SOFIYA SEMENOVA

SENIOR THESIS

Contents

Introduction

My senior thesis focuses on leveraging machine learning methods to develop cost-effective performance management services for cloud databases. To that end, the first part of my thesis consists of developing a demo for WiSeDB, an end-to-end, machine learning based solution for workload management for cloud databases. The second part consists of developing a machine learning approach to data-driven fragmentation and fragment distribution.

Motivation

With the movement of database applications from on-premises data centers to cloud environments, database application developers are responsible for addressing resource distribution, workload distribution, and query scheduling, all while meeting specific performance criteria. To leverage the full benefits of cloud databases, an end-to-end solution that addresses all the above concerns and maintains different cost and performance constraints would, many times, create better solutions than a database developer could.

Further, the cost of using IaaS cloud databases could be greatly decreased with custom data-driven fragmentation, and

Previous research

Previous research has addressed resource provisioning, workload distribution, and query scheduling independently, but before WiSeDB, an end-to-end solution that addresses all the problems had not yet been implemented.

Summary of work done

The demo

The Demo

WiSeDB

THE DEMONSTRATION is built on WiSeDB, a machine-learning driven system for cost and performance management for cloud databases. WiSeDB is comprised of two learning-based approaches: a supervised learning approach, and a reinforcement learning approach.

The supervised learning approach

Demo

THE DEMO ALLOWS the user to select a subset of provided query templates, specify an SLA type and value, and

Economic Methods

Introduction