Relational Database Administration Notes

Chuks Okoli

18 December, 2023

Week 1

Overview of Database Management Tasks

Typical day for a database administrator:

- Checking the state of the database. Database should be running with no errors.
- Check that scheduled backups have completed and resolve issues.
- Responding to support tickets.
- Meeting with developers and other stakeholders.
- Monitoring database activity.
- Determine appropriate server resources needed for planning a new database.

Database Management Lifecycle
The database lifecycle is shown in **Fig. 1**.

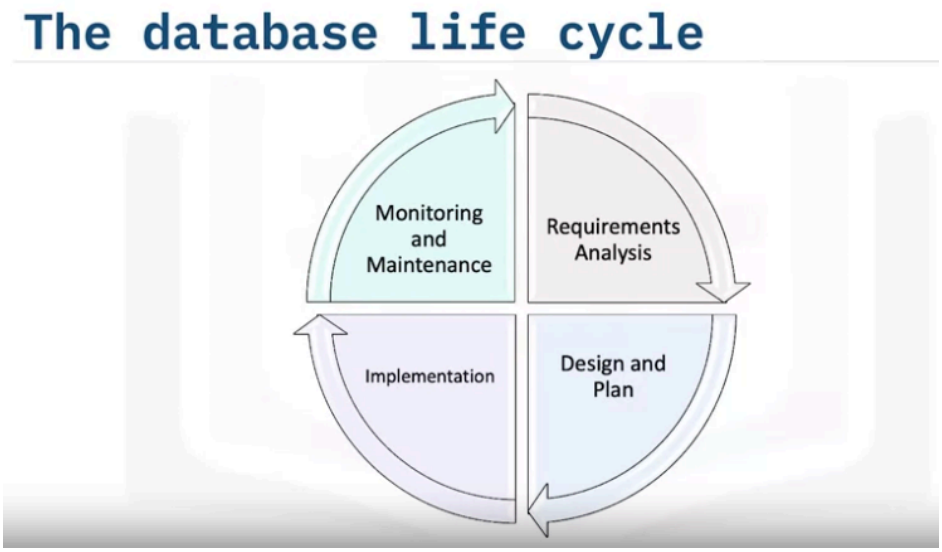

Figure 1: The database lifecycle

Requirement analysis

- Understand purpose and scope of the database
  - Establish what data is involved
  - Talk to the users and producers of the data, and develop samples of how users will use the data such as reports and dashboards

- Work with stakeholders such as Developers, Data Engineers, DBAs, Technology Managers, End-Users and determine requirements

Design and Plan

- Develop a plan for implementing the database by working with database objects such as instances, databases, tables, and indexes.
  - Develop database model showing which instances contains which databases and tables, how tables relate to each other, how users access data and so on.
  - Design Entity Relationship Diagrams or ERDs
- Determine appropriate server resources like storage space, memory, processing power, and log file size.
- Plan for how database objects are physically stored.

Implementation

- Create and configure database objects like instances, databases, tables, views, and indexes.
- Configuring database security, granting access for database users, groups, and roles
- Automates repeated database tasks such as backups, restores, and deployments to improve efficiency
- Import data from other databases, export data based on a query from a different source, or migrate projects from one environment to another, such as moving a project from the Application Development environment to the Production environment.

Monitoring and Maintenance

- Looks after the daily operations of the database
- Monitor the system for long-running queries and help end-users optimize them to run faster and not overuse system resources
- Review report and monitor activity, identify expensive queries, resource waits
- Apply upgrades and security patches to database software.
- Recommend and implement emerging database technologies
- Automate deployments and routine tasks such as backups whenever possible to keep processes working efficiently.
- Reviews logs and alerts, looking for failed logins and data access attempts to identify potential threats and vulnerabilities.
- Maintains database user and application permissions – revoking access to users and groups who should no longer have access and adding new users and roles as required to perform their jobs.

Server Objects and Hierarchy

Database Objects

- An instance is a logical boundary for a database or set of databases where you organize database objects and set configuration parameters.
- Common database objects are items that exist within the database such as tables, constraints, indexes, keys, views, aliases, triggers, events, and log files.
- Different RDBMSs use different names for their system objects. Most use the terms system schema, system tables, catalog, or directory.
- Database storage is managed through logical database objects and physical storage.

Week 2

Backup and Restore Databases

Backup and restore refer to the process of backing up data for protection purposes-restoring it after data loss from an unplanned shutdown, accidental deletion, or data corruption.

Backup and Restore scenarios

- Saving a copy of data for protection

- Recovering from data loss
  - After unplanned shutdown
  - Accidental deletion
  - Data corruption
- Move to a different database system
- Share data with business partners
- Use a copy of the data, e.g., in development or test environment

Physical vs. logical backups

- Logical backup
  - Contains DDL and DML commands to recreate database
  - Can reclaim wasted space
  - Slow and may impact performance
  - Granular
  - Backup/restore, import/export, dump & load utilities
- Physical backup
  - Copy of physical files, including logs, and configuration
  - Smaller and quicker
  - Less Granular
  - Can only restore to similar RDBMS
  - Common for specialized storage and cloud

What to backup

- Database
- Schema
- Tables
- Subset of data
- Other objects

Key considerations

- Check that your backup is valid
- Check that your restore plan works
- Ensure that your back up files are secure

Types of Backup

- Full backups - complete copy of all of the data in the object or objects that you are backing up
- Point-in-time backups - uses logged transactions to restore to an earlier point in time
- Differential backups - a copy of any data that has changed since the last full backup was taken
- Incremental backup - a copy of any data that has changed since the last backup of any type was taken

Backup Policies
Hot backup or online backup - are performed while the data is in use, while in Cold backup, the data is offline. Many decisions have to be made about backup policies such as:

- Physical or logical
- Full, differential, or incremental
- Hot or cold
- Compression
- Encryption
- Frequency:
  - Is data regularly changing or being added?
  - Is the existing table large?
- Schedule:
  - Is the data accessed equally across the 24-hour day?
  - Is it accessed at weekends?

- Automated backups
- Consider cloud backups

Summary

- The types of backups are full, point-in-time, differential, and incremental.
- The difference between physical backups and logical backups, and between hot backups and cold backups.
- Your backup policy should be determined from your recovery needs and your data usage.
- Database transaction logs keep track of all activities that change the database structure and record transactions that insert, update, or delete data in the database.

Week 3

Monitoring and Optimization

- A shell script is a program that begins with a 'shebang' directive and is used to run commands and programs. Scripting languages are interpreted rather than compiled.
- Filters are shell commands. The pipe operator '|' allows you to chain filter commands.
- Shell variables can be assigned values with '=' and listed using 'set.' Environment variables are shell variables with extended scope, and you can list them with 'env.'
- Metacharacters are special characters that have meaning to the shell.
- Quoting specifies whether the shell should interpret special characters as metacharacters or 'escape' them.
- Input/Output, or I/O redirection, refers to a set of features used for redirecting.
- You can use command substitution to replace a command with its output.
- Command line arguments provide a way to pass arguments to a shell script.
- In concurrent mode, multiple commands can run simultaneously.
- You can schedule cron jobs to run periodically at selected times. 'm h dom mon dow command' is the cron job syntax.
- You can edit cron jobs by running 'crontab -e,' and 'crontab -l' lists all cron jobs in the cron table.

Week 4

Troubleshooting and Automation