Education and Training

EACEA
Education, Audiovisual & Culture
Executive Agency

Education and Culture DG)G
Lifelong Learning Programme

Introducing the DBTech EXT Virtual Laboratories

With the support of the EC LLP Transversal Programme of the European Union

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

Databases are needed everywhere in the modern information society, building the basis of all business critical applications requiring reliable, secure and scalable data access. The lack of knowledgeable database professionals in the European labour market is already a productivity problem that has led to failed and delayed application development projects in the industry.

The DBTech EXT project of DBTechNet, the European database teacher network (www.DBTechNet.org), has extended the work of the successful LdV project DBTech Pro of years 2002-2005, now as part of the EU LLP Transversal Programme KAY3: ICT. The Project was scheduled for the time from 1 Jan 2009 to 31 Dec 2010. In the following we introduce part of the project deliverables.

To support students and database professionals already in industry, we have prepared **virtual laboratories** suitable as self-study laboratories or live exercises on database courses. Instead of learning just from "PaperWare", we believe in **learning by doing**, and especially **learning by verifying** using the real DBMS products, also used by the industry. Some of the virtual laboratories presented here have been organized as **virtual laboratory workshops** (VLW) in the **DBTech Portal** at http://dbtech.uom.gr

Public Deliverables of the DBTech EXT

(updated on 10 March 2011)

The **ref** numbers below are internal reference numbers of the deliverables in the DBTech EXT project plan.

DBTLab on Linux (ref 19)

To make at least part of the materials of the popular DBTech Pro live workshops available to public, to schools, individual students and database professionals in their own Life-Long Programmes, we provide downloadable private labs built of free or open source software and free professional, mainstream database systems, such as DB2 Express-C of IBM and Oracle XE of Oracle, available on multiple operating system platforms, including free Linux platforms.   SQL Server Express of Microsoft is also freely available, but since it is available only on the proprietary Windows platforms of Microsoft, we cover its features and examples in our tutorials, but users need to build the SQL Server labs of their own.

These are the current downloadable Database Labs:

- http://www.DBTechNet.org/download/VirtualBox_dbtech_debian.zip
  This is the latest virtual lab providing Java, Eclipse, DB2 Express-C 9.7, Oracle 10g XE, MySQL 5, with JBoss 6

- http://www.DBTechNet.org/download/VMware_Ubuntu_DB2Lab.zip
  DB2 Express-C V9.5 on Ubuntu 7.10 Linux in VMware virtual machine image prepared for us by IBM Finland. In the ZIP archive we have included also the free VMware Player for Windows, installing notes, learning materials provided by IBM, and the excellent SQL Cookbook of Graeme Birchall.
  The latest version of the SQL Cookbook is available on Graeme's home pages at
  http://mysite.verizon.net/Graeme_Birchall/
  The latest version of DB2 Express-C is available at IBM's web site
  http://www-01.ibm.com/software/data/db2/express/download.html

- http://www.DBTechNet.org/download/VMware_SUSE_DB2Lab.zip
  DB2 Express-C V9.7 on SUSE 10 Linux in VMware image based on the free virtual computer with DB2 V9.5 version available on IBM Web site. We have updated the DB2 Express-C version into V9.7 and installed there our Index Design Lab. For more details see our Index Design Lab tutorial below.

All DBTLabs can be used as private labs on experimenting with the appropriate examples and exercises of the following tutorials.

As a remote lab for experimenting with various SQL implementations we recommend http://SQLzoo.net

Virtual Laboratories covering topics of the DBTech Pro Framework
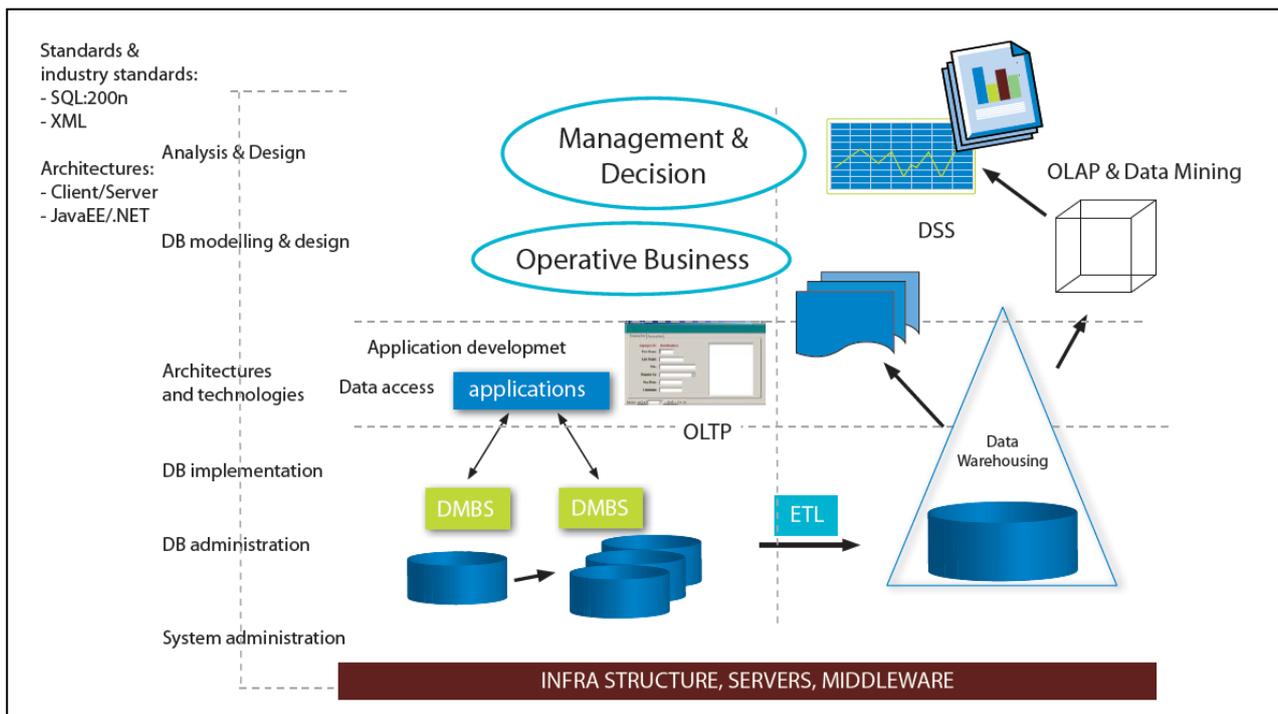
Concurrency Control (ref 22)

CCLab1 - Isolation Levels and Concurrency Technologies
SQL Concurrency Technologies tutorial presents overview on the concurrency topics covered in traditional textbooks, explaining the concurrency control concepts in terms of the SQL standard, and explaining how the real mainstream database systems, such as DB2, Oracle, and SQL Server, behave and what concurrency control services are available in context of SQL transaction processing and also context of cursor processing. A virtual laboratory based on this tutorial is available in DBTech Portal at http://dbtech.uom.gr

CCLab2 - TPCA2 Benchmark

In this lab, based on the old TPC-A benchmark of Transaction Processing Council (www.tpc.org), we will experiment with workload of concurrent transactions by multiple client sessions on accessing a simplified database of a bank application using the tutorial packed in http://www.dbtechnet.org/labs/ccr_lab/TPCA2.zip. We can experiment with

- different DBMS systems
- data access patterns of using transaction per database connection or using the same connection per client session (simulating the effect of connection pooling)
- different isolation levels
- data access pattern of accessing the server with separate SQL commands, or having the commands implemented as a stored procedure

The original TPC-A benchmark was defined by Jim Gray in 1985. We have adapted our TPCA benchmark from the book "Client /Server Programming with JAVA and CORBA" by Orfali and Harkey.

CCLab3 - Data Access Patterns and RVV Discipline

Database textbooks and product manuals often forget the real data access needs of applications on implementing user transactions.   A user transaction, based on use case requirements, will often need an interrelated sequence of SQL transactions, which may need to use different isolation levels. In the tutorial "RVV Paper" we will focus on use case transaction programming discipline of row version verifying, avoiding so called Blind Writes.

Database Administration and Tuning (ref 23)

Database Administration

Database implementations, security management, database backups and restore, monitoring and tuning operations are among the essential topics to understand and manage by a database administrator (DBA).   Our tutorial on these topics is available at http://www.DBTechNet.org/labs/dba_lab/DBALabs.pdf

Backup and Recovery (moved from ref 22)

are among the most important duties of a DBA. We present an overview of generic database server instance and its database(s) with transaction logs, principles of database backups and recovery in case of hardware or software crashes, or other failures. Hands-on exercises will need use of some local SQL Server instance. Tutorial is available at http://www.DBTechNet.org/labs/ccr_lab/RCLabs.pdf

Index Design

Based on the DBTech Pro **Index Design** workshop, which we created with Tapio Lahdenmäki at Malaga University in 2004, and the book "Relational Database Index Design and the Optimizers" by Tapio Lahdenmäki and Michael Leach (Wiley 2005), we have prepared the workshop in the form of self-study VLW. Tutorial part of this workshop is available at http://www.DBTechNet.org/labs/idp_lab/IDPLabs.pdf
and the VMware based lab using DB2 Express-C on SUSE platform can be downloaded from http://www.DBTechNet.org/download/VMware_SUSE_DB2Lab.zip

Business Intelligence (ref 24)

Business Intelligence (BI) & Knowledge Discovery from Databases (KDD)

An introduction to the practice of online analytical processing (OLAP), data warehousing, and knowledge discovery from databases (KDD). As outlined in the relevant DBTech EXT presentation, the BI & KDD virtual laboratory workshop (VLW) consists of two parts: (a) OLAP & Data Warehousing, and (b) Knowledge Discovery from Databases.
The OLAP & Data Warehousing part covers the ETL process, the data warehousing architectures and their implementations, OLAP operations, OLAP indexing, etc. The hands-on section of the VLW utilizes the PALO and Pentaho Mondrian OSS software.
The BI & KDD part focuses on the utilization of data mining techniques in the process of extracting information from large databases. After differentiating data from information, and after realizing the difference of applying SQL and data mining operations for information extraction purposes, the workshop participant learns about, and has the opportunity to apply in practice, a number of data mining techniques and algorithms (namely: association rules, decision trees, and clustering). In the course of the VLW, s/he also conducts the necessary data preparation operations, thus acquiring practical experience on the data (pre)processing stage which usually takes nearly 80% of the overall effort in the KDD process.

The hands-on section of the VLW involves the utilization of both free ([WEKA](#)) and proprietary ([IBM DB2 DWE Intelligent Miner](#)) software. In relation with the latter, the VLW elaborates on the IBM-developed "[Mining Your Business in Retail](#)" tutorial. A virtual laboratory based on these tutorials is available in DBTech Portal at http://dbtech.uom.gr

```
** KDD (association rules)*** :
```
http://www.dbtechnet.org/labs/kdd_lab/AssociationRules.pdf Association rule mining (theory)
http://www.dbtechnet.org/labs/kdd_lab/AnswersAssociationRulesWeka.pdf Association rule mining using Weka
In addition, the virtual VMware lab containing SUSE Linux, DB2 Express-C configured to work with Weka and some data. This and the instruction file (containing username/password) will be available at http://www.dbtechnet/download/

Business Intelligence Workshop
OLAP and Data Mining using SQL Server 2008 Analysis Services
To be available on downloadable virtual lab, which needs license activations (for example MSDN AA of the institution)

XML and Databases (ref 25)

XML and the Big Three
http://www.DBTechNet.org/labs/xml_lab/XMLandDatabasesTutorial.pdf
provides overview of the relevant XML technologies in the context of relational databases, the SQL/XML part of the SQL Standard, and the current XML implementations of the three mainstream database systems: DB2, Oracle and SQL Server. A virtual laboratory based on this tutorial is available in DBTech Portal at http://dbtech.uom.gr

Database Modeling and Semantics (ref 26)

Database Modeling
tutorial on the methodology of designing and implementing a relational database using UML data modeling by Rational Rose Data Modeler 7.0.0 at
http://http://www.DBTechNet.org/labs/dbm_lab/Database_Modeling.pdf
with exercises at
http://http://www.DBTechNet.org/labs/dbm_lab/Database_Modeling_lab1.pdf

Database Reverse Engineering
http://www.DBTechNet.org/labs/dre_lab/DRE_LAB_Theory_Slides.pdf
Database Modeling is a critical process in software engineering. In the industry Database Reverse Engineering is used for creating documentation on existing databases (e.g., for improvement and maintenance of legacy systems) and porting database structures between different DBMS environments.

The tutorial consists of an overview and hands-on exercises on using a Database Reverse Engineering tool to port a database structure from one environment to another, modify the database structure, and reverse engineer the database to its conceptual schema.

Semantic Modeling with RDF
Semantic modeling means that the concepts of a certain domain are described in a formal way. In this paper, we shall explain semantic modeling using RDF. The acronym RDF comes from the terms Resource Description Framework. RDF is a data model and a World Wide Web Consortium (W3C) specification. The syntax of RDF is often expressed using XML. However, because RDF is a data model, also other syntactic representations besides XML are possible.
http://www.dbtechnet.org/labs/sem_lab/SemanticModelingUsingRDF_Tutorial.pdf
http://www.dbtechnet.org/labs/sem_lab/SemanticModelingAunimo.pdf `(lecture slides)`
http://www.dbtechnet.org/labs/sem_lab/RDF_hands-on.pdf `(laboratory exercises)`

SPARQL and OWL
Tutorial introduction to semantic web and SPARQL and OWL specifications.
Sample demo on SPARQL using the Pyrrho DBMS implementation
see http://pyrrhodb.uws.ac.uk - Sample Code - SPARQL
http://www.dbtechnet.org/labs/sem_lab/SemanticWebAndDatabases.pdf (6.7 MB)
http://www.dbtechnet.org/labs/sem_lab/SemanticWebAndDatabases.rar (rar archived,1.5 MB)

Database Access Patterns and OR-mapping (ref 27)

The materials of these live workshops are available in DBTech Portal at http://dbtech.uom.gr

ORM - Object-Relational Mapping
Tutorial at
http://www.dbtechnet.org/labs/dae_lab/Orm.pdf
and instructions at
http://www.dbtechnet.org/labs/dae_lab/ORMenvironment.pdf

RVV Paper
www.DBTechNet.org/papers/RVV_Paper.pdf
The Row Version Verification (RVV) Discipline, also called as "optimistic locking" in older literature, extends the concurrency context to series of SQL transactions in a use case (a user transaction) for avoiding Blind Overwriting problems.

Distributed, Mobile, and Embedded Databases (ref 28)
The materials of these live workshops are available in DBTech Portal at http://dbtech.uom.gr

<< >>

Terms of Use and Responsibility of Authors
We have made our best to keep the downloadable labs and materials as reliable as possible. Since these materials are not commercial deliverables, and not aimed for production use, we will not accept any responsibility of potential problems raised on use of these materials. All materials are made as it is. In future we may setup discussion board were the authors may intervene the discussions for providing limited support and improving the materials.

Copyrights and Trademarks
We acknowledge all copyrights and trademarks, mentioned on these pages and materials, as copyrights or trademarks of their owners, such as Canonical Ltd., International Business Machines, Microsoft, Oracle, Sun, etc

For more information contact the coordinator of the project

HAAGA-HELIA University of Applied Sciences
att: DBTech / M Laiho
Ratapihantie 13
00520 HELSINKI, Finland
web : www.haaga-helia.fi