Supplementary Material for: Automatic configuration of the Cassandra database using irace

Moisés Silva-Muñoz[*1], Alberto Franzin[†1], and Hugues Bersini[‡1]

[1]IRIDIA, Université Libre de Bruxelles (ULB), Belgium

December 2020

1 Tuning 5 parameters for all the six YCSB workloads (scenario 1

[*]moises.silva.munoz@ulb.be
[†]afranzin@ulb.ac.be
[‡]bersini@ulb.be

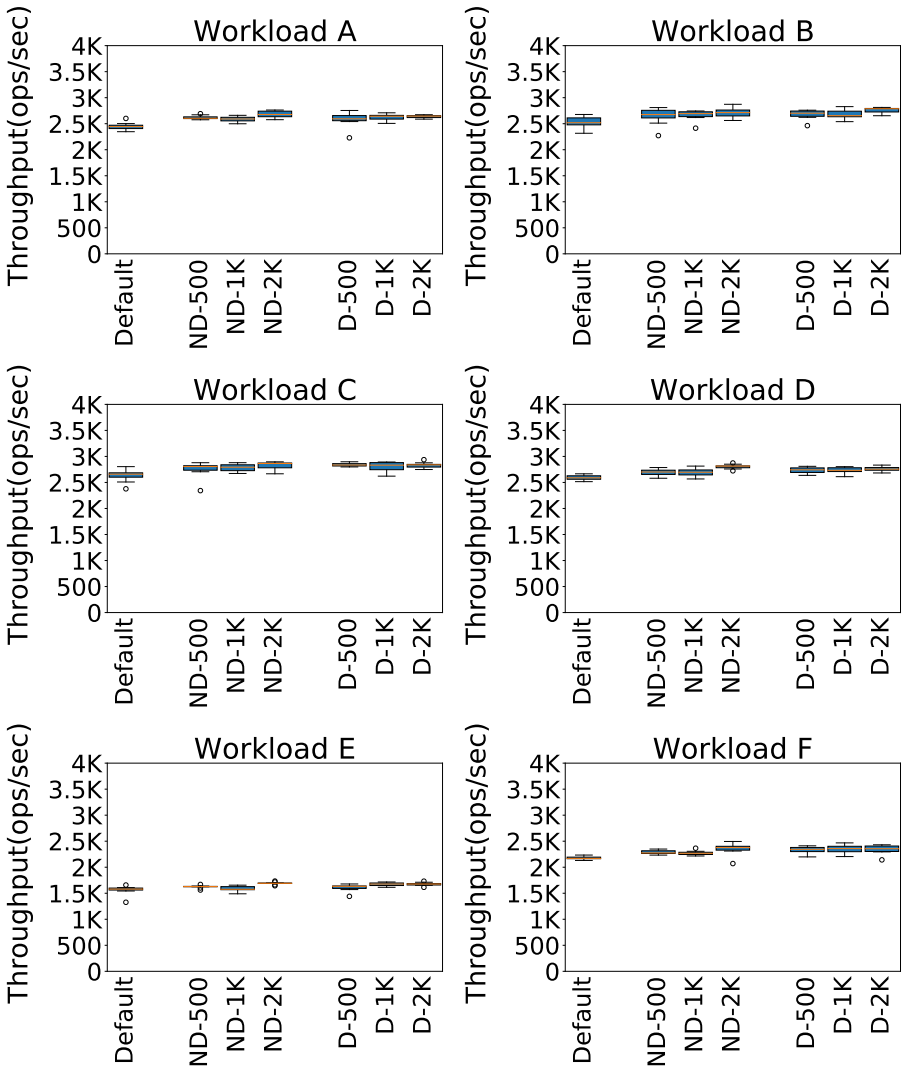

Figure 1: Throughput obtained tuning 5 parameters for all the six YCSB workloads (scenario 1). The results reported represent the throughput measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

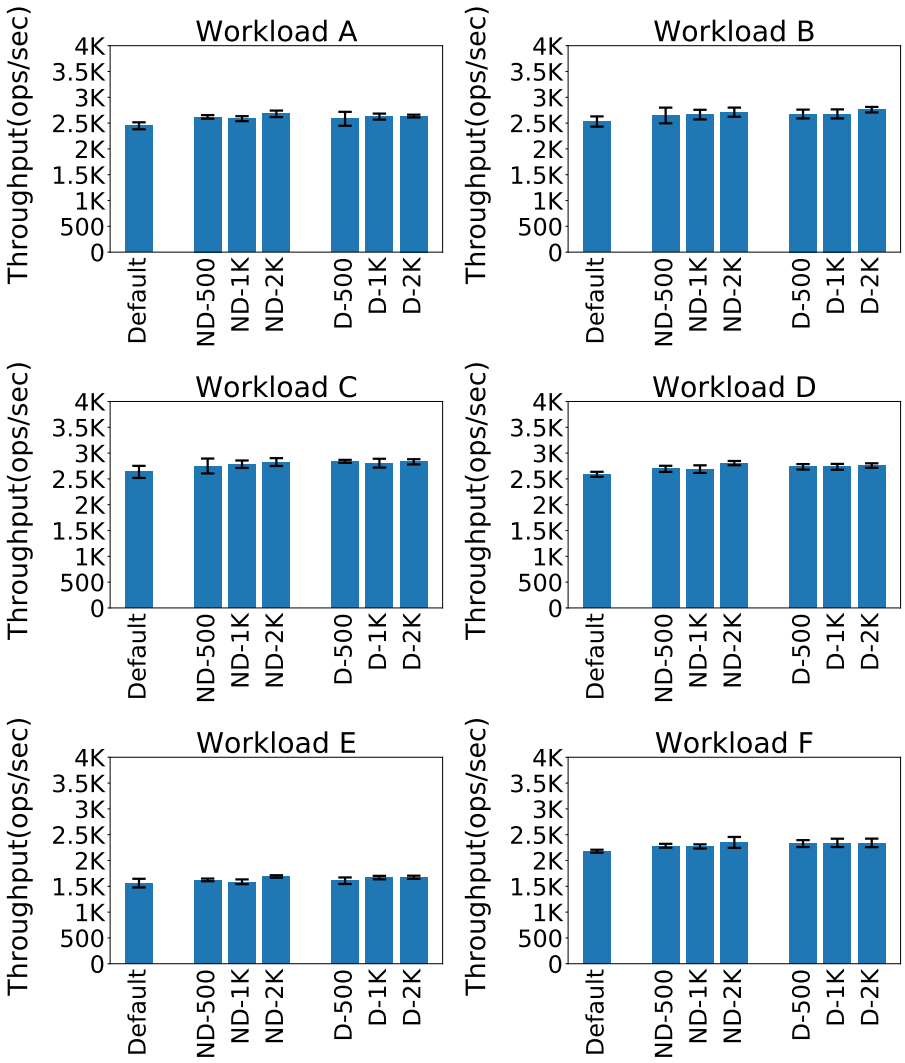

Figure 2: Throughput obtained tuning 5 parameters for all the six YCSB workloads (scenario 1). The results reported represent the throughput measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

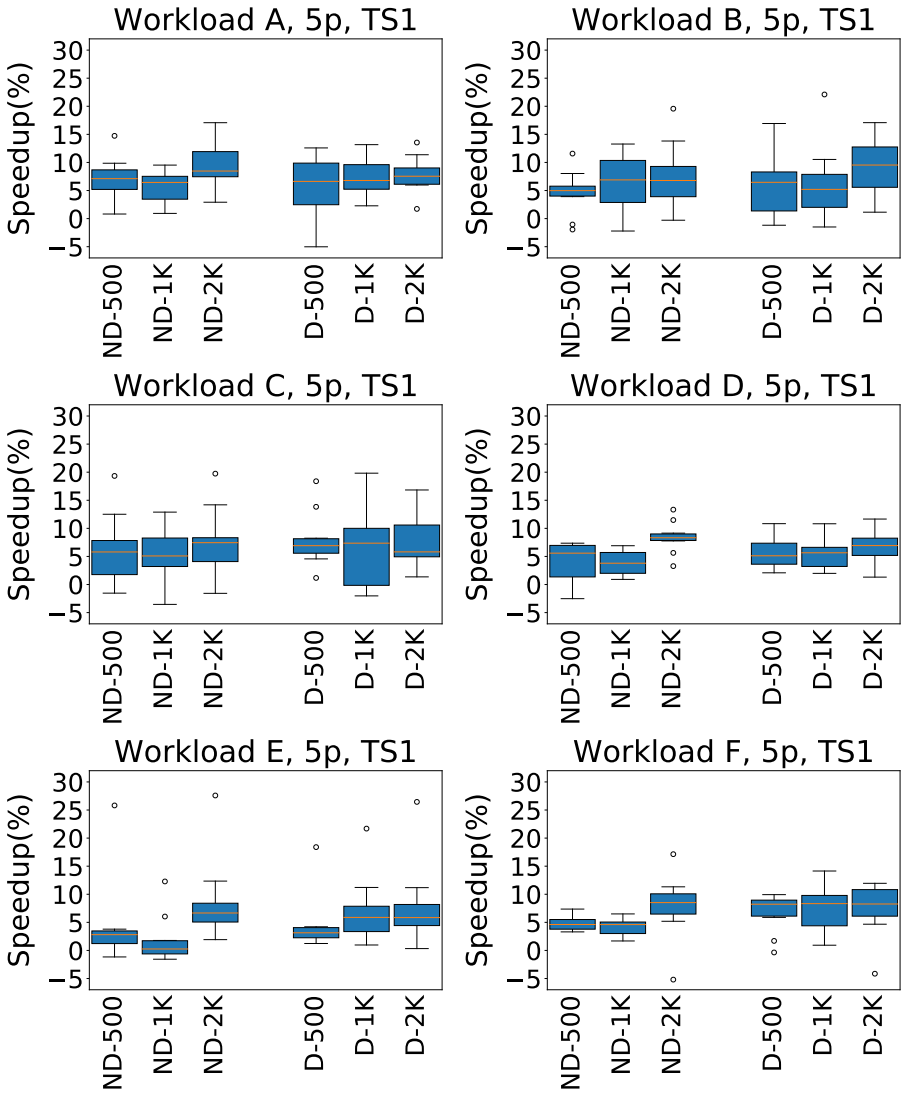

Figure 3: Speedup with respect to the default configuration obtained tuning 5 parameters for all the six YCSB workloads (scenario 1). The results reported represent the throughput measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the speedup obtained by, from left to right, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

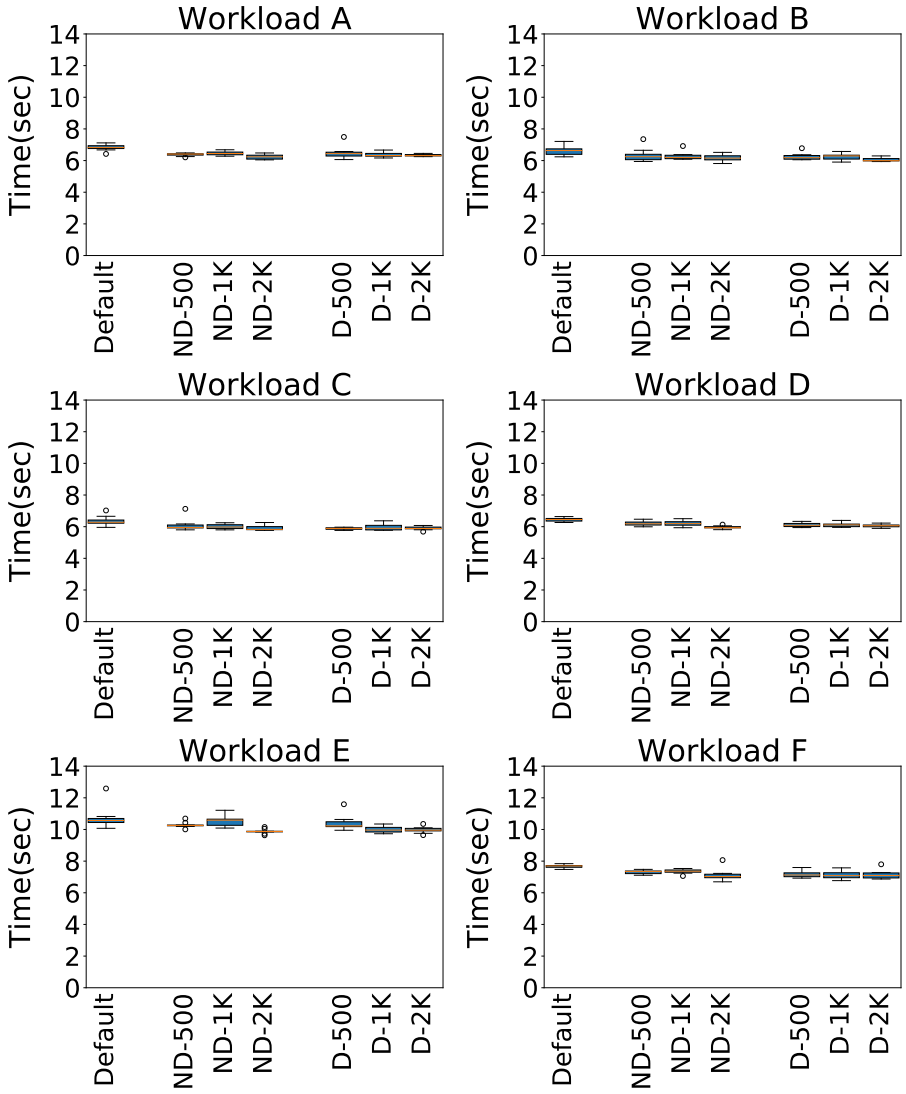

Figure 4: Time tuning 5 parameters and the 6 Workloads treated as instance in scenario 1. The results reported represent the time measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the time obtained by, from left to right, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

2 Tuning 5 parameters for all the six YCSB workloads (scenario 2)

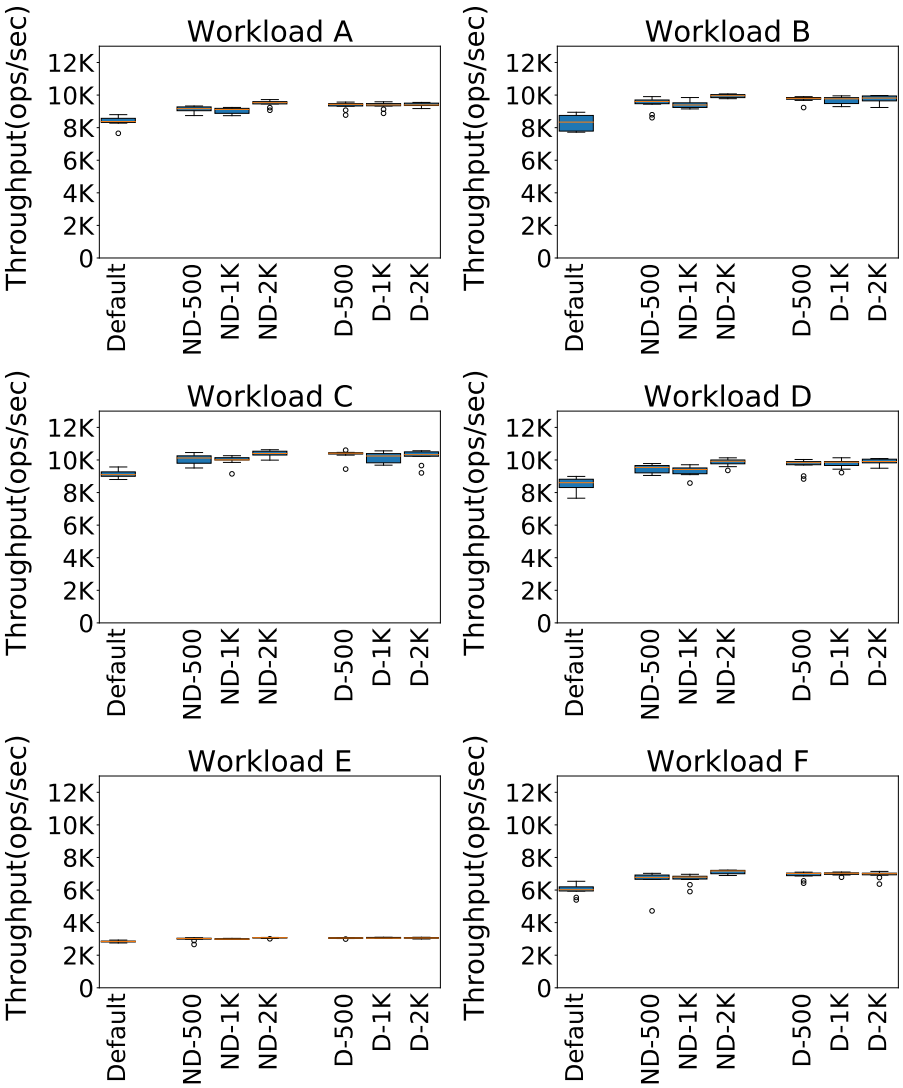

Figure 5: Throughput obtained tuning 5 parameters for all the six YCSB workloads (scenario 2). The results reported represent the throughput measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

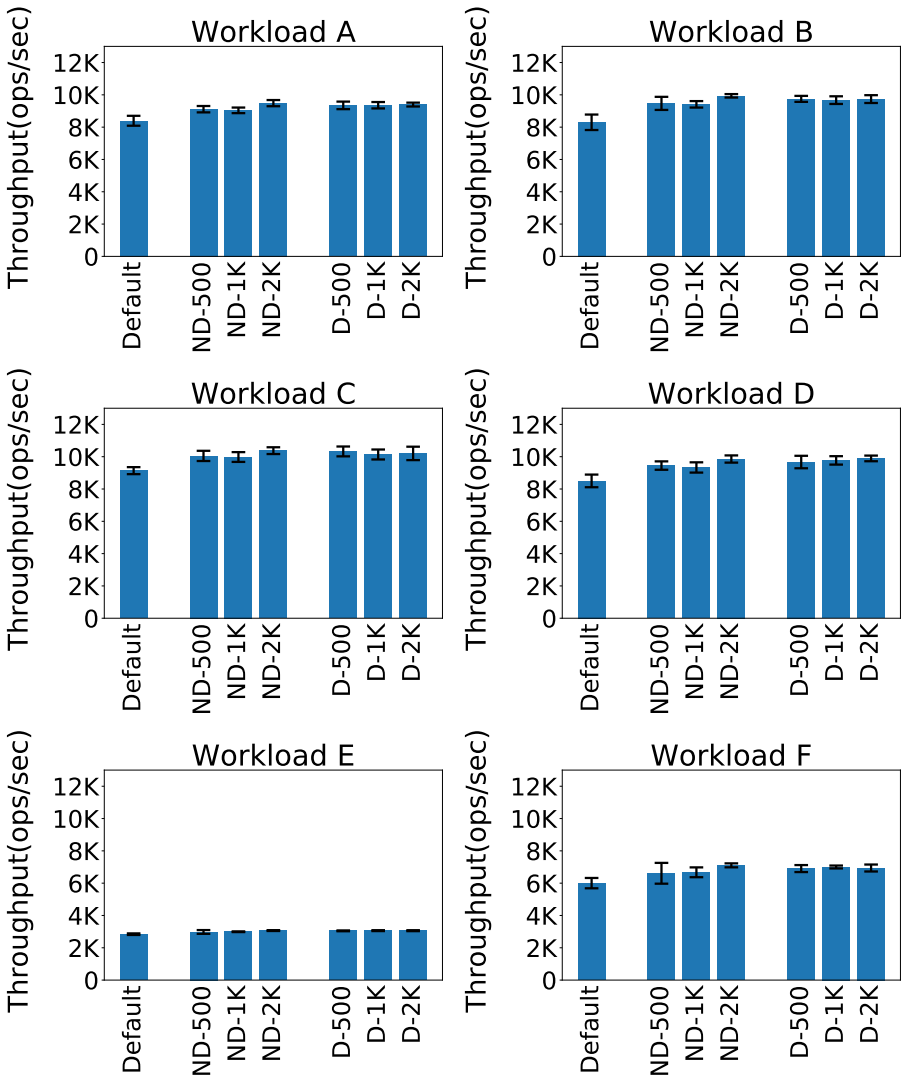

Figure 6: Throughput obtained tuning 5 parameters for all the six YCSB workloads (scenario 2). The results reported represent the throughput measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

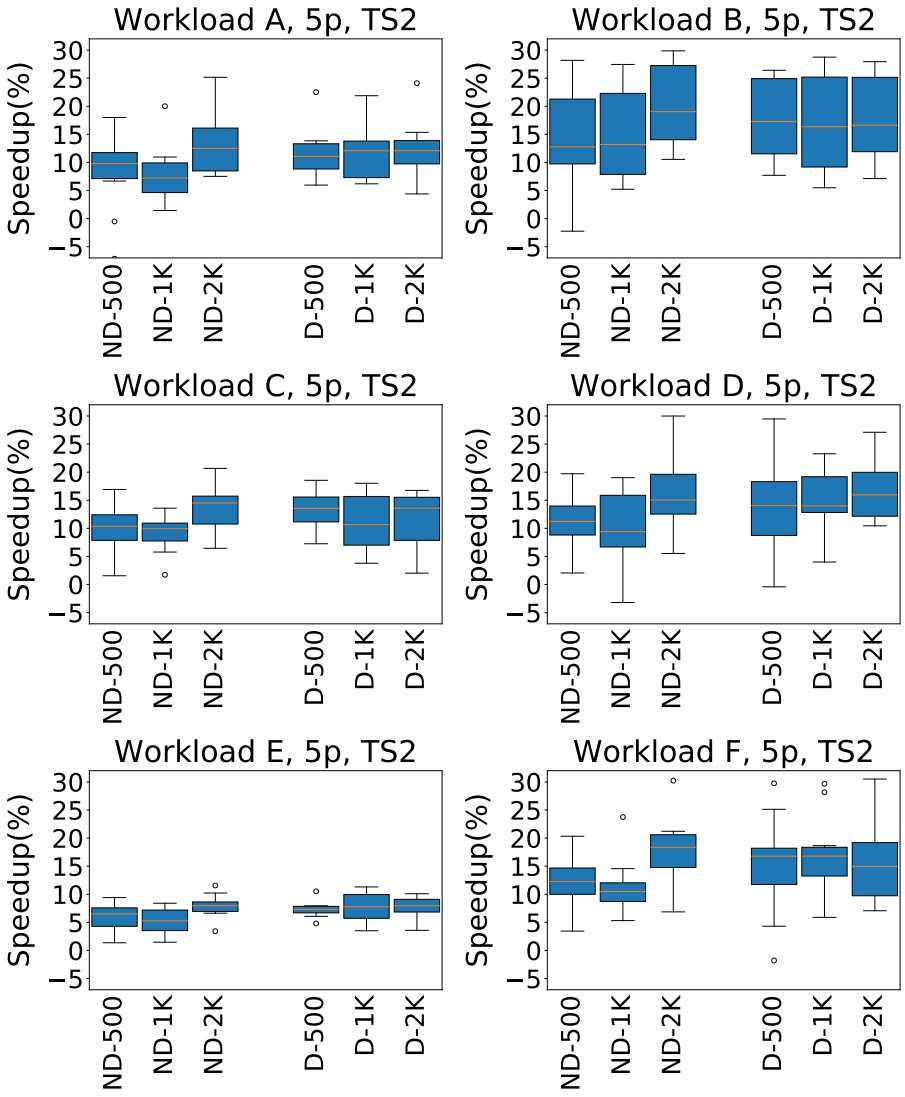

Figure 7: Speedup with respect to the default configuration obtained tuning 5 parameters for all the six YCSB workloads (scenario 2). The results reported represent the throughput measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the speedup obtained by, from left to right, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

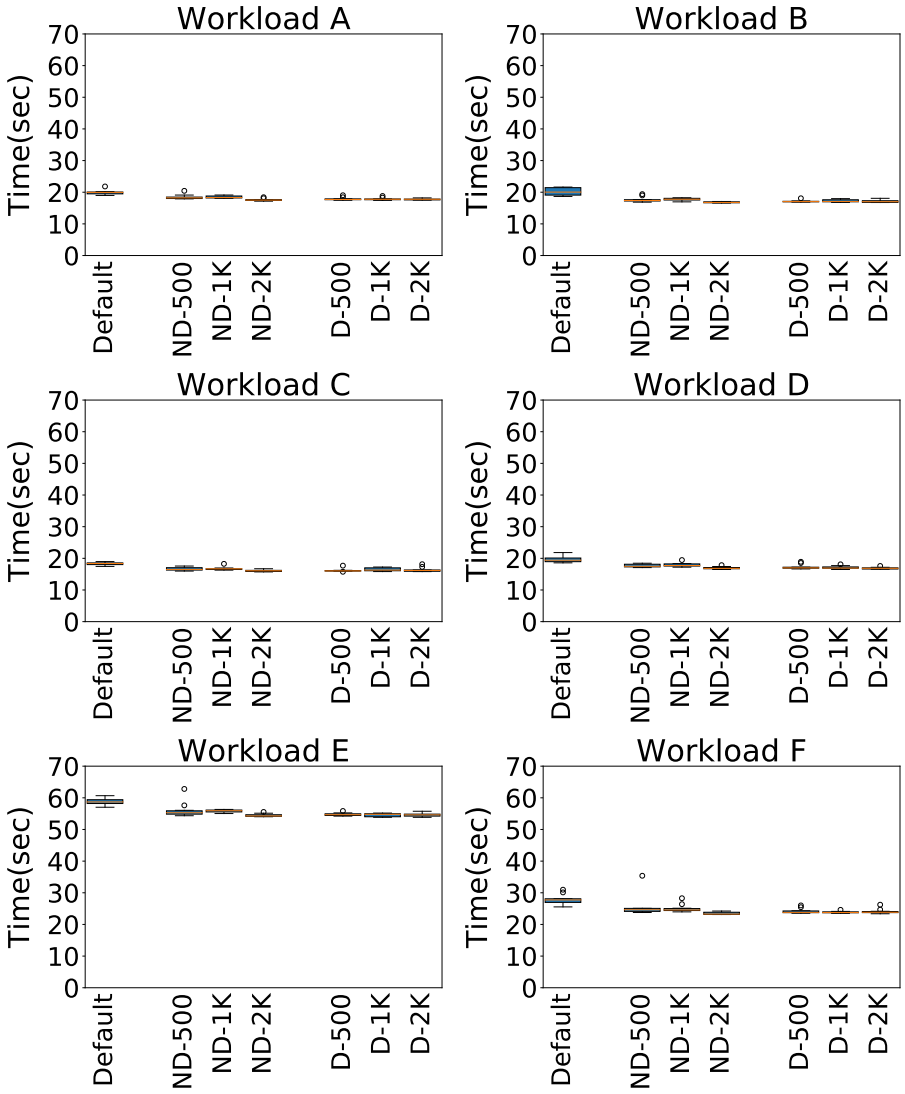

Figure 8: Time tuning 5 parameters and the 6 Workloads treated as instance in scenario 2. The results reported represent the time measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the time obtained by, from left to right, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

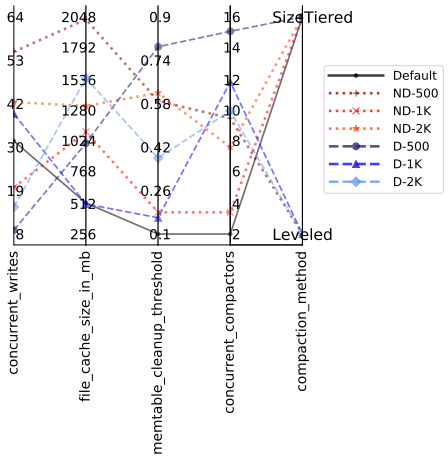

Figure 9: parallel coordinates plot WL6 and 5 parameters

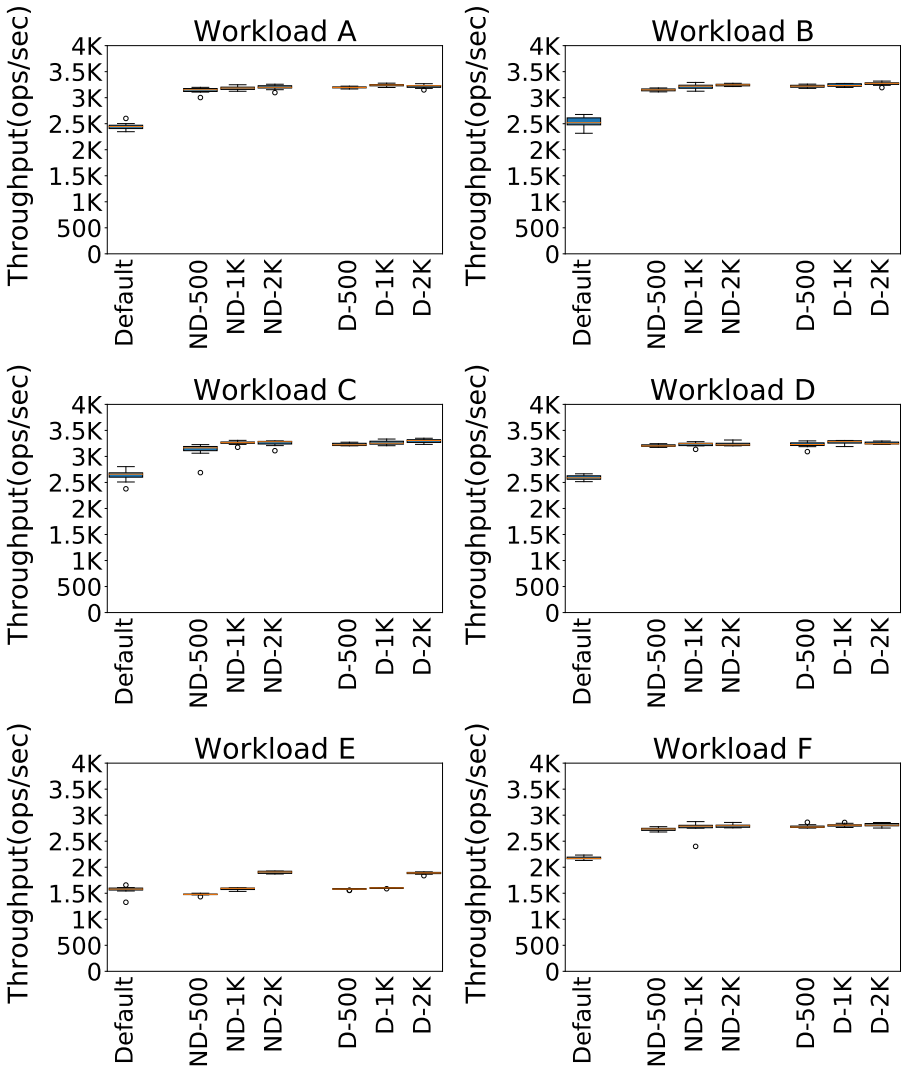

Figure 10: Throughput obtained tuning 23 parameters for all the 6 YCSB workloads (scenario 1). The results reported represent the throughput measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

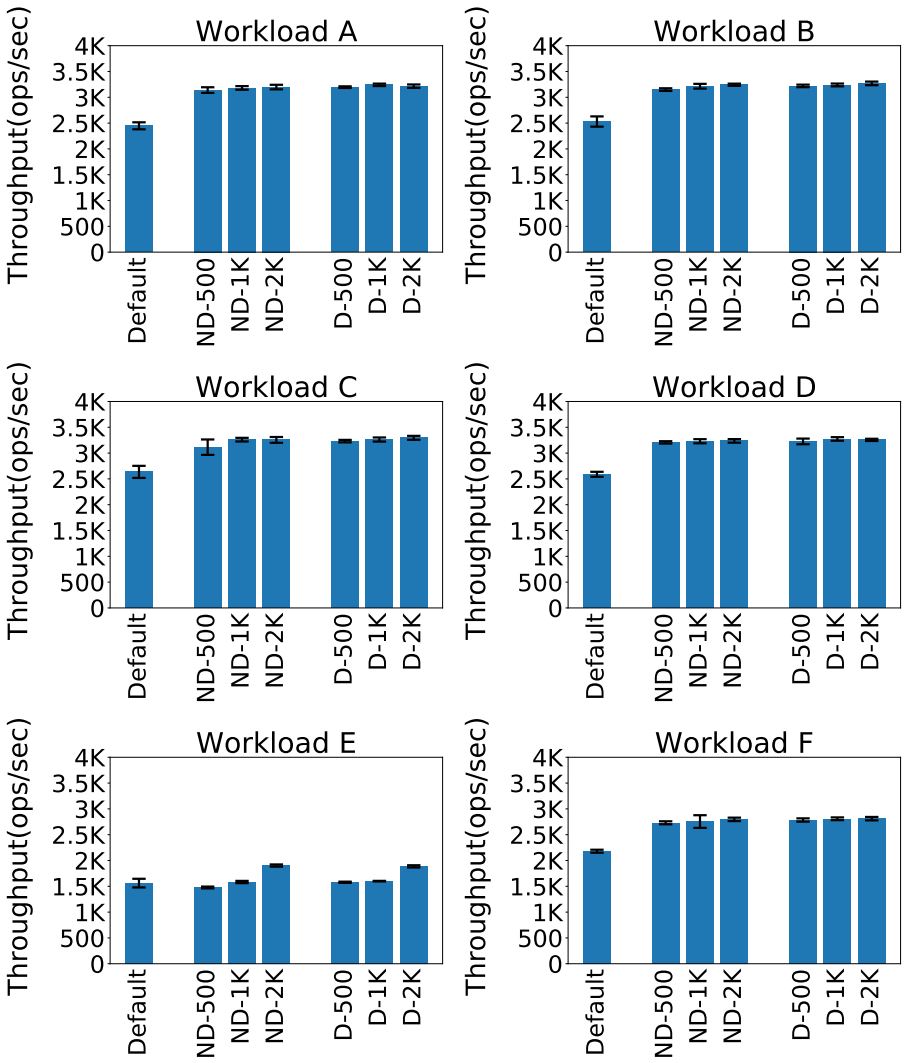

Figure 11: Throughput obtained tuning 23 parameters for all the 6 YCSB workloads (scenario 1). The results reported represent the throughput measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

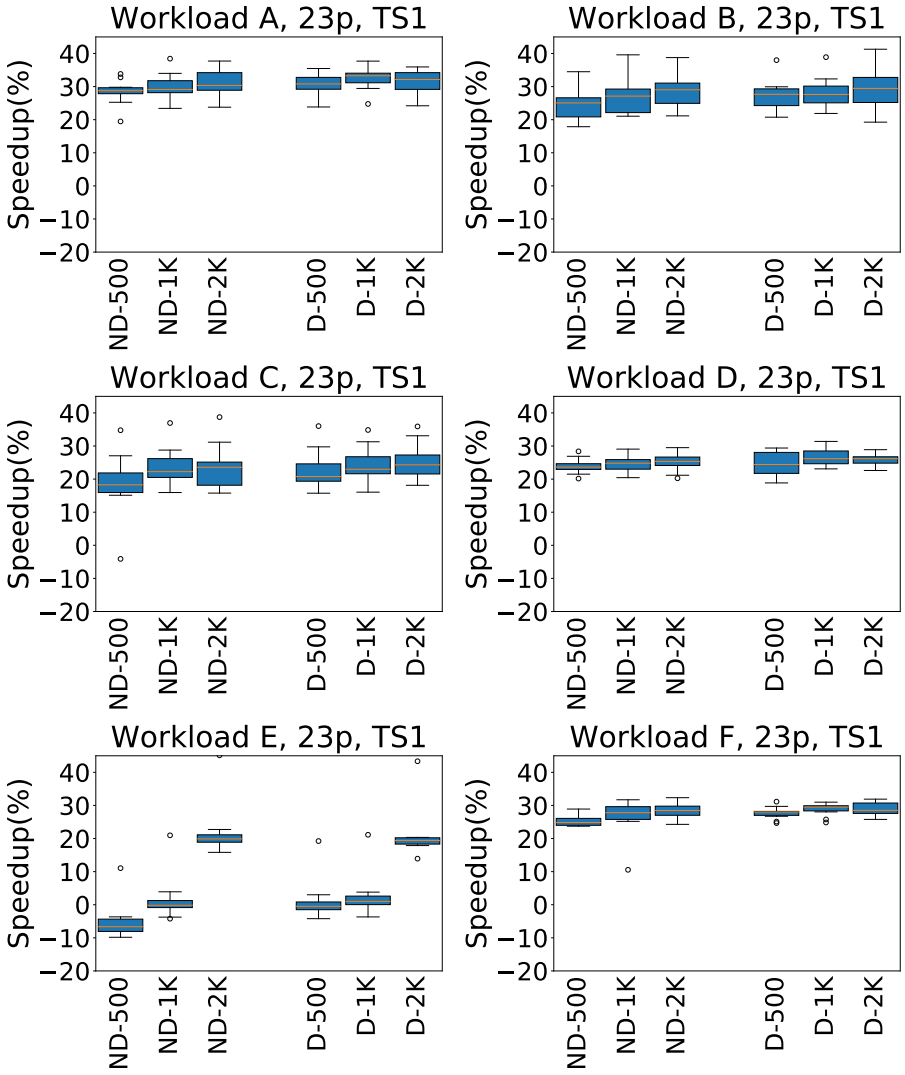

Figure 12: Speedup with respect to the default configuration obtained tuning 23 parameters for all the 6 YCSB workloads (scenario 1). The results reported represent the throughput measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the speedup obtained by, from left to right, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

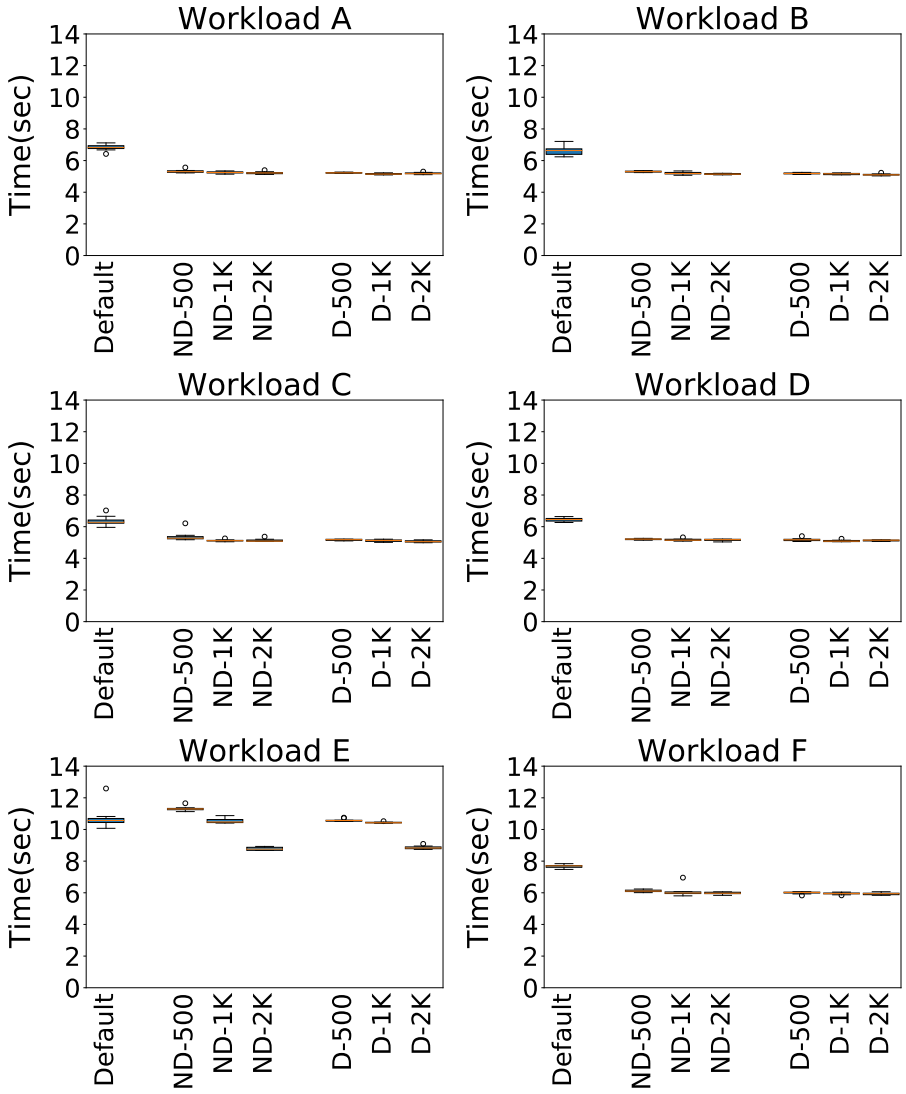

Figure 13: Time tuning 23 parameters and the 6 Workloads treated as instance in scenario 1. The results reported represent the time measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the time obtained by, from left to right, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

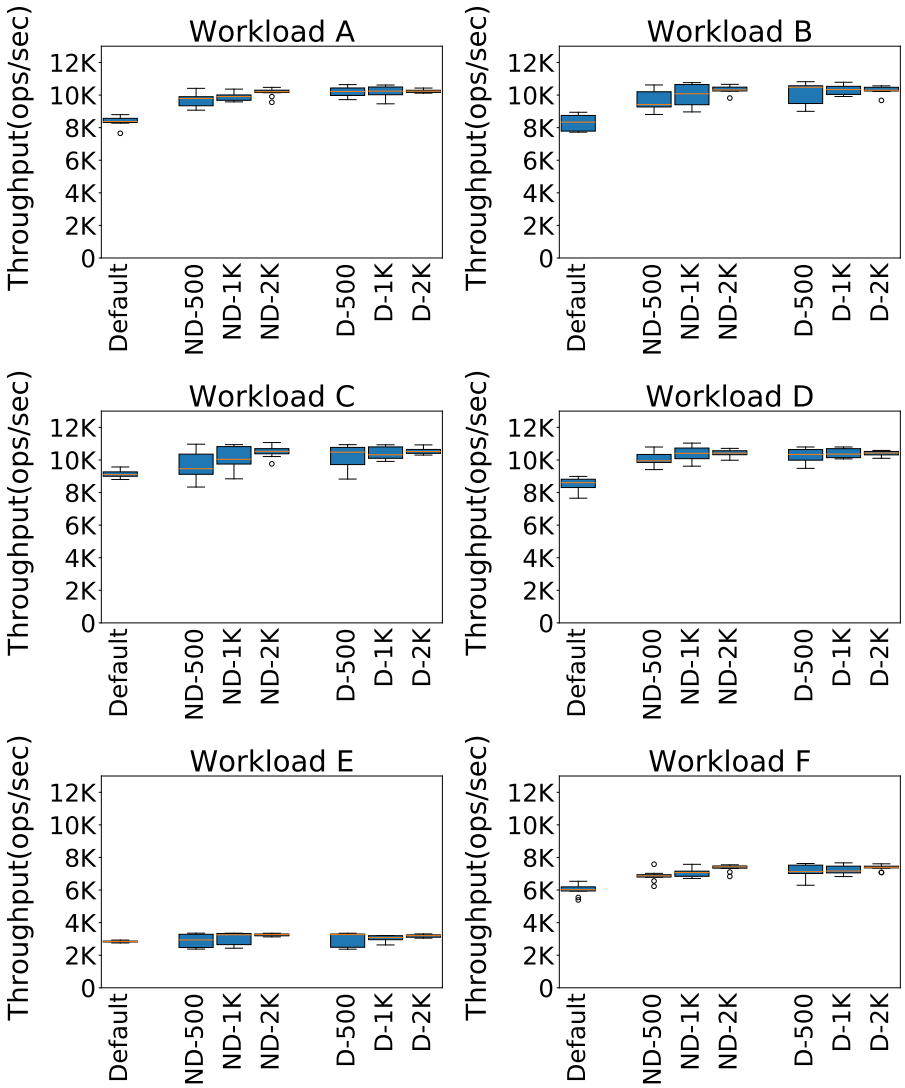

Figure 14: Throughput obtained tuning 23 parameters for all the 6 YCSB workloads (scenario 2). The results reported represent the throughput measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

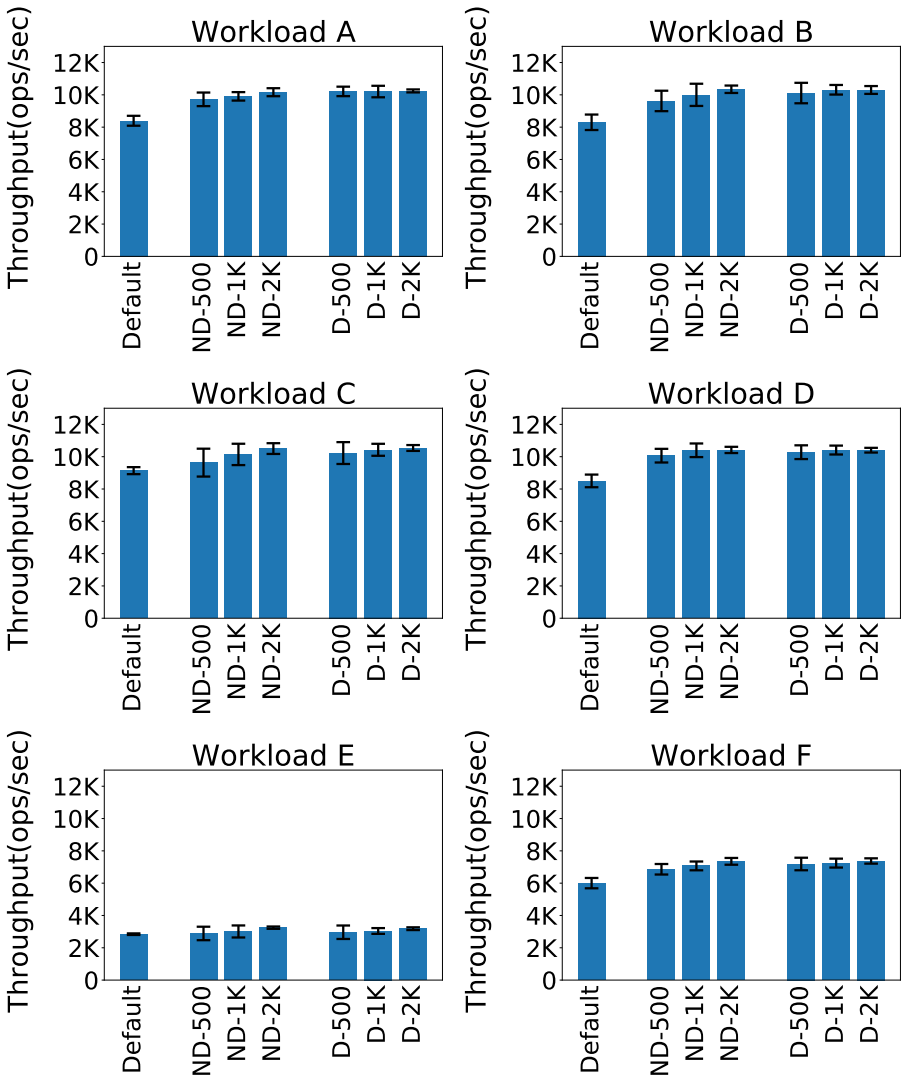

Figure 15: Throughput obtained tuning 23 parameters for all the 6 YCSB workloads (scenario 2). The results reported represent the throughput measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

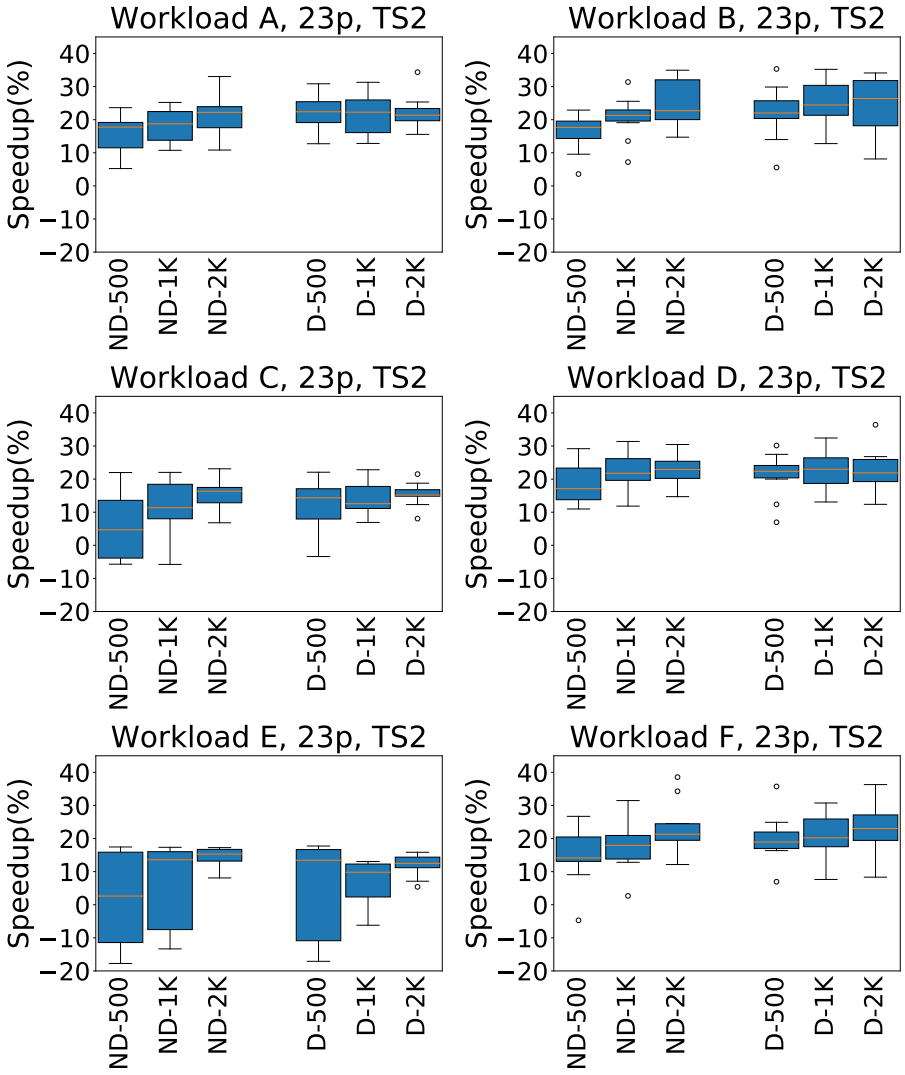

Figure 16: Speedup with respect to the default configuration obtained tuning 23 parameters for all the 6 YCSB workloads (scenario 2). The results reported represent the throughput measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the speedup obtained by, from left to right, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

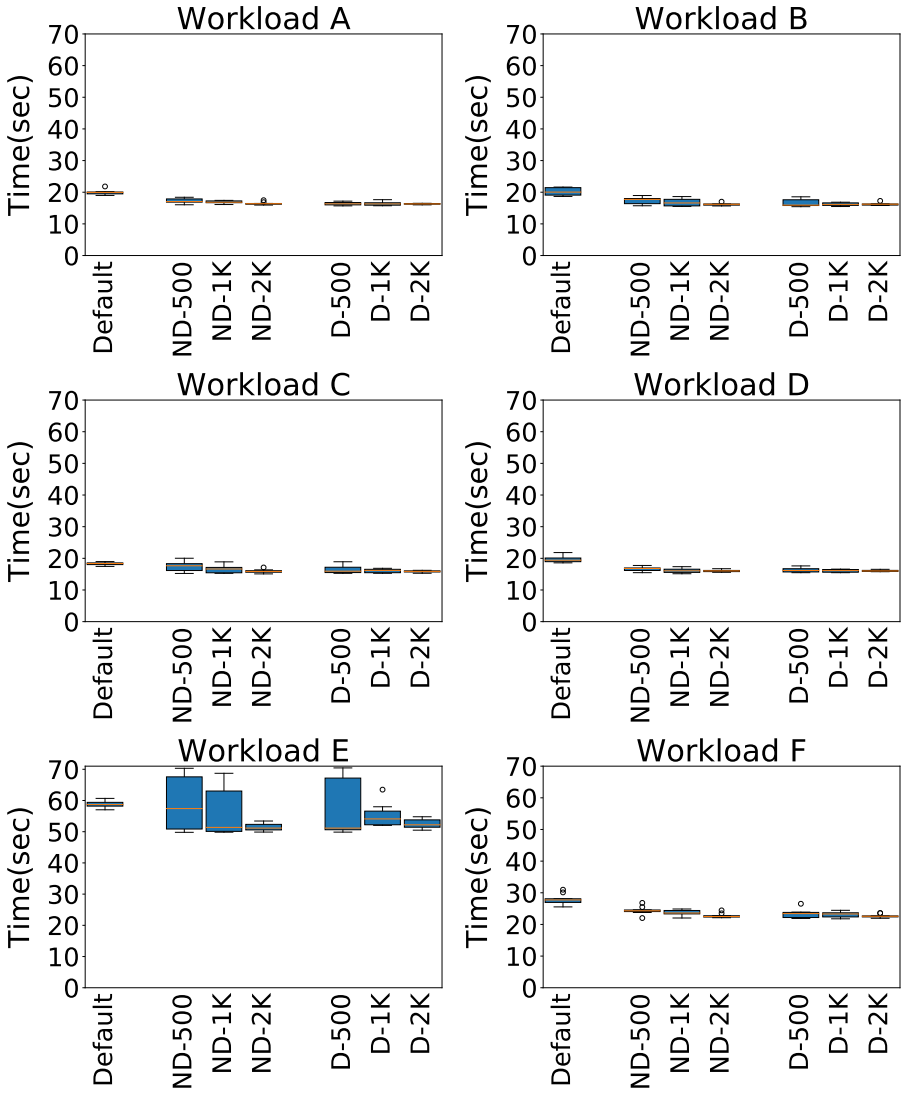

Figure 17: Time tuning 23 parameters and the 6 Workloads treated as instance in scenario 2. The results reported represent the time measured testing the final configuration over each workload, repeating each experiment ten times. The boxplots report the time obtained by, from left to right, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments).

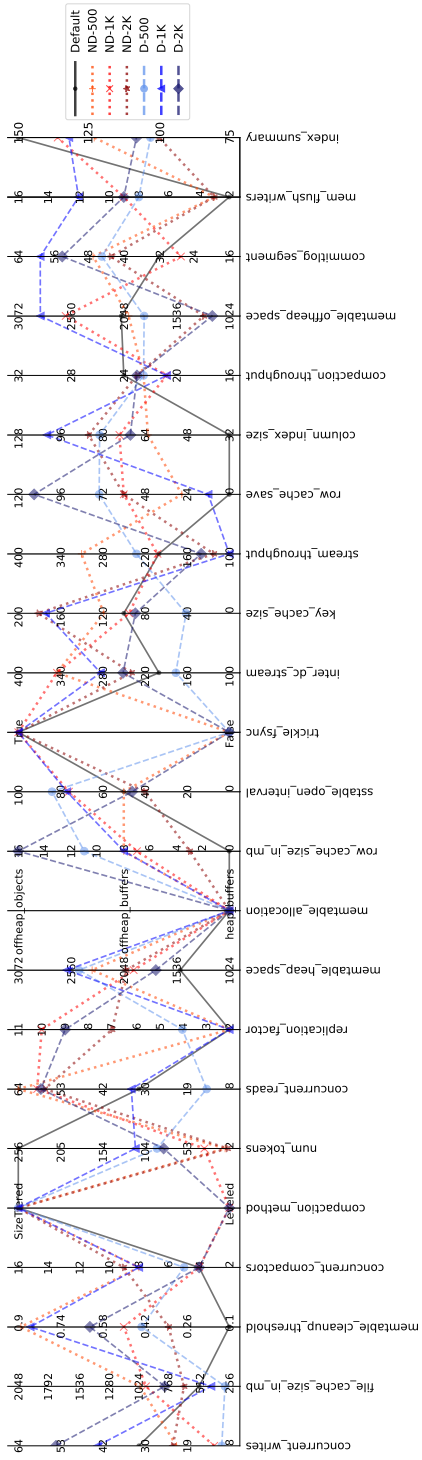

Figure 18: parallel coordinates plot WL6 and 23 parameters

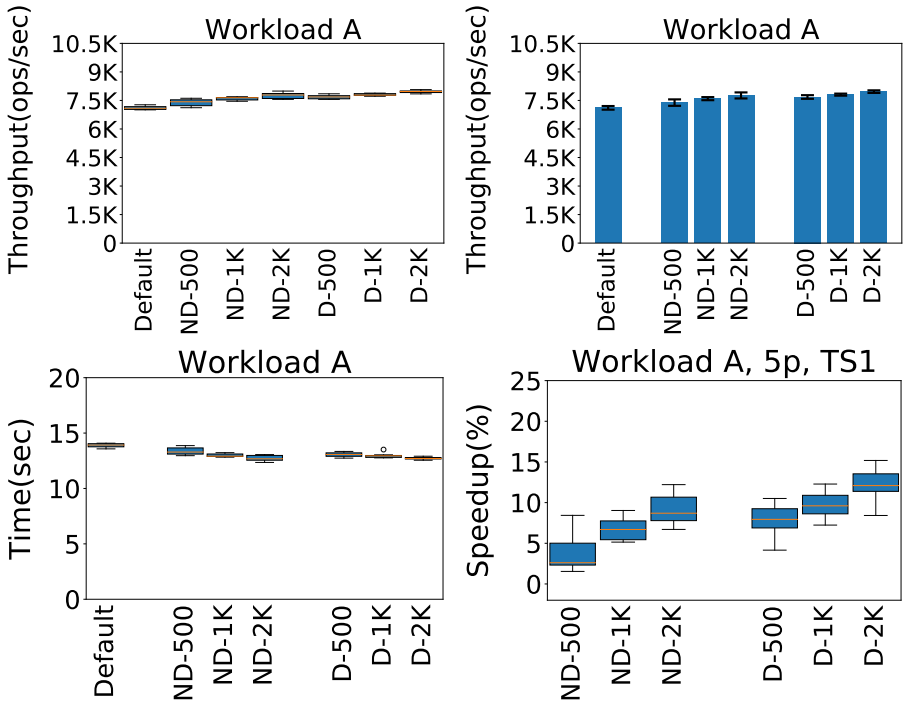

Figure 19: Results obtained when tuning the YCSB workload A in test scenario 1 tuning 5 parameters. The results reported the boxplot and the histogram of the throughput in addition to the boxplot of the speedup and the time measured the final configurations, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments). The speedup is computed with respect to the default configuration.

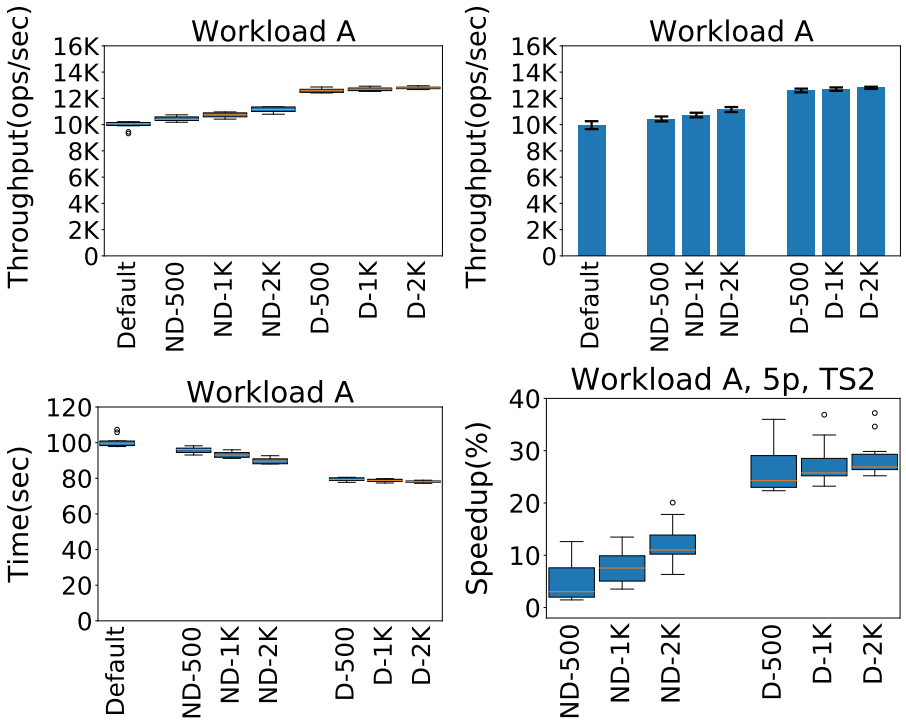

Figure 20: Results obtained when tuning the YCSB workload A in scenario 2 tuning 5 parameters. The results reported the boxplot and the histogram of the throughput in addition to the boxplot of the speedup and the time measured the final configurations, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments). The speedup is computed with respect to the default configuration.

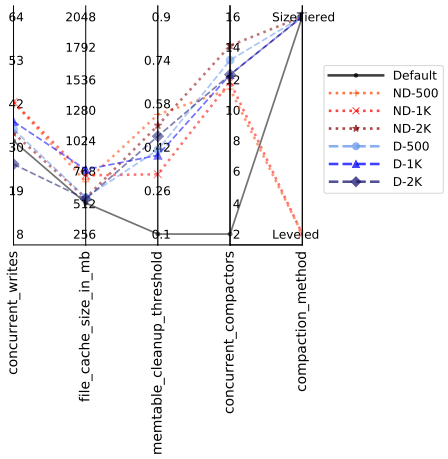

Figure 21: parallel coordinates plot WLA and 5 parameters

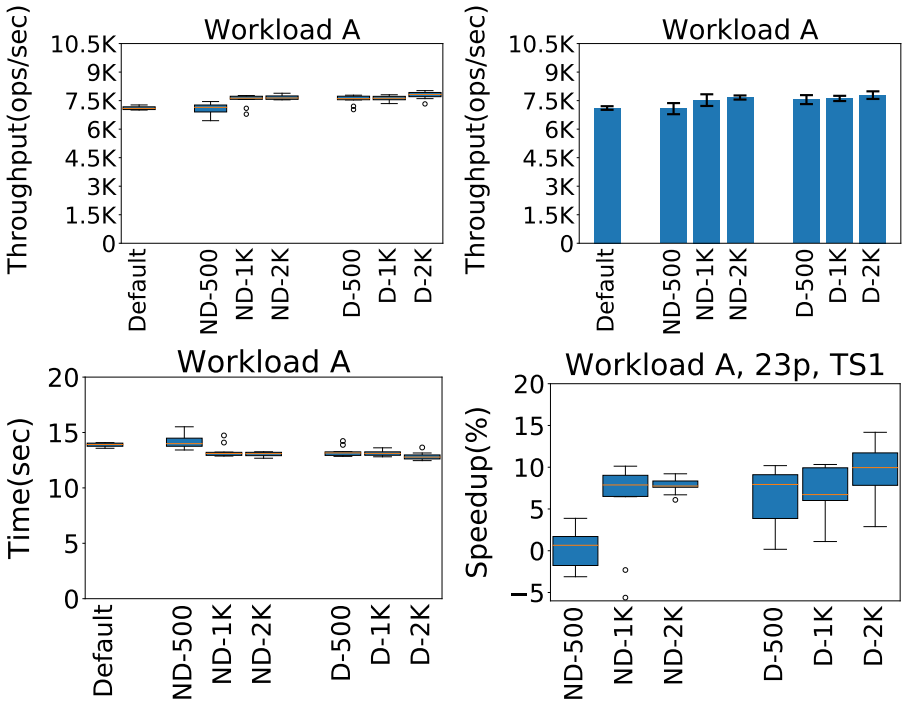

Figure 22: Results obtained when tuning the YCSB workload A with 23 parameters (scenario 1). The results reported the boxplot and the histogram of the throughput in addition to the boxplot of the speedup and the boxplot of the time measured in the final configurations in scenario 1, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments). The speedup is computed with respect to the default configuration.

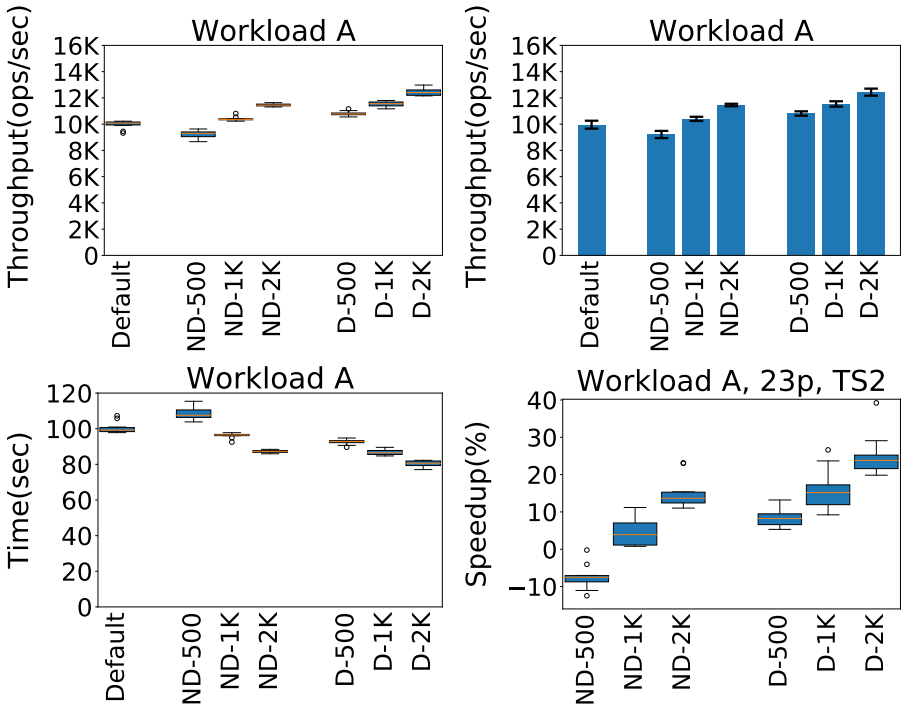

Figure 23: Results obtained when tuning the YCSB workload A with 23 parameters (scenario 2). The results reported the boxplot and the histogram of the throughput in addition to the boxplot of the speedup and the time measured the in the final configurations in scenario 2, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments). The speedup is computed with respect to the default configuration.

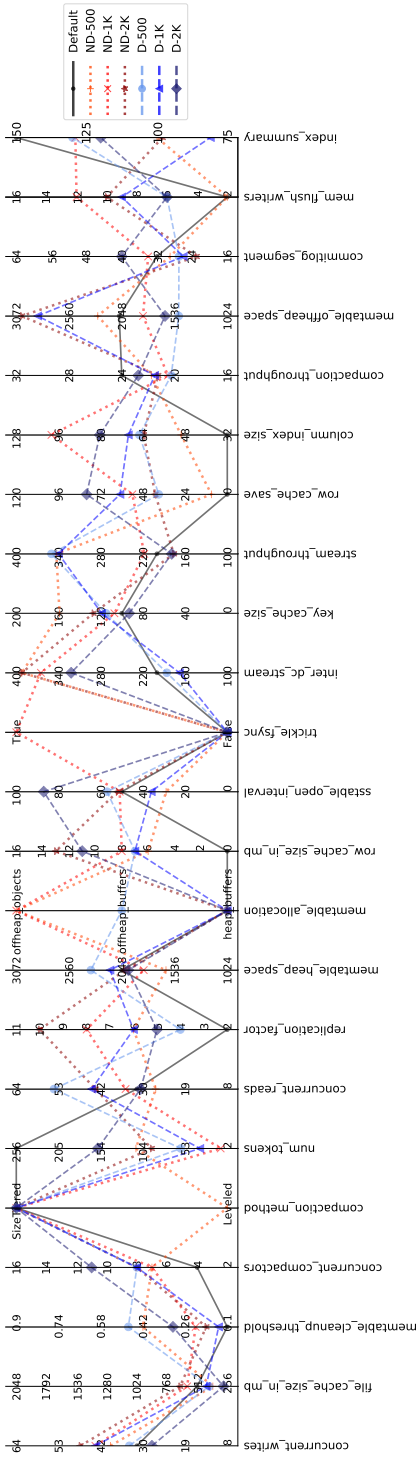

Figure 24: parallel coordinates plot WLA and 23 parameters

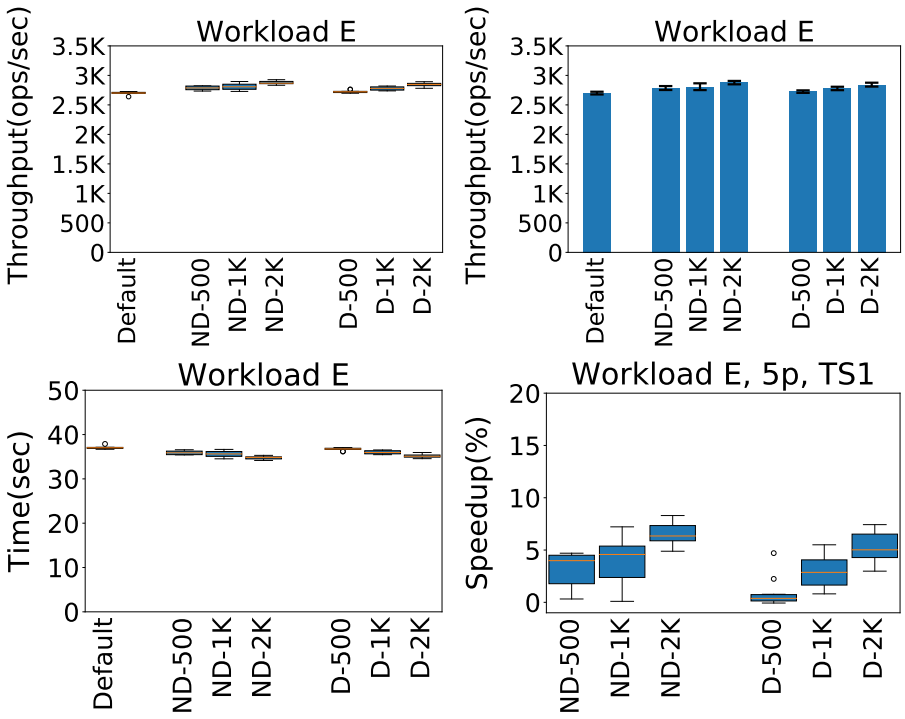

Figure 25: Results obtained when tuning the YCSB workload E with 5 parameters (scenario 1). The results reported the boxplot and the histogram of the throughput in addition to the boxplot of the speedup and the boxplot of the time measured in the final configurations in scenario 1, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments). The speedup is computed with respect to the default configuration.

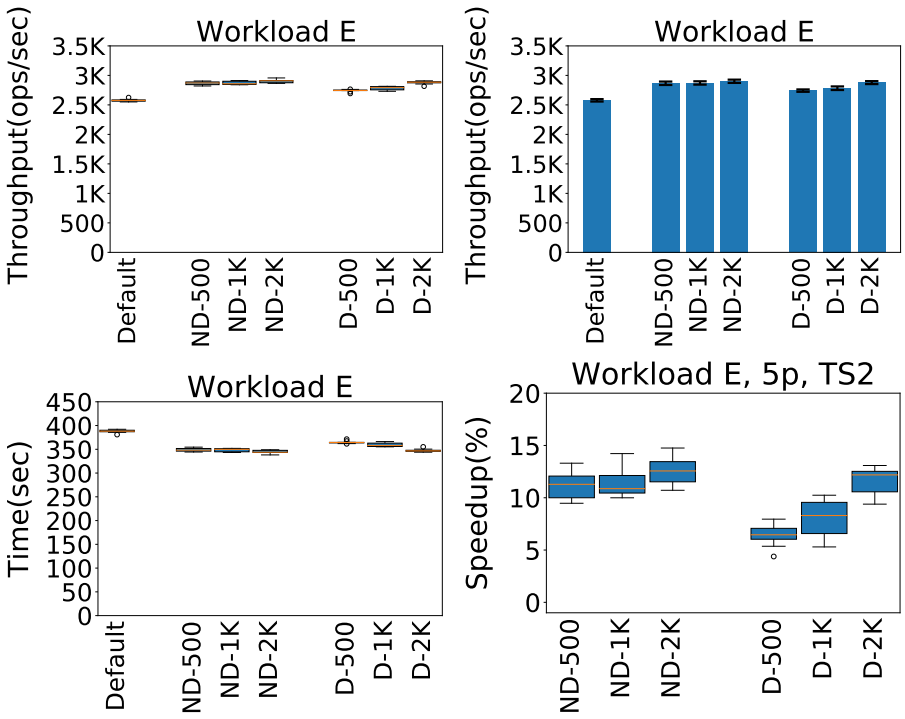

Figure 26: Results obtained when tuning the YCSB workload E with 5 parameters (scenario 2). The results reported the boxplot and the histogram of the throughput in addition to the boxplot of the speedup and the time measured the in the final configurations in scenario 2, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments). The speedup is computed with respect to the default configuration.

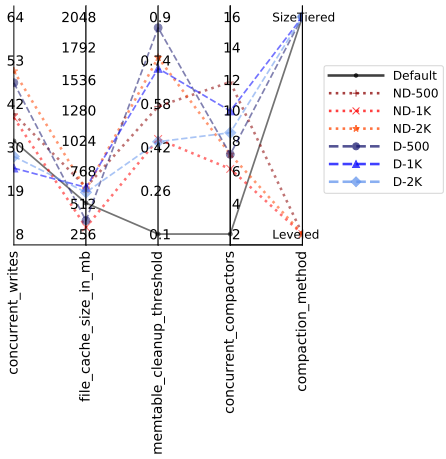

Figure 27: parallel coordinates plot WLE and 5 parameters

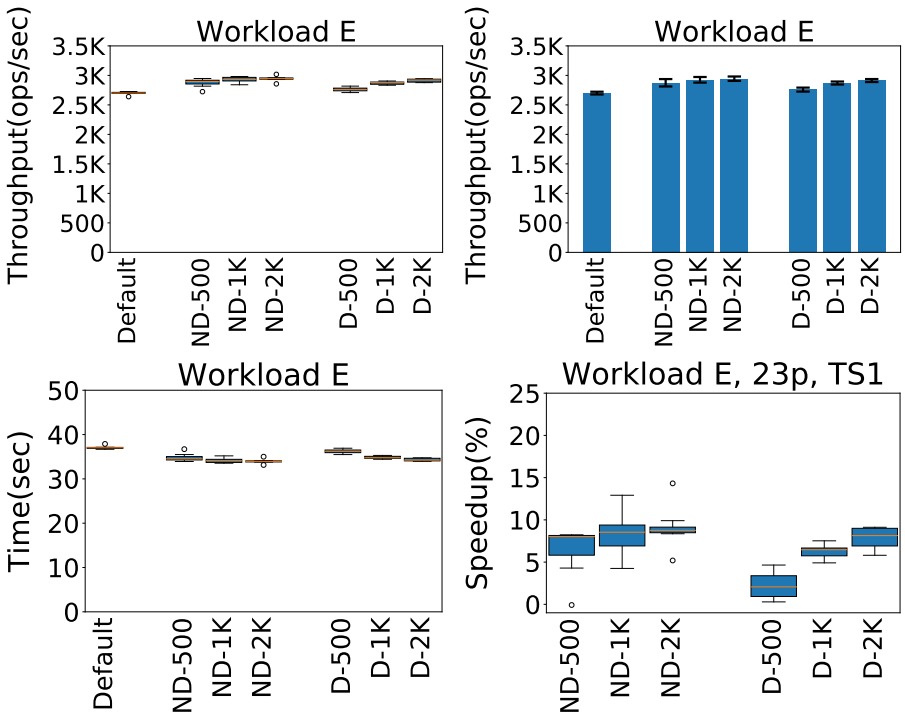

Figure 28: Results obtained when tuning the YCSB workload E with 23 parameters (scenario 1). The results reported the boxplot and the histogram of the throughput in addition to the boxplot of the speedup and the boxplot of the time measured in the final configurations in scenario 1, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments). The speedup is computed with respect to the default configuration.

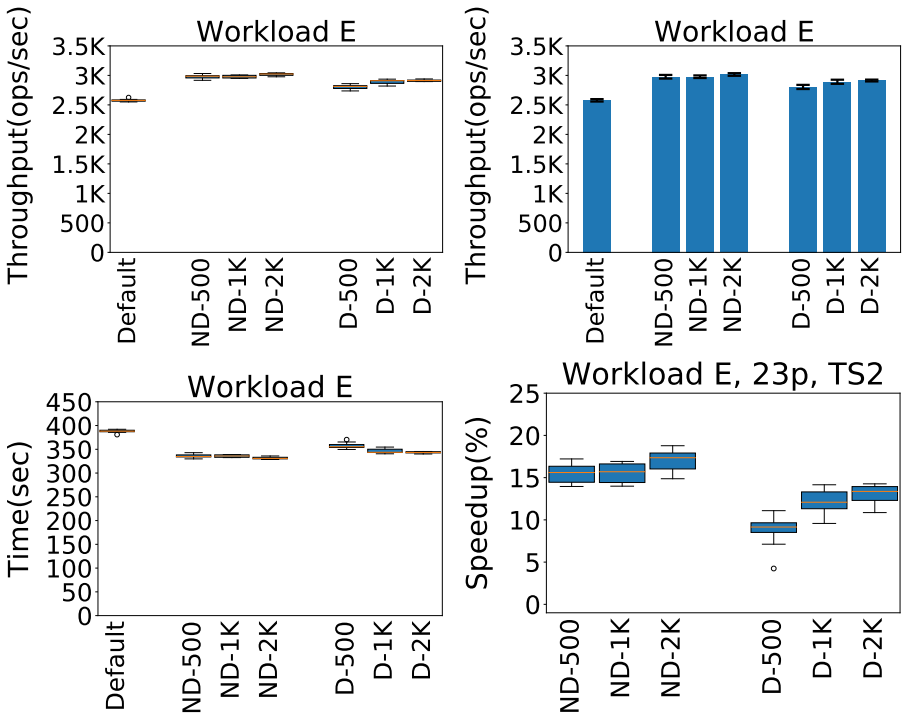

Figure 29: Results obtained when tuning the YCSB workload E with 23 parameters (scenario 2). The results reported the boxplot and the histogram of the throughput in addition to the boxplot of the speedup and the time measured the in the final configurations in scenario 2, repeating each experiment ten times. The boxplots report the results obtained by, from left to right, the default configuration, irace without the default configuration (budgets of 500, 1000 and 2000 experiments), and irace with the default configuration (budgets of 500, 1000 and 2000 experiments). The speedup is computed with respect to the default configuration.

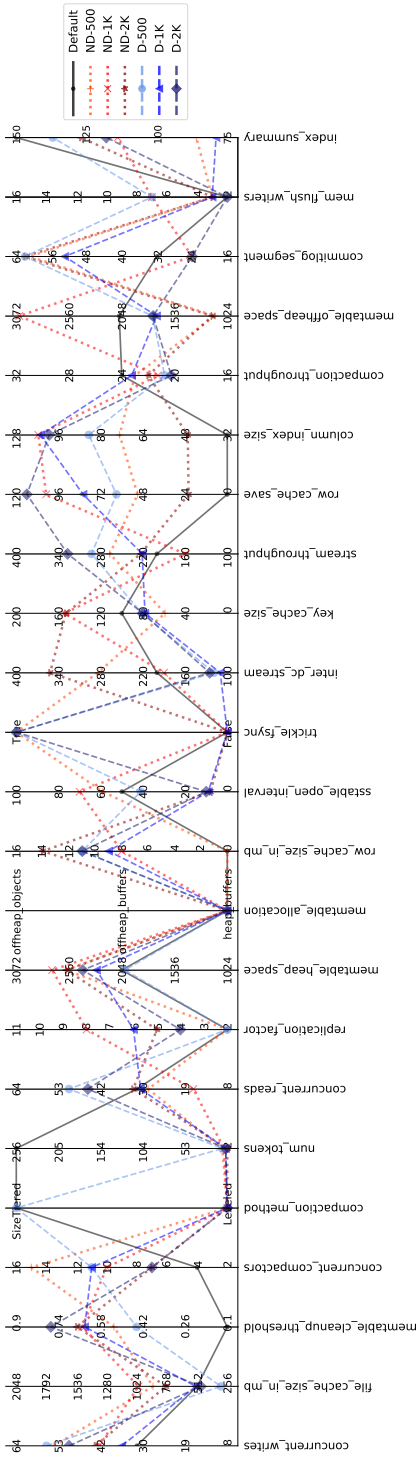

Figure 30: parallel coordinates plot WLE and 23 parameters