75	2560.48	0.53	2560.48	1.68
BOCA		2563.28	0.87	2562.96	0.85	2562.29	0.58	2561.21	1.74
AdMMO		2557.44	1.27	2555.31	1.17	2554.62	1.14	2553.25	1.67