www.DBTechNet.org

Study Guide for Concurrency Lab 1

(CCLab1_StudyGuide)

Disclaimers

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Trademarks of products mentioned are trademarks of the product vendors.

Foreword

Properly designed and managed databases provide the reliable storage and access services for applications, but only if applications use these database services properly. The most critical of these services are transaction processing and concurrency control services. Database textbooks typically present concurrency topics according to the traditional serializability theory and ISO SQL standard, whereas behaviors of DBMS products used for applications vary, and contain bugs. The challenge of the "Art of Transaction Programming" in developing reliable applications is to cope with all these issues.

The purpose of this lab is to focus on concurrency management from point of view of application developers studying use of the mainstream DBMS products: DB2, Oracle and SQL Server, used by the ICT industry. The traditional serializability theory and research papers are out scope of this lab. The motto of this lab is "don't believe in what you have read, but in what you have verified yourself". Too many application developers only believe they know SQL and the DBMS which they use, without verifying the real behavior of the DBMS. Quoting Fekete et al[1] "All major database system products are delivered with default nonserializable isolation levels .. and we suspect that numerous isolation errors occur each day at many large sites because of this, leading to corrupt data sometimes noted in data warehouse applications". No developer should be allowed to develop applications which maintain data in a production database, unless the developer knows the transaction processing and concurrency disciplines of the DBMS system used by the application. The purpose of this lab is to help application developers, students, and their teachers in understanding the basic concurrency problems and the behavior of the modern RDBMS products, and how to verify this behavior, for example studying the locks in problematic situations.

[1] Fekete et al "Making Snapshot Isolation Serializable", at ACM Transactions on Database Systems. Vol. 30, No. 2, June2005, pages 492-528

This lab is based on the DBTech tutorial on "SQL Concurrency Technologies" available at http://www.DBTechNet.org/papers/SQL_ConcurrencyTechnologies.pdf
which covers the basic theory and concurrency control mechanisms used in the commercial mainstream RDBMS systems, from point of view of application developers and database professionals working in the ICT industry. So the lab will introduce relevant knowledge for use of concurrency services of the RDBMS, which knowledge is critical for establishing best practices and in developing reliable applications for use of industry.

The lab introduces single-connection flat SQL transactions focusing on transaction concurrency problems (anomalies) and solutions. The problems are tested using the concurrency control services of the real DBMS systems.

Pre-Requisites
Learners using this lab should know the basics of SQL and have some hands-on experience on using some RDBMS product. Recommended SQL tutorials in Web include

- SQL Tutorial of w3school at http://www.w3schools.com/sql/default.asp and
- Interactive SQL Tutorial of Andrew Cumming at http://SQLzoo.net

Recommended Reading
An easy reading on databases and SQL is the free e-Book "Database Fundamentals" by Neejar Sharma et al available from Web site of "DB2 on Campus" programme of IBM, although it does not cover transaction programming and concurrency.
Recommended theory-oriented database textbooks with some touch of DBMS products include the following

- T. Connolly and C. Begg, "Database Systems", 5th edition, Addison-Wesley, Oracle-oriented parts
- R. Ramakrishnan and J. Gehrke, "Database Management Systems", 3rd edition, McGraw-Hill, referring also to the mainstream DBMS products.

Learning Objective(s)
Learning objectives of this lab include

- Understanding the concept of an ACiD single-connection flat SQL transaction
- Understanding the basic concurrency problems of multi-user databases
- Understanding the basic concurrency control services of the mainstream DBMS products
- Knowledge of how to experiment with and monitor typical concurrency conflicts
- Knowledge of what means of SQL the DBMS product provides for application developer to use properly the concurrency control services of the DBMS product

Tutorial
The tutorial http://www.DBTechNet.org/papers/SQL_ConcurrencyTechnologies.pdf
can be used as course reading material or self-study material. It tries to bring some realism and order in the overloaded or loosely defined concepts on concurrency technologies presented in literature. The important chapter is the "Concurrency Technologies of Real DBMS systems".

Review Questions
The Review Questions are included in the tutorial

Hands-On Laboratory
We are focusing on ISO SQL standard and the SQL implementations in the mainstream DBMS systems used by industry nowadays. Luckily all mainstream DBMS vendors – IBM, Microsoft and Oracle - provide free Express editions of the products. With the exception of SQL Server, which only runs on Windows platforms, this has made it possible to create free and portable virtual laboratories which consists of a free operating system, typically a Linux variant, and free DBMS with its basic tools. For hands-on experimenting and verifying the presented topics, we provide free virtual computer image for VirtualBox having Debian Linux, DB2 Express-C 9.7 and Oracle 10g XE packed in ZIP archive at http://www.dbtechnet.org/download/VirtualBox_dbtech_debian.zip

For experiments with Microsoft SQL Server we recommend downloading the free SQL Server 2008 R2 Express with its free administration tools from the Web sites of www.Microsoft.com/sql

Sample Hands-On Exercises
Hands-on exercises for experimenting with the DBMS products like DB2 and Oracle are provided in document

CCLab1.pdf

Self-Assessment as MCQs
Multi-Choice Questions are provided in the document

MCQuestionsOnConcurrency.pdf

Materials

DBTech EXT VL (VLW) Checklist:	**CCLab1 files:**
Learner's guide	this document
Tutorial introduction material (in digital video and/or written text form)	Tutorial: SQL_ConcurrencyTechnologies.pdf Digital videos: not planned
Sample exercises	ReviewQuestions.pdf
Self-study, hands-on laboratory practicing material	CCLab1.pdf
Hands-on laboratory model	CCLab1_SQL.pdf
Self-assessment (MCQs)	CCLab1_MCQ.pdf
DBTech EXT (webropol) evaluation	(not during the project)
Recommended bibliography	see References part in the tutorial
Accompanying software with installation guide(s), and/or virtual machine with pre-installed software with user/learner guide(s)	VirtualBox_debian-dbtech.zip readMe.doc included