JYOTI NIVAS COLLEGE
PG CENTER

DEPARTMENT OF MCA

I YEAR

TECH-ON-TOP

E-JOURNAL ON

DATABASE MANAGEMENT SYSTEMS

Issue 8, April 2020

INDEX

INSIGHT INTO TEMPORAL AND SPATIAL DATABASE

CHAITHRA R (19MCA04)

A database is collection of data which are related to each other. Everyone involves in day to day use of database, like for instance our brain requires some data to store in and retrieve (remember). Some of real world examples of database are mobile data backups, Google services like text searches, maps, Gmail etc. Temporal and Spatial database are also one of the types of database.

A **temporal database** stores time related data, it provides data types and stores information related to time i.e., past, present and future time. There are two types of time one is validate time-time period when a fact is true in real world and other is transaction time-time period when a fact stored in database was known and combination of both is a bi-temporal data. This type of database is used for insurance (accident and claims histories), banking (credit and debit histories), health care(stock prices) and scientific data(data about experiments) which are all time dependent data.

A **spatial database** is optimized for storing and querying of data related to objects in a geometric space. There are two types of data one is raster data-represented in the form of grid of pixels and other is the vector data-graphical representation of real world i.e., points, lines and polygons. A real world example is maps in which data stored in the form of points, lines and polygons for representing roads, areas, city, landmarks. Spatial database widely used in Geographical Information System (MAP),multimedia database(videos, audio, image), remote sensing, medical imaging etc.

Combinations of both temporal and spatial data are called as spatiotemporal database, which contains/manages space and time information. For data analysis and tracking of moving objects like for example rotation of earth and other several planets from satellites and storing it in a database.

REFERENCES:

https://www.docsity.com/it/introduction-to-spatial-databases-temporal-and-spatial-databases-lecture-10-slides-computer-science/57959/

HADOOP DATABASE

ABHILASHA D (19MCA01)

DEEPIKA N (19MCA06)

INTRODUCTION:

Hbase is called the Hadoop database because it has a No Sql. Database that runs on top of Hadoop Distributed File System (HDFS), it is well suited for sparse data sets, which is common in big data. It is a open source, distributed database developed by Apache Software foundations. At initial stage it was named Google Big Table, later it was re-named as Hbase and is primarily written in Java.

Hbase:

Hbase is a scalable big data store which can store massive amount of data from terabytes to petabytes in the form of table. It is built for low latency operations and is used extensively for read and write operation. It can handle structured, unstructured and semi-structured data. It groups rows into regions that define how data is split over multiple nodes in a cluster. If a region gets too large it is automatically split to share the load across more servers.

Hbase store data in the table like format with the ability to store millions of rows with the million of columns. Columns can be grouped together in a column family which allows physical distribution of row values on different cluster nodes. The data stored in h base does not need to fit into grade scheme like with an RDBMS making it ideal for storing unstructured or semi structured data. Hbase is architecture to have strongly consistent read and writes as opposed to other no SQL databases like Cassandra that are eventually consistent. Once a right has been perform all read request for the data will return the same value. Hbase table are replicated for failover. It was created to host when large tables for interactive and batch analytics, making it a great choice to store multiple structured for sparse data as we can use Apache has base when we need random real-time read or write access to big data h base is negatively integrated with Hadoop and can work seamlessly with other data access engines such as Apache spark Apache Hive and mapR database.

An h base column represents an attribute of an object if the table is storing diagnostic logs from serverbin to our environment each row might be log record and a typical column could be in a timestamp of when the log record was written or the server name where the record originated.

REFRENCES:

https://en.wikipedia.org/wiki/Apache_HBase

https://www.informit.com/articles/article.aspx?p=2253412

ORIENT DB

V SANTHIYA (19MCA16)

S JENIFER (19MCA07)

INTRODUCTION:

OrientDb is an open source and NoSql database system. It is multi model database which supports graph, document, keys and object model. The direct connection between records and graph database is managed. It supports schema-less, schema-full, schema mixed modes. The security profiling system is based on users and roles. There are several indexing mechanism based on B-tree, Extendible hashing and last one is known as hash index.

DESCRIPTION:

It is built with a multi-model graph/document engine. Orient DB feels like a graph database but there's no reason the key-value store can't be used on its own. Orient DB includes a SQL layer, the support for edges effectively means that these may be used to traverse relationships. It handles every record / document as an object and the linking between objects / documents is not through references, it's direct linking. This leads to quick retrieval of related data as compared to joins in an RDBMS.

FEATURES:

- Quick installation:It can be installed and running in less than 60 seconds.
- Fully transactional: Its supports ACID transactions guaranteeing that all database transactions are processed reliably and in the event of a crash all pending documents are recovered and committed.
- Graph structured data model: native management of graphs. Fully compliant with the Apache TinkerPop· open source graph computing framework.
- SQL: It supports SQL queries with extensions to handle relationships without SQL join, manage trees, and graphs of connected documents.
- Distributed: full support for multi-master replication including geographically distributed clusters.

APPLICATIONS:

- Banking
- Big data
- Fraud prevention
- Non-coding RN human interaction database
- Social networking

REFERENCES:

https://en.wikipedia.org/wiki/OrientDB

ORACLE AUTONOMOUS DATABASE

PRIYANKA RANI (19MCA12)

OVERVIEW OF ORACLE AUTONOMOUS DATABASE:

Oracle Autonomous Database is a cloud-based technology designed to automate many of the routine tasks required to manage Oracle databases, which Oracle says can free up database administrators (DBAs) to do higher-level and more strategic work.The Oracle Autonomous Database is the integration of the Oracle Database running on the Exadata platform with our complete infrastructure automation and our fully automated data center operations. Automated data sections include provisioning, patching, upgrading, and online backups, monitoring, scaling, diagnosing, performance tuning, optimizing, testing, and change management of complex applications and workloads, and automatically handling failures and errors.

FEATURES OF ORACLE AUTONOMOUS DATABASE:

- Because of its machine learning functionality, Oracle Autonomous Database can assimilate the information that it needs to take care of itself. For example, the autonomous software provisions databases on its own, finding, allocating and configuring all the necessary hardware and software for users.
- Oracle Autonomous Database also doesn't require manual tuning to optimize performance; the technology tunes itself, including automatic creation of database indexes to help improve application performance.
- It also automatically applies database updates and security patches, backs up databases and encrypts data to protect information against unauthorized access.
- The system patches itself on a regular quarterly schedule, although users can override this feature and reschedule the automatic patches if desired. Oracle Autonomous Database can also apply out-of-cycle security updates when necessary.
- Oracle Autonomous Database can scale itself up as needed; it also monitors capacity limits and bottlenecks in key system components in an effort to avoid performance problems

BENEFITS OF USING AUTONOMOUS DATABASE:

- Oracle Autonomous Database is likely to change the way that Oracle DBAs function in organizations that adopt the technology. Because many of the more mundane tasks that DBA’s now handle will be automated, Oracle says they'll be able to focus on things like data security, data lifecycle management, data architecture and data modeling.
- The technology could also minimize data loss and human error in Oracle databases, while cutting back substantially on both planned and unplanned downtime.

REFERENCES:

https://docs.oracle.com/en/cloud/paas/atp-cloud/index.html

https://docs.cloud.oracle.com/en-us/iaas/Content/Database/Concepts/adboverview.htm

https://searchoracle.techtarget.com/definition/Oracle-Autonomous-Database

CLOUD DATABASE

YOGAPRIYA P (19MCA21)

INTRODUCTION:

A **cloud database** is a database that typically runs on a cloud computing platform, and access to the database is provided as-a-service. Database services take care of scalability and high availability of the database. Database services make the underlying software-stack transparent to the user.

CLOUD DATABASE:

The design and development of typical systems utilize data management and relational databases as their key building blocks. Advanced queries expressed in SQL work well with the strict relationships that are imposed on information by relational databases. However, relational database technology was not initially designed or developed for use over distributed systems. This issue has been addressed with the addition of clustering enhancements to the relational databases, although some basic tasks require complex and expensive protocols, such as with data synchronization.

Modern relational databases have shown poor performance on data-intensive systems, therefore, the idea of NoSQL has been utilized within database management systems for cloud based systems. Within NoSQL implemented storage, there are no requirements for fixed table schemas, and the use of join operations is avoided. "The NoSQL databases have proven to provide efficient horizontal scalability, good performance, and ease of assembly into cloud applications." Data models relying on simplified relay algorithms have also been employed in data-intensive cloud mapping applications unique to virtual frameworks. It is also important to differentiate between cloud databases which are relational as opposed to non-relational or NoSQL.

Most database services offer web-based consoles, which the end user can use to provision and configure database instances. Database services consist of a database-manager component, which controls the underlying database instances using a service API. The service API is exposed to the end user, and permits users to perform maintenance and scaling operations on their database instances. Underlying software-stack stack typically includes the operating system, the database and third-party software used to manage the database. The service provider is responsible for installing, patching and updating the underlying software stack and ensuring the overall health and performance of the database.

Scalability features differ between vendors – some offer auto-scaling, others enable the user to scale up using an API, but do not scale automatically. There is typically a commitment for a certain level of high availability. This is achieved by replicating data and failing instances over to other database instances.

BEST OF CLOUD DATABASE:

Amazon Relational Database, Microsoft Azure SQL Database, Oracle Database, IBM Db2 on Cloud, Google Cloud SQL.

REFRENCE:

https://en.wikipedia.org/wiki/Cloud_database

POSTGRESQL

PRIYANKA S (19MCA13)

PostgreSQL also referred as Postgres is a general purpose object relational database management system and the most advanced open source system, which was designed to run on UNIX like platform. It was developed based on PostgreSQL 4.2 at Berkeley Computer Science Department, University of California. PostgreSQL is the first database management system before Oracle, that implements multi version concurrency control (MVCC). It is designed to be extensible. In this we can define our own data type, index, functional languages. If we don't like any part of the system, we can develop a custom plugin to enhance it to meet our requirements. It allows us to add custom functions which are developed using programming languages like C, C++, Java etc. It is an open source relational database management system (DBMS) which is developed by volunteers worldwide and controlled by any corporation or private entities. One of the main reasons for PostgreSQL to gradually take over MySQL users is that it is the only free and open-source RDBMS. It requires minimum maintenance efforts because of its stability, thus if we develop application based on PostgreSQL, the cost of ownership will be low compared to other database management system. Postgres document and key-value store technologies provide the advantages of NoSQL technologies plus the ability to integrate them into traditional corporate data infrastructures without operational complexity, loss of data integrity, and risk. Postgres has combined the leading JSON (JavaScript Object Notation) implementation with its proven ACID (Atomicity, Consistency, Isolation, Durability) transaction engine. This unique combination gives users industry leading workload flexibility on the most cost effective and industry proven platform.

Many companies like Apple, Red Hat, Cisco, Jupiter Network etc. have built products and solutions using PostgreSQL.

PostgreSQL has many advanced features likeuser-defined type, table inheritance, sophisticated locking mechanism, foreign key referential integrity, nested transactions (savepoints) and asynchronous replication.

The recent version of PostgreSQL supports features like native Microsoft windows server version, tablespaces, point-in-time recovery etc.

PostgreSQL offers many advantages such as diverse community, open source DBMS, ACID and transaction, diverse kinds of replication, diversified extension functions.

REFERENCE:

https://www.educba.com/sql-server-vs-postgresql/

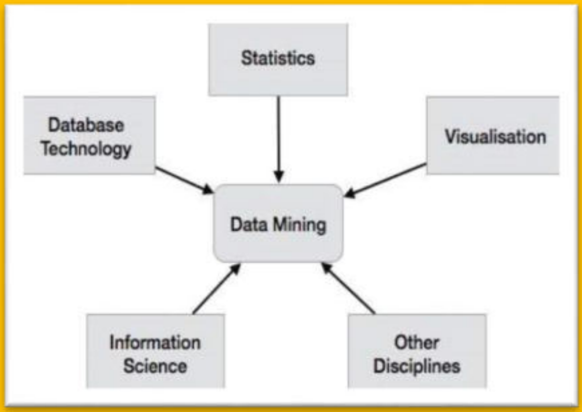
https://www.postgresqltutorial.com/

DATA MINING

NISHANTHI K (19MCA09)

INTRODUCTION OF DATA MINING:

Data mining is looking for hidden, valid, and potentially useful patterns in huge data sets. Data Mining is all about discovering unsuspected/ previously unknown relationships amongst the data. It is a multi-disciplinary skill that uses machine learning, statistics, AI and database technology. The insights derived via Data Mining can be used for marketing, fraud detection, and scientific discovery, etc,. Data mining is also called knowledge discovery, Knowledge extraction, data/pattern analysis, information harvesting, etc.



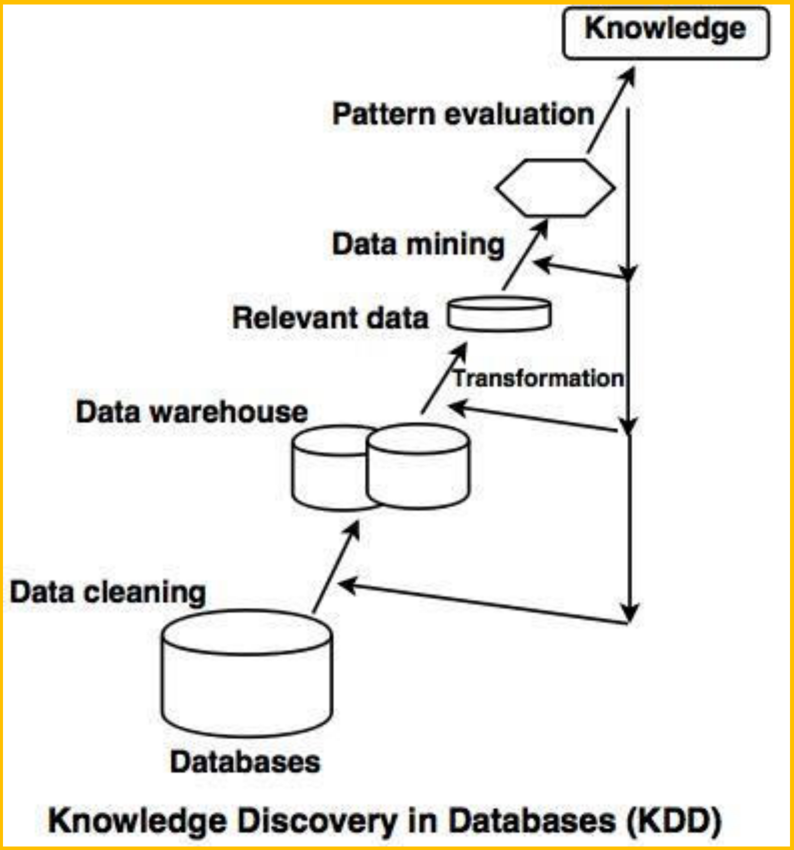
What is Data mining?

- Data mining is the process of extracting the useful information stored in the large database.
- It is the extraction of hidden predictive information.
- Data Mining is the practice of automatically searching the large stores of data to discover patterns.
- Data Mart is a powerful new technology with great potential that helps organization to focus on the most important information in their data warehouse.
- It uses mathematical algorithms to segment the data and evaluates the probability of future events.
- Data mining is a powerful tool used to retrieve the useful information from available data warehouses.
- Data mining can be applied to relational databases, object-oriented databases, data warehouses, structured-unstructured databases etc.
- Data mining is also known as **Knowledge Discovery in Databases (KDD).**

TYPES OF DATA:

Data mining can be performed on following types of data

- Relational databases
- Data warehouses
- Advanced DB and information repositories
- Object-oriented and object-relational databases
- Transactional and Spatial databases
- Heterogeneous and legacy databases
- Multimedia and streaming database
- Text databases
- Text mining and Web mining



Different steps of KDD as per the above diagram are:
1. Data cleaning removes irrelevant data from the database.
2. Data integration: The heterogeneous data sources are merged into a single data source.
3. Data selection retrieves the relevant data to the analysis process from the database.

4. Data transformation: The selected data is transformed in forms which are suitable for data mining.
5. Data mining: The various techniques are applied to extract the data patterns.
6. Pattern evaluation evaluates different data patterns.
7. Knowledge representation: This is the final step of KDD which represents the knowledge.

REFERENCES:

https://www.tutorialride.com/dbms/data-mining.htm.

https://www.geeksforgeeks.org/data-mining/

MONGO DB

PRAJWALA S REDDY (19MCA11)

INTRODUCTION:

MongoDB is a NoSQL database which stores the data in form of key-value pairs. It is an Open Source, Document Database which provides high performance and scalability along with data modelling and data management of huge sets of data in an enterprise application.

MongoDB also provides the feature of Auto-Scaling. Since, MongoDB is a cross platform database and can be installed across different platforms like Windows, Linux etc.

BRIEF HISTORY OF MONGO DB:

MongoDB was developed by Eliot Horowitz and Dwight Merriman in the year 2007, when they experienced some scalability issues with the relational database while developing enterprise web applications at their company DoubleClick. According to Dwight Merriman, one of the developers of MongoDB, this name of the database was derived from the word humongous to support the idea of processing large amount of data.

KEY FEATURES OF MONGO DB:

Apart from most of the NoSQL default features, MongoDB does bring in some more, very important and useful features:

1. MongoDB provides high performance. Input/Output operations are lesser than relational databases due to support of embedded documents(data models) and Select queries are also faster as Indexes in MongoDB supports faster queries.
2. MongoDB has a rich Query Language, supporting all the major CRUD operations. The Query Language also provides good Text Search and Aggregation features.
3. Auto Replication feature of MongoDB leads to High Availability. It provides an automatic failover mechanism, as data is restored through backup copy if server fails.
4. MongoDB supports multiple Storage Engines. When we save data in form of documents(NoSQL) or tables(RDBMS) who saves the data? It's the Storage Engine. Storage Engines manages how data is saved in memory and on disk.

REFERENCE:

https://en.wikipedia.org/wiki/MongoDB

SOLARWINDS® DATABASE PERFORMANCE ANALYZER (DPA)

AMINA S N (19MCA 02)
SYEDA SANIYA (19MCA20)

INTRODUCTION:

Cross-platform support for both cloud and on-premises databases. Anomaly detection powered by machine learning gets smarter over time to help you identify the bottlenecks slowing down your applications.

The intuitive interface visualizes every aspect of the system affecting performance: SQL statements, execution plans, blocking, deadlocks, and more. Down-to-the-second data collection with both real-time and historical analysis allows you to pinpoint problems to find the root cause of performance issues.

SolarWinds use

SolarWinds Orion Platform 2018.4 and below only support hosting a Microsoft SQL Server database

DPA SERVER:

DPA database management software is the performance tool of choice providing you with database and SQL query performance monitoring, analysis and tuning. ... Monitor all of your databases in hybrid environments—on-premises, virtualized, and in the cloud--without being forced to use different tools.

How to access SolarWinds database?

Search for Database Manager. You can also start the Database Manager in the Advanced Features folder by clicking Start > All Programs > SolarWinds Orion > Database Manager. When the application loads, click Add Default Server. To connect to a non-default server, click Add Server and enter the data manually

How is database performance measured?

The following sections provide tips for interpreting the various types of database statistics to measure database performance:

1. Using Hit Ratios.
2. Using Wait Events with Timed Statistics.
3. Using Wait Events without Timed Statistics.
4. Using Idle Wait Events.
5. Comparing Database Statistics with Other Factors

In order to keep your database running smoothly, you can actively monitor metrics covering four areas of performance and resource utilization:

- Query throughput.
- Query execution performance.
- Connections.
- Buffer pool usage

Why is database monitoring important?

Database monitoring is a critical part of any application's maintenance. Finding database issues in time can help the application remain healthy and accessible. Without solid monitoring in place, database outages can go unnoticed until it's too late and the business is losing money and customers.

Five Best Practices for Proactive Database Performance Monitoring

1. Monitor Availability and Resource Consumption. This first step of proactive monitoring is to check whether all databases are online at regular intervals.
2. Measure and Compare Throughput.
3. Monitor Expensive Queries.
4. Track Database Changes.
5. Monitor Logs.

REFERENCES:

1. https://www.google.com/search?q=solarwinds+database+performance+analyzer&rlz=1C5CHFA_enIN879IN882&oq=sol&aqs=chrome.0.69i59j69i57j0l3j46j69i60l2.13469j0j7&sourceid=chrome&ie=UTF-8
2. https://www.solarwinds.com/database-performance-monitoring-software
3. https://support.solarwinds.com/SuccessCenter/s/article/Support-for-hosting-the-Orion-Platform-database-on-MySQL-Oracle-DB2
4. https://www.google.com/search?q=solarwinds+database+performance+analyzer&rlz=1C5CHFA_enIN879IN882&oq=sol&aqs=chrome.0.69i59j69i57j0l3j46j69i60l2.13469j0j7&sourceid=chrome&ie=UTF-8
5. https://support.solarwinds.com/SuccessCenter/s/article/Use-Database-Manager-to-view-the-SolarWinds-database
6. https://support.solarwinds.com/SuccessCenter/s/database-performance-analyzer-dpa

TRANSACTION PROCESS OF DATABASE

A SARUMATHY (19MCA17)

Transaction Process is a logical unit of work that represents real-world events of any organization or an enterprise whereas concurrency control is the management of concurrent transaction execution. Transaction processing systems execute database transactions with large databases and hundreds of concurrent users,
for example, Railway and Air Reservations Systems, Banking system, Credit card Processing.

A Transaction Process can be defined as an action or series of actions that is carried out by a single user or application program to perform operations for accessing the contents of the database. The operations can include retrieval, (Read), insertion (Write), deletion and modification. A transaction must be either completed or aborted. It can either be embedded within an application program or can be specified interactively via a high-level query language such as SQL. Its execution preserves the consistency of the database. Each transaction should access shared data without interfering with the other transactions and whenever a transaction successfully completes its execution; its effect should be permanent.

TRANSACTION PROPERTIES:

A Transaction Process must have the following four properties, called ACID properties (also called ACIDITY of a transaction), to ensure that a database remains stable state after the transaction is executed:

1.**A**tomicity.

2.**C**onsistency.

3.**I**solation.

4.**D**urability.

ATOMICITY:

The atomicity property of a transaction requires that all operations of a transaction be completed, if not, the transaction is aborted. In other words, a transaction is treated as single, individual logical unit of work. Therefore, a transaction must execute and complete each operation in its logic before it commits its changes. As stated earlier, the transaction is considered as one operation even though there are multiple read and writes. Thus, transaction completes or fails as one unit.

The atomicity property of transaction is ensured by the transaction recovery subsystem of a DBMS. In the event of a system crash in the midst of transaction execution, the recovery techniques undo any effects of the transaction on the database.

CONSISTENCY:

Database consistency is the property that every transaction sees a consistent database instance. In other words, execution of a transaction must leave a database in either its prior stable state or a new stable state that reflects the new modifications (updates) made by the transaction. If the transaction fails, the database must be returned to the state it was in prior to the execution of the failed transaction. If the transaction commits, the database must reflect the new changes. Thus, all resources arealways in a consistent state. The preservation of consistency is generally the responsibility of the programmers who write the database programs or of the DBMS module that enforces integrity constraints. A database program should be written in a way that guarantees that, if the database is in a consistent state before executing the transaction, it will be in a consistent state after the complete execution of the transaction, assuming that no interference with other transactions occur. In other words, a transaction must transform the database from one consistent state to another consistent state.

ISOLATION:

Isolation property of a transaction means that the data used during the execution of a transaction cannot be used by a second transaction until the first one is completed. This property isolates transactions from one another. In other words, if a transaction T1 is being executed and is using the data item X, that data item cannot be accessed by any other transaction (T2……….Tn) until T1 ends. The isolation property is enforced by the concurrency control subsystem of the DBMS.

DURABILITY:

The durability property of transaction indicates the performance of the database's consistent state. It states that the changes made by a transaction are permanent. They cannot be lost by either a system failure or by the erroneous operation of a faulty transaction. When a transaction is completed, the database reaches a consistent state and that state cannot be lost, even in the event of system's failure. Durability property is the responsibility of the recovery subsystem of the DBMS.

REFERENCES:

https://www.bloorresearch.com/technology/transaction-processing-dbms/

https://en.wikipedia.org/wiki/Transaction_processing

COUCH DATABASE

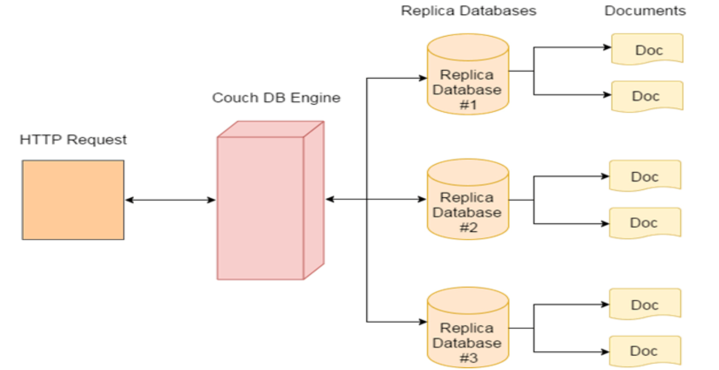
KAVANA C (19MCA08)
SUDIPTA BALA (19MCA18)

DEFINITION:

CouchDB is an open source database developed by Apache software foundation. The focus is on the ease of use, embracing the web. It is a NoSQL document store database. It uses JSON, to store data (documents), java script as its query language to transform the documents, http protocol for api to access the documents, query the indices with the web browser.

USES OF COUCH DATABASE:

- CouchDB takes whatever you pass emit() function and puts it into a list. Each row includes key and value.
- CouchDB is designed to avoid any extra cost. It only runs through all documents once, when you first query your view.
- If a document is changed,the map function is only run once to recomputed the keys and values for the simgle document.
- The view result is stored in a B-tree, just like the structure that is responsible for holding your document.
- View B-tree are stored in their own fill,so that for high performance cause DB usage.
- CouchDB provides easy-to-use replication, using which you can copy, share, and synchronize the data between databases and machines.



REFERENCES:

- https://www.tutorialspoint.com/couchdb/index.htm
- https://en.wikipedia.org/wiki/Apache_CouchDB

MONGO DATABASE

ROSHNI RATHORE (19MCA14)

INTRODUCTION:

MongoDB, the most popular NoSQL database, is an open-source document-oriented database. The term ‘NoSQL’ means ‘non-relational’. It means that MongoDB isn’t based on the table-like relational database structure but provides an altogether different mechanism for storage and retrieval of data. This format of storage is called BSON (similar to JSON format) A simple MongoDB document Structure:

```
{
title: ‘Geeksforgeeks’,
by: ‘Harshit Gupta’,
url: ‘https://www.geeksforgeeks.org’,
type: ‘NoSQL’
}
```

SQL databases store data in tabular format. This data is stored in a predefined data model which is not very much flexible for today’s real-world highly growing applications. Modern applications are more networked, social and interactive than ever. Applications are storing more and more data and are accessing it at higher rates. Relational Database Management System(RDBMS) is not the correct choice when it comes to handling big data by the virtue of their design since they are not horizontally scalable. If the database runs on a single server, then it will reach a scaling limit. NoSQL databases are more scalable and provide superior performance. MongoDB is such a NoSQL database that scales by adding more

FEATURES OF MONGO DB:

- Document Oriented: MongoDB stores the main subject in the minimal number of documents and not by breaking it up into multiple relational structures like RDBMS.
- Indexing: Without indexing, a database would have to scan every document of a collection to select those that match the query which would be inefficient. So, for efficient searching Indexing is a must and MongoDB uses it to process huge volumes of data in very less time.
- Scalability: MongoDB scales horizontally using sharding (partitioning data across various servers.

REFERENCES:

https://www.geeksforgeeks.org/mongodb-an-introduction/
https://www.mongodb.com/nosql-explained

BUSINESS INTELLIGENCE AND DATABASE

SWATI KUMARI (19MCA19)

What is business intelligence?

BI is a framework that allow a business to transform data into information, information into knowledge, and knowledge into wisdom (coronel & morries, 2015, p.560).

According to Forrestor Research BI" a set of methodologies, processes, architectures, technologies, and tools that transform data into meaning full and useful information used to enable more effective strategic, tactical and operational insight and decision making."

BI EVOLUTION:

- Hans Peter's luhn(1958).A Business intelligence system. IBM Journal of research and development.
- BI evolved from the decision support systems(DSS) in 1960s
- Howard Dresner of gartner used BI as a marketing brand andStrategy in1989.
- Then BI has been increasing attention since late 1900s.

BI BENEFITS:

- Improve decision making
- Integrate architecture
- Common user interfaces for data reporting and analysis
- Common database or data warehouse
- Improve organizational performance
- Get more competitive advantage.

BI TECHNOLOGY AND TOOLS:

- Contemporary BI technologies and tools for consolidating, analyzing, providing access to big data to help user make better business decisions and improve business performance .
- Decision support systems (DSS).
- Multidimensional data analysis.
- Database, data warehouse, data mining, data marks.
- ETL (extract, transform and load).
- Big data analytics and business analytics.
- Top BI tools: Informatica, Microsoft BI (MSBI),IBM Congos ,Oracles etc.
- , mathematical, and artificial intelligence techniques to extract and identify useful

DATABASE:

- A shared, integrated computer structure that stores a collection of data which mainly are
- End-user data - Raw facts of interest to end user

- Metadata: Data about data, which the end-user data are integrated and managed
- (Relational) database is a collection of relations or tables.

DATABASE MANAGEMENT SYSTEM (DBMS):

- Is a software system that
- Enables the creation and maintenance of databases
- Manages the database structure
- Controls access to data in the database
- provide decision support and transaction processing
- Relational DBMS
- 1970-3, E.F. Codd of IBM
- Represent data as two-dimensional tables called relations or files
- Each table contains data on entity and attributes
- Many DBMS products available:

DATA WAREHOUSES (DW):

- Bill Inmon, the father of DW
- An integrated, subject-oriented, time-variant, nonvolatile collection of data that provides support for decision making
- A data warehouse is a collection of integrated databases for DM.
- Data warehouse
- Stores current and historical data from many operational transaction systems
- Consolidates and standardizes information for use across organization
- Provides analysis and reporting tools

DATA MINING:

- Data mining is a process that uses statistical information from large databases
- Introduced in 1973
- KDD: Knowledge discovery from a large database
- Example: customer buying patterns
- Data mining can be applied to:
- Uncover hidden trends, patterns, and relationships in large databases or big data
- Form computer models to stimulate and explain the findings
- Use the models to support business decision making

REFERENCES:

- https://www.google.com/amp/s/www.cio.com/article/2439504/business-intelligence-definition-and-solutions.amp.html
- https://en.m.wikipedia.org/wiki/Business_intelligence

CLOUD DATABASE

C H V S P LAVANYA (19MCA05)

INTRODUCTION:

A cloud database is a database service built and accessed through a cloud platform. It serves many of the same functions as a traditional database with the added flexibility of cloud computing. It enables enterprise users to host databases without buying dedicated hardware. It can be deployed either as virtual machine image or as database as service.

METHODS TO RUN A DATABASE IN A CLOUD:

1) **Virtual machine image:** Cloud platforms allow users to purchase virtual-machine instances for a limited time, and one can run a database on such virtual machines. Users can either upload their own machine image with a database installed on it, or use ready-made machine images that already include an optimized installation of a database.

2) **Database-as-a-service (DBaaS):** With a database a service model, application owners do not have to install and maintain the database themselves. Instead, the database service provider takes responsibility for installing and maintaining the database, and application owners are charged according to their usage of the service. This is a type of Saas – Software as a Service.

ARCHITECTURE AND COMMON CHARACTERISTICS:

- Most database service offer web-based consoles, which the end user can use to provision and configure database instances.
- Database services consist of a database manager component, which controls the underlying database instances using a service API. The service API is exposed to the end user, and permits users to perform maintenance and scaling operations on their database instances.
- Underlying software-stack stack typically includes the operating system, the database and third-party software used to manage the database. The service provider is responsible for installing, patching and updating the underlying software stack and ensuring the overall health and performance of the database.
- Scalability features differ between vendors some offer auto-scaling, other enable the user to scale up using an API, but do not scale automatically.
- There is typically a commitment for a certain level of high availability. This is achieved by replicating data and failing instances over to other database instances.

REFRENCES:

http://www.devx.com/
https://www.ibm.com/in-en