



Catanduanes State University
College of Information and Communications Technology
Virac, Catanduanes

LEARNING MATERIALS AND COMPILATION OF LECTURES/ACTIVITIES



CC105

INFORMATION MANAGEMENT

DISCLAIMER

This learning material is used in compliance with the flexible teaching-learning approach espoused by CHED in response to the pandemic that has globally affected educational institutions. Authors and publishers of the contents are well acknowledged. As such, the college and its faculty do not claim ownership of all sourced information. This learning material will solely be used for instructional purposes, not for commercialization.

College of Information and Communications Technology

CC105

Information Management

Chapter 2 : Database Query Languages

DISCLAIMER

This learning material is used in compliance with the flexible teaching-learning approach espoused by CHED in response to the pandemic that has globally affected educational institutions. Authors and publishers of the contents are well acknowledged. As such, the college and its faculty do not claim ownership of all sourced information. This learning material will solely be used for instructional purposes, not for commercialization.

College of Information and Communications Technology

OVERVIEW

This module introduces the Introduction to MySQL data manipulation how the performance tuning and optimization in creating database. It will also introduce the XQuery and Path, reports and query example and query optimization.

LEARNING OUTCOMES

At the end of the lesson, students shall be able to:

1. Write SQL queries for data manipulation and definition in database.
2. Understand the performance of tuning and optimizing in MySQL
3. Identify the XQuery and Path and their example
4. Recognize the functional dependencies and their relationship to keys of MySQL

KEY TERM

RDBMS
MySQL

Structured Query Language (SQL)
Extensible Markup Language (XML)

INTRODUCTION

MySQL is an open source relational database management system (RDBMS) with a client-server model. RDBMS is a software or service used to create and manage databases based on a relational model. MySQL is indeed not the only (R)DBMS on the market, but it is one of the most popular ones and only second to Oracle Database when scored using critical parameters like the number of mentions in search results, professional profiles on LinkedIn, and frequency of technical discussions on internet forums.

For the IT students they must know how to create basic information in creating database because they can use this for all the subjects that they are taken because even in some fields if you know to create a database it will be easily for you to manage all the information to your company or to your system.

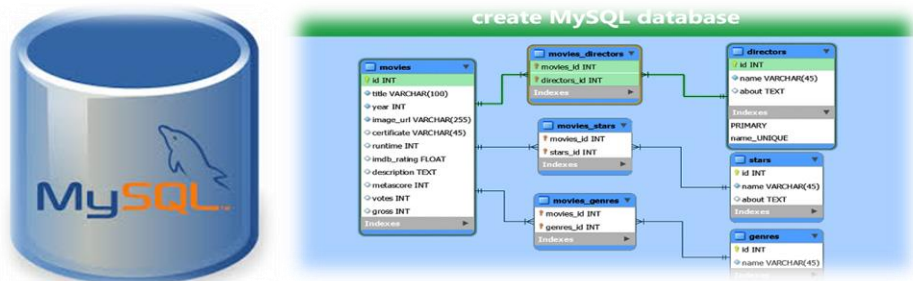
In this module, we will be focusing on developing DBMS Database Management System. You will understand the basic concepts and fundamentals of a database. Topics include:

1. Introduction to SQL data manipulation
2. SQL performance tuning/optimization
3. XQuery and XPath
4. Reports
5. Query by Example
6. Query optimization

LESSON 1

Introduction to MYSQL data manipulation

What is MySQL?



MySQL is an open source relational database management system (RDBMS) with a client-server model. RDBMS is a software or service used to create and manage databases based on a relational model.

MySQL was created by a Swedish company, MySQL AB, founded by David Axmark, Allan Larsson and Michael "Monty" Widenius. Original development of MySQL by Widenius and Axmark began in 1994. The first version of MySQL appeared on 23 May 1995. It was initially created for personal usage from mSQL based on the low-level language ISAM, which the creators considered too slow and inflexible. They created a new SQL interface, while keeping the same API as mSQL. By keeping the API consistent with the mSQL system, many developers were able to use MySQL instead of the proprietary licensed mSQL antecedent.

Database

A database is a place in which data is stored and organized. The word "relational" means that the data stored in the dataset is organized as tables. Every table relates in some ways. If the software doesn't support the relational data model, just call it DBMS.

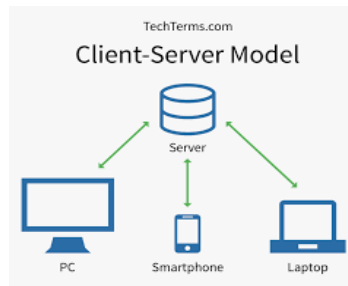


Open source



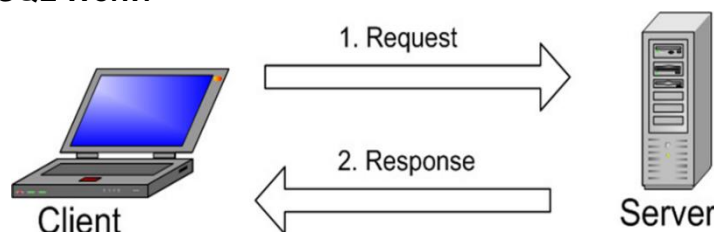
Open source means that you're free to use and modify it. Anybody can install the software. You can also learn and customize the source code to better accommodate your needs. However, The GPL (GNU Public License) determines what you can do depending on conditions. The commercially licensed version is available if you need more flexible ownership and advanced support.

Client-server model



Computers that install and run RDBMS software are called clients. Whenever they need to access data, they connect to the RDBMS server. That's the "client-server" part. MySQL is one of many RDBMS software options. RDBMS and MySQL are often thought to be the same because of MySQL's popularity. A few big web applications like Facebook, Twitter, YouTube, Google, and Yahoo! all use MySQL for data storage purposes. Even though it was initially created for limited usage, it is now compatible with many important computing platforms like Linux, macOS, Microsoft Windows, and Ubuntu.

How Does MySQL Work?



The image explains the basic structure of the client-server structure. One or more devices clients connect to a server through a specific network. Every client can make a request from the graphical user interface (GUI) on their screens, and the server will produce the desired output, as long as both ends understand the instruction. Without getting too technical, the main processes taking place in a MySQL environment are the same, which are:

1. MySQL creates a database for storing and manipulating data, defining the relationship of each table.
2. Clients can make requests by typing specific SQL statements on MySQL.
3. The server application will respond with the requested information and it will appear on the clients' side.

That's pretty much it. From the clients' side, they usually emphasize which MySQL GUI to use. The lighter and more user-friendly the GUI is, the faster and easier their data management activities will be. Some of the most popular MySQL GUIs are MySQL WorkBench, SequelPro, DBVisualizer, and the Navicat DB Admin Tool. Some of them are free, while some are commercial, some run exclusively for macOS, and some are compatible with major operating systems. Clients should choose the GUI depending on their needs. For web database management, including a WordPress site, the most obvious go-to is phpMyAdmin.

Why is MySQL so Popular?

341 systems in ranking, December 2018

Rank			DBMS	Database Model	Score		
Dec 2018	Nov 2018	Dec 2017			Dec 2018	Nov 2018	Dec 2017
1.	1.	1.	Oracle	Relational DBMS	1283.22	-17.89	-58.32
2.	2.	2.	MySQL	Relational DBMS	1161.25	+1.36	-156.82
3.	3.	3.	Microsoft SQL Server	Relational DBMS	1040.34	-11.21	-132.14
4.	4.	4.	PostgreSQL	Relational DBMS	460.64	+20.39	+75.21
5.	5.	5.	MongoDB	Document store	378.62	+9.14	+47.85
6.	6.	6.	IBM Db2	Relational DBMS	180.75	+0.87	-8.83
7.	7.	↑ 8.	Redis	Key-value store	146.83	+2.66	+23.59
8.	8.	↑ 10.	Elasticsearch	Search engine	144.70	+1.24	+24.92
9.	9.	↓ 7.	Microsoft Access	Relational DBMS	139.51	+1.08	+13.63
10.	10.	↑ 11.	SQLite	Relational DBMS	123.02	+0.31	+7.82

MySQL is indeed not the only (R)DBMS on the market, but it is one of the most popular ones and only second to Oracle Database when scored using critical parameters like the number of mentions in search results, professional profiles on LinkedIn, and frequency of technical discussions on internet forums. The fact that many major tech giants rely on it further solidifies the well-deserved position. Why so? Here are the reasons:

❖ Flexible and easy to use

You can modify the source code to meet your own expectations, and don't need to pay anything for this level of freedom, including the options for upgrading to the advanced commercial version. The installation process is relatively simple, and shouldn't take longer than 30 minutes.

❖ High performance

A wide array of cluster servers' backs MySQL. Whether you are storing massive amounts of big e-Commerce data or doing heavy business intelligence activities, MySQL can assist you smoothly with optimum speed.

❖ An industry standard

Industries have been using MySQL for years, which means that there are abundant resources for skilled developers. MySQL users can expect rapid development of the software and freelance experts willing to work for a smaller wage if they ever need them.

❖ Secure

Your data should be your primary concern when choosing the right RDBMS software. With its Access Privilege System and User Account Management, MySQL sets the security bar high. Host-based verification and password encryption are both available.

Download MySQL Installer

If you want to install MySQL on the Windows environment, using MySQL installer is the easiest way. MySQL installer provides you with an easy-to-use wizard that helps you to install MySQL with the following components:

- MySQL Server
- All Available Connectors
- MySQL Workbench with Sample Data Models
- MySQL Notifier
- Tools for Excel and Microsoft Visual Studio
- MySQL Sample Databases
- MySQL Documentation

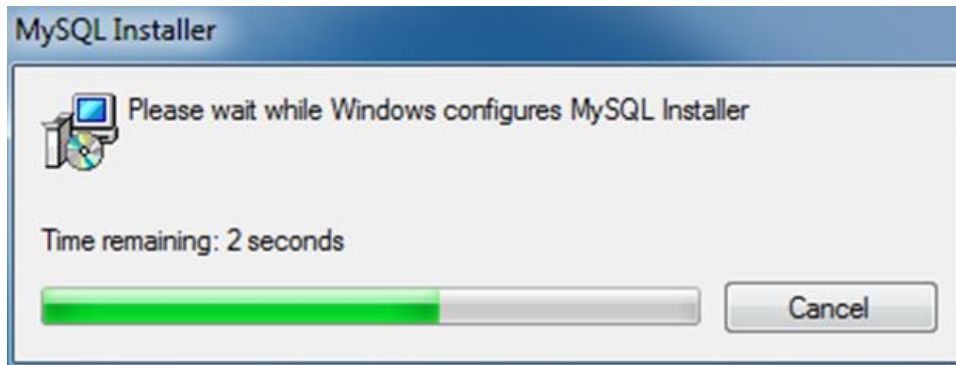
To download MySQL installer, go to the following link <http://dev.mysql.com/downloads/installer/>. There are two installer files:

- If you are connecting to the internet while installing MySQL, you can choose the online installation version `mysql-installer-web-community-<version>.exe`.
- In case you want to install MySQL offline, you can download the `mysql-installer-community-<version>.exe` file.

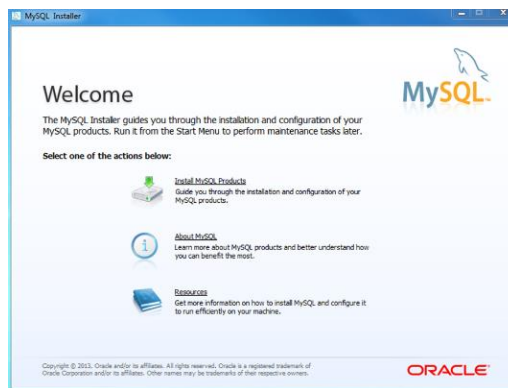
Install MySQL via MySQL Installer

To install MySQL using the MySQL installer, double-click on the MySQL installer file and follow the steps below:

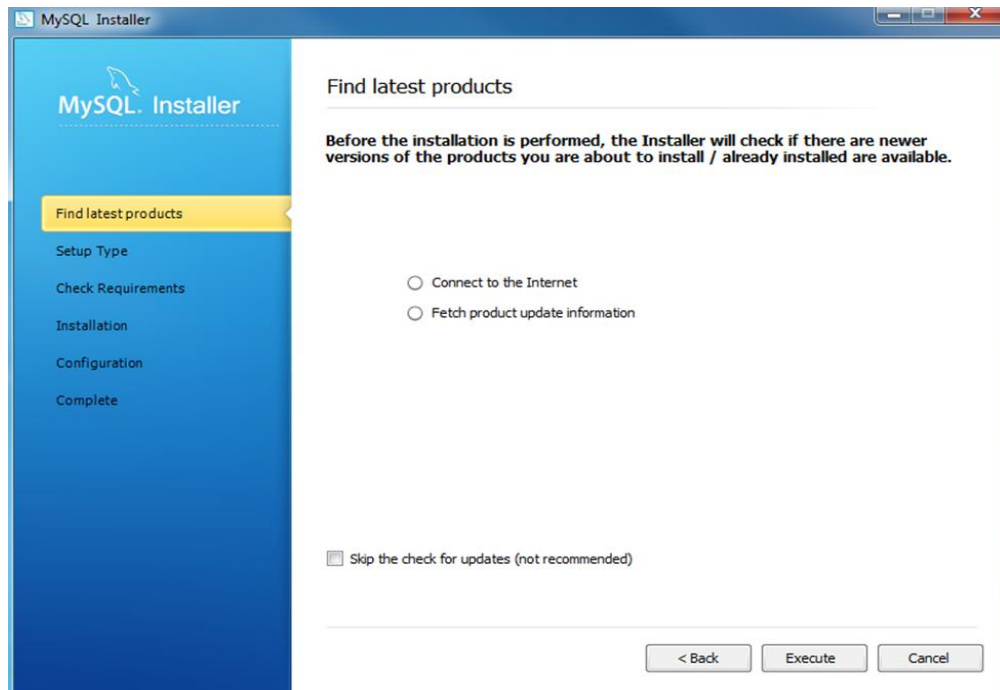
- ❖ Install MySQL Step 1: Windows configures MySQL Installer



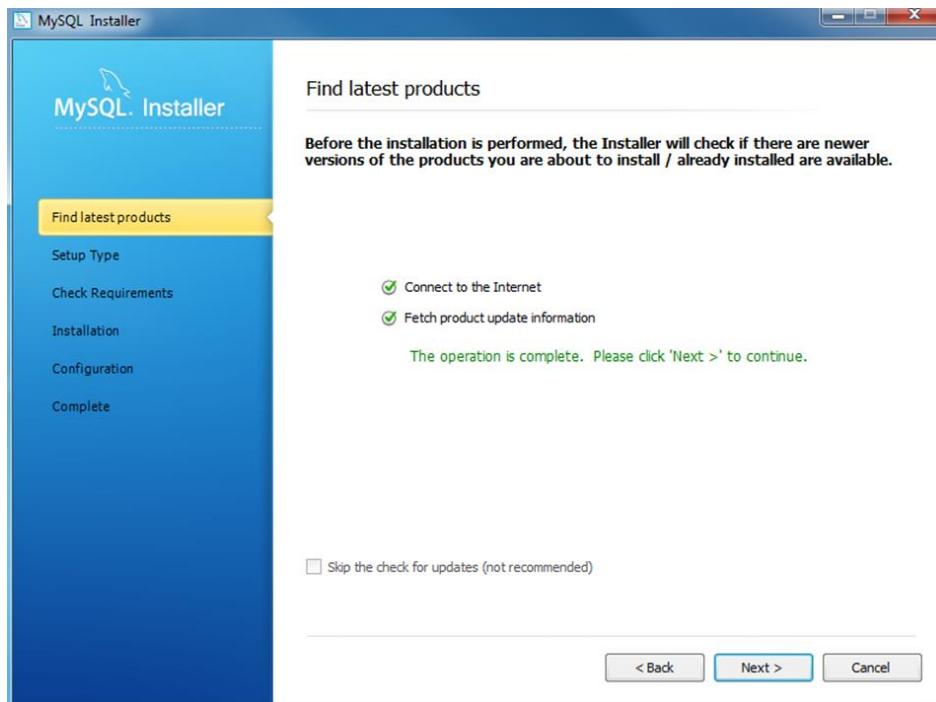
- ❖ Install MySQL Step 2 – Welcome Screen: A welcome screen provides several options. Choose the first option: Install MySQL Products



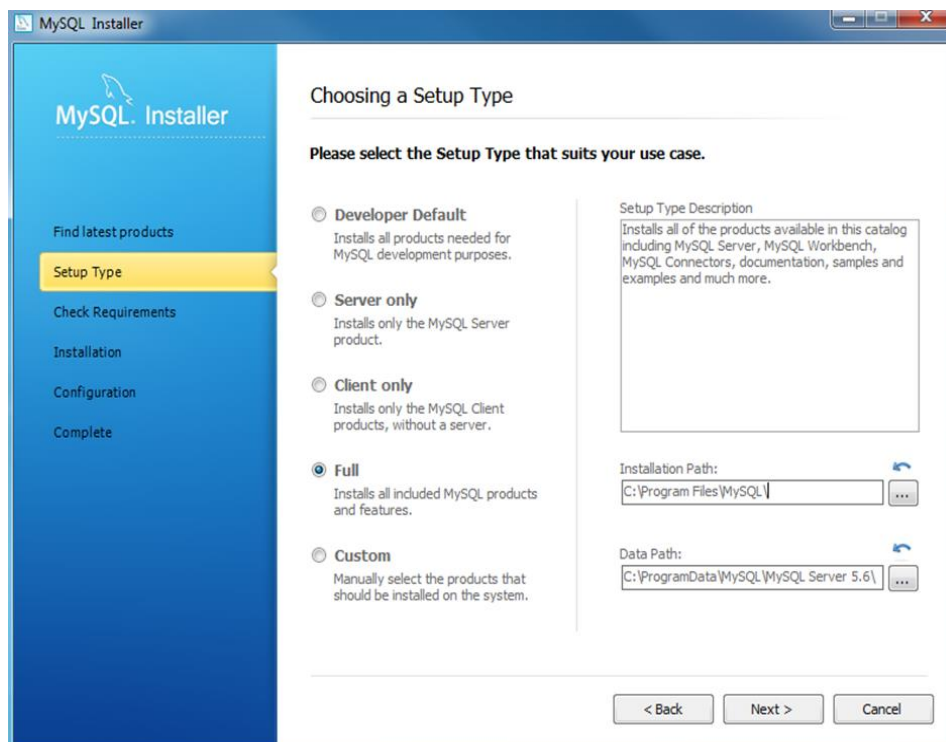
- ❖ Install MySQL Step 3 – Download the latest MySQL products: MySQL installer checks and downloads the latest MySQL products including MySQL server, MySQL Workbench



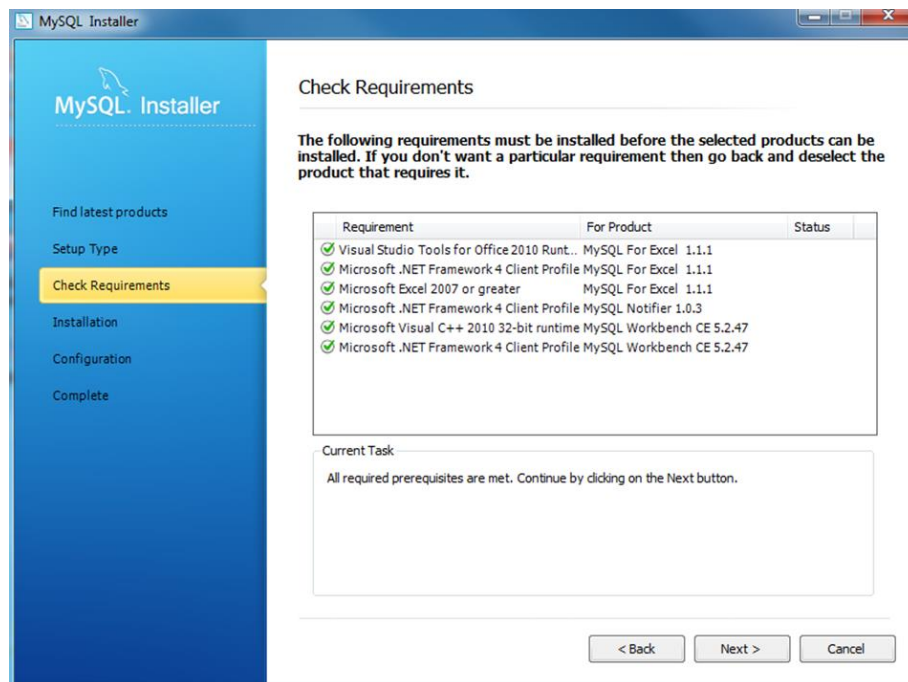
- ❖ Install MySQL Step 4: Click the Next button to continue



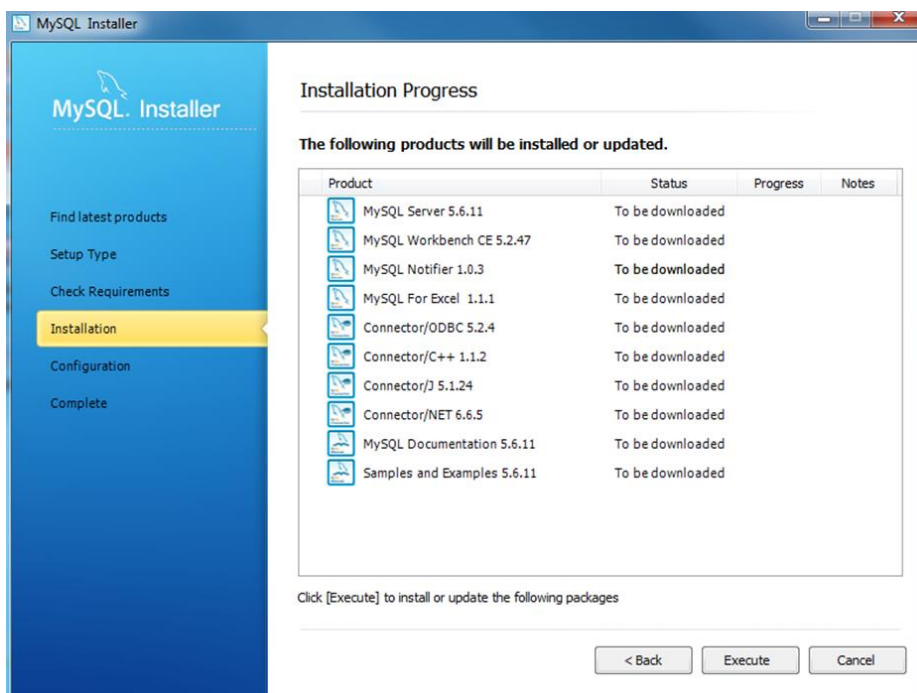
- ❖ Install MySQL Step 5 – Choosing a Setup Type: there are several setup types available. Choose the Full option to install all MySQL products and features.



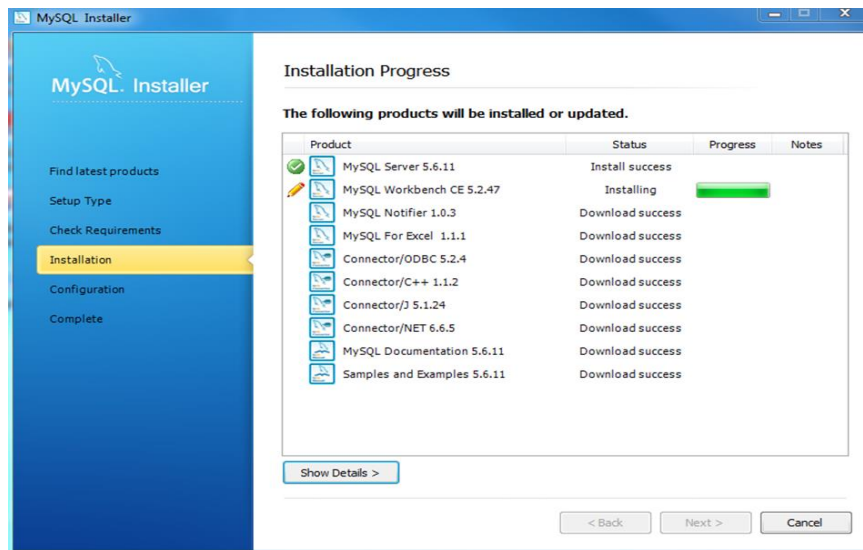
❖ Install MySQL Step 6 – Checking Requirements



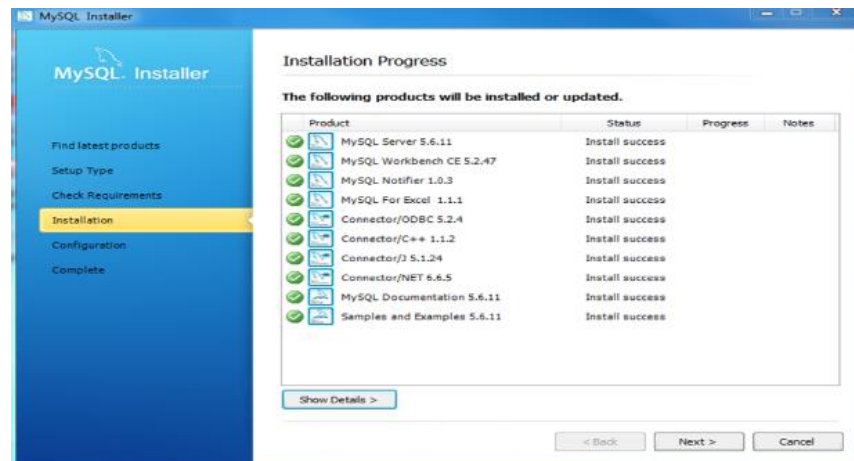
❖ Install MySQL Step 7 – Installation Progress: MySQL Installer downloads all selected products. It will take a while, depending on which products you selected and the speed of your internet connection.



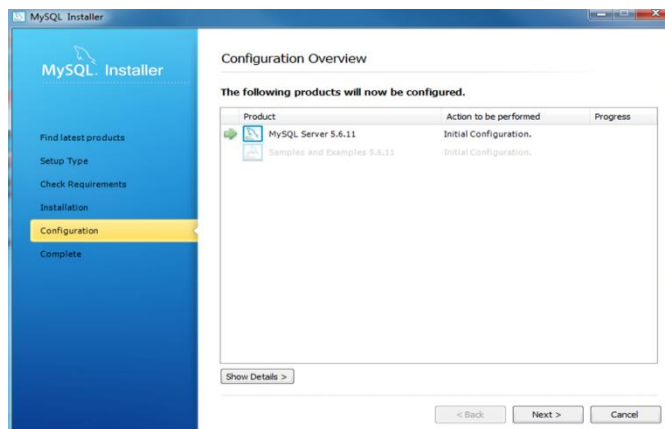
- ❖ Install MySQL Step 8– Installation Progress: downloading Products in progress.



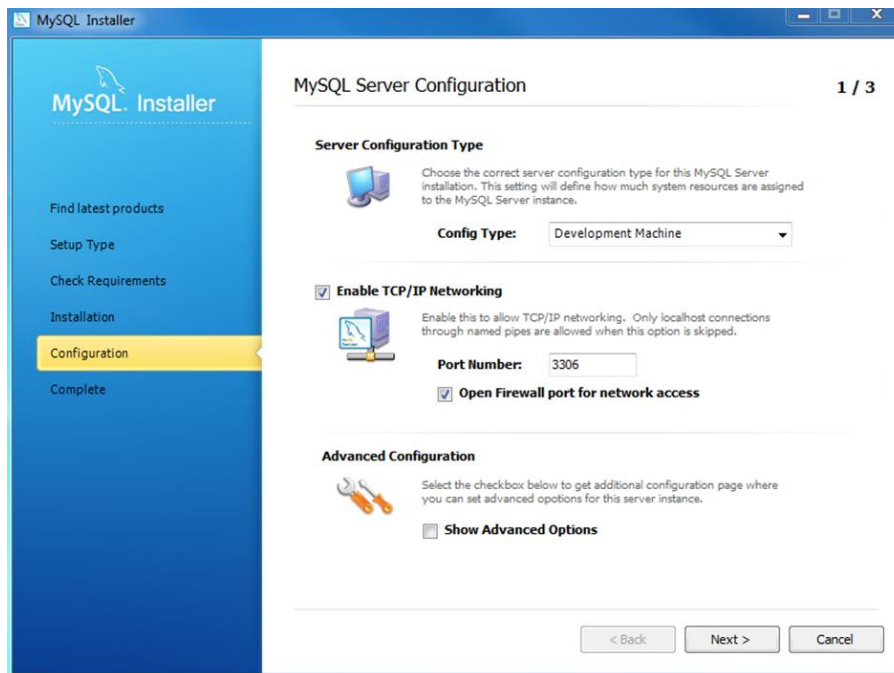
- ❖ Install MySQL Step 9 – Installation Progress: Complete Downloading. Click the Next button to continue.



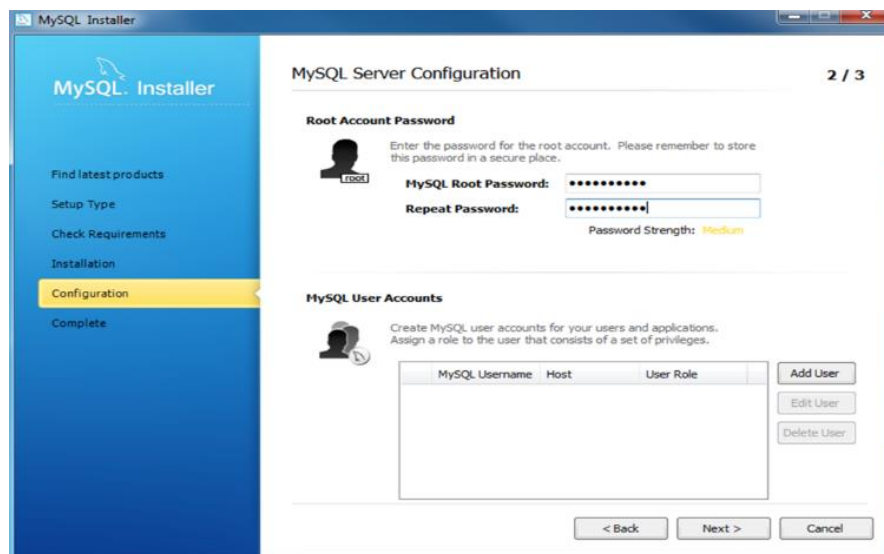
- ❖ Install MySQL Step 10 – Configuration Overview. Click the Next button to configure MySQL Database Server



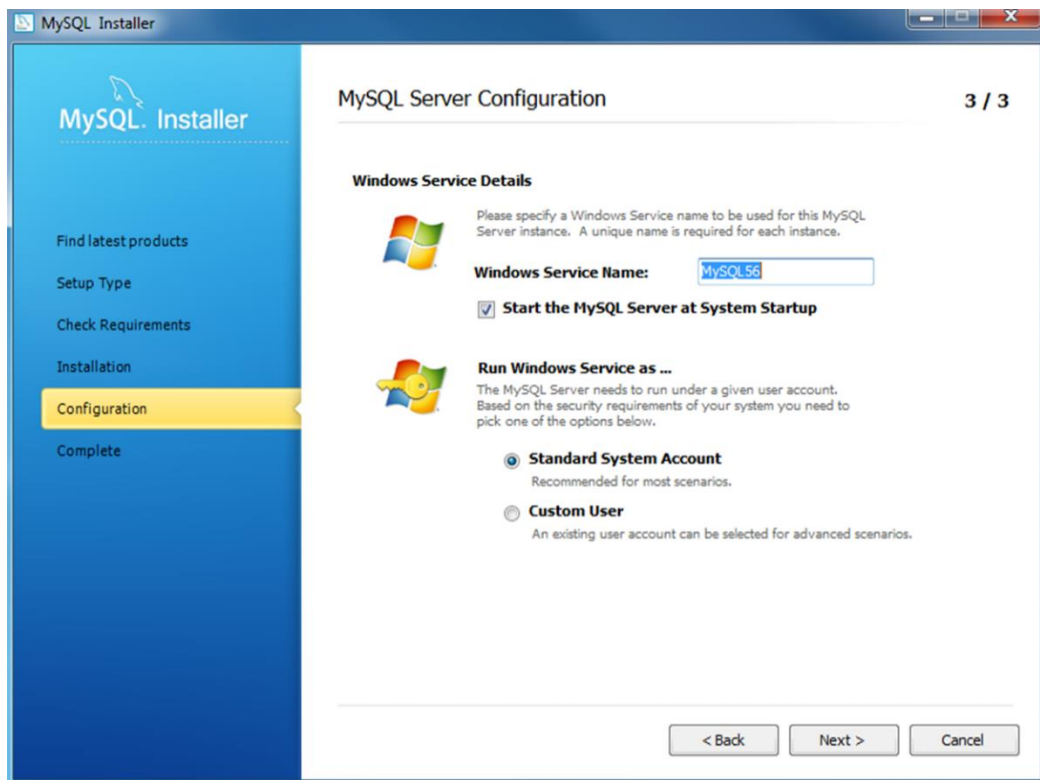
- ❖ Install MySQL Step 11 – MySQL Server Configuration: choose Config Type and MySQL port (3006 by default) and click Next button to continue.



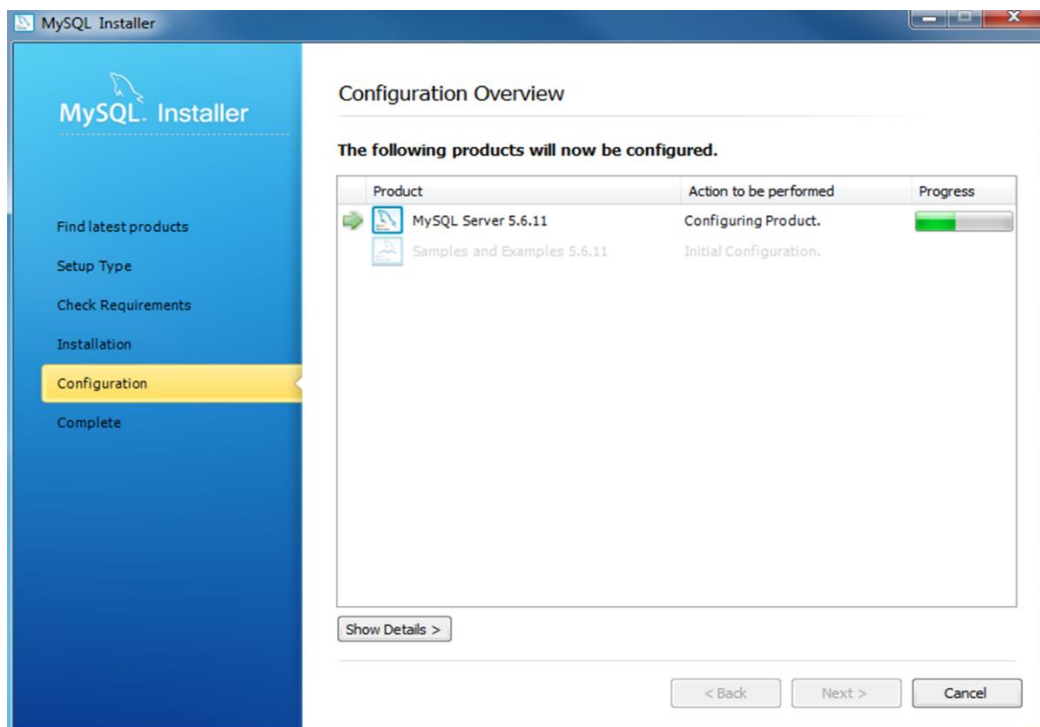
- ❖ Install MySQL Step 12 – MySQL Server Configuration: choose a password for the root account. Please note the password download and keep it securely if you are installing MySQL database server on a production server. If you want to add a more MySQL user, you can do it in this step.



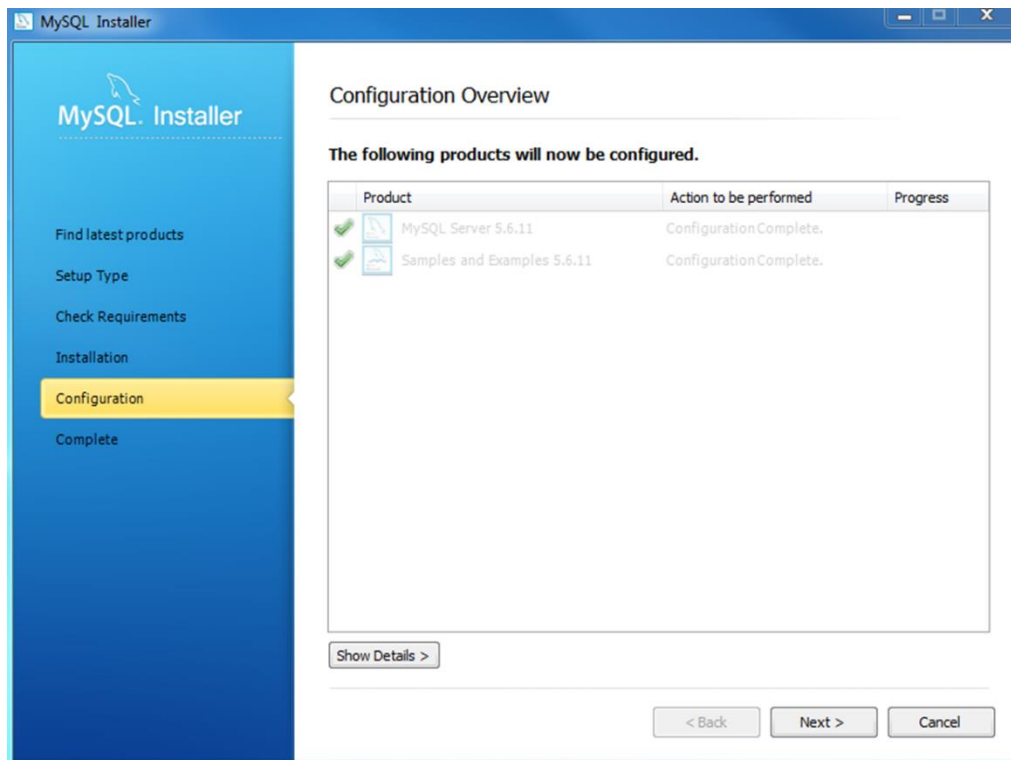
- ❖ Install MySQL Step 13 – MySQL Server Configuration: choose Windows service details including Windows Service Name and account type, then click Next button to continue.



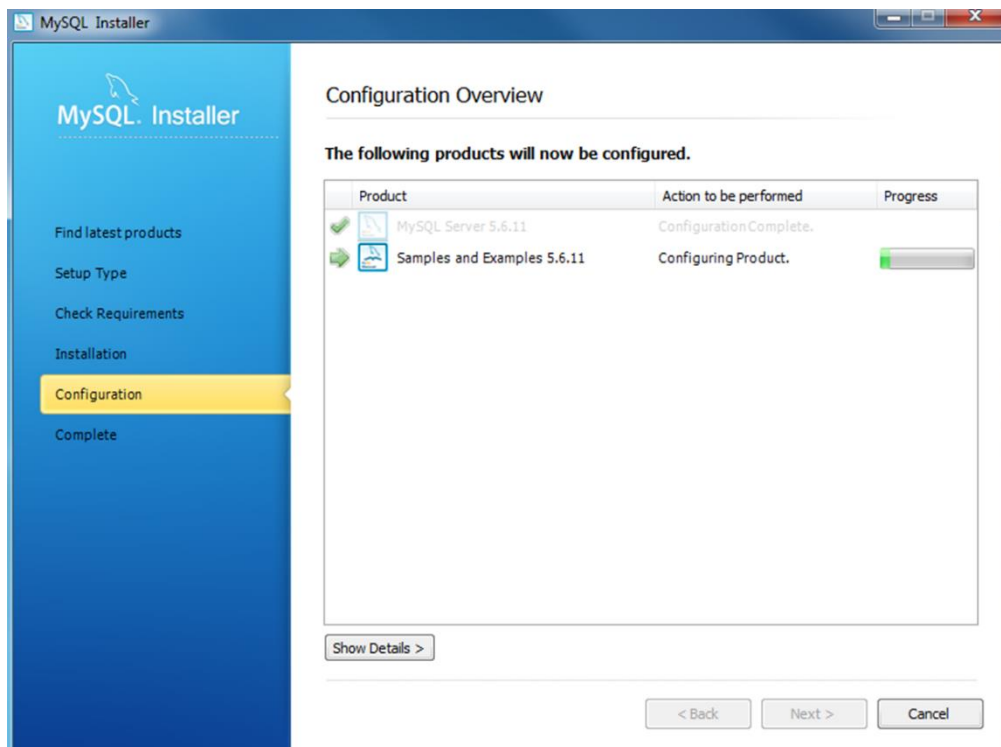
- ❖ Install MySQL Step 14 – MySQL Server Configuration – In Progress: MySQL Installer is configuring MySQL database server. Wait until it is done and click the Next button to continue.



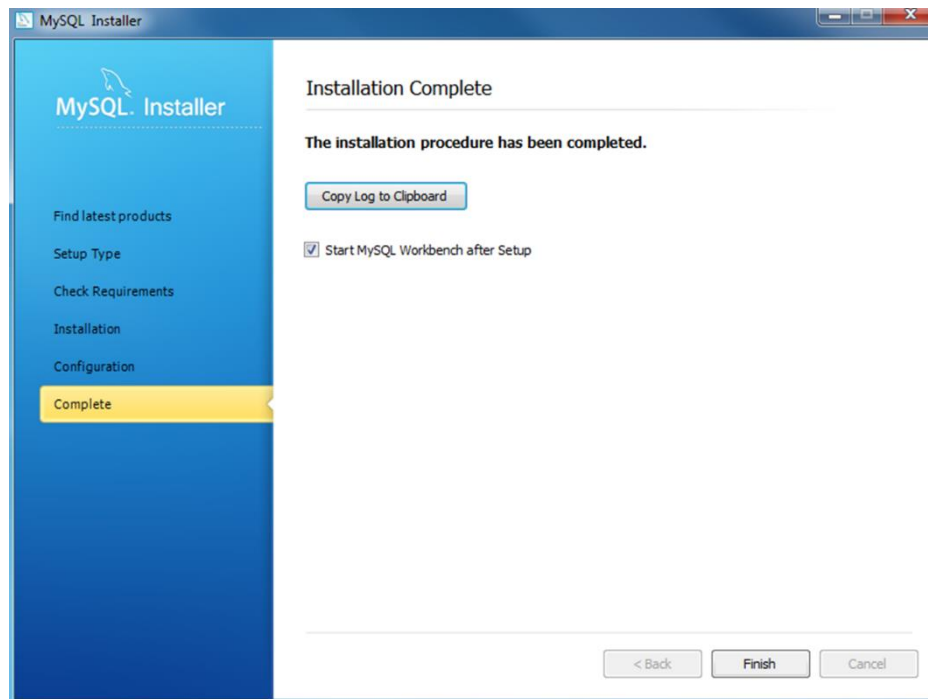
- ❖ Install MySQL Step 15 – MySQL Server Configuration – Done. Click the Next button to continue.



- ❖ Install MySQL Step 16 – Configuration Overview: MySQL Installer installs sample databases and sample models.



- ❖ Install MySQL Step 17 – Installation Completes: the installation completes. Click the Finish button to close the installation wizard and launch the MySQL Workbench



SUPPLEMENTARY LEARNING RESOURCES

What is MySQL: MySQL explained for beginners. (2020). Hostinger Tutorials. <https://www.hostinger.ph/tutorials/what-is-mysql>

Install MySQL on Windows using MySQL installer. (2020, April 11). MySQL Tutorial. <https://www.mysqltutorial.org/install-mysql/>

LESSON 2

MySQL Performance Tuning/Optimization

MySQL Performance Tuning and Optimization



MySQL performance tuning refers to a set of upkeep processes designed to maintain the efficiency, usefulness, and accuracy of databases.

The amount of information generated by end users each day means the amount and degree of information contained in a database can vary widely over a given period something potentially affecting database capacity and load, system performance, and availability of resources. By making MySQL database tuning a regular habit, network administrators can help ensure databases are properly provisioned and prepared to deal with the demands of changing workloads, expanding data volumes, and the need for maintained accessibility and service.

MySQL performance optimization can have positive effects to other departments in an organization. Prioritizing MySQL update performance tuning and optimization can help ensure the overall infrastructure meets the needs of the company without being too much of a drain on organizational resources and finances, ultimately helping save businesses money.

MySQL performance tuning tools help database professionals quickly identify bottlenecks, target insufficient operations through review of query execution plans and eliminate any guessing games.

The Benefits of MySQL Performance Tuning for database allows you to avoid over-provisioning and reducing cost by right sizing your servers. It also gives you insights into whether moving data storage or adding server capacity will bring improvement in performance or not and if so, then how much will it be.

MySQL Query Optimization Guidelines

❖ Avoid using functions in predicates

The database doesn't use index if it has some function predefined in the column.

For example: `SELECT * FROM TABLE1 WHERE UPPER(COL1)='ABC'`Copy

Because of the UPPER() function, the database doesn't utilize the index on COL1. If there isn't any way to avoid that function in MySQL, you will have to create a new function-based index or have to generate custom columns in database in order to improve performance.

❖ Avoid using wildcard (%) in the beginning of a predicate

The predicate LIKE '%abc' causes full table scan.

For example: `SELECT * FROM TABLE1 WHERE COL1 LIKE '%ABC'`Copy

In most cases this wildcard usage brings major performance limitation.

❖ Avoid unnecessary columns in SELECT clause

Instead of using 'SELECT *', always specify columns in the SELECT clause to improve MySQL performance. Because unnecessary columns cause additional load on the database, slowing down its performance as well whole systematic process.

❖ Use inner join, instead of outer join if possible

Use outer join only when it is necessary. Using it needlessly not only limits database performance but also limits MySQL query optimization options, resulting in slower execution of SQL statements.

❖ Use DISTINCT and UNION only if it is necessary

Using UNION and DISTINCT operators without any major purpose causes unwanted sorting and slowing down of SQL execution. Instead of UNION, using UNION ALL brings more efficiency in the process and improves MySQL performance more precisely.

❖ The ORDER BY clause is mandatory in SQL if you expect to get a sorted result

The ORDER BY keyword sorts the result-set in predefined statement columns. Though the statement brings advantage for the database admins for getting the sorted data, but it also produces bit performance impact in the SQL execution.

SUPPLEMENTARY LEARNING RESOURCES

What is MySQL: MySQL explained for beginners. (2020). Hostinger Tutorials. <https://www.hostinger.ph/tutorials/what-is-mysql>

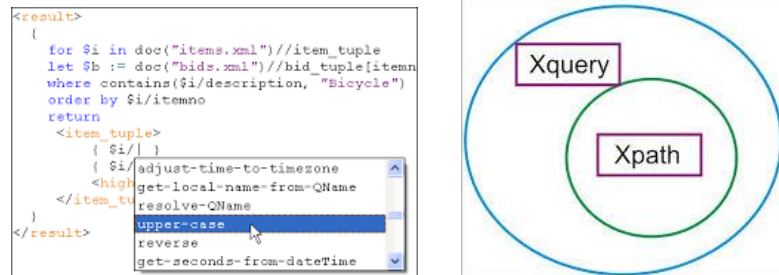
Install MySQL on Windows using MySQL installer. (2020). MySQL Tutorial. <https://www.mysqltutorial.org/install-mysql/>

MySQL performance tuning and database optimization tool. (n.d.). IT Management Software & Remote Monitoring Tools | SolarWinds. <https://www.solarwinds.com/database-performance-analyzer/use-cases/mysql-performance-tuning>

LESSON 3

XQuery and XPath

XQuery



A functional programming and query language that is used to query a group of XML(Extensible Markup Language) data. Also, XQuery is used to extract and manipulate data from either xml documents or relational databases and MS office documents that support an xml data source.

It represented in the form of a tree model with seven nodes, namely processing instructions, elements, document nodes, attributes, namespaces, text nodes, and comments.

Extensible Markup Language (XML) is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable.

All values are referred to as sequences. Even a single value is considered as a sequence of length one. The sequence can consist of either nodes or atomic values like integers, strings, or Booleans.

It has the following features that are used for the transformation of XML data:

- *Side effect free.*
- *Logical/physical data independence.*
- *Strongly typed.*
- *High level.*
- *Declarative.*

XPath

The XML Path Language that is used for selecting nodes from an XML document using queries. It can also compute values like strings, numbers, or Boolean type from another XML document. It is represented as a tree structure with the ability of XPath to navigate it by selecting different nodes.

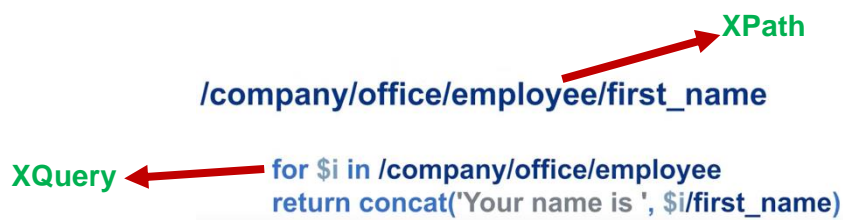
XPath has the following features:

- XPath defines the syntax for an XML document.
- It has the capability to navigate path expressions in XML documents.
- It has its own library defining standard functions.
- It is a major component of XSLT.

Differences between XPath and XQuery

- ❖ **XPath** is viewed as a regular expression whereas **XQuery** is like a C-programming language w.r.t. XML documents.
- ❖ **XPath** is a filter for an XML dataset and is the transformational component of XSLT. **XQuery** is used to select several nodes from an XML document for the purpose of processing using different queries.

XQuery and XPath Example



XPath
`/company/office/employee/first_name`

XQuery
`for $i in /company/office/employee
return concat("Your name is ", $i/first_name)`

XQuery has an expression called a FLWOR expression, which is similar to a SQL Select statement that has From and Where clauses. FLWOR is pronounced "flower", and is an acronym for the keywords used to introduce each clause for, let, where, order by, and return.

For - selects a sequence of nodes

Let - binds a sequence to a variable

Where - filters the nodes

Order by - sorts the nodes

Return - what to return (gets evaluated once for every node)

The XML Example Document

```

▼<bookstore>
  ▼<book category="cooking">
    <title lang="en">Everyday Italian</title>
    <author>Giada De Laurentiis</author>
    <year>2005</year>
    <price>30.00</price>
  </book>
  ▼<book category="children">
    <title lang="en">Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</year>
    <price>29.99</price>
  </book>
  ▼<book category="web">
    <title lang="en">XQuery Kick Start</title>
    <author>James McGovern</author>
    <author>Per Bothner</author>
    <author>Kurt Cagle</author>
    <author>James Linn</author>
    <author>Vaidyanathan Nagarajan</author>
    <year>2003</year>
    <price>49.99</price>
  </book>
  ▼<book category="web" cover="paperback">
    <title lang="en">Learning XML</title>
    <author>Erik T. Ray</author>
    <year>2003</year>
    <price>39.95</price>
  </book>
</bookstore>

```

We will use the "books.xml" document in the examples below

How to Select Nodes From "books.xml" With FLWOR

Look at the following path expression:

```
doc("books.xml")/bookstore/book[price>30]/title
```

The expression above will select all the title elements under the book elements that are under the bookstore element that have a price element with a value that is higher than 30.

The following FLWOR expression will select exactly the same as the path expression above:

```
for $x in doc("books.xml")/bookstore/book
where $x/price>30
return $x/title
```

The result will be:

```
<title lang="en">XQuery Kick Start</title>
<title lang="en">Learning XML</title>
```

With FLWOR you can sort the result:

```
for $x in doc("books.xml")/bookstore/book
where $x/price>30
order by $x/title
```

The **for** clause selects all book elements under the bookstore element into a variable called \$x. The **where** clause selects only book elements with a price element with a value greater than 30. The **order by** clause defines the sort-order. Will be sort by the title element. The **return** clause specifies what should be returned. Here it returns the title elements.

The result of the XQuery expression above will be:

```
<title lang="en">Learning XML</title>
<title lang="en">XQuery Kick Start</title>
```

SUPPLEMENTARY LEARNING RESOURCES

What is MySQL: MySQL explained for beginners. (2020). Hostinger Tutorials. <https://www.hostinger.ph/tutorials/what-is-mysql>

Difference between XQuery and XPath - *javatpoint.* (n.d.). [www.javatpoint.com. https://www.javatpoint.com/difference-between-xquery-and-xpath](https://www.javatpoint.com/difference-between-xquery-and-xpath)

LESSON 4

Reports

Database Reporting



A database report is a report created from a culmination of queried data visualized for the purposes of analysis, data discovery, and decision-making.

Database reports can be created through traditional BI platforms and embedded BI platforms through front-end calls to a backend database. These calls are often formulated as SQL queries.

A traditional BI platform may run a call to an internal data resource to then visualize a database report through a centralized front-end BI system.

Database reporting tools

Allow you to create reports based on the data stored in your database or data warehouse. Two of the most popular types of databases are relational and NoSQL.

Relational databases store data in tabular relations and they are managed through a database management system. Database reporting and database reporting tools rely on connections to a relational database management system (RDBMS), typically through JDBC or ODBC. The most popular types of relational database management systems are MySQL, Oracle, Microsoft SQL Server, IBM DB2 and PostgreSQL. Structured Query Language (SQL) is used by database reporting tools to query and manage data in relational database management systems. Once a connection is established to a RDBMS, database reporting tools then can present data in reports and dashboards.

On the other hand, if you require database reporting on large volume, unstructured, or real-time data you may need a big data storage solution. NoSQL databases are well-enabled at storing and retrieving big data which the tabular relations and structure found in relational databases simply cannot. Database reporting on NoSQL data sources requires either a customized SQL connector, ETL tool, or another native way of accessing data.

Database Reporting with Logi Report

Logi Report comes with additional reporting features which allow the representation of data to be taken several steps forward. Traditional reporting features like cascading parameters, which allows you to define restrictions on different data elements such as selecting a data range, as well as filtering and sorting capabilities which allow you to manipulate data within a given report, and fully interactive and ad hoc reporting capabilities allows you to do much more with a database report. Logi Report also allows you define row and column level security parameters meaning the output parameters are restricted based on user, role, and organization levels. These types of additional data discovery and analysis capabilities are often categorized in the larger umbrella term “business intelligence”.

Logi Report, a leader in BI reporting tools for embedded use cases, empowers you to access different database types and mash up multiple databases to create sophisticated, interactive reports and dashboards with ease. Logi Report’s embedding technology allows you to embed and white label reports, dashboards, and visualizations into any application whether it be an internal business portal or a full-fledged customer facing SaaS application. Logi Report supports many different types of data

sources, from relational databases, flat files, web services, big data sources, and even user-defined data sources.

Benefits of Database Reporting with Logi Report:

- **Enhance Your Applications Reporting-** Logi Report allows you to embed advanced reporting capabilities such as ad hoc reporting, report scheduling, report bursting, interactive reports, etc. so your end users can make quicker decisions.
- **Implement and Deploy Quickly-** When compared to building a reporting solution in your application from scratch or even using open source BI tools as a starting point, Logi Report can be embedded into your application quickly and seamlessly.
- **Flexibility, Control, and Customization-** Logi Report's full suite of JavaScript and Java APIs, along with its advanced white-labeling capabilities, allows you to fully customize not only the look and feel of your reporting capabilities, but all allow you to match the business logic already present in your current application UX. Our flexible security options also allow you to securely embed with ease.
- **Save Money, Time, and Effort-** The functional improvements seen from utilizing a 3rd party embedded BI platform is not the only advantages. Quicker time to market and less dedicated development resources also mean you can spend more time on the core capabilities of our products value proposition and spend less time working on the database reporting aspects of your suite.
- **Scale with Ease-** Another huge advantage of utilizing a 3rd party embeddable business intelligence platform is the ability to scale with ease. Logi Report's architecture is designed to scale up and down as you need it with load balancing and failover capabilities.
- **Future Proof Your Database Reporting-** Utilizing a third-party product like Logi Report also helps future-proof your application's database reporting functions because instead of having to spend additional development cycles creating new capabilities to keep up with your market, Logi Report does that for you.

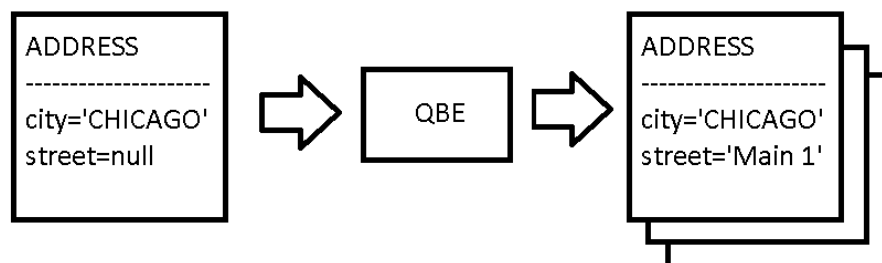
SUPPLEMENTARY LEARNING RESOURCES

Database reporting tools: Everything you need to know. (2017). *Logi Report*. <https://www.jinfonet.com/resources/bi-defined/database-reporting>

LESSON 5

Query by Example

Query by Example (QBE)



If we talk about normal queries we fire on the database they should be correct and in a well-defined structure which means they should follow a proper syntax if the syntax or query is wrong definitely we will get an error and due to that our application or calculation definitely going to stop. So, to overcome this problem QBE was introduced. QBE stands for Query by Example and it was developed in 1970 by Moshe Zloof at IBM.

It is a graphical query language where we get a user interface and then we fill some required fields to get our proper result.

In SQL we will get an error if the query is not correct but in the case of QBE if the query is wrong either we get a wrong answer or the query will not be going to execute but we will never get any error.

Note:

In QBE we don't write complete queries like SQL or other database languages it comes with some blank so we need to just fill that blanks and we will get our required result.

Example:

Consider the example where a table 'SAC' present in the database with Name, Phone_Number and Branch fields. And we want to get the name of SAC-Representative name who belongs to the MCA Branch. If we write this query in MySQL we have to write it like this.

```
SELECT NAME
FROM SAC
WHERE BRANCH = 'MCA' "
```

And definitely we will get our correct result. But in the case of QBE, it may be done as like there is a field present and we just need to fill it with "MCA" and then click on SEARCH button we will get our required result.

Points about QBE:

- Supported by most of the database programs.
- It is a Graphical Query Language.
- Created in parallel to SQL development.

SUPPLEMENTARY LEARNING RESOURCES

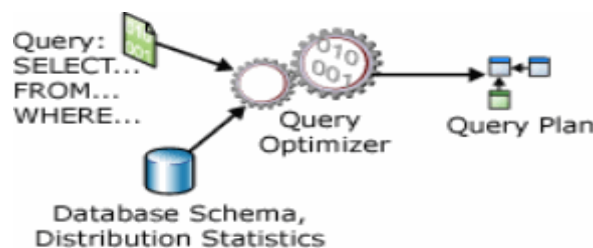
Query by example (QBE). (2020). GeeksforGeeks. <https://www.geeksforgeeks.org/query-by-example-qbe/>

(n.d.). YouTube. <https://www.youtube.com/watch?v=X9vyzpdUWHs>

LESSON 6

Query optimization

Query Optimization



The part of the query process in which the database system compares different query strategies and chooses the one with the least expected cost.

There are two approaches to optimization:

- ❖ **Cost based** - This was developed by IBM. The optimizer estimates the cost of each processing method of the query and chooses the one with the lowest estimate.
- ❖ **Heuristic** - Rules are based on the form of the query. Oracle used this at one point.

Query Optimization with MySQL

MySQL parses queries to create an internal structure (the parse tree), and then applies a variety of optimizations. These can include rewriting the query, determining the order in which it will read tables, choosing which indexes to use, and so on. You can pass hints to the optimizer through special keywords in the query, affecting its decision-making process. You can also ask the server to explain various aspects of optimization.

This lets you know what decisions the server is making and gives you a reference point for reworking queries, schemas, and settings to make everything run as efficiently as possible.

Here are some types of optimizations MySQL knows how to do:

- ❖ **Reordering joins**

Tables don't always have to be joined in the order you specify in the query. Determining the best join order is an important optimization.

- ❖ **Converting OUTER JOINs to INNER JOINs**

An OUTER JOIN doesn't necessarily have to be executed as an OUTER JOIN. Some factors, such as the WHERE clause and table schema, can actually cause an OUTER JOIN to be equivalent to an INNER JOIN. MySQL can recognize this and rewrite the join, which makes it eligible for reordering.

- ❖ **Applying algebraic equivalence rules**

MySQL applies algebraic transformations to simplify and order expressions. It can also fold and reduce constants, eliminating impossible constraints and constant conditions. For example, the term (5=5 AND a>5) will reduce to just a>5. Similarly, (a<b AND b=c) AND a=5 becomes b>5 AND b=c AND a=5. These rules are very useful for writing conditional queries.

- ❖ **COUNT(), MIN(), and MAX() optimizations**

Indexes and column nullability can often help MySQL optimize away these expressions. For example, to find the minimum value of a column that's leftmost in a B-Tree index, MySQL can just request the first row in the index. It can even do this in the query optimization stage, and treat the value as a constant for the rest of the query. Similarly, to find the maximum value in a B-Tree index, the server reads the last row. Likewise, COUNT(*) queries without a WHERE clause can often be optimized away on some storage engines (such as MyISAM, which keeps an exact count of rows in the table at all times).

- ❖ **Covering indexes**

MySQL can sometimes use an index to avoid reading row data, when the index contains all the columns the query needs it returns the data from index itself.

- ❖ **Subquery optimization**

MySQL can convert some types of subqueries into more efficient alternative forms, reducing them to index lookups instead of separate queries.

- ❖ **Early termination**

MySQL can stop processing a query (or a step in a query) as soon as it fulfills the query or step. The obvious case is a LIMIT clause, but there are several other kinds of early termination. For instance, if MySQL detects an impossible condition, it can abort the entire query. You can see this in the following example:

```
mysql> EXPLAIN select * from users where id = -1;
```

id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	NULL	NULL	NULL	NULL	NULL	NULL	NULL	Impossible WHERE noticed after reading const tables

1 row in set (0.05 sec)

This query stopped during the optimization step, but MySQL can also terminate the execution early in some other cases. The server can use this optimization when the query execution engine recognizes the need to retrieve distinct values, or to stop when a value doesn't exist.

❖ **IN() list comparisons**

In many database servers, IN() is just a synonym for multiple OR clauses, because the two are logically equivalent. Not so in MySQL, which sorts the values in the IN() list and uses a fast-binary search to see whether a value is in the list. This is $O(\log n)$ in the size of the list, whereas an equivalent series of OR clauses is $O(n)$ in the size of the list much slower for large lists.

SUPPLEMENTARY LEARNING RESOURCES

Srivastava, S. (2020, May 3). Query optimization with MySQL. Medium. <https://medium.com/@shashwat12june/mysql-query-optimizer-fc158b3de623>

REFERENCES

What is MySQL: MySQL explained for beginners. (2020). Hostinger Tutorials. <https://www.hostinger.ph/tutorials/what-is-mysql>

Install MySQL on Windows using MySQL installer. (2020). MySQL Tutorial. <https://www.mysqltutorial.org/install-mysql>

MySQL performance tuning and database optimization tool. (n.d.). IT Management Software & Remote Monitoring Tools | SolarWinds. <https://www.solarwinds.com/database-performance-analyzer/use-cases/mysql-performance-tuning>

Difference between XQuery and XPath - javatpoint. (n.d.). www.javatpoint.com. <https://www.javatpoint.com/difference-between-xquery-and-xpath>

Database reporting tools: Everything you need to know. (2017). Logi Report. <https://www.jinfonet.com/resources/bi-defined/database-reporting>

Query by example (QBE). (2020). GeeksforGeeks. <https://www.geeksforgeeks.org/query-by-example-qbe/>

(n.d.). YouTube. <https://www.youtube.com/watch?v=X9vyzpdUWHs>

Srivastava, S. (2020, May 3). Query optimization with MySQL. Medium. <https://medium.com/@shashwat12june/mysql-query-optimizer-fc158b3de623>