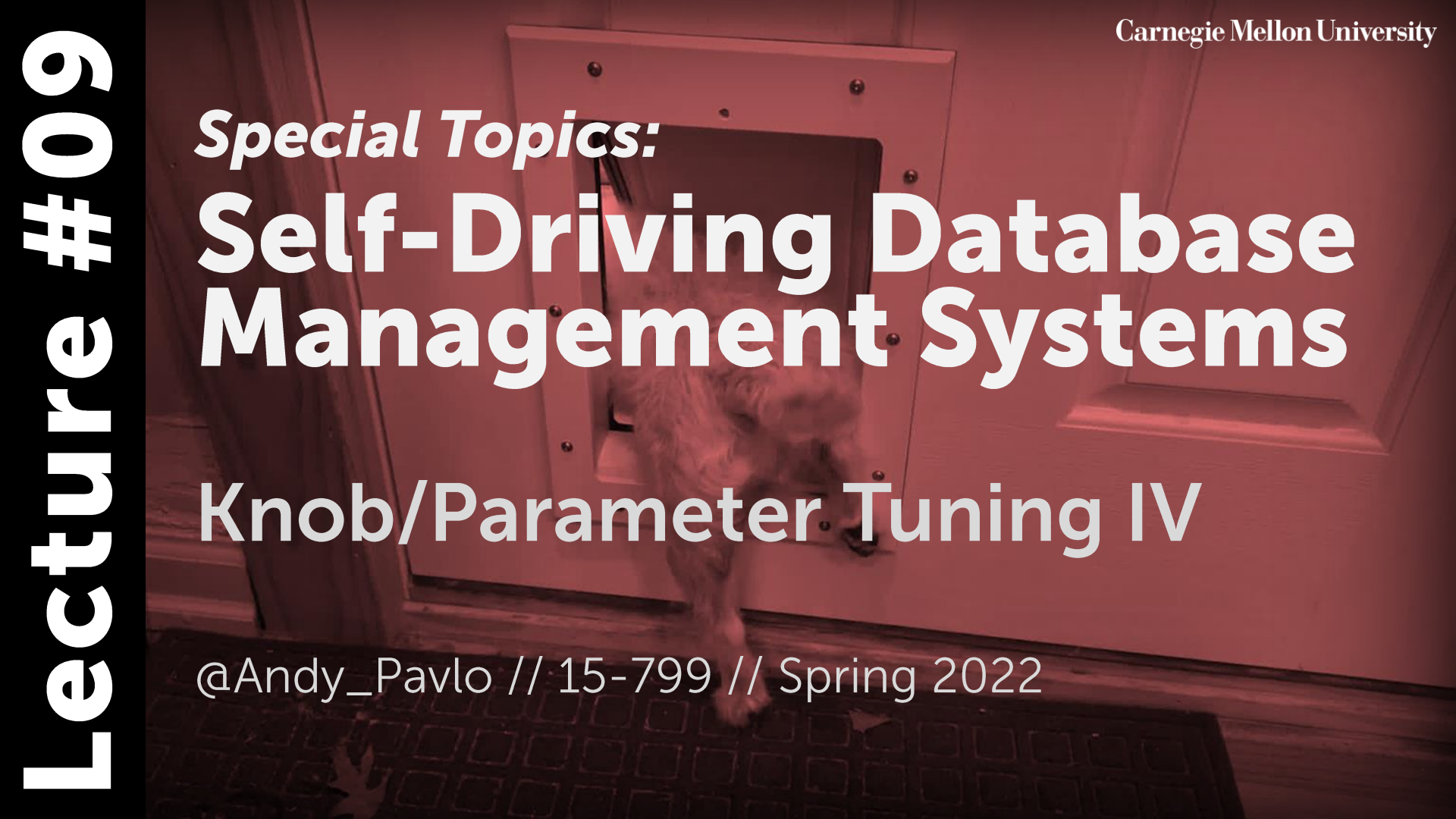
Lecture #09

Carnegie Mellon University

Special Topics:

Self-Driving Database Management Systems

Knob/Parameter Tuning IV



@Andy_Pavlo // 15-799 // Spring 2022

ADMINISTRIVIA

DB Speaker Mon Feb 21 @ 4:30pm ET (Zoom)
→ Venky Raghavan (Greenplum Orca)

Andy will release the remaining P1 workload this weekend.

LAST CLASS

Last three lectures have been about different optimization methods for DBMS knobs

→ OtterTune – Bayesian Optimization
→ CDBTune – DDPG
→ ResTune – Meta-Learning + Bayesian Optimization

Beyond the core search algorithm, there are additional design choices that are tacked on to each approach…

KNOB TUNING DESIGN CHOICES

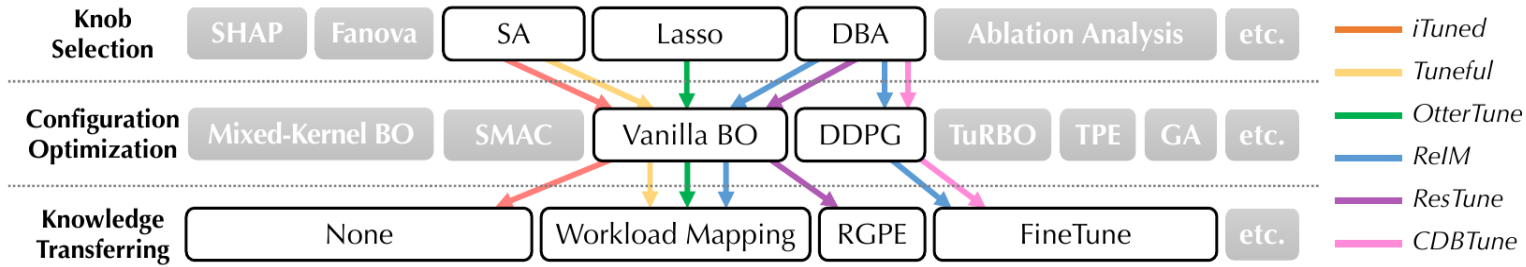


Figure S1: Detailed Intra-algorithm Designs of Configuration Tuning Systems (The Full Picture): Black boxes denote the algorithms adopted by existing database tuning systems (indicated by colored paths), and grey boxes denote the algorithms in the HPO field. SA denotes sensitivity analysis based on Gini score and GA denotes Genetic algorithm.

TODAY'S AGENDA

Taxonomy
Knob Selection
Optimization Algorithm
Knowledge Transfer

KNOB SELECTION

Automatically identify which knobs to tune.
→ Varies per DBMS / version
→ Varies per workload / database

Increasing the number of knobs to optimize affects how long the optimization will take to converge.

KNOB SELECTION

Table 2: Importance Measurements in Knob Selection.

Measure	Category	Brief Description
Lasso [79]	Variance based	Based on coefficient of linear regression, effective for when existing irrelevant features.
Gini score [63]	Variance based	Based on the times a feature is used in tree splits, successful in high-dimensional feature selection.
fANOVA [35]	Variance based	Decomposing the variance of target function, commonly used in the HPO field.
Ablation analysis[12]	Tunability based	Comparing feature difference between configurations, straightforward and intuitive.
SHAP [53]	Tunability based	Decomposing the performance change additively, solid theoretical foundation derived from game theory.

Source: Xinyi Zhang + Zhuo Chang

GINI INDEX

Measures the degree or probability of a variable being wrongly classified in decision tree when it is randomly chosen.

→ Input: Knob Values

→ Output: Target Objective

Knobs that have a greater impact on the objective will be higher up in a decision tree, therefore the algo should tune them first.

TO TUNE OR NOT TO TUNE?: IN SEARCH OF OPTIMAL CONFIGURATIONS FOR DATA ANALYTICS
SIGKDD 2020

GINI INDEX

Measures the degree o… probability of a variable being wrongly classifi… randomly chosen.

→ Input: Knob Values

→ Output: Target Object…

Knobs that have a gre… will be higher up in a… algo should tune them first.

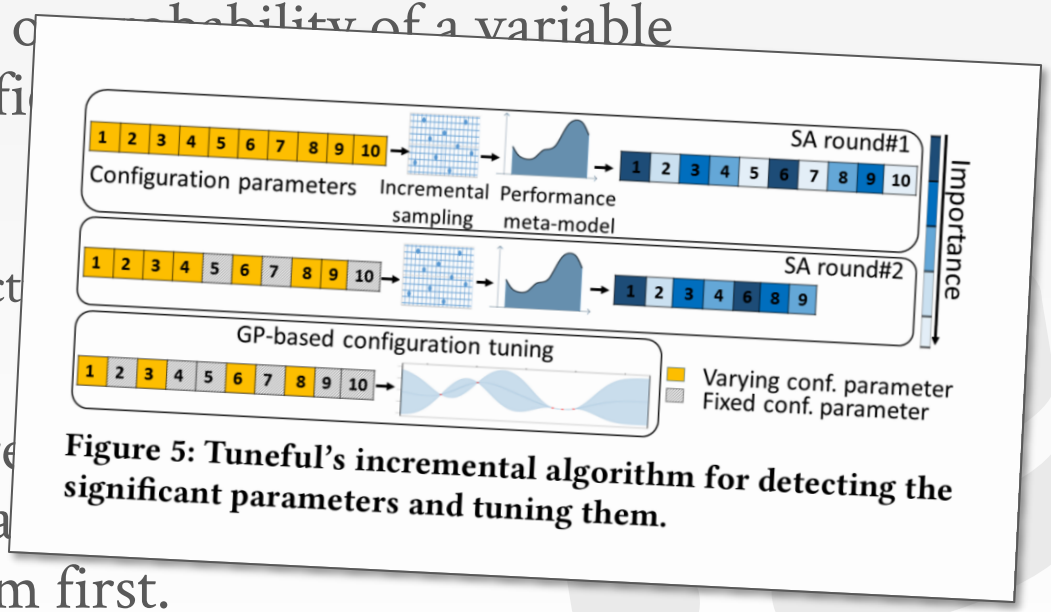
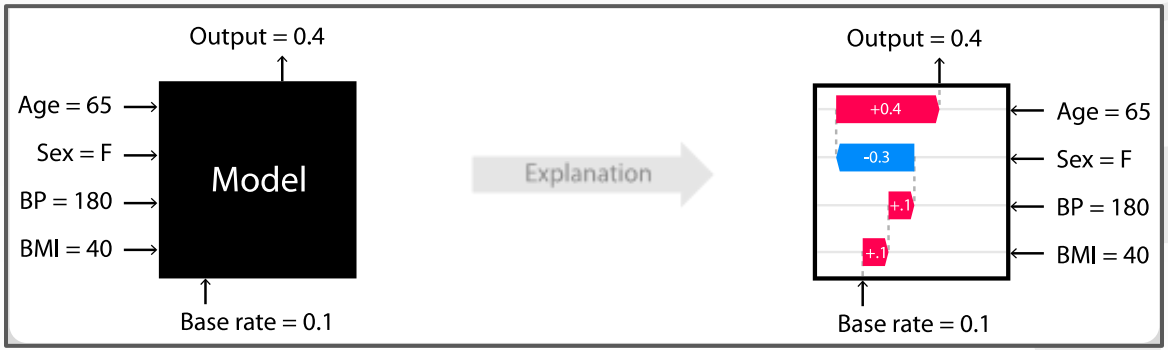


Figure 5: Tuneful's incremental algorithm for detecting the significant parameters and tuning them.

TO TUNE OR NOT TO TUNE?: IN SEARCH OF OPTIMAL CONFIGURATIONS FOR DATA ANALYTICS
SIGKDD 2020

SHAP

SHAP (SHapley Additive exPlanations) is a game theoretic approach to explain the output of any machine learning model.



Source: Scott Lundberg

KNOB SELECTION COMPARISON

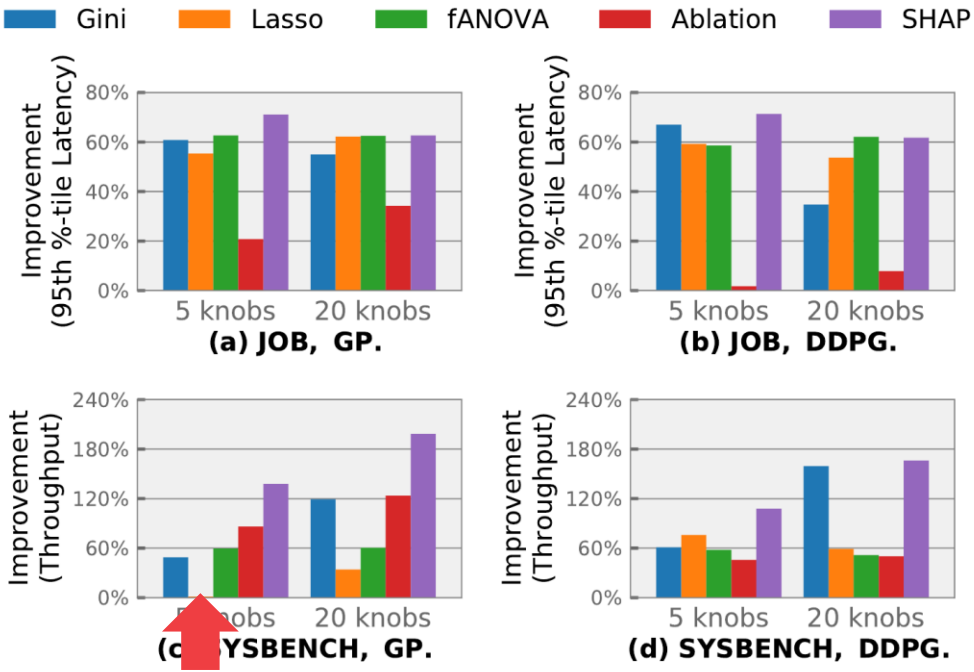


Figure 3: Performance improvement when tuning top 5/20 knobs ranked by different importance measurements.

Source: Xinyi Zhang + Zhuo Chang

KNOB SELECTION COMPARISON

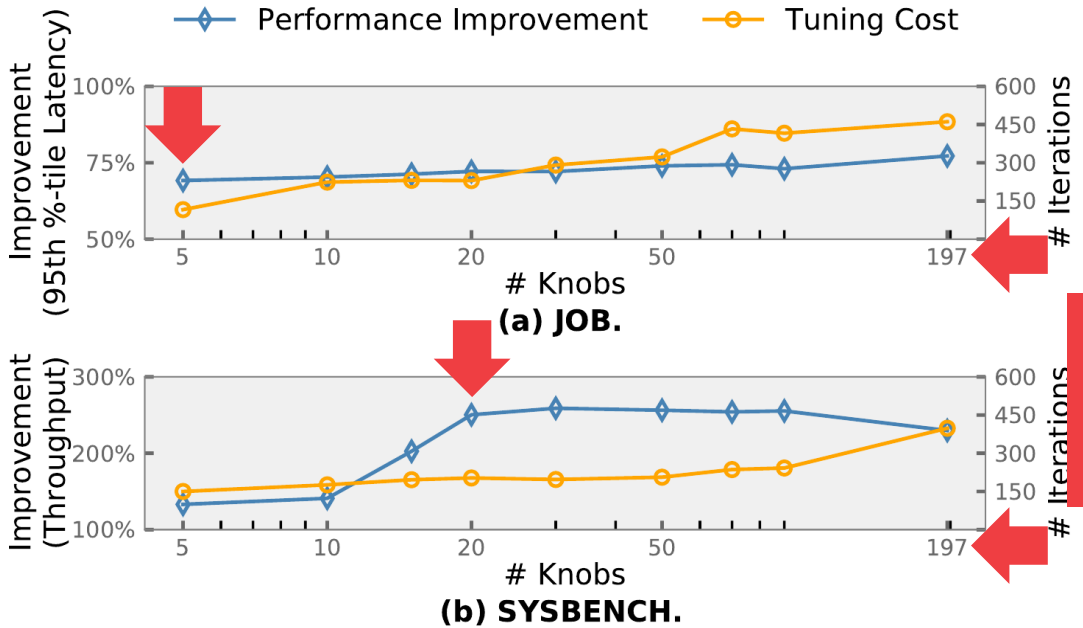


Figure 5: Performance improvement and tuning cost when increasing the number of tuning knobs.

Reminder: There are not 197 MySQL knobs to autotune!

Source: Xinyi Zhang + Zhuo Chang

MYSQL V5.7 KNOBS

Table S3: The Top-10 impacting knobs with high tunability for OLTP workloads

Knob	Type	Dynamic	Module	Description
innodb_thread_concurrency	Integer	Yes	Concurrency	The maximum number of threads permitted inside of InnoDB.
innodb_log_file_size	Integer	No	Logging	The size in bytes of each log file in a log group.
max_allowed_packet	Integer	Yes	Replication	The upper limit on the size of any single message between the MySQL server and clients.
innodb_io_capacity_max	Integer	Yes	IO	The maximum number of IOPS performed by InnoDB background tasks
tmp_table_size	Integer	Yes	Memory	The maximum size of internal in-memory temporary tables.
query_prealloc_size	Integer	Yes	Memory	The size in bytes of the persistent buffer used for statement parsing and execution.
max_heap_table_size	Integer	Yes	Memory	The maximum size to which user-created memory tables are permitted to grow.
innodb_doublewrite	Categorical	No	Memory	Whether the doublwrite buffer is enabled.
transaction_alloc_block_size	Interger	Yes	Memory	The amount in bytes by which to increase a per-transaction memory pool which needs memory.
join_buffer_size	Interger	Yes	Memory	The minimum size of the buffer that is used for joins.

× DBMS can reject queries from application.

Source: Xinyi Zhang + Zhuo Chang

MYSQL V5.7 KNOBS

Table S3: The Top-10 impacting knobs with high tunability for OLTP workloads

Knob	Type	Dynamic	Module	Description
innodb_thread_concurrency	Integer	Yes	Concurrency	The maximum number of threads permitted inside of InnoDB.
innodb_log_file_size	Integer	No	Logging	The size in bytes of each log file in a log group.
max_allowed_packet	Integer	Yes	Replication	The upper limit on the size of any single message between the MySQL server and clients.
innodb_io_capacity_max	Integer	Yes	IO	The maximum number of IOPS performed by InnoDB background tasks
tmp_table_size	Integer	Yes	Memory	The maximum size of internal in-memory temporary tables.
query_prealloc_size	Integer	Yes	Memory	The size in bytes of the persistent buffer used for statement parsing and execution.
max_heap_table_size	Integer	Yes	Memory	The maximum size to which user-created memory tables are permitted to grow.
innodb_doublewrite	Categorical	No	Memory	Whether the doublwrite buffer is enabled.
transaction_alloc_block_size	Interger	Yes	Memory	The amount in bytes by which to increase a per-transaction memory pool which needs memory.
join_buffer_size	Interger	Yes	Memory	The minimum size of the buffer that is used for joins.

× DBMS can abort queries due to memory limitation.

Source: Xinyi Zhang + Zhuo Chang

MYSQL V5.7 KNOBS

Table S3: The Top-10 impacting

Knob	Type	Dynamic	Module
innodb_thread_concurrency	Integer	Yes	Concurrency
innodb_log_file_size	Integer	No	Logging
max_allowed_packet	Integer	Yes	Replication
innodb_io_capacity_max	Integer	Yes	IO
tmp_table_size	Integer	Yes	Memory
query_prealloc_size	Integer	Yes	Memory
max_heap_table_size	Integer	Yes	Memory
innodb_doublewrite	Categorical	No	Memory
transaction_alloc_block_size	Interger	Yes	Memory
join_buffer_size	Interger	Yes	Memory

- `innodb_doublewrite`

Command-Line Format	`--innodb-doublewrite[={OFF|ON}]`
System Variable	`innodb_doublewrite`
Scope	Global
Dynamic	No
`SET_VAR` Hint Applies	No
Type	Boolean
Default Value	`ON`

The `innodb_doublewrite` variable controls whether the doublwrite buffer is enabled. It is enabled by default in most cases. To disable the doublewrite buffer, set `innodb_doublewrite` to 0 or start the server with `--skip-innodb-doublewrite`. You might consider disabling the doublewrite buffer if you are more concerned with performance than data integrity, as may be the case when performing benchmarks, for example.

× Improves performance but can potentially cause lost data.

Source: Xinyi Zhang + Zhuo Chang

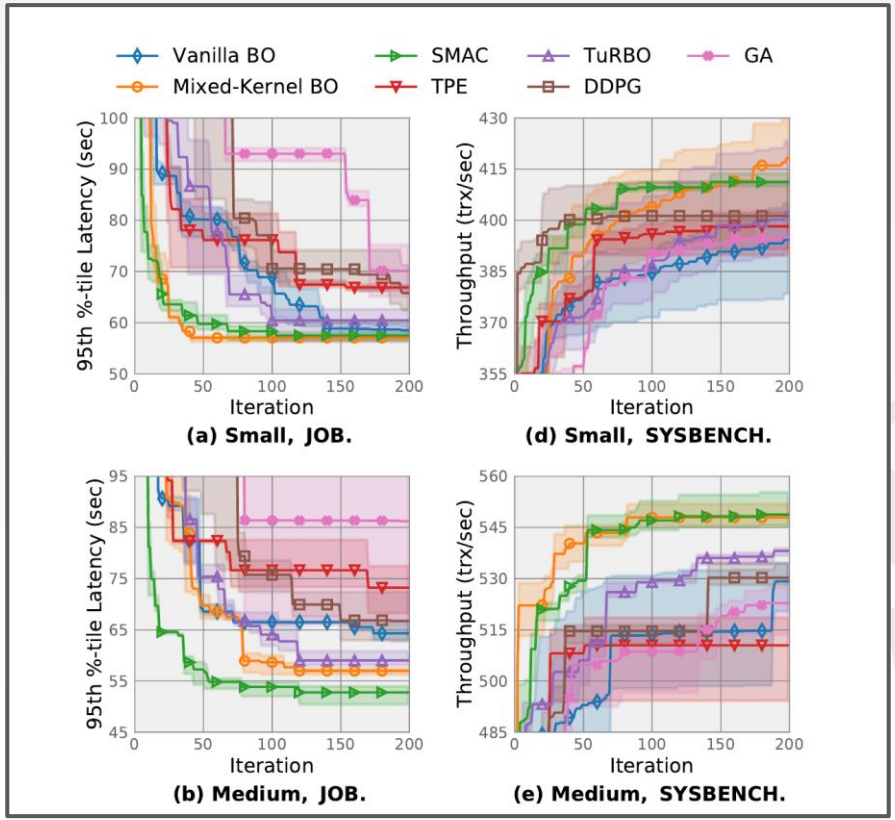
OPTIMIZATION ALGORITHM

Table 3: Algorithms for Optimizers. ✓means the optimizer has specific design for the case in the column and – means the optimizer does not have such design.

Algorithm	High-dimensionality	Heterogeneity
Vanilla BO	–	–
Mixed-Kernel BO [41]	–	✓
SMAC [34]	✓	✓
TPE [9]	–	✓
TurBO [22]	✓	–
DDPG [52]	✓	–
GA [46]	–	✓

Source: Xinyi Zhang + Zhuo Chang

OPTIMIZATION ALGORITHM



Source: Xinyi Zhang + Zhuo Chang

KNOWLEDGE TRANSFER

Reduce the optimization time by reusing data collected from previous tuning sessions.

Need to account for both **workload** and **hardware** variations in training data.

KNOWLEDGE TRANSFER

Workload Mapping:

→ Match target workload to previous one using Euclidean distance of DBMS metrics.

Ranking-Weight GP Ensemble:

→ Maintain multiple GP models that are weighted by their similarity to target workload.

Fine-Tune:

→ Start with a global model and then adjust its weights based on the target workload.

KNOWLEDGE TRANSFER

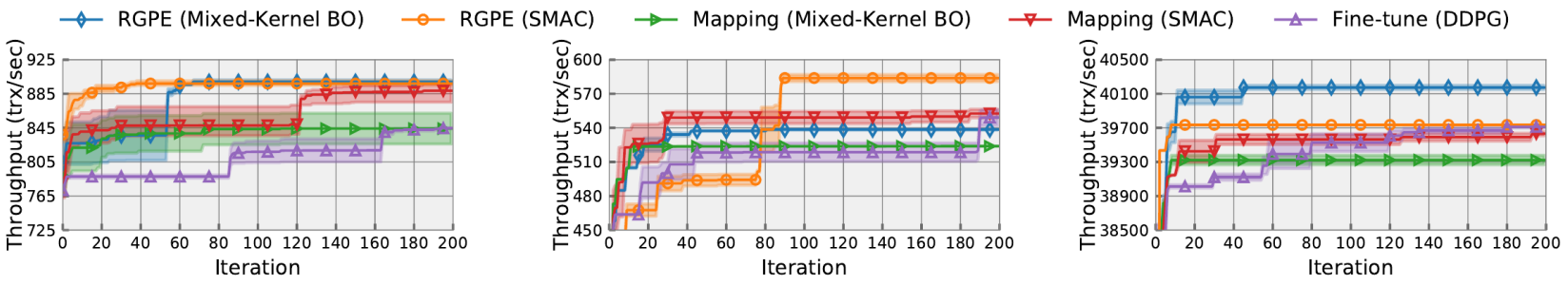


Figure S2: The absolute performance over iteration of each combination of transfer framework and base learner.

Source: Xinyi Zhang + Zhuo Chang

OPEN QUESTIONS

Under what conditions should the DBMS reoptimize its current knobs configuration?

How to optimize knobs when it is not possible to create a snapshot / workload trace?

How to automatically determine which knobs could cause correctness problems?

OPEN QUESTIONS

Can the tuning algorithm reuse training data across different versions of the DBMS?

→ Major Versions: MySQL v5.7 ➞ MySQL v8.0

→ Minor Versions: PostgreSQL v10.1 ➞ PostgreSQL v10.2

What about tuning other knobs that matter for the DBMS?

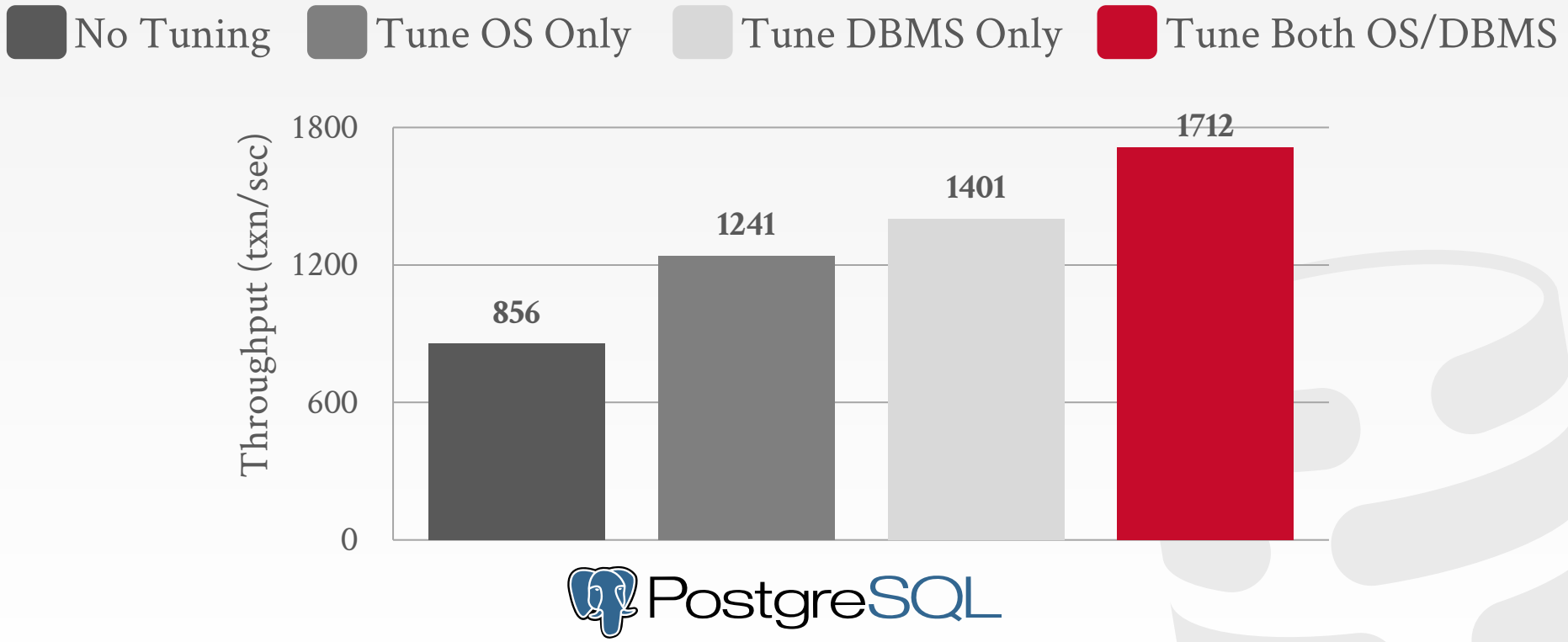
OS KERNEL TUNING

We extended OtterTune to support tuning OS kernel parameters together with DBMS.

Linux kernel parameters:

→ `vm.swappiness`
→ `kernel.sched_migration_cost_ns`
→ `kernel.sched_autogroup_enabled`
→ `vm.dirty_background_bytes`
→ `vm.dirty_bytes`

OS KERNEL TUNING: TPC-C



PARTING THOUGHTS

Andy thinks that ResTune is probably the best approach for knob tuning right now.

The approaches that we have discussed tune knobs without considering other design issues.

They also assume that tuning is a one-time activity. There may be better approaches if tuning occurs over weeks / months / years.

NEXT CLASS

One last paper on automatically selecting partitioning schemes for tables.

We will then switch to the auxiliary problems for a holistic autonomous DBMS.

→ Workload Forecasting
→ Behavior Modeling