QUERY AWARE DATABASE TUNING SYSTEM WITH DEEP REINFORCEMENT LEARNING

DZ1933026 王国畅

KNOB TUNING

- Knobs are parameters of database configuration
- Knob tuning aims to achieve optimal configuration for performance
  - High throughput and low latency
  - Hybrid transaction/analytical processing(HTAP)

BACKGROUND

- **Knob tuning is NP-hard, database usually have hundreds of knobs:**
  - **MySQL: 215**
  - **PostgreSQL: 247**
  - **MongoDB: 132**
- **Existing methods have limitations**

PRIOR WORKS

- **BestConfig**
- **OtterTune**
- **CDBTune**

BESTCONFIG

- **Basic Idea: Use a heuristic method to search for the optimal configuration from the history**
- **But: may NOT find good knob values if there is no similar configuration in the history**
- **Can’t guarantee a systematically optimal configuration**

OTTERTUNE

- **Basic Idea: Find optimal configuration by learning DBA's experience from historical data through Machine Learning**
- **But: A large number of high-quality training data(DBA experience data), which is hard to obtain, is necessary for a well-performing model**
- **Hard to realize**

CDBTUNE

- **Basic Idea: Use deep reinforcement learning (DRL) to tune the database by using a try-and-error strategy**
- **But:**
  - **Time-Consuming: Run a SQL query workload multiple times for evaluation**
  - **Coarse-grained only: e.g. read-only/write-only/read-write workload**
  - **Pretrained-model: Based on reconfiguring actions**

STATE OF ART

- **Three granularities: query-level/workload-level/cluster-level**
- **Easy to realize: evaluated under PostgreSQL/MySQL/MongoDB**
- **High performance: outperform state-of-art tuning methods before**

QTune: A Query-Aware Database Tuning System with Deep Reinforcement Learning

Guoliang Li†, Xuanhe Zhou†, Shifu Li‡, Bo Gao‡
† Department of Computer Science,Tsinghua University, Beijing, China ‡ Huawei Company
liguoliang@tsinghua.edu.cn, zhouxuanhe@bupt.edu.cn, {gaobo15,lishifu}@huawei.com

DATASET FEATURIZATION

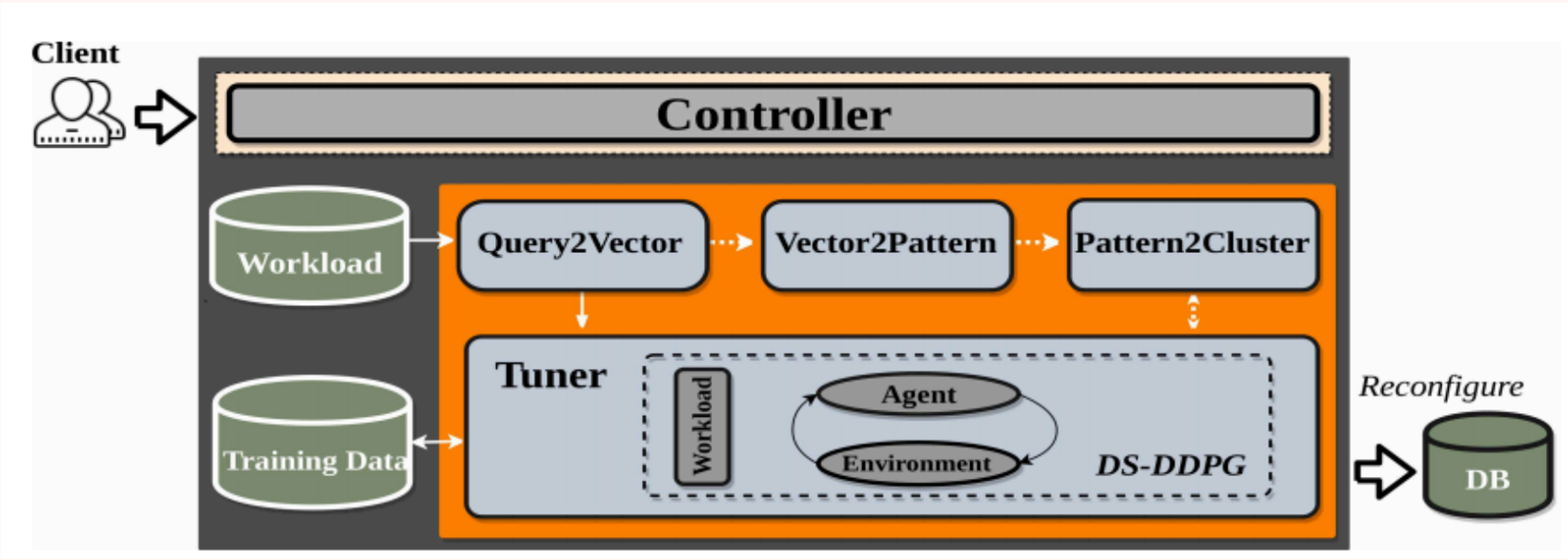
- Featurize SQL queries
- Considering rich query features:
  - Query type
  - Tables
  - Query cost
- Output a vector

MODEL

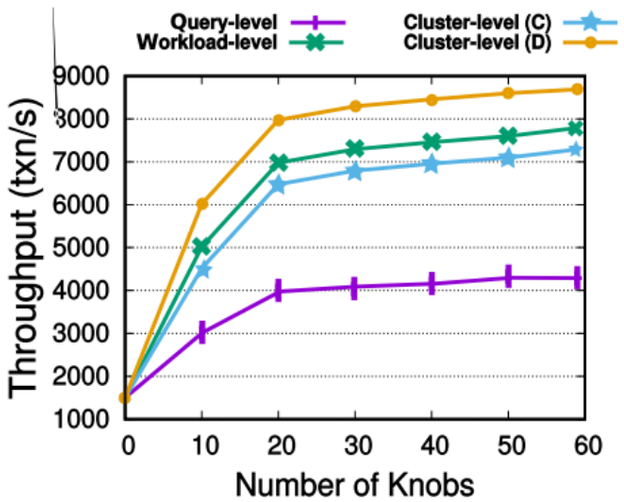
- **A Double-State Deep Deterministic Policy Gradient(DS-DDPG) Model**
- **Using actor-critic networks**
- **A DS-DDPG Model embeds the query features and utilizes the actor-critic algorithm to learn the relations among queries, database state and configurations to tune database configurations**

OVERVIEW

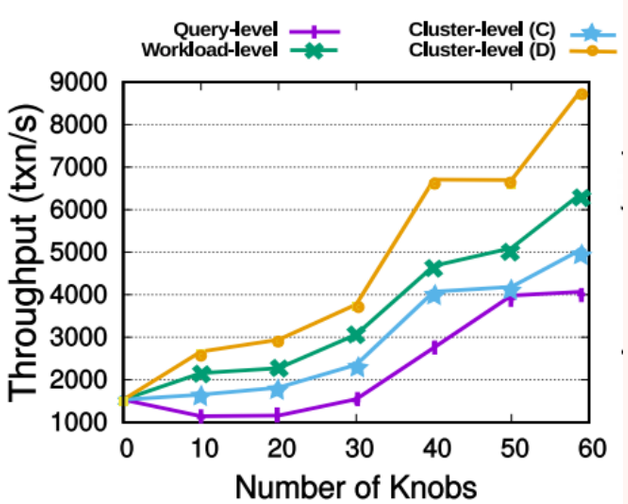
- Query2Vector
- Vector2Pattern
- Pattern2Cluster
- Tuner



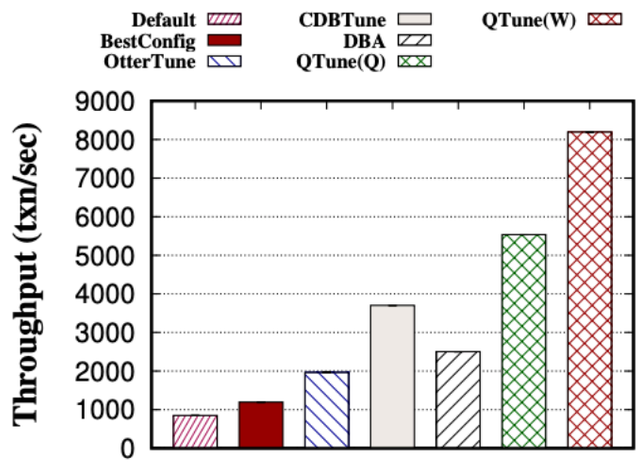
EVALUATION



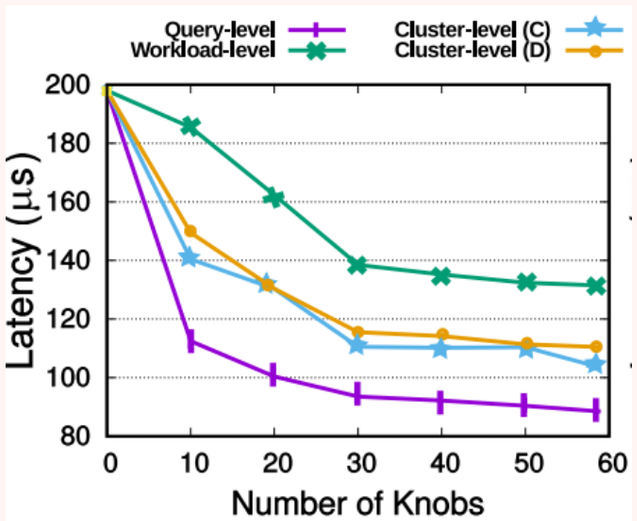
(a) IF-Throughput



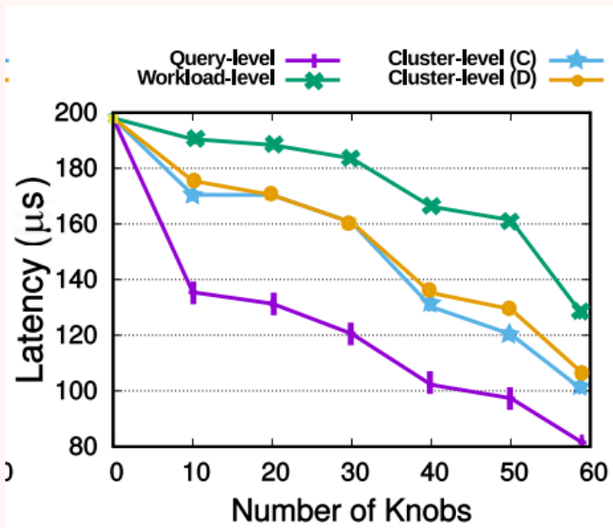
(b) RC-Throughput



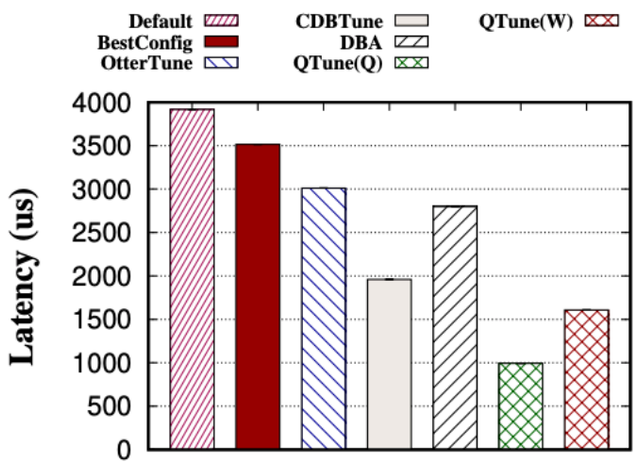
(a) Sysbench (RW)



(c) IF-Latency



(d) RC-Latency



(d) Sysbench (RW)

THANK YOU