MetaTune : Towards Workload Specific Configuration Tuning via Meta-Learning for RocksDB

연세대학교 컴퓨터과학과 염찬호

2024년 7월

과제명: IoT 환경을 위한 고성능 플래시 메모리 스토리지 기반 인메모리 분산 DBMS 연구개발

과제번호: 2017-0-00477

Table of contents

Introduction

Introduction

- Database has numerous parameters (knob)
  - Significantly affect performance metrics (e.g., throughput, latency, space amplification, write amplification, ect.)

- Database configuration tuning is a significant effort for database administrators (DBAs)
  - The interrelationships between the knobs are complex
  - Depending on the workload
  - Requires additional tuning each time

- Automatic database configuration tuning is required to replace traditional experience-based tuning approaches

Introduction

Limitations of previous research

1) Execute database for every samples
2) Workload mapping act as noise
3) GP is hard to reflect the complex knob space

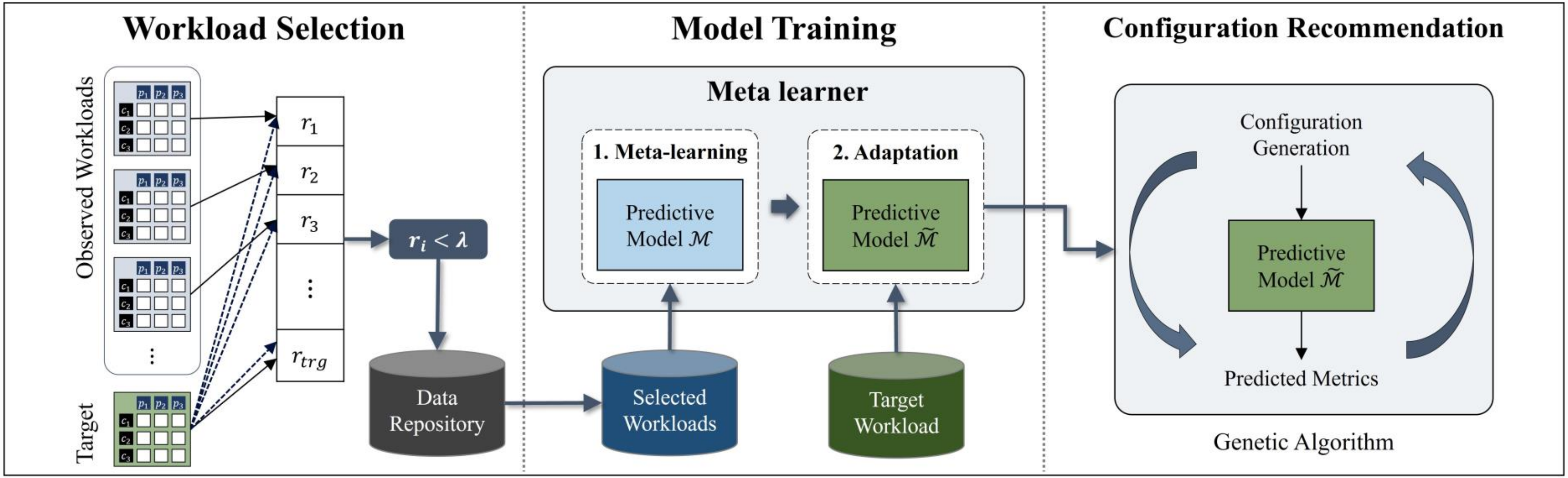
MetaTune

1) Use a predictive model instead of running a database
2) Workload selection & MAML → effectively select data and utilize it
3) GA can broadly search knob space by crossover, mutation step

Methodology

- System Overview
- Workload Selection
- Model Training
- Configuration Recommendation

System Overview

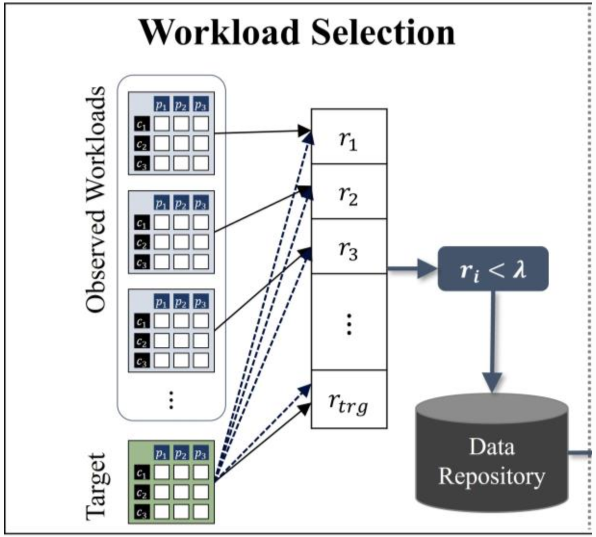


Workload Selection

- Additional data
  - → Improve accuracy of model
- When Data similarity is far from the target
  - → Act as noise
- ➢ Need to choose a workload similar distribution to target workload

Workload Selection

- For each workload, calculate the Mahalanobis distance to the target workload
- Select workloads that are similar to the target workload

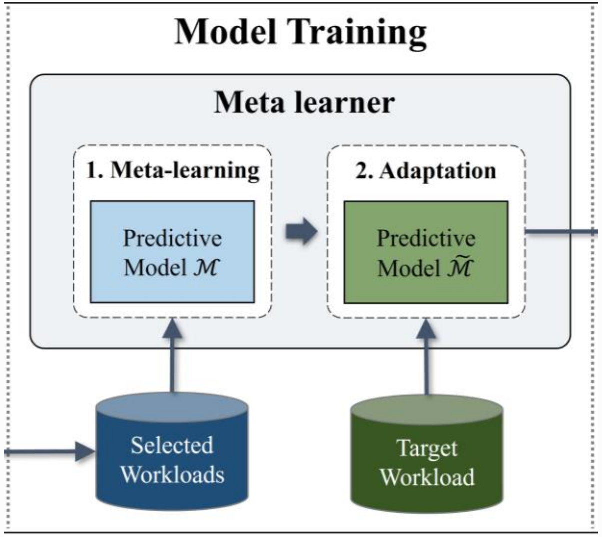


➢ Mahalanobis distance can reflect the variance of the data

Model Training

Predictive model

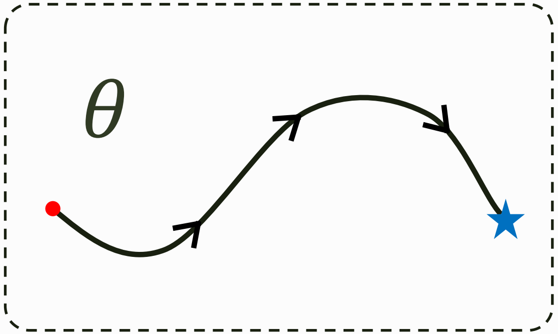
- ConvS2S (Convolutional sequence-to-sequence)
- Input
  - Configuration
- Output
  - External metrics of RocksDB (e.g., TIME, RATE, WAF, SAF)



Model Training

General training

- Initialize model parameter (θ)
- Update θ by model training
- <u>Need large amount of data to convergence</u>



Parameter space

Model-Agnostic Meta-Learning

MAML

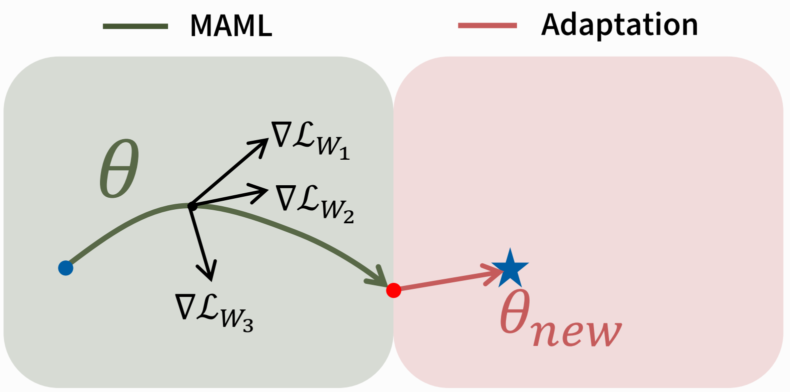
Learn the initial θ)of predictive model
that can quickly adapt to new tuning session
with small amount of data

θ

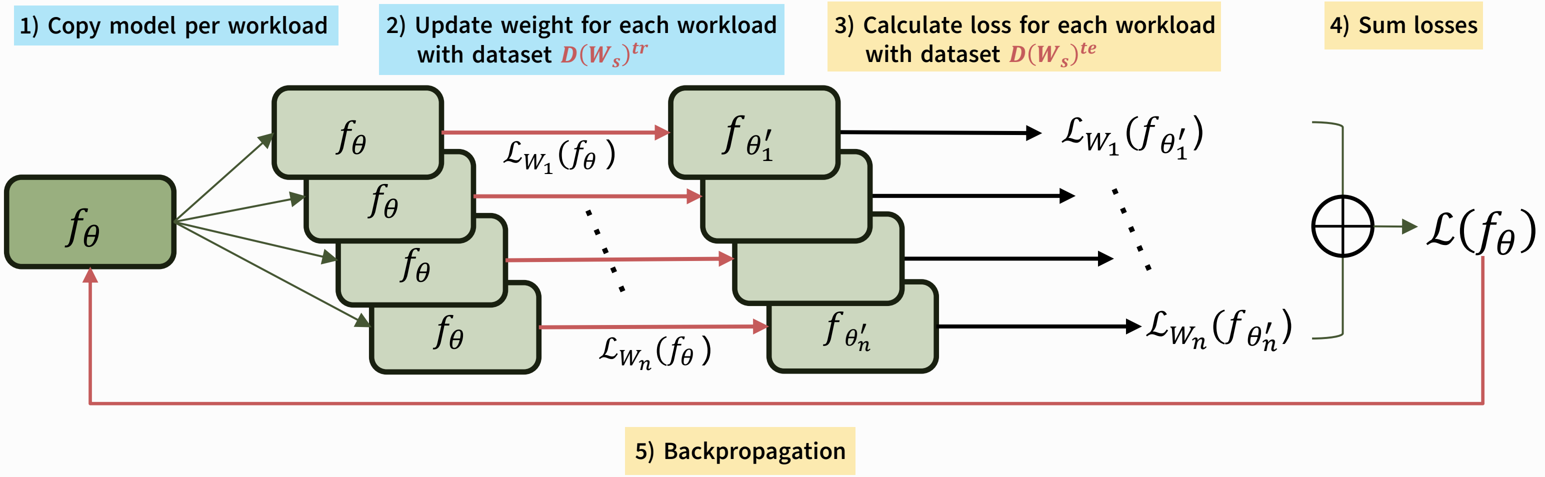
Parameter of predictive model being meta-learned with selected workload data

θ_{new}

Optimal parameter of predictive model for new tuning session with small amount of data



Model-Agnostic Meta-Learning



Inner loop (Step 1, 2)

- For each workload data, perform a gradient descent method update based on f_θ

Outer loop (Step 3~5)

- Based on the trained models ($f_{\theta'_i}$) for each workload, update the weight parameters of f_θ

Configuration Recommendation

Genetic Algorithm

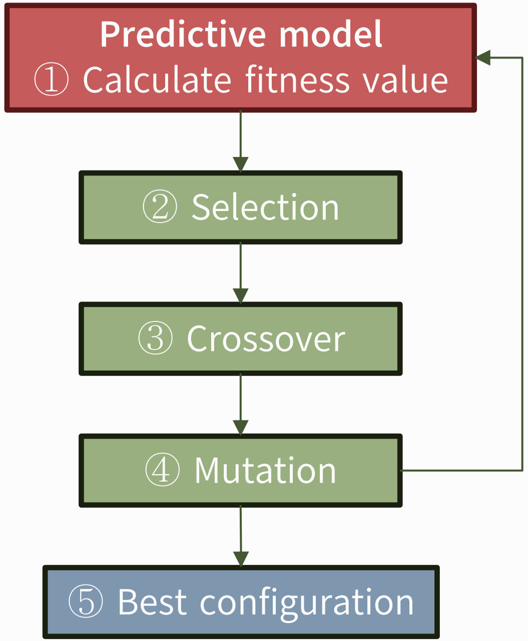
① Calculate and rank the fitness value of each configuration

② Choose superior configurations

③ Exchange knob value between configuration

④ Mutant a part of configuration

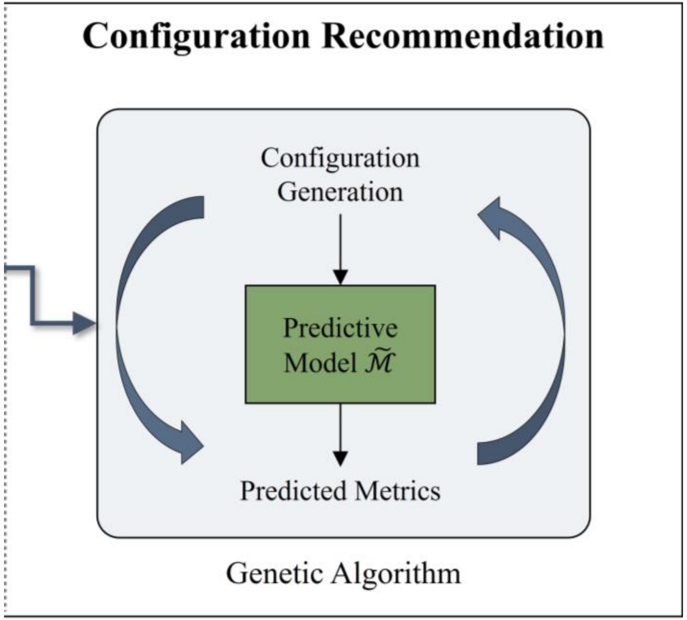
⑤ Recommend the best configuration



Configuration Recommendation

Advantages of Genetic Algorithm

- GA can broadly search knob space by crossover, mutation step
- We can use a predictive model to calculate fitness value
  - Don't need to execute a database for each configuration
  - Parallel computation is possible
- Comparing relative fitness values is more important rather than comparing exact values
  - Alleviate the requiring a large amount of data



Experiments

- Experiment Setup
- Comparison with baseline models
- Effectiveness of Meta-Learning
- Effectiveness of Workload Selection

Workload setting

TABLE I

OBSERVED WORKLOAD SETTING FOR TRAIN DATASET.

Workload	Value size(B)	Benchmark Option
R90W10_1	1024	readrandomwriterandom (read 90%, write 10%)
R90W10_4	4096	
R90W10_16	16384	
R90W10_64	65536	
R50W50_1	1024	readrandomwriterandom (read 50%, write 50%)
R50W50_4	4096	
R50W50_16	16384	
R50W50_64	65536	
R10W90_1	1024	readrandomwriterandom (read 10%, write 90%)
R50W50_4	4096	
R50W50_16	16384	
R50W50_64	65536	
UPDATE_1	1024	updaterandom
UPDATE_4	4096	
UPDATE_16	16384	
UPDATE_64	65536	

TABLE II

TAGET WORKLOAD SETTING FOR TEST DATASET.

Workload	Value size (B)	Benchmark Options
R70W30_8	8192	readrandomwriterandom (read 70%, write 30%)
R30W70_8		readrandomwriterandom (read 30%, write 70%)
Update_8		updaterandom
R70W30_32	32768	readrandomwriterandom (read 70%, write 30%)
R30W70_32		readrandomwriterandom (read 30%, write 70%)
Update_32		updaterandom
fs-rr	32768	fillseq, readrandom
fs-rww		fillseq, readwhilewriting
fr-rr		fillrandom, readrandom
fr-rww		fillrandom, readwhilewriting
fr-rr-sr		fillrandom, readrandom, seekrandom
fs-rr-ow		fillseq, readrandom, overwrite

External Metrics

TIME (↓)

- Cumulative time consumed across diverse database operations

RATE (↑)

- Number of operations executed per second

WAF (Write Amplification Factor) (↓)

- Proportion of data in the storage compared with the data generated by write operations

SAF (Space Amplification Factor) (↓)

- Proportion of the actual storage capacity used by RocksDB compared with the space consumed by the storage

Evaluation Metrics

Predictive model

- **PCC** (Pearson Correlation Coefficient)
  - Correlation coefficient that measures linear correlation between two sets of data two variables
  - e.g., predicted and observed values
- **RMSE** (Root Mean Square Error)
  - Difference between the predicted and observed values

DBMS performance

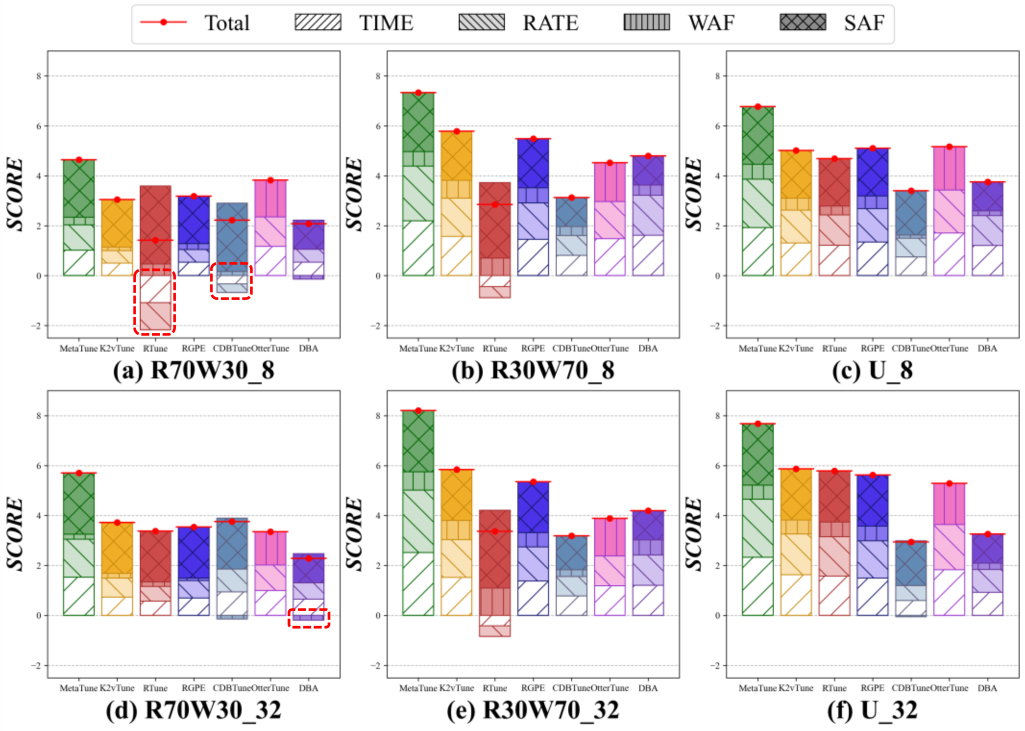
- $\mathbf{Score} = \log\left(\frac{TIME}{TIME_d}\right) + \log\left(\frac{RATE_d}{RATE}\right) + \log\left(\frac{WAF}{WAF_d}\right) + \log\left(\frac{SAF}{SAF_d}\right)$
  - Score is 0 when default configuration setting
  - Higher score indicates better tuning performance

- $\boldsymbol{TIME_d, RATE_d, WAF_d, SAF_d}$
  - Each external metric at default setting
- $\boldsymbol{TIME, RATE, WAF, SAF}$
  - Each external metric at recommended configuration

Comparison with baselines

- MetaTune is Best performance across all target workloads

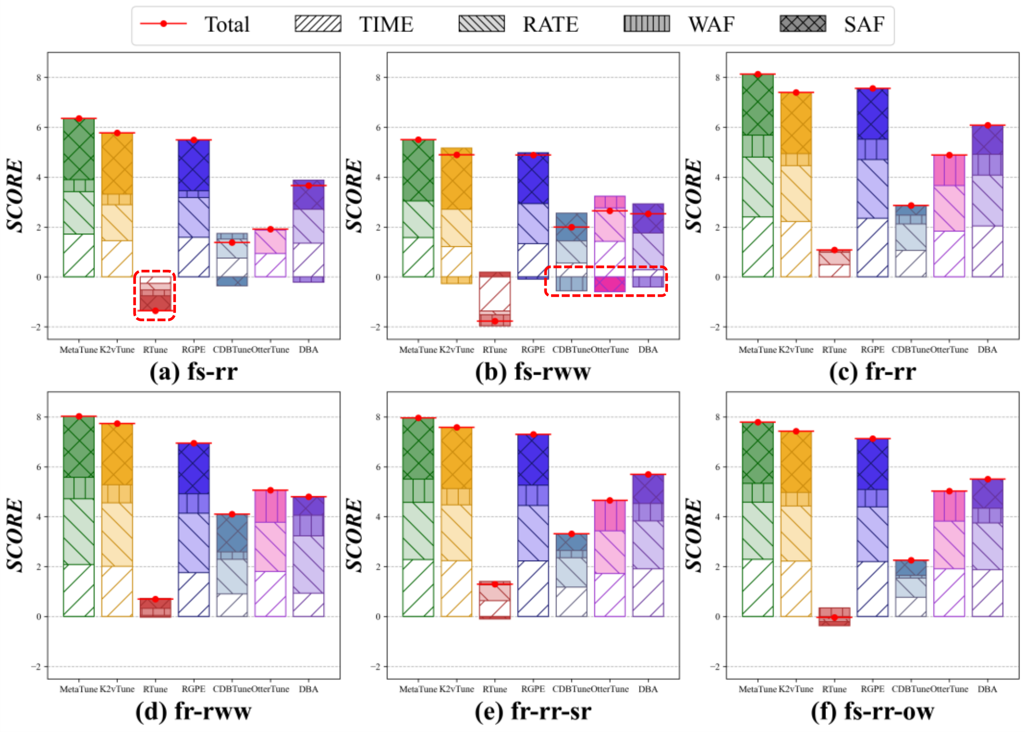
- For other baselines, there are cases where the ***score*** of a specific external metric is **negative**, indicating a lower performance than before tuning. In contrast, all ***scores*** for MetaTune are positive.



Comparison with baselines

- <u>MetaTune</u> is <u>Best performance</u> across all target workloads

- For other baselines, there are cases where the ***score*** of a specific external metric is <u>**negative**</u>, indicating a lower performance than before tuning. In contrast, all ***scores*** for <u>MetaTune</u> are positive.



Effectiveness of Meta Learing

- MAML outperformed in terms of both the predictive accuracy and performance score across all target workloads

- MAML approach learns the initial parameters of the predictive model at a balanced point across multiple workloads
  - Allowing for better adaptation to the target workload

Method	Workload	Score (↑)	PCC(↑)	RMSE(↓)
MAML	R70W30_8	**4.6458**	**0.8703**	**6.9853**
	R30W70_8	**7.3344**	**0.8152**	**14.3024**
	UPDATE_8	**6.7671**	**0.9199**	**19.0432**
	R70W30_32	**5.7071**	**0.855**	**8.8562**
	R30W70_32	**8.2065**	**0.8609**	**15.0312**
	UPDATE_32	**7.6783**	**0.8813**	**19.0681**
Workload Mapping	R70W30_8	3.9427	0.7013	49.7579
	R30W70_8	6.6367	0.6491	76.7732
	UPDATE_8	6.1166	0.7283	103.6948
	R70W30_32	5.1850	0.7216	50.6117
	R30W70_32	7.9496	0.6886	76.3425
	UPDATE_32	6.1315	0.6936	102.2122

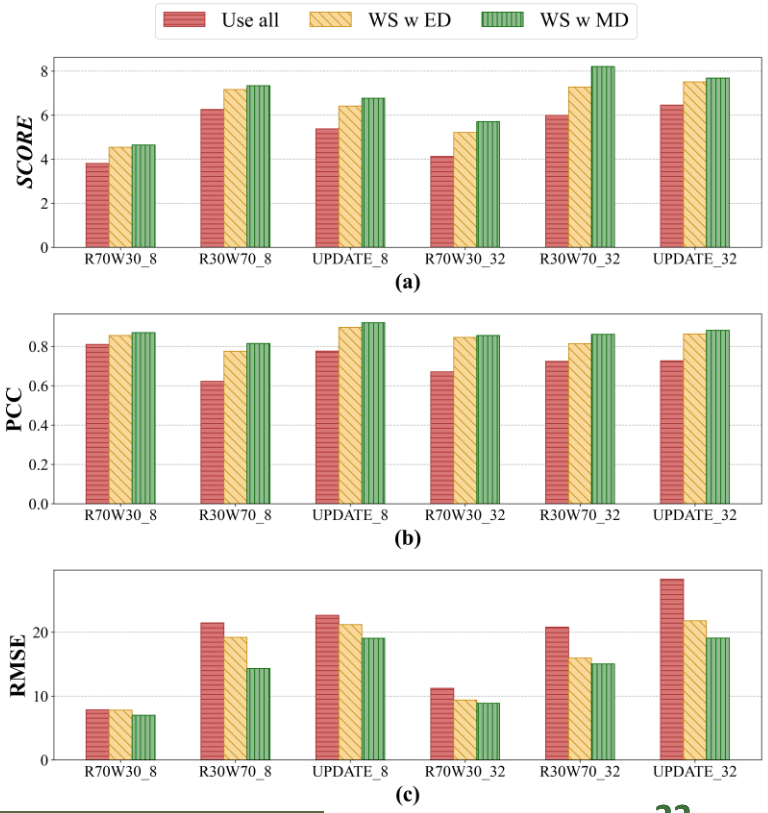
Effectiveness of Workload Selection

- Performance score

W.S w M.D > W.S w E.D > w/o W.S

- Even data obtained through database configuration tuning can lead to poor performance if it is dissimilar to the target workload

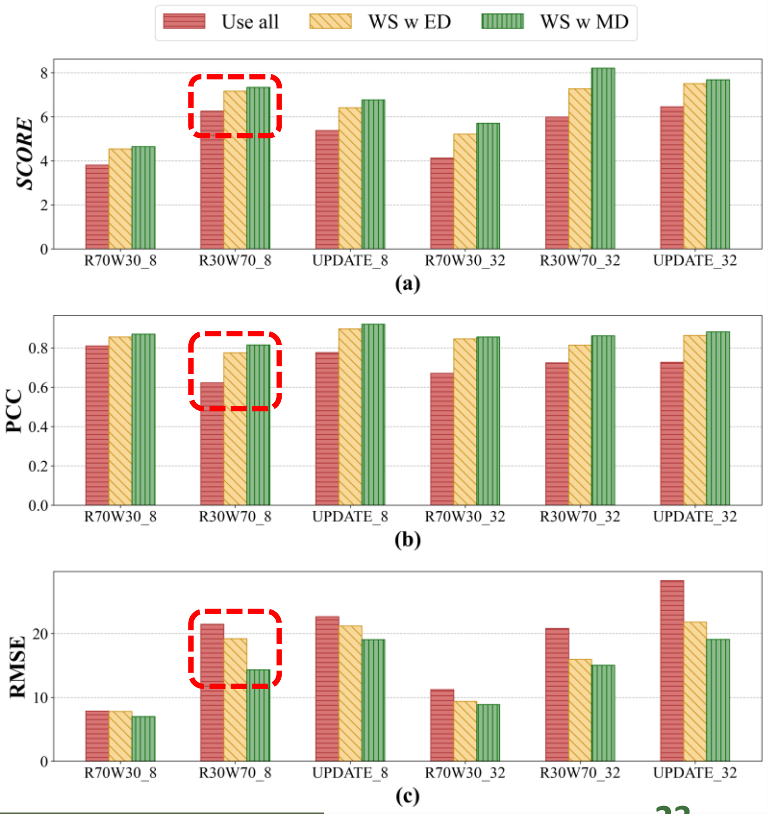
- Mahalanobis distance, which considers data distribution, correlations between variables is more effective than the Euclidean distance



- Use all — Without workload selection (= using all workload data)
- W.S w E.D — Workload selection with Euclidean Distance
- W.S w M.D — Workload selection with Mahalanobis Distance

Effectiveness of Workload Selection

- The **variance in performance socre** is more similar to the **variance in PCC** than RMSE
  - PCC is a more explanatory metric for performance variance in database tuning
  - Capturing the trend of performance variance (PCC) than to predict the exact performance values (RMSE) is more crucial for the tuning performance
  - This reduces the burden of requiring extensive training data



Conclusion

Conclusion

MetaTune

- In addition to general tuning, RocksDB also needs to improve the performance of WAF and SAF
  - Propose an automatic RocksDB tuning system that operates on a multi-objective optimization approach → Tuning TIME, RATE, WAF, SAF simultaneously
- Proposed workload selection to effectively select data and this improves tuning performance
- Improved performance of predictive model on target workloads by applying **meta-learning**
- Utilize GA and **predictive models** together → Reduce the training burden of the predictive model

Future work

- Experiment with extending our proposed model to other DBMSs
- Increase the accuracy of predictive model with smaller amount of data

Q & A

Thank you