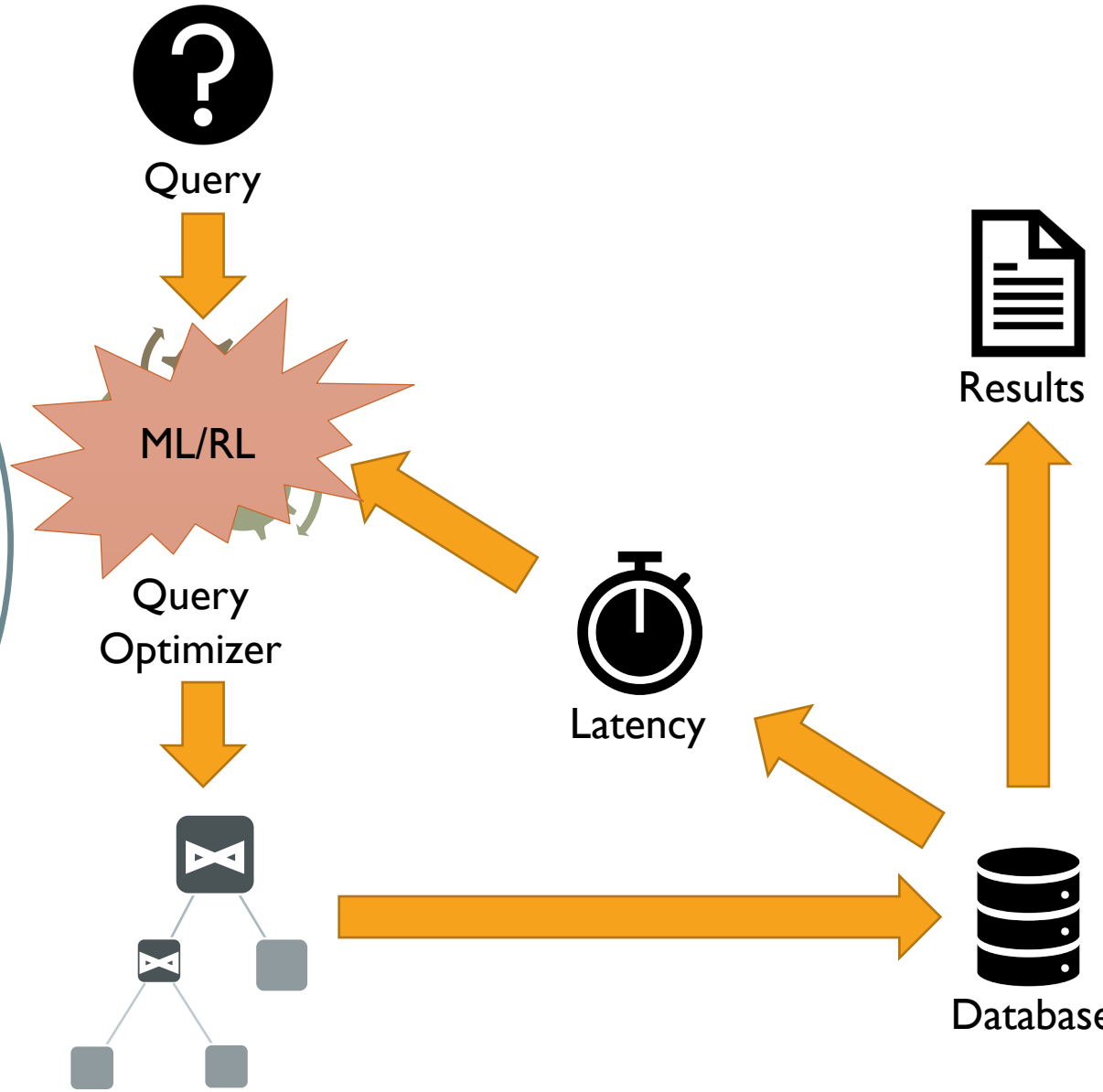
BAO: MAKING LEARNED QUERY OPTIMIZATION PRACTICAL

Authored by: Ryan Marcus, Parimarjan Negi, Hongzi Mao, Nesime Tatbul, Mohammad Alizadeh, Tim Kraska

Presented by: Mattie Nejati

March, 2022

LEARNED QUERY OPTIMIZATION



ISSUES WITH PRIOR APPROACHES AND BENEFITS OF BAO

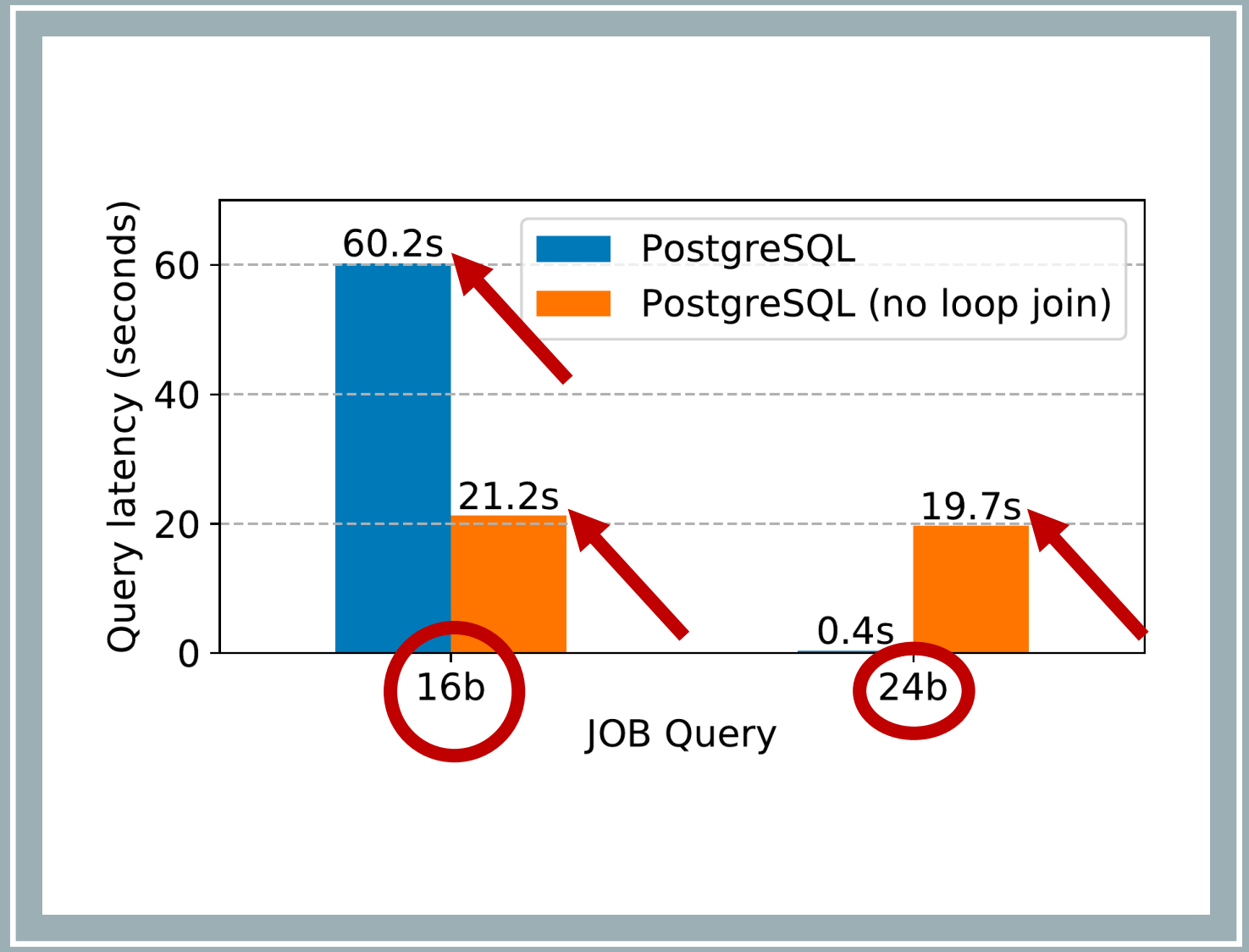
PRIOR APPROACHES

1. Long training time
2. Inability to adjust to database and workload changes
3. Tail catastrophe
4. Black-box decisions
5. Integration cost

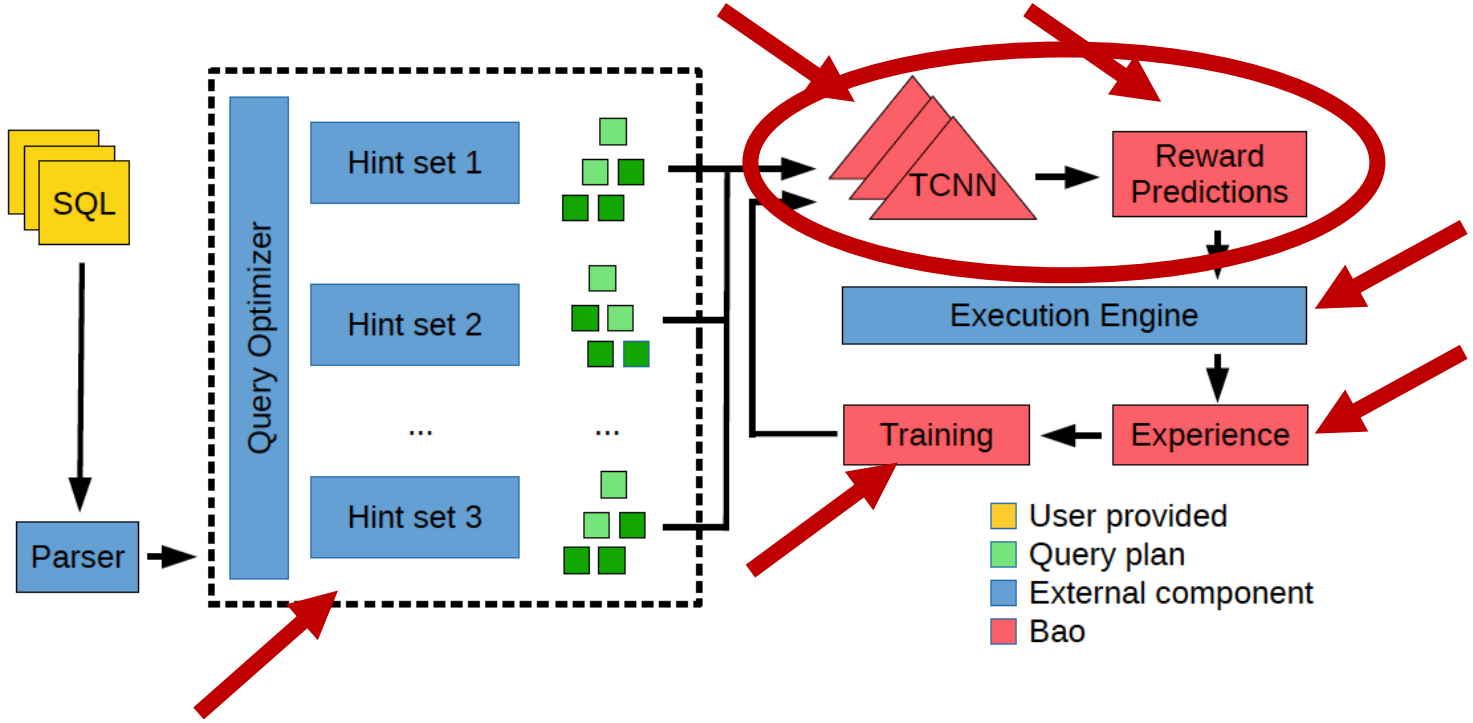
BAO

1. Short training time
2. Robustness to schema, data, and workload changes
3. Improved tail latency
4. Interpretability and easier debugging
5. Low integration cost

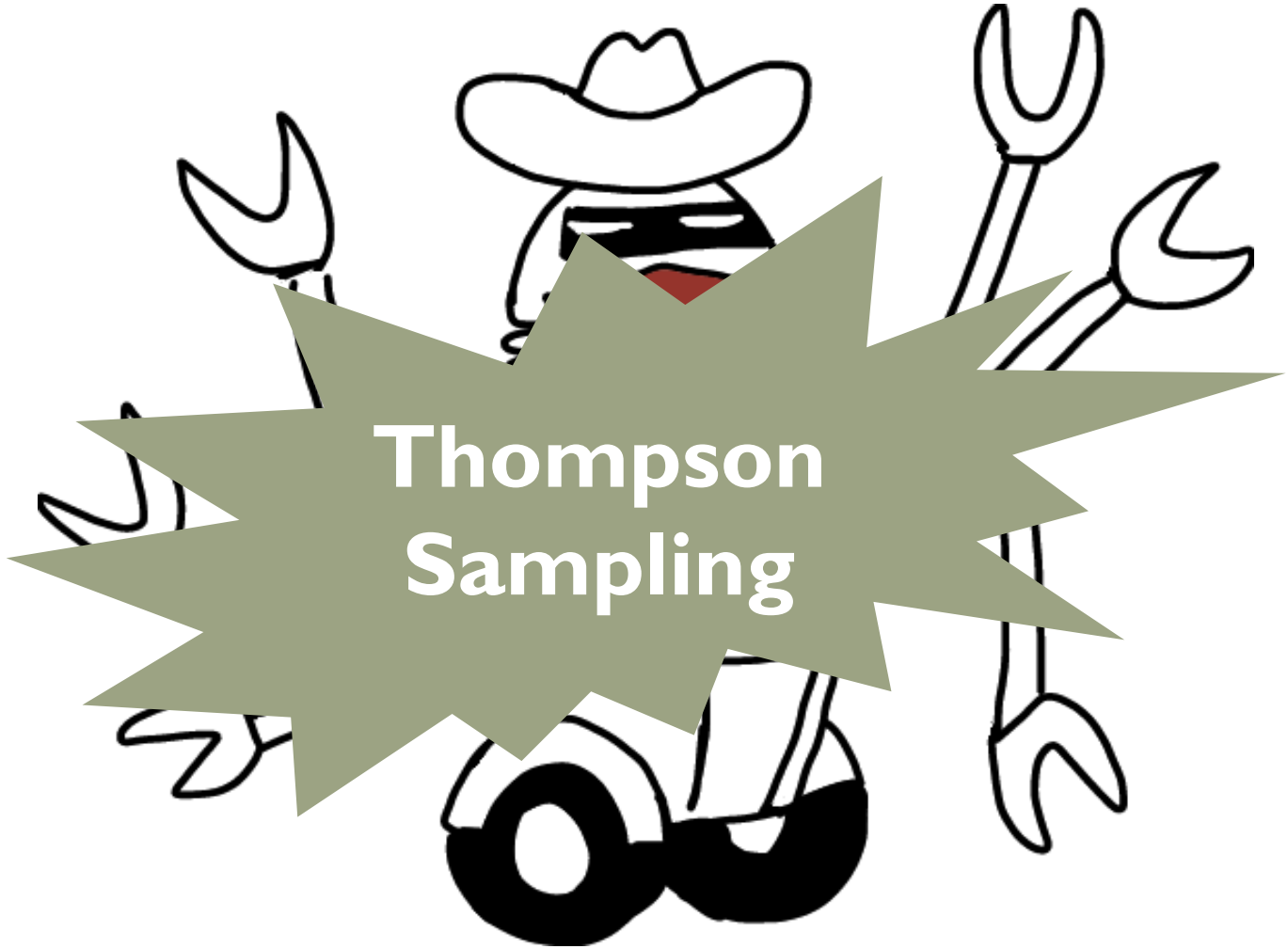
THE BIG IDEA

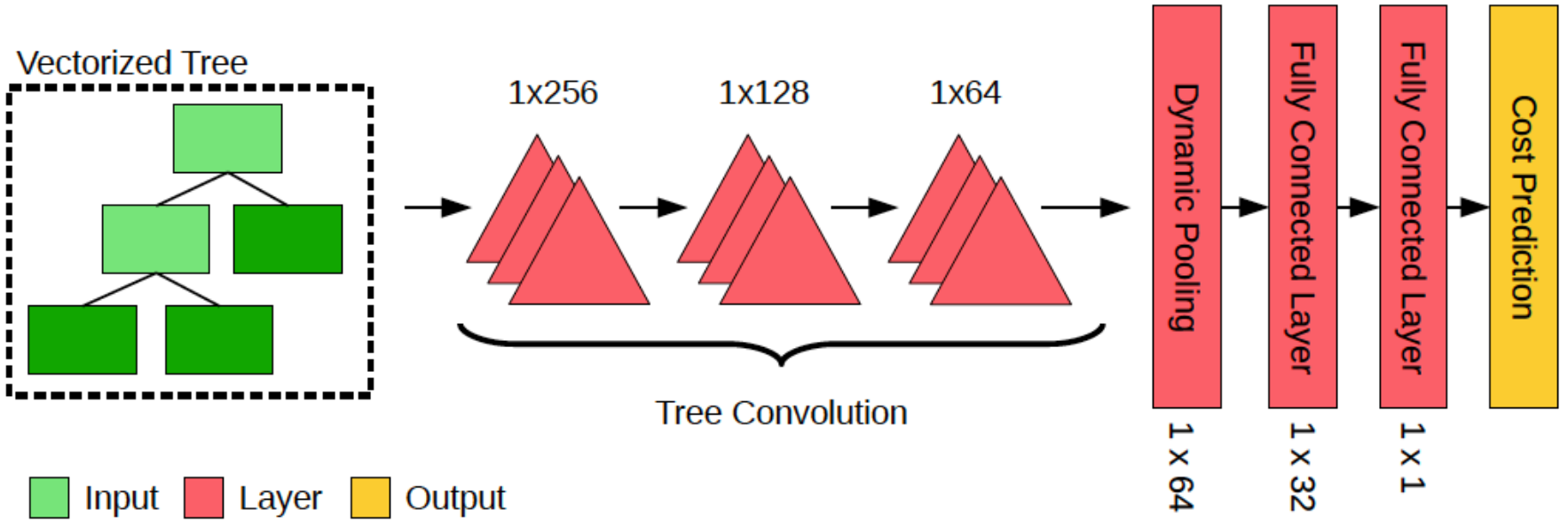

Query latency (seconds)
0
20
40
60
60.2s
21.2s
0.4s
19.7s
PostgreSQL
PostgreSQL (no loop join)
16b
24b
JOB Query

BAO, THE BANDIT OPTIMIZER

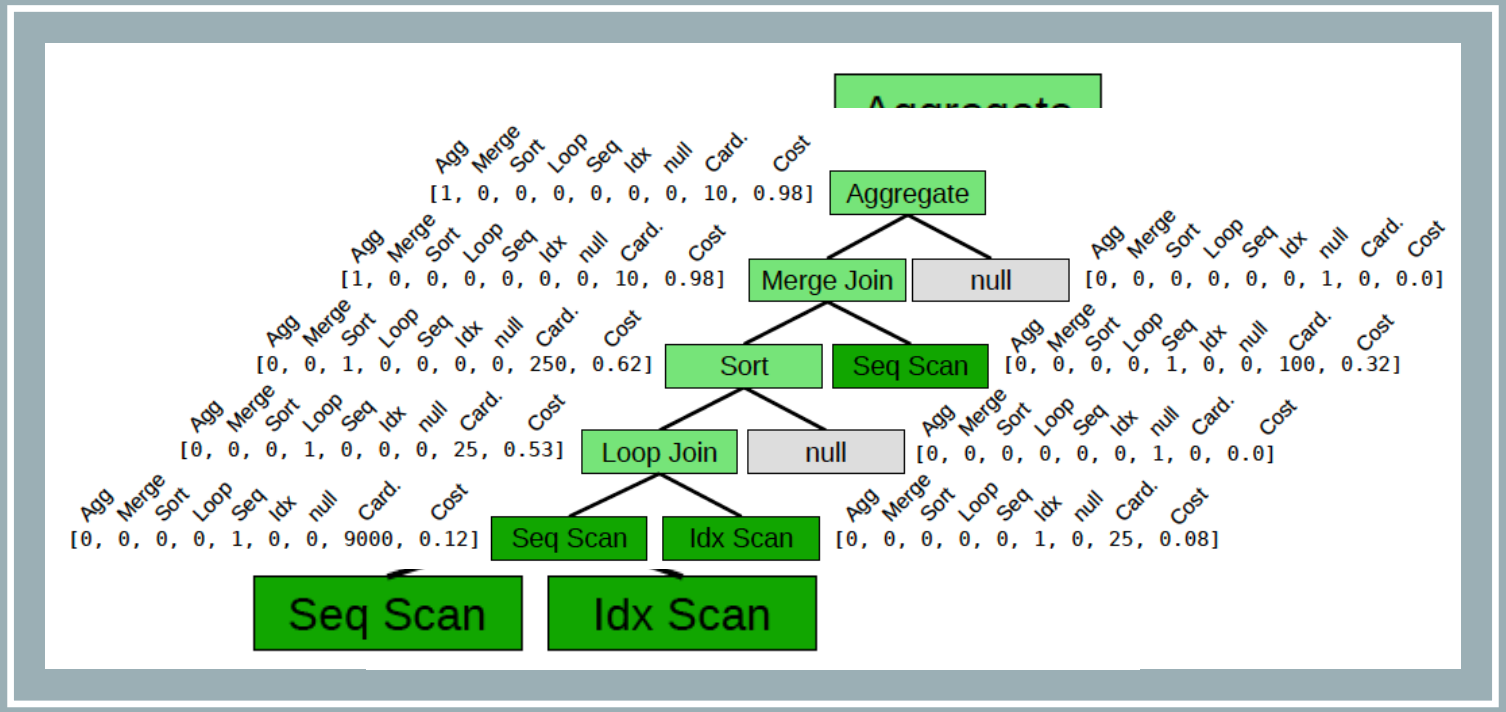


SELECTING QUERY PLAN

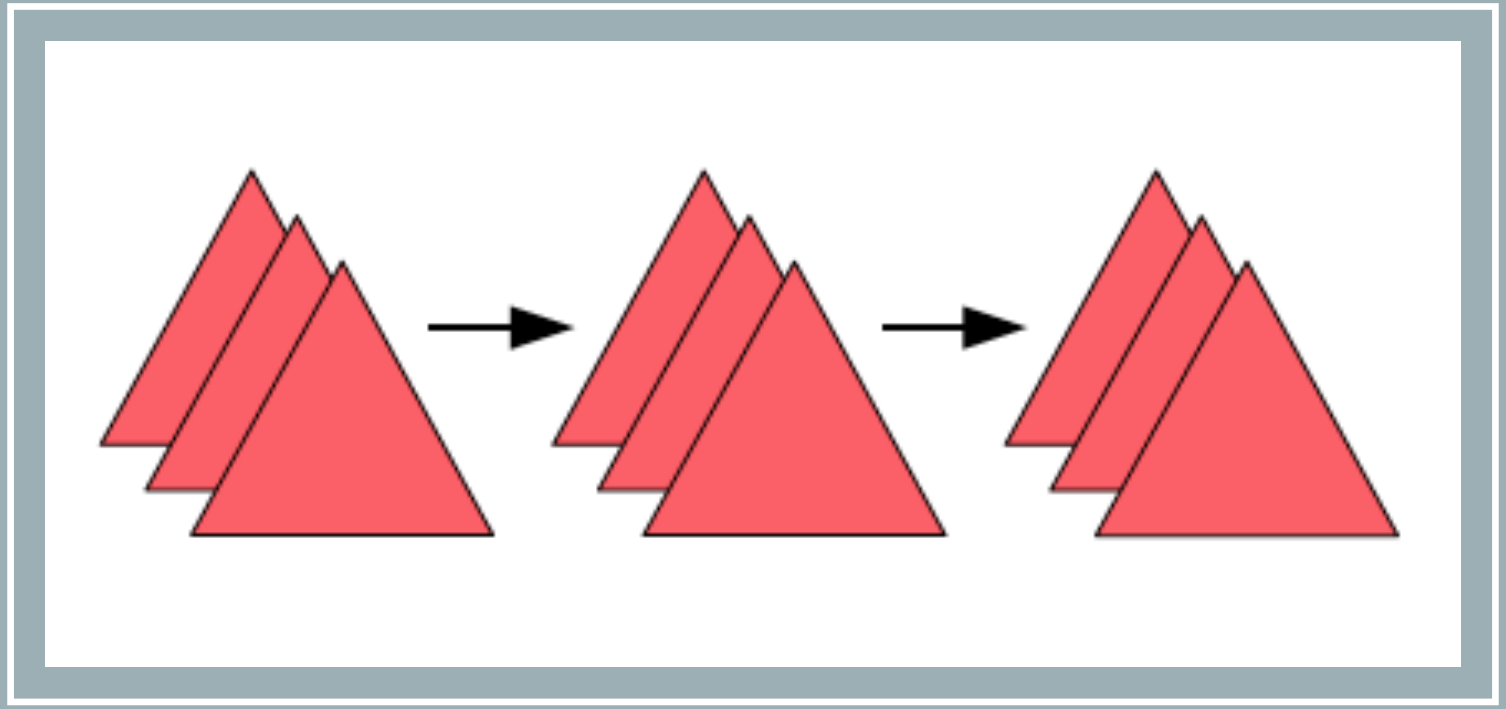




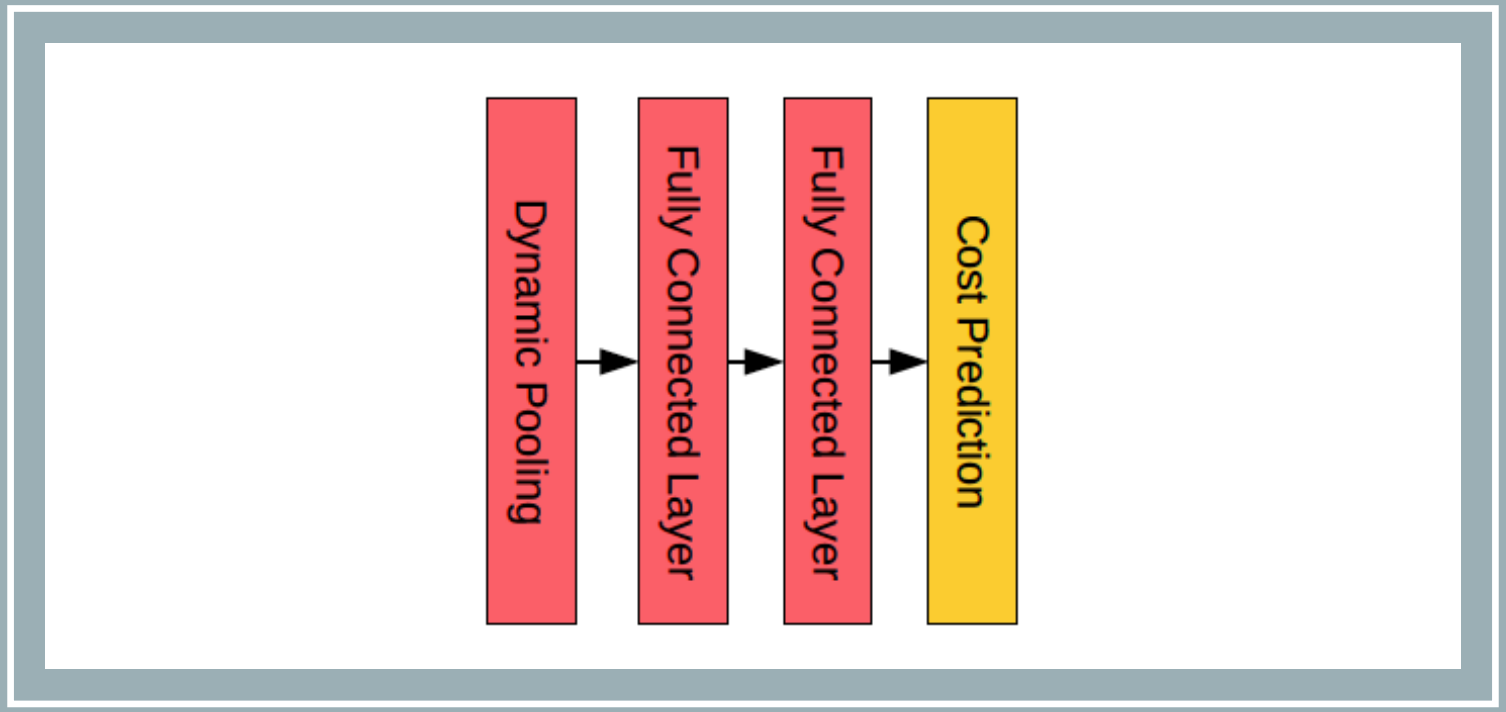
THE PREDICTIVE MODEL



THE PREDICTIVE MODEL (VECTORIZING QUERY PLAN TREES)

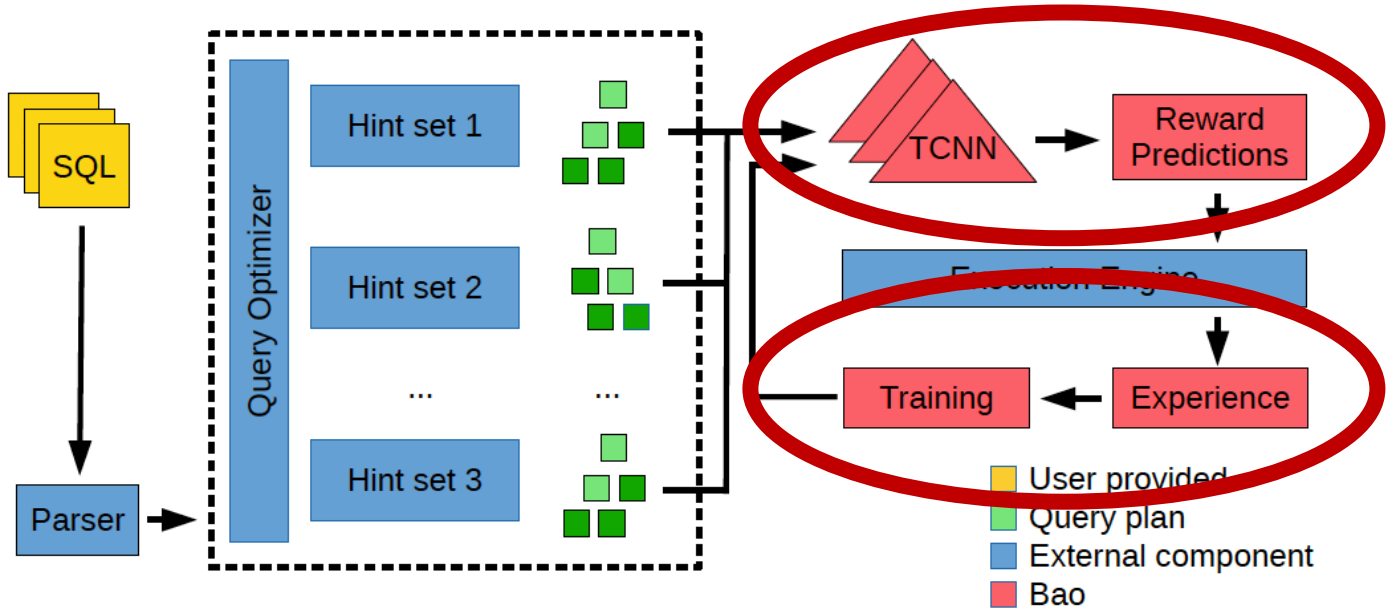


THE PREDICTIVE MODEL (TREE CONVOLUTIONAL NEURAL NETWORKS)



THE PREDICTIVE MODEL (PREDICTING THE PERFORMANCE)

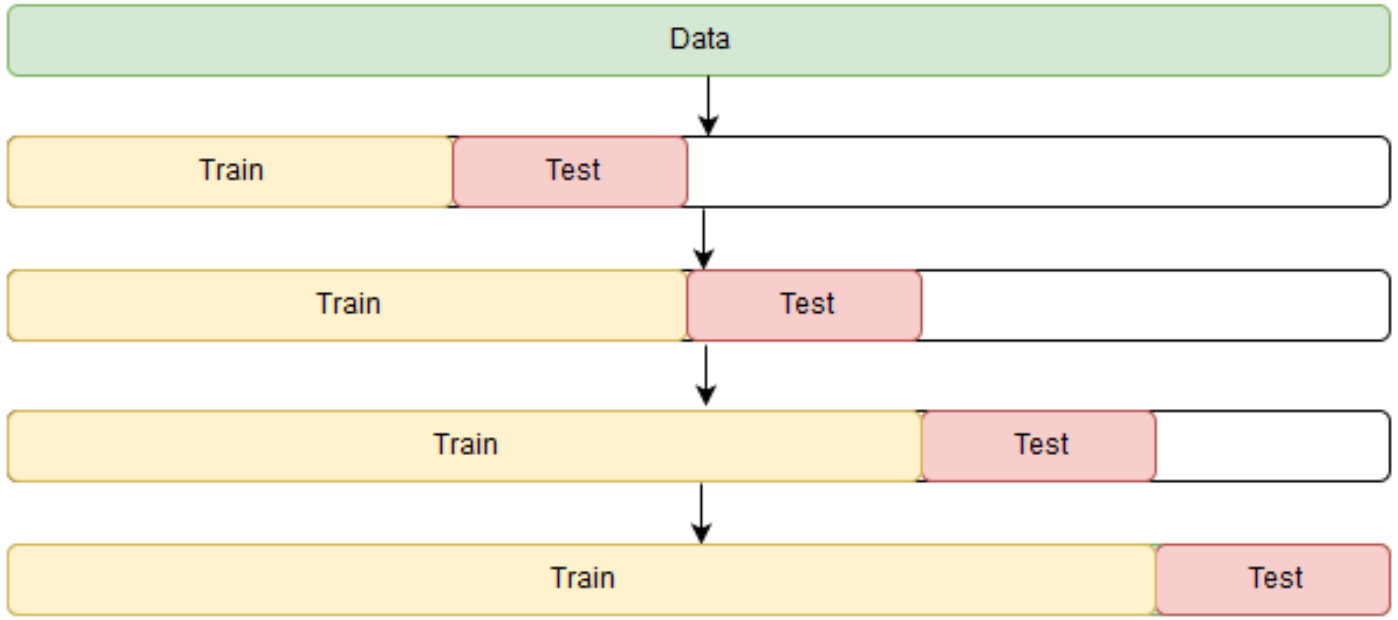
BAO, THE BANDIT OPTIMIZER



EXPERIMENTS

	Size	Queries	WL	Data	Schema
IMDb	7.2 GB	5000	Dynamic	Static	Static
Stack	100 GB	5000	Dynamic	Dynamic	Static
Corp	1 TB	2000	Dynamic	Static[a]	Dynamic

DATA

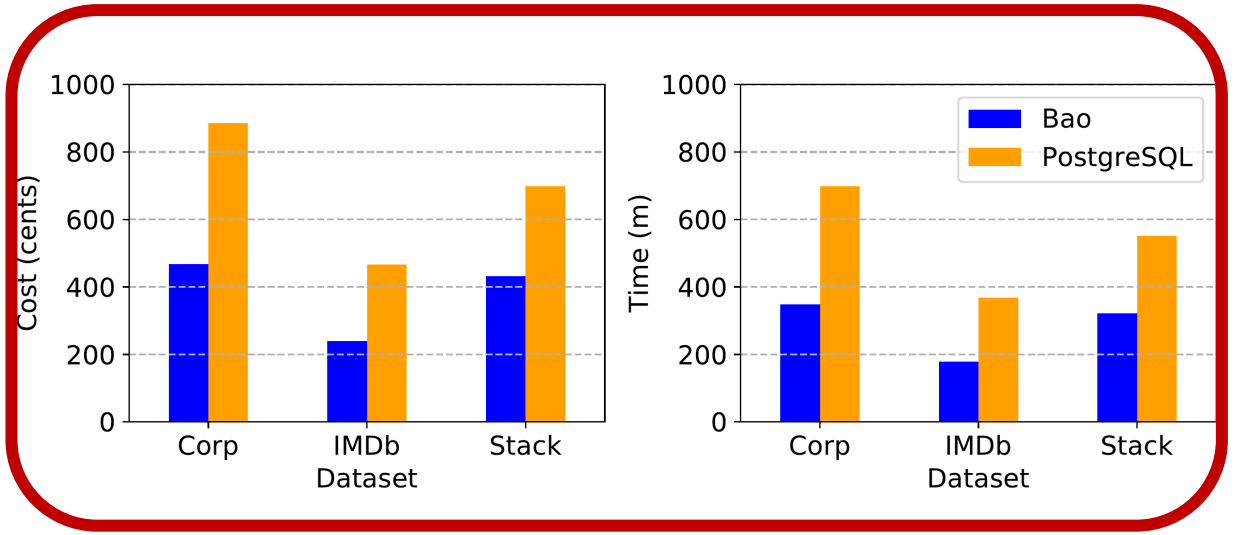


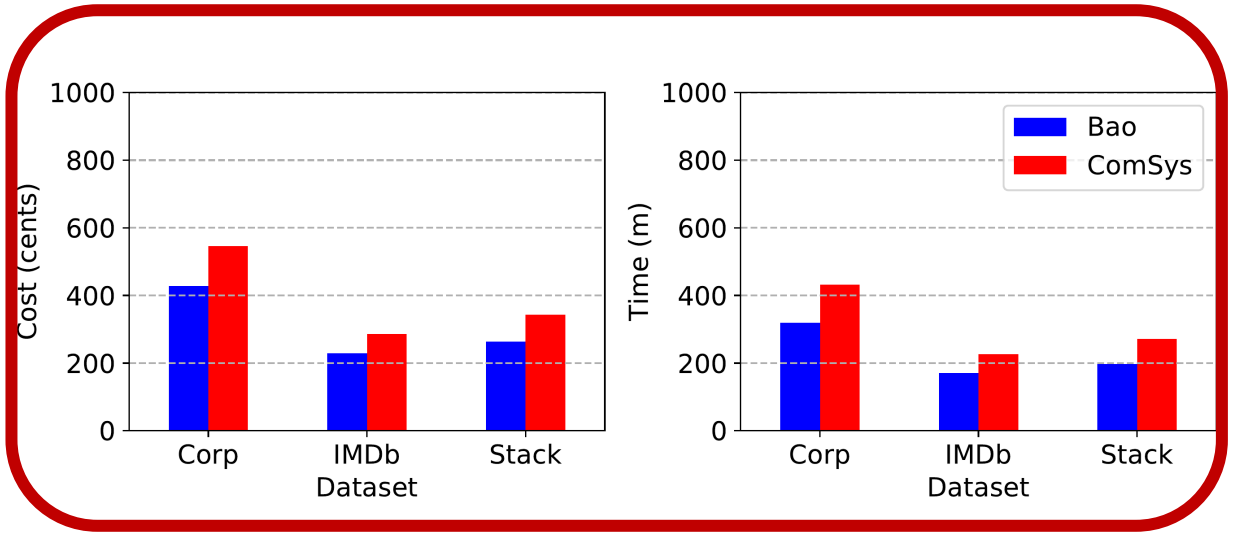
METHODS

IS BAO PRACTICAL?

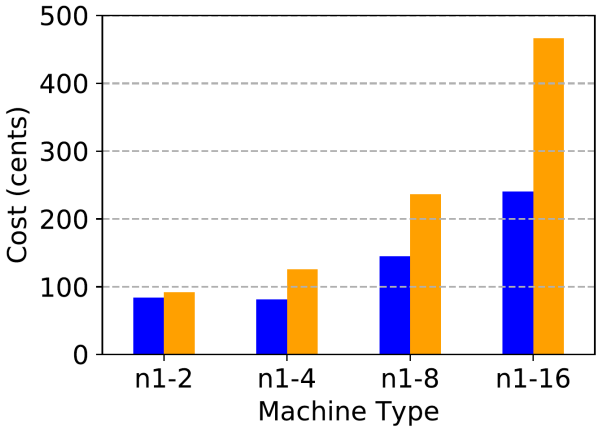
1. Cost and performance on the cloud
2. Cost and performance on varied hardware
3. Tail latency analysis
4. Training time and convergence
5. Query regression analysis
6. Query optimization analysis
7. Prior learned optimizers

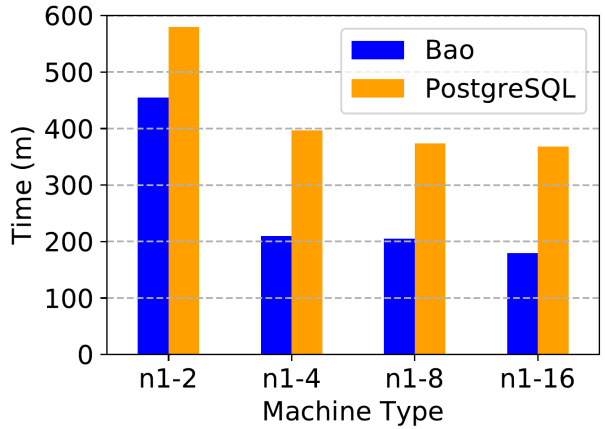
1. COST AND PERFORMANCE ON THE CLOUD

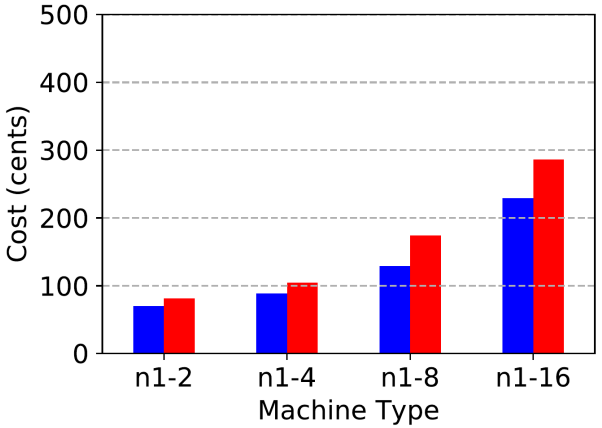


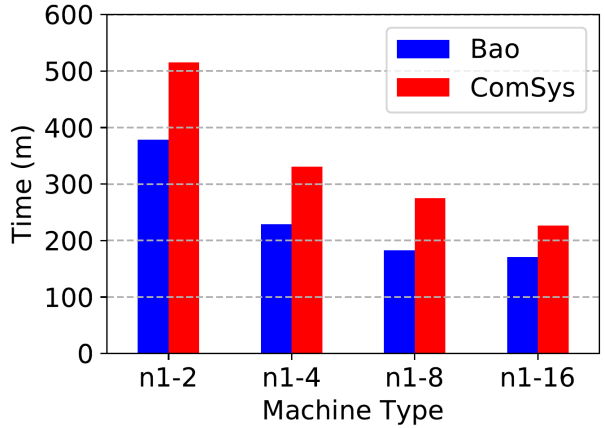


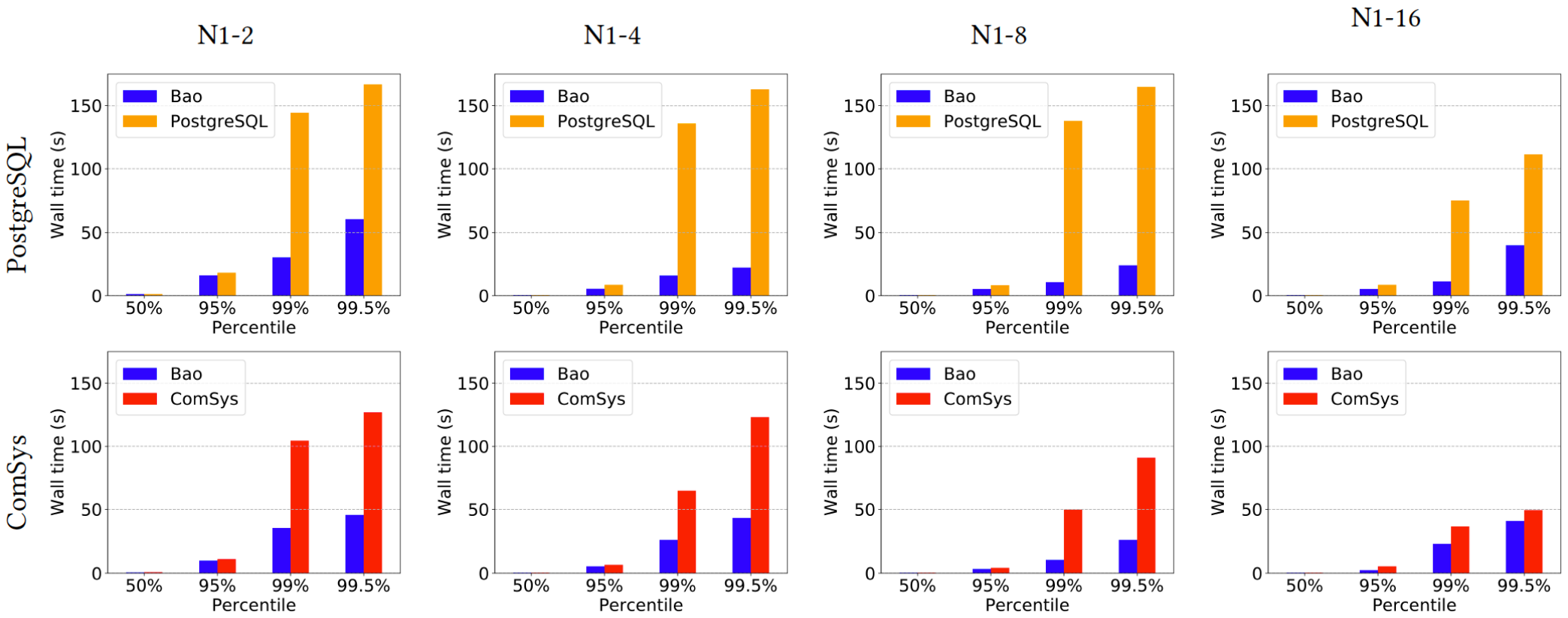
2. COST AND PERFORMANCE ON VARIED HARDWARE



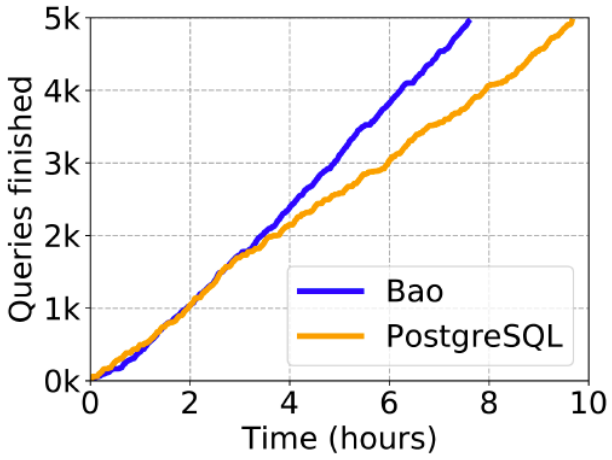




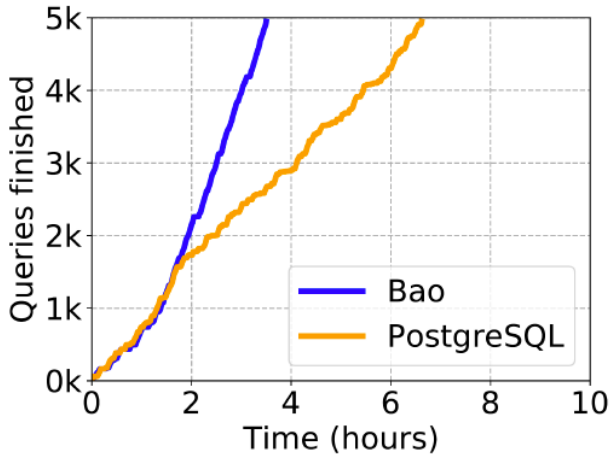




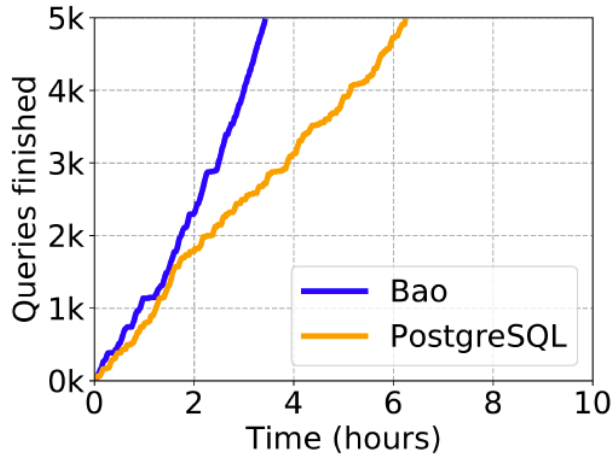
3. TAIL LATENCY ANALYSIS



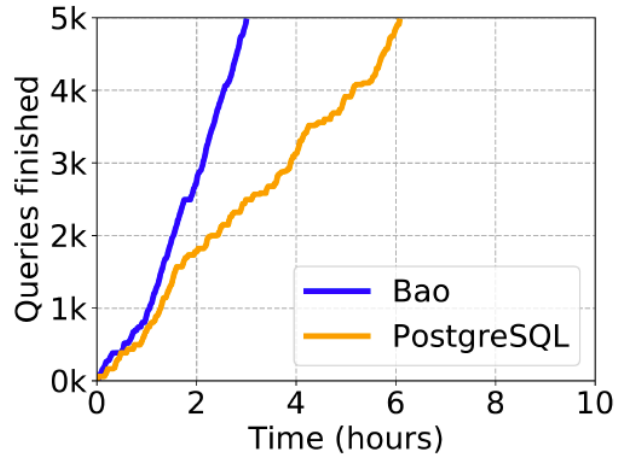
(a) VM type N1-2



(b) VM type N1-4

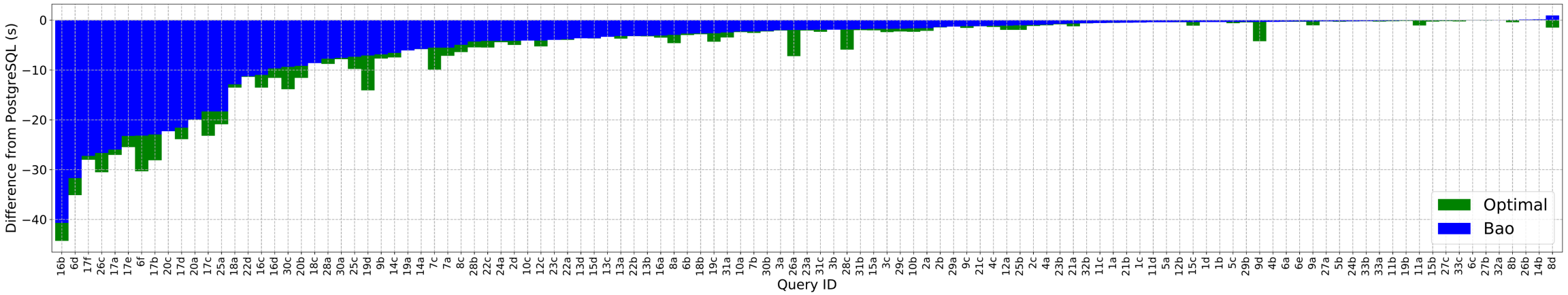


(c) VM type N1-8



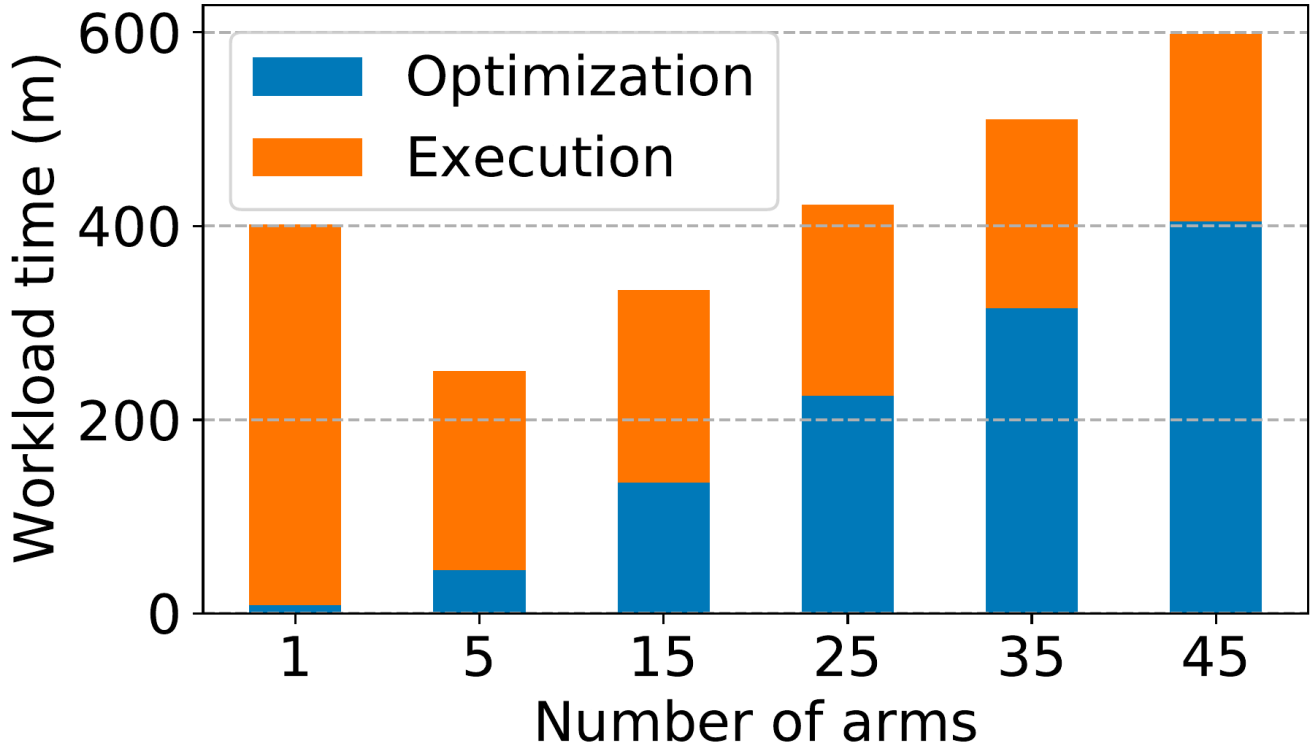
(d) VM type N1-16

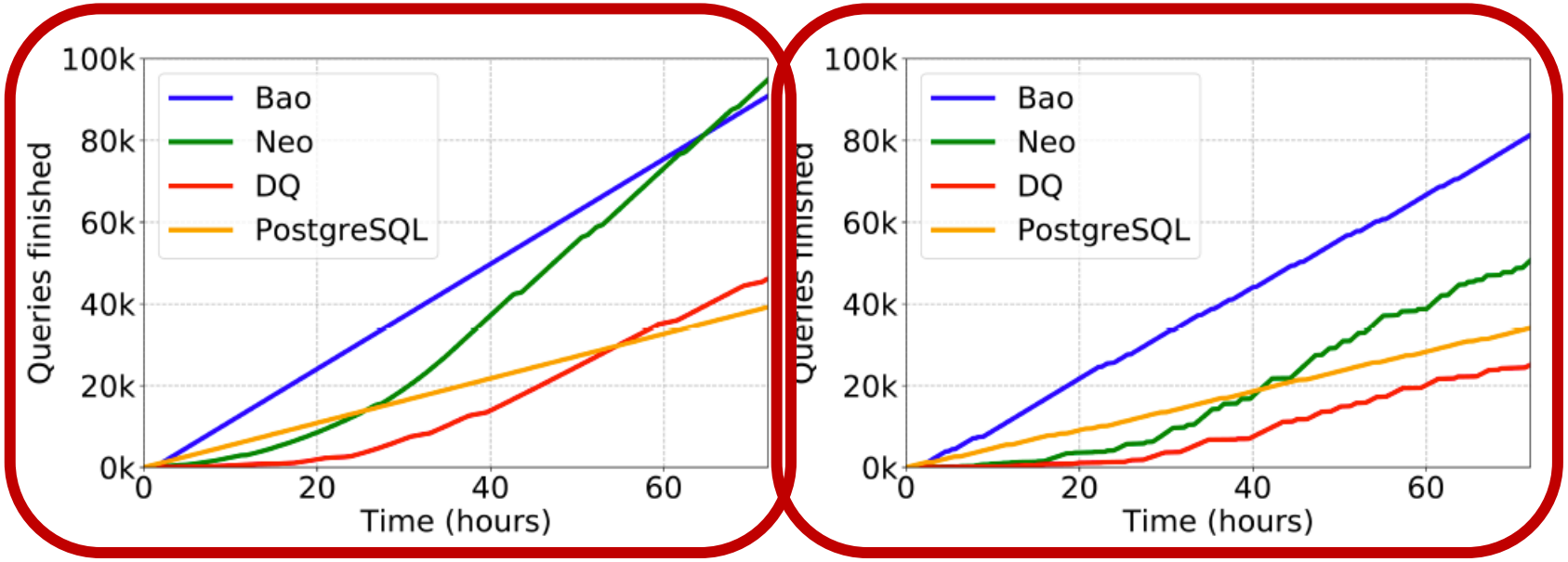
4. TRAINING TIME AND CONVERGENCE



5. QUERY REGRESSION ANALYSIS

6. QUERY OPTIMIZATION TIME





(a) Stable query workload **(b) Dynamic query workloads**

7. PRIOR LEARNED OPTIMIZERS

01

Do we need Neural Networks?

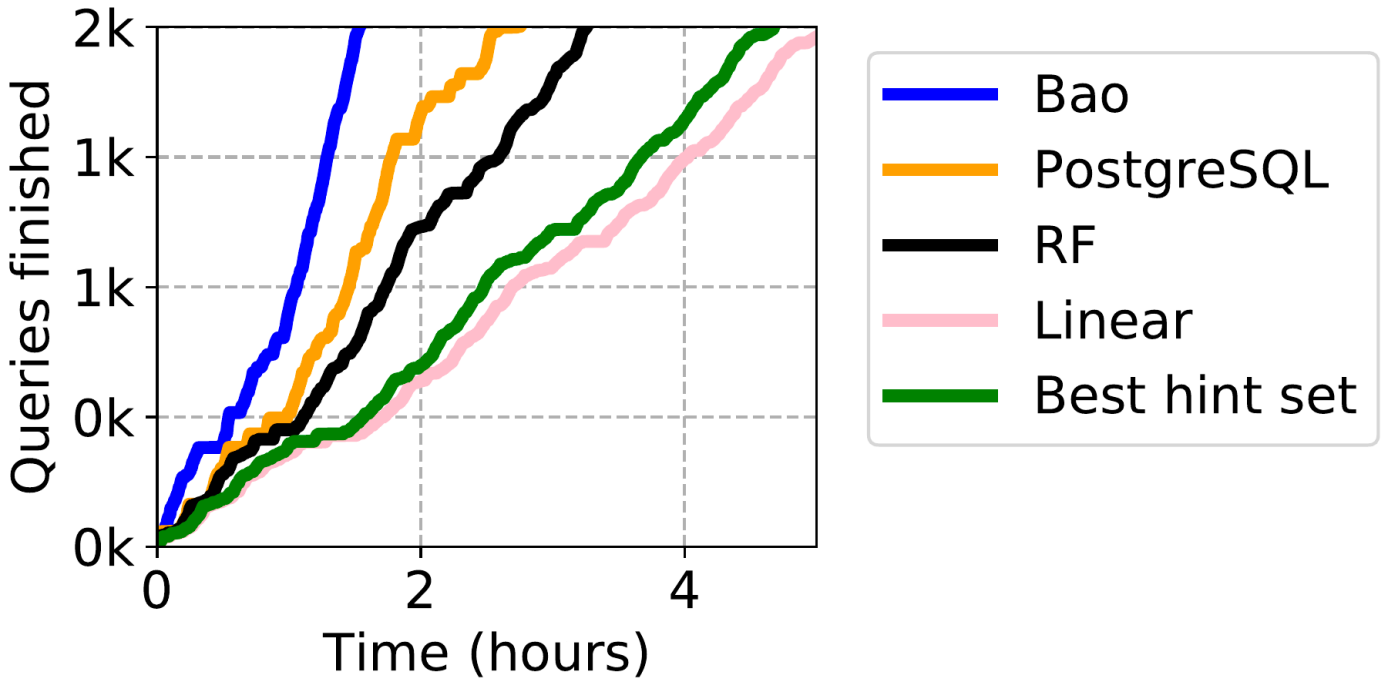
02

How accurate is BaO's predictive model?

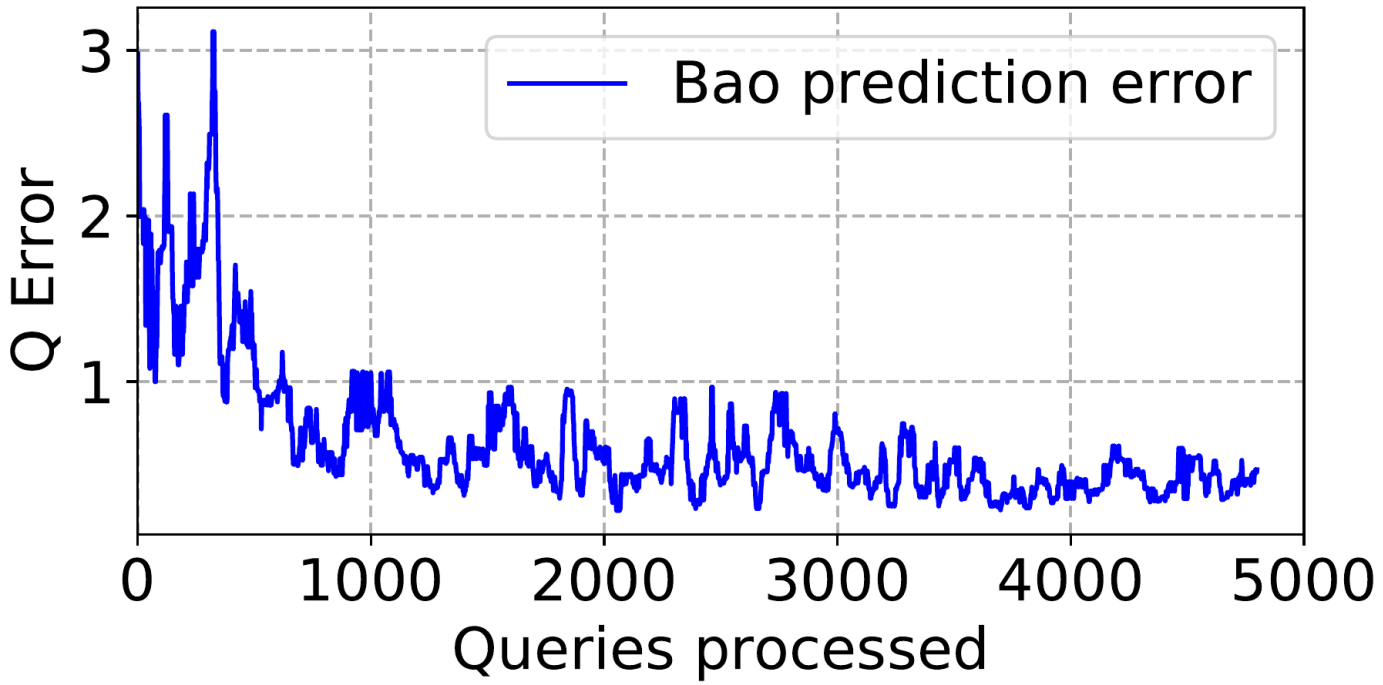
03

How long does training on GPU take?

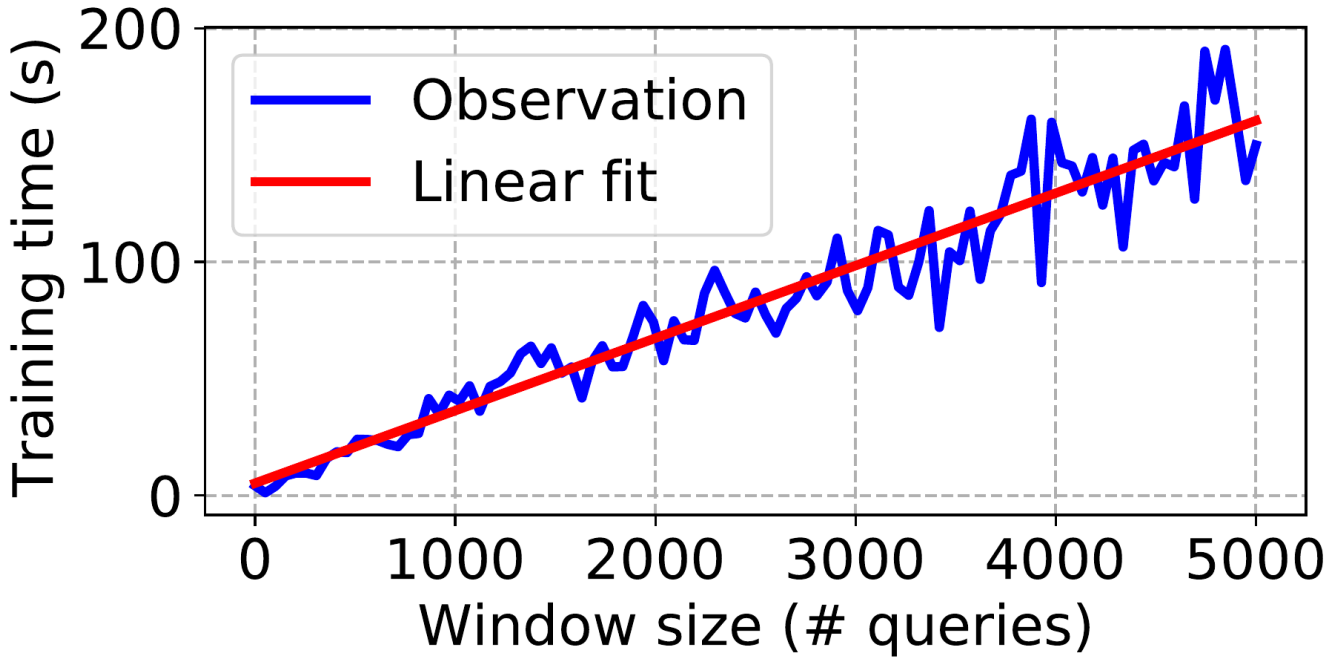
BAO'S MACHINE LEARNING MODEL



1. DO WE NEED NEURAL NETWORKS?

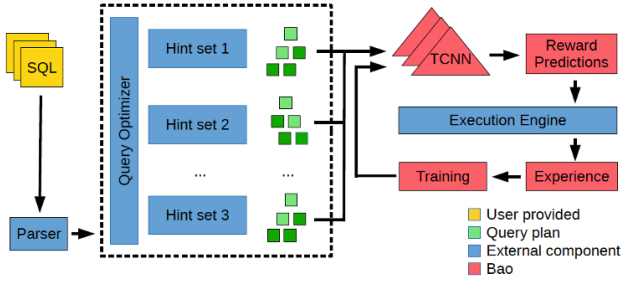


2. HOW ACCURATE IS BAO'S PREDICTIVE MODEL?

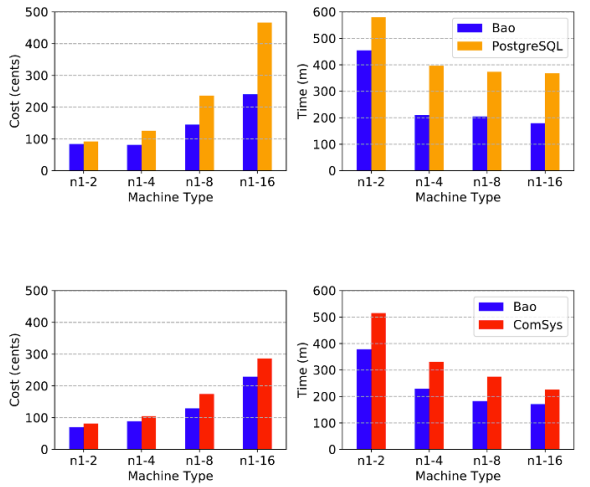


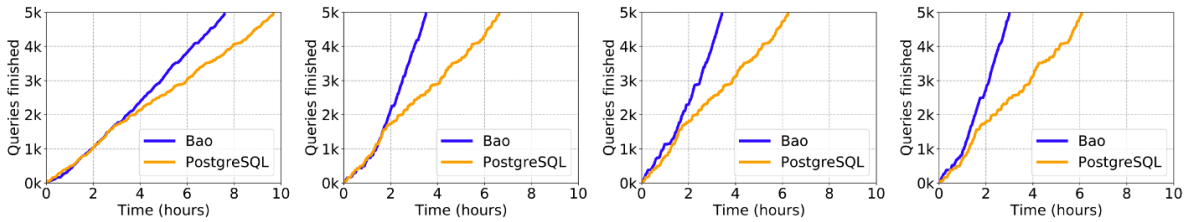
3. HOW LONG DOES TRAINING ON GPU TAKES?

BAO, THE BANDIT OPTIMIZER



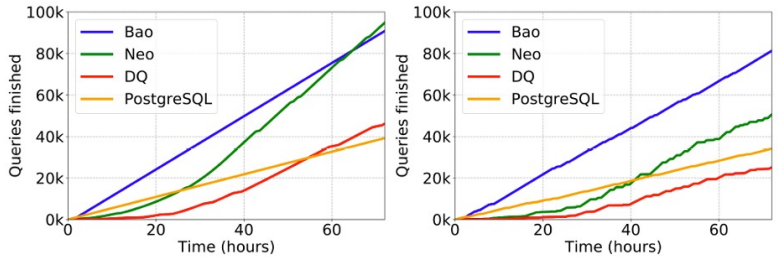
2. COST AND PERFORMANCE ON VARIED HARDWARE





(a) VM type N1-2 (b) VM type N1-4 (c) VM type N1-8 (d) VM type N1-16

4. TRAINING TIME AND CONVERGENCE



(a) Stable query workload (b) Dynamic query workloads

7. PRIOR LEARNED OPTIMIZERS