**AZDBLab: A Laboratory Information System for Large-Scale Empirical DBMS Studies**

**Young-Kyoon Suh**, Richard T. Snodgrass*, and Rui Zhang*

*University of Arizona and *Dataware Ventures

---

### Introduction

Scientific methodology in the database field can provide a deep understanding of DBMS query optimizers, for better engineered designs.

Few **DBMS-centric** labs are available for scientific investigation; prior labs have focused on networks and smartphones.

---

**AZDBLab (AriZona DataBase Laboratory)**

- Has been in development for seven years.
- Assists database researchers to conduct **large-scale empirical** studies across **multiple** DBMSes.
- Runs massive experiments with thousands or millions of queries on multiple DBMSes.
- Supports as experiment subjects seven relational DBMSes supporting SQL and JDBC.
- Provides robustness to collect data over 8,277 hours running about **2.4 million** query executions.
- Conducts automated analyses on multiple query execution runs.

---

**Contributions**

- Novel research infrastructure, dedicated for large-scale empirical DBMS studies
- Seamless data provenance support
- Several decentralized monitoring schemes: phone apps, web apps, and watcher
- Reusable GUI
- Extensibility through a variety of plugins: labshelf, analysis, experiment subject, and scenario

---

**AZDBLab Architecture**

---

**Demonstration**

Step 1: Choose a labshelf, add a user, and create a notebook, a paper, and a study in the paper on the Observer GUI.

Step 2: Load an experiment specification into the notebook.

Step 3: Schedule an experiment run on a particular DBMS.

Step 4: Monitor the run status via Observer, a web app, and a mobile app, and wait for the experiment to be done.

Step 5: Add the completed experiment run to the study and conduct a timing protocol analysis for the study.

Step 6: Produce LaTeX/PDF documents containing the analysis results.