

DB2 Exam 610 Summary

Zeyuan Hu

December 30, 2015

Contents

1	Planning	3
1.1	Objectives	3
1.2	Database workloads	3
1.3	OLTP vs. Data Warehousing	3
1.4	DB2 pureScale - IBM solution for OLTP	4
1.5	InfoSphere Warehouse - IBM solution for Data warehousing	5
1.6	Notable DB2 features & products	5
1.7	DB2 offering	7
1.8	Large Objects (LOB)	7
1.9	XML data	8
2	Security	9
2.1	Objectives	9
2.2	Authentication	9

1 Planning

1.1 Objectives

- Knowledge of DB2 products (z/OS vs LUW vs pureScale - at a high-level; different products and what they do)
- Knowledge of database workloads (appropriate DB2 product to use - OLTP vs warehousing)
- Knowledge of non-relational data concepts (XML data, LOB data)

1.2 Database workloads

Two main types of database application workloads:

- online transactional processing (OLTP)
- data warehousing
 - reporting
 - online analytical processing (OLAP)
 - data mining applications
 - decision support

1.3 OLTP vs. Data Warehousing

An OLTP system is typical of a web order system, where you perform transactions over the web (such as ordering a product). Online transaction processing (OLTP) systems features:

- Support day-to-day, mission-critical business activities (ie. web-based order entry, stock trading) [*current data*]
- Support hundreds to thousands of users issuing millions of transactions per day against databases that vary in size [*Frequent updates, Granular transactions*]
- Response time requirements tend to be subsecond [*Sub-second response time*]
- Queries:
 - tend to be a mix of real-time insert, update, and delete operations against current-as opposed to historical-data
 - single-row lookups with logic that likely updates a small number of records

Data warehousing system typically consist of:

- Store and manage large volumes of data that is often historical in nature and is used primarily for analysis [*Voluminous historical data*]
- Optimized for queries

- Heterogeneous data sources
- Queries: (ie. [Summarized queries that perform aggregations and joins])
 - bulk load operations
 - short-running queries
 - long-running complex queries
 - random queries
 - occasional updates to data
 - execution of online utilities

Example 1. A database will be used primarily to identify sales patterns for products sold within the last three years and to summarize sales by region, on a quarterly basis. In case, a Data warehouse system is needed.

Remark. Different by *queries that are typically used to access the data (aka workloads)*.

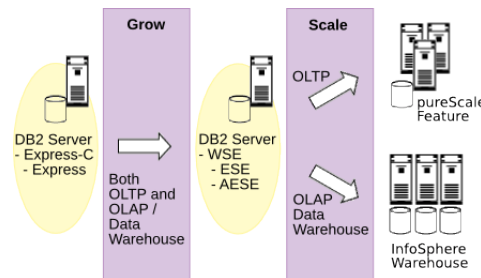


Figure 1: DB2 products, OLTP, Data Warehouse

1.4 DB2 pureScale - IBM solution for OLTP

System highlights:

- Best suited for OLTP workloads
- Enables a DB2 for LUW database to continuously process incoming requests, even if multiple system components fail simultaneously, which makes it ideal for OLTP workloads where high availability is crucial
- Provides a database cluster solution for nonmainframe platforms
- Can **ONLY** work with the General Parallel File System (GFPS) file system

Usage:

- Can be used with DB2 Workgroup Server Edition (WSE), DB2 Enterprise Server Edition (ESE), DB2 Advanced Enterprise Server Edition (AESE)

- Can **ONLY** be installed on IBM p Series or x Series servers that are running either the AIX (p Series) or the Linux (x Series) operating system
- **CANNOT** be installed on IBM mainframes running z/OS, IBM p Series server running Linux, or IBM x Series servers running Windows

1.5 InfoSphere Warehouse - IBM solution for Data warehousing

System highlights:

- is a complete data warehousing solution that contains components that facilitate data warehouse construction and administration, as well as tools that enable embedded data mining and multidimensional online analytical processing (OLAP)

1.6 Notable DB2 features & products

IBM Data Studio

- is an Eclipse-based, integrated development environment (IDE) that can be used to perform instance and database administration, routine (SQL procedure, SQL functions, etc.) and application development, and performance-tuning tasks.
- replaces the **DB2 Control Center** as the standard GUI tool for DB2 database administration and application development.
- allows users to connect to a DB2 database using a wizard; however, users are required to provide login credentials before a connection will be established.
- components:
 - **IBM Data Studio administration client**
 - * can be installed on servers running Red Hat Linux, SUSE Linux, and Windows
 - * **CANNOT** be installed on AIX servers
 - **IBM Data Studio full client**
 - * can be installed on servers running Red Hat Linux, SUSE Linux, and Windows
 - **IBM Data Studio web console**
 - * can be installed on servers running Red Hat Linux, SUSE Linux, and Windows
 - * can be installed on servers running the AIX operating system as well

IBM Workload Manager (WLM)

- is a comprehensive workload management feature that can help identify, manage, and control database workloads to maximize database server throughput and resource utilization
- customize execution environments for the purpose of controlling system resources so that no single workload can control and consume all of the system resources available. (This prevents any one department or service class from overwhelming the system.)

IBM InfoSphere Optim Performance Manager Extended Edition

- can be used to identify, diagnose, solve, and prevent performance problems in DB2 products and associated applications including Java and DB2 Call Level Interface (CLI) applications.

Self-Tuning Memory Manager (STMM)

- responds to significant changes in a database's workload by dynamically distributing available memory resources among several different database memory consumers

Connection Concentrator

- improves the performance of applications that require frequent, but relatively transient, simultaneous user connections by allocating host database resources only for the duration of an SQL transaction,

IBM InfoSphere Data Architect

- A complete solution for designing, modeling, discovering, relating, and standardizing data assets.
- You can use it for data modeling, transformation, and DDL generation, and to build, debug, and manage database objects such as SQL stored procedures and functions.

IBM InfoSphere Optim Query Tuner (Query Tuner)

- can analyze and make recommendations on ways to tune existing queries, as well as provide expert advice on writing new queries.

IBM InfoSphere Optim pureQuery Runtime

- Lets you deploy advanced pureQuery applications that use static SQL for a wide range of benefits.
- Bridges the gap between data and Java technology by harnessing the power of SQL within an easy-to-use Java data access platform.
- Increases security of Java applications helping to prevent threats like SQL injection.

DB2 for i

- combines with IBM BLU Acceleration to handle Analytical workloads
- formerly known as DB2 for i5/OS, is an advanced, 64-bit Relational Database Management System that leverages the high performance, virtualization, and energy efficiency features of IBM's Power Systems
- its self-managing attributes, security, and built-in analytical processing functions make DB2 for i an ideal database server for applications that are analytical in nature

DB2 pureXML

- offers a simple and efficient way to create a "hybrid" DB2 database that allows XML data to be stored in its native, hierarchical format.

Data Partitioning Feature (DPF)

- provides the ability to divide very large databases into multiple parts (known as partitions) and store them across a cluster of inexpensive servers.

1.7 DB2 offering

DB2 for z/OS

- full-function database management system that has been designed specifically for z/OS, IBM's flagship mainframe operating system.
- Tightly integrated with the IBM mainframe, **DB2 for z/OS** leverages the strengths of System z 64-bit architecture to provide, among other things, the ability to support complex data warehouse.

1.8 Large Objects (LOB)

LOB data types-**not LOB locators**-are used to store binary data values in a DB2 database.

- By default, LOB data is stored in separate LOB storage objects.
- Changes to LOB data are not recorded in transaction log files.

Inline LOBs

- improve query performance by storing LOB data in the same data pages as the rest of a table's rows, rather than in a separate LOB storage object. Thus, no additional I/O is needed to store and access this type of data.
- is eligible for compression.
- When a table contains columns with inline LOBs, fewer rows can fit on a page.
- transactions that modify inline LOB data are always logged. Consequently, the use of inline LOBs can **increase** logging overhead.
- are created by appending the **INLINE LENGTH** clause to a LOB column's definition.

LOB locator

- represents a value for a LOB resource that is stored in a database
- is a simple token value that is used to refer to a much bigger LOB value
- is a mechanism that refers to a LOB value from within a transaction
- is **NOT** a data type, nor is it a database object
- **do NOT** store copies of LOB data-they store a description of a base LOB value, and the actual data that a LOB locator refers to is only materialized when it is assigned to a specific location, such as an application host variable or another table record
- they behave as a snapshot of a piece of an LOB value, and not as a pointer to a row or a location in the database

1.9 XML data

- with `pureXML`, XML documents are stored in tables that contain one or more columns that have been defined with the XML data type.
- `CREATE TABLE employee (empid INT, resume XML)`

2 Security

2.1 Objectives

- Knowledge of restricting data access
- Knowledge of different privileges and authorities
- Given a DCL SQL statement, knowledge to identify results (grant/revoke/connect statements)
- Knowledge of Row and Column Access Control (RCAC)
- Knowledge of Roles and Trusted Contexts

Three levels of security:

- level 1: control access to the instance under which a database was created
- level 2: control access to the database itself
- level 3: control access to the data and data objects reside within the database

2.2 Authentication

- is the process by which a system verifies a user's identity.
- normally, an external facility (ie. OS, DCE) that is not part of DB2 performs the authentication.
- *authentication type* for a server is a database manager configuration parameter to decide how and where users are authenticated.

Type	Description
SERVER	→ Authentication occurs on the server
SERVER_ENCRYPT	→ Authentication occurs on the server → Passwords are encrypted at the client machine before being sent to the server
CLIENT	→ Authentication occurs at the client workstation, with no further checks on the server
KERBEROS	→ Authentication is performed by the Kerberos security software
KRB_SERVER_ENCRYPT	→ Authentication is performed by Kerberos security software if the client's authentication type is set to KERBEROS. Otherwise, SERVER_ENCRYPT is used
DATA_ENCRYPT	→ SERVER_ENCRYPT → user data are encrypted
DATA_ENCRYPT_CMP	→ DATA_ENCRYPT → Use SERVER_ENCRYPT if DATA_ENCRYPT not supported
GSSPLUGIN	→ Authentication is controlled by an external GSS-API plugin
GSS_SERVER_ENCRYPT	→ GSSPLUGININ → Use SERVER_ENCRYPT if GSSPLUGIN not supported