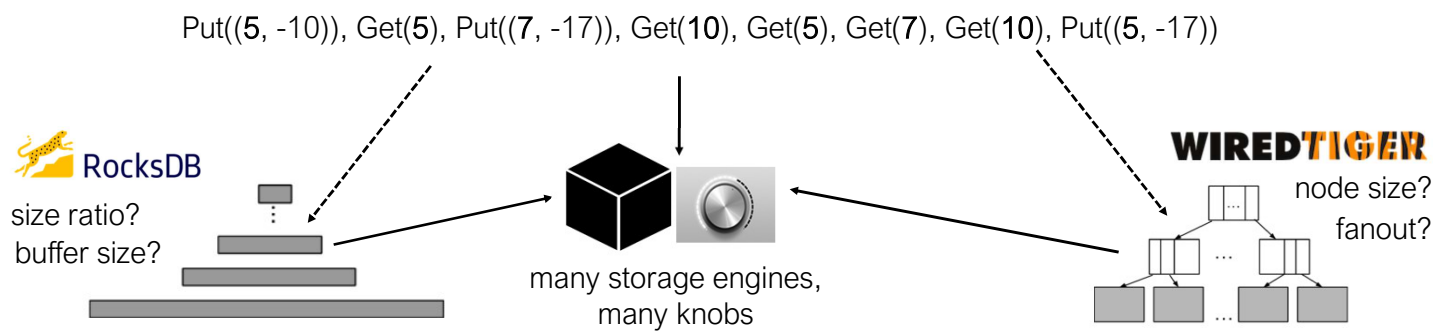
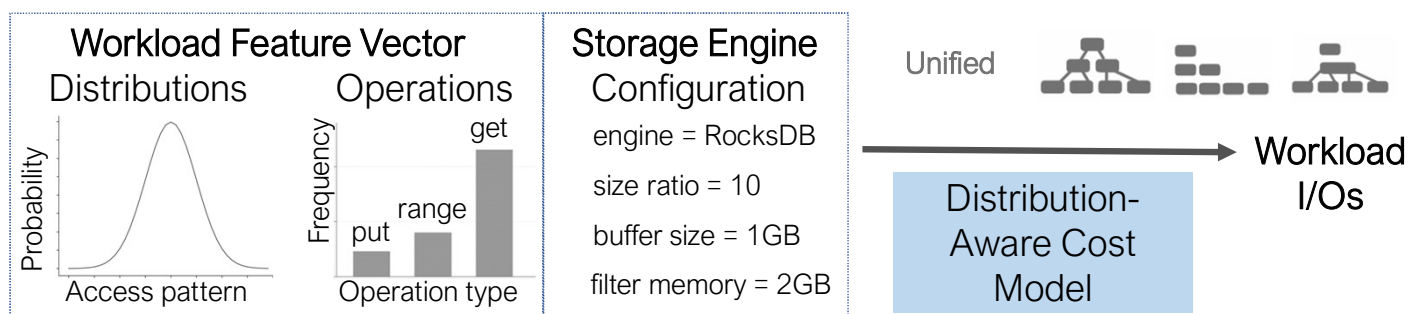




## THE GOAL: Which storage engine and tuning is best for an application?

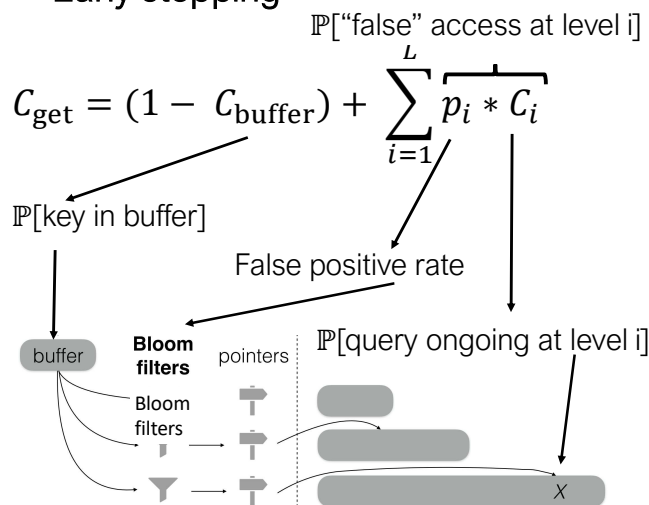


## AVERAGE-CASE ANALYSIS: A Distribution-Aware Framework

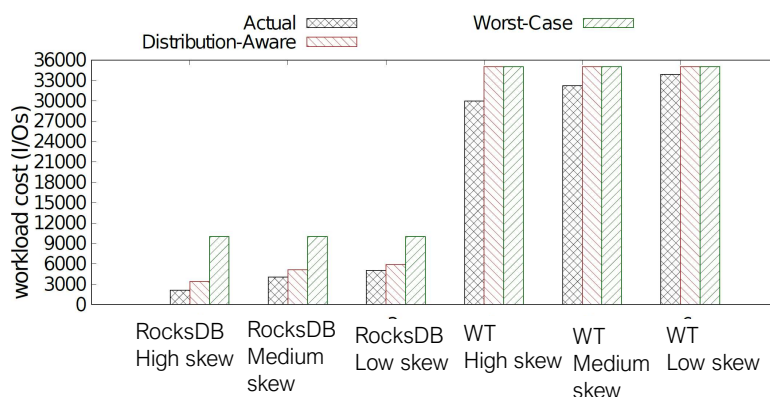


## EXAMPLE: Get Cost on LSM

### "Early stopping"



## VERIFICATION: Key-Value Stores



**Distribution-aware model captures workload I/Os!**

10 mil keys; workload (20,000 ops): 50% reads/50% writes

## USE CASES: How can these models be utilized?



**Compare** designs



**Automate** selection



**Search** design space



**Create** self-designing storage engine

