Data(Base) Administration

- Data Administration
- Database Administration
- Database Life Cicle: Administration responsabilities
- Data/Database Administration in the Cloud
- Other Data Management Roles

Changes in Traditional Roles

- This is being driven by rapid changes in
  - Technology
  - Platforms (e.g., Micro vs. Mainframe vs. Server vs. Cloud)
  - Organizational Structure
  - Database Models
  - Information requirements

- We will focus on the core functions and tasks of these roles (traditional or current)

Responsibilities

- Data Administration

  Responsibility for the overall management of data resources within an organization, including maintaining corporate-wide definitions and standards

- Database Administration

  Responsibility for physical database design and technical issues in database management such as security enforcement, database performance, and backup and recovery

- Data Warehouse Administration

  Responsibility for integration and coordination of metadata/data across many data sources: support decision applications, manage data warehouse growth, establish service level agreements regarding data warehouses and data marts, etc

- These roles combined or overlapping

Terms and Concepts

- DA
  - Data Administrator - person responsible for the Data Administration function in an organization
  - Sometimes may be the CIO or CDO
- DBA
  - Database Administrator - person responsible for the Database Administration Function
- DWA
  - Data Warehouse Administrator - person responsible for the Data Warehouse Administration Function

Database System Life Cycle
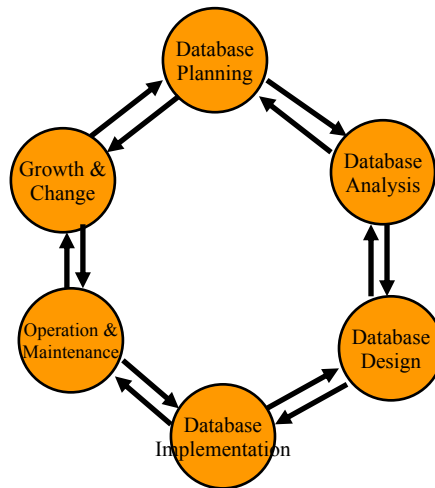
Database Planning

- Development of a strategic plan for database development that supports the overall organization's business plan
- DA supports top management in development of this plan
- The result of this stage is an *enterprise data model*

Database Planning: DA & DBA functions

- Develop corporate database strategy (DA)
- Develop enterprise model (DA)
- Develop cost/benefit models (DA)
- Design database environment (DA)
- Develop data administration plan (DA)

Database Analysis

- This is the process of identifying data entities currently used by the organization, precisely defining those entities and their relationships, and documenting the results in a form that can support the follow-on design phase
- Must also identify new data elements or changes to data elements that will be required in the future
- The result of this phase is the Conceptual Data Model -- represented as an ER or UML diagram

Database Analysis: DA & DBA functions

- Define and model data requirements (DA)
- Define and model business rules (DA)
- Define operational requirements (DA)
- Maintain corporate Data Dictionary (DA)

Database Design

- Purpose of the design phase is the development of the **logical database design** that will serve the needs of the organization and the **physical design** implementing the logical design
- In relational systems the outcome is normalized relations, and the data definition for a particular database systems (including indexes, etc.)

Physical Database Creation

- Development of the Physical Model of the Database
    - data formats and types
    - determination of indexes, etc.
- Load a prototype database and test
- Determine and implement security, privacy and access controls
- Determine and implement integrity constraints

Database Design: DA &DBA functions

- Perform logical database design (DA)
- Design external models (subschemas) (DBA)
- Design internal model (Physical design) (DBA)
- Design integrity controls (DBA)

Database Implementation

- Database design gives you an empty database
- Load data into the database structure
- Convert existing data sets and applications to use the new database
    - May need programs, conversion utilities to convert old data to new formats.
- Outcome is the actual database with its data

Database Implementation: DA & DBA functions

- Specify database access policies (DA & DBA)
- Establish Security controls (DBA)
- Supervise Database loading (DBA)
- Specify test procedures (DBA)
- Develop application programming standards (DBA)
- Establish procedures for backup and recovery (DBA)
- Conduct User training (DA & DBA)

Operation and Maintenance: Operation

- Users are responsible for updating the database, DA and DBA are responsible for developing procedures that ensure the integrity and security of the database during the update process.
- Specific responsibility for data collection, editing and verification must be assigned
- Quality assurance must be practiced to protect and audit the database quality.

Operation and Maintenance: Maintenance

- The ongoing process of updating the database to keep it current
  - adding new records
  - deleting obsolete records
  - changing data values in particular records
  - modifying relation structures (e.g. adding new fields)
- Privacy, security, access control must be in place.
- Recovery and Backup procedures must be established and used

Operation and Maintenance: DA & DBA functions

- Monitor database performance (DBA)
- Tune and reorganize databases (DBA)
- Enforce standards and procedures (DBA)
- Support users (DA & DBA)

Growth & Change

- Change is a way of life
  - Applications, data requirements, reports, etc. will all change as new needs and requirements are found
  - The Database and applications and will need to be modified to meet the needs of changes to the organization and the environment
  - Database performance should be monitored to maintain a high level of system performance

Growth & Change: DA & DBA functions

- Implement change control procedures (DA & DBA)
- Plan for growth and change (DA & DBA)
- Evaluate new technology (DA & DBA)

DBA in the Cloud

Cloud has shifted much of DBA duties from de Cloud Consumer to the Cloud Provider:

- Installation
- Backup
- Recovery
- Replication
- Stand by Servers
- Fail over Servers

How to maintain the enterprise DBA's presence unique as always?

DBA should acquire new skills !!!!!

DBA in the Cloud

DBA requires new requirements and skills.

DBA should concentrate more on business needs and to give support to enterpise how to use Cloud Plataform.

New Skills:

- Improve on NonSQL database usage and skills
- Capacity Planning in terms of On-Demand Usage Cost
- Multi Tennancy and Data Partitioning
- Monitoring and virtuality (tools)
- Data In and Data Out between Cloud and Data Centers

Other Data Management Roles

Data Analyst: Extracts data from business systems, analyses it and creates reports and dashboards to highlight trends and other business information for the users.

Data Modeler: Documents business data definitions and rules and creates conceptual, logical, and physical data models to translate them into database designs.

BI Engineer: Documents business needs on data analysis and reporting and develops BI systems, reports and dashboards to support them; also works to integrate data in BI systems.

Data Scientist: Gathers and analyses data from databases and other source systems, running machine learning algorithms and predictive models; creates data visualizations for business users.

Other Data Management Roles

Data Quality Engineer: Finds and cleanses bad data in databases; works with application developers and business users to create processes to prevent data quality problems.

Data Architect: Designs and implements database systems, data models and other data architecture components; also evaluates and recommends purchases of data management technologies.

Data Engineer: Creates data pipelines, big data platforms and data integrations in databases, data warehouses and data lakes, working with various cloud and on-premises technologies.

Database Administrator: Manages database systems, ensuring that they meet service-level agreements on availability, performance and cost and are properly secured and backed up.

Bibliography

- University of California, Berkeley. School of Information, Course of Database Management IS257, 2008.

- University of Surrey, CS263 Lecture 15.

- Mullins, C. S. *Database Administration: The Complete Guide to Practices and Procedures*. Addison Wesley, 2002.

- http://cloudcomputing.sys-con.com/node/1660119

- https://searchdatamanagement.techtarget.com/