Name :Nvar Hilal

Group:C1

Building a Data Warehouse

----------------------------------------What Is a Point of Sale (POS) System?

A point of sale system, or POS, is the place where your customer makes a payment for products or services at your store. Simply put, every time a customer makes a purchase at your store, they’re completing a point of sale transaction.

The POS serves as the central component for your business; it’s the hub where everything—like sales, inventory and customer management—merges.

As evident as the benefits of a POS system are, we found that 56 percent of single-store retailers are still not using one. Instead, we found, many are still using a combination of manual methods, cash registers, QuickBooks and Excel for bookkeeping.

So why have retailers not taken that step to POS yet? To begin with, implementing new technology—especially technology that’s central to your business process—can be scary and overwhelming. Retailers need to consider the negative consequences of failing to have a POS in place.

------------------------------------------------------------------

Enterprise resource planning (ERP)

Enterprise resource planning (ERP) refers to a type of software that organizations use to manage day-to-day business activities such as accounting, procurement, project management, risk management and compliance, and supply chain operations. A complete ERP suite also includes enterprise performance management, software that helps plan, budget, predict, and report on an organization’s financial results.

ERP systems tie together a multitude of business processes and enable the flow of data between them. By collecting an organization’s shared transactional data from multiple sources, ERP systems eliminate data duplication and provide data integrity with a single source of truth.

Today, ERP systems are critical for managing thousands of businesses of all sizes and in all industries. To these companies, ERP is as indispensable as the electricity that keeps the lights on.

-----------------------------------------------------------------------

What does Legacy System?

A legacy system, in the context of computing, refers to outdated computer systems, programming languages or application software that are used instead of available upgraded versions.

Legacy systems also may be associated with terminology or processes that are no longer applicable to current contexts or content, thus creating confusion. In theory, it would be great to be able to have immediate access to use the most advanced technology. But in reality, most organizations have legacy systems - to some extent. A legacy system may be problematic, due to compatibility issues, obsoletion or lack of security support.

A legacy system is also known as a legacy platform

---------------------------------------------------------------------

What is On-Line Transaction Processing (OLTP) System

refers to the system that manage transaction oriented applications. These systems are designed to support on-line transaction and process query quickly on the Internet.

For example: POS (point of sale) system of any supermarket is a OLTP System.

Every industry in today’s world use OLTP system to record their transactional data. The main concern of OLTP systems is to enter, store and retrieve the data. They covers all day to day operations such as purchasing, manufacturing, payroll, accounting, etc.of an organization. Such systems have large numbers of user which conduct short transaction. It supports simple database query so the response time of any user action is very fast.

The data acquired through an OLTP system is stored in commercial RDBMS, which can be used by an OLAP System for data analytics and other business intelligence operations.

------------------------------------------------------------------------

What is a External Web Document?

web document is web based document which opens with a browser whether connected to network or not. Owner of the document can change and modify that document anytime Just like a html document.. Most web documents are basically connected to a host server which creates a link between document and protocol.

---------------------------------------------------------------------

Operational Data Store (ODS)

An operational data store (ODS) is a type of database that collects data from multiple sources for processing, after which it sends the data to operational systems and data warehouses

It provides a central interface or platform for all operational data used by enterprise systems and applications.

An ODS is used to store short term data or data currently in use by operational systems or applications, prior to storage in a data warehouse or data repository. Thus, it serves as an intermediate database.

An ODS helps clean and organize data and ensure that it meets business and regulatory requirements. It only supports low level data and allows for the application of limited queries.

--------------------------------------------------------------------

What is Electronic Data Interchange (EDI)?

Electronic data interchange (EDI) is the electronic transmission of structured data by agreed message standards from one computer system to another without human intervention. It is a system for exchanging business documents with external entities.

EDI refers to a family of standards and does not specify transmission methods, which are freely agreed upon by the trading partners.

The wide adoption of EDI in the business world facilitates efficiency and cost reduction. EDI is used in such diverse business-to-business relationships as

Interchanges between health care providers and insurers:

\*Travel and hotel bookings

\*Education

\*Supply chain management

\*Administration

\*Tax reporting

------------------------------------------------------------------------

What is Extraction Transformation(ETL)?

ETL is short for extract transform load three database functions that are combined into one tool to pull data out of one database and place it into another database.

\*Extract is the process of reading data from a database. In this stage, the data is collected, often from multiple and different types of sources.

\*Transform is the process of converting the extracted data from its previous form into the form it needs to be in so that it can be placed into another database. Transformation occurs by using rules or lookup tables or by combining the data with other data.

\*Load is the process of writing the data into the target database.

How ETL Works

Data from one or more sources is extracted and then copied to the data warehouse. When dealing with large volumes of data and multiple source systems, the data is consolidated. ETL is used to migrate data from one database to another, and is often the specific process required to load data to and from data marts and data warehouses, but is a process that is also used to to large convert (transform) databases from one format or type to another.

------------------------------------------------------------------------

What does Replication?

Replication is the continuous copying of data changes from one database (publisher) to another database (subscriber). The two databases are generally located on a different physical servers, resulting in a load balancing framework by distributing assorted database queries and providing failover capability. The server for the subscriber database may be configured as a backup in the event of failure of the server for the publisher database.

When implementing replication, database systems use a variety of methods but use the same concept. If there is a way to frequently copy data from one database to a second, or even to a third version of the same database, two main database administration goals may be achieved: failover capability and load balancing.

After data is distributed between two or more servers, querying users are not required to connect to a primary database. If replication is set up designating the primary database as the publisher, some queries can be redirected to the subscriber. Known as load balancing, this process is useful for databases that have high user traffic.

Replication can also provided failover capability. When data is held on two separate servers, a subscriber can be set up as the default primary if the publisher becomes unavailable.

There are three different replication methods:

Snapshot Replication: Data is copied from the publisher to the subscribers.

Note: Subscriber changes must come from the publisher. Thus, only the subscriber may be queried; but its data cannot be edited by users.

Merge Replication: Data is combined from two or more databases into one super master database, which is trickier to implement than snapshot replication.

Transactional Replication: The full database is copied, followed by publisher updates which are periodically copied to the subscriber(s).

If important data is accidentally deleted from the publisher, replication deletes the same data from its subscriber(s). Replication is not a substitute for backups.

----------------------------------------------------------------

What is a Metadata Repository?

A Metadata Repository is a software tool that stores descriptive information about the data model used to store and share metadata. Metadata Repositories combine diagrams and text, enabling metadata integration and change. The Metadata Repository’s power lies with the easily accessible way people can view and navigate its contents.

Metadata Management tool allowing an organization to create “a high-level conception or map of its data,” consequently providing better data usage across systems.

------------------------------------------------------------------------

What is Enterprise Data Warehouse

An enterprise data warehouse is a strategic repository that provides analytical information about the core operations of an enterprise. It is distinct from traditional data warehouses and marts, which are usually limited to departmental or divisional business intelligence.

The enterprise data warehouse (EDW) is “by far the largest and most computationally intense business application” in a typical enterprise. EDW systems consist of huge databases, containing historical data on volumes from multiple gigabytes to terabytes of storage Successful EDW systems face two issues regarding the workload of the system: first, they experience rapidly increasing data volumes and application workloads and, second, an increasing number of concurrent users In order to meet the performance requirements, EDW systems are implemented on large-scale parallel computers, such as massively parallel processing (MPP) or symmetric multiprocessor (SMP) system environments and clusters and parallel database software. In fact, most medium- to large-size data warehouses could not be implementable without larger-scale parallel hardware and parallel database software to support them.

In order to handle the requested workload, there is more required than parallel hardware or parallel database software. The logical and physical design of the databases has to be optimized for the expected data volumes.

------------------------------------------------------------------------

Introduction to federated data warehouse

corporate usually has a set of heterogeneous system landscape that contains transaction systems and business intelligence tools which provide analytical capabilities for each individual department needs.

Each department views a business model from their own perspective. For example, a product in Sales can be defined as a material in Manufacturing and equipment in Service Management. In order to integrate those heterogeneous systems that aim to provide analytic capabilities across the different functions and departments, the federated data warehouse was invented.

A federated data warehouse is a practical approach to achieving the “single version of the truth” across the organization. The federated data warehouse is used to integrate key business measures and dimensions. The foundations of the federated data warehouse are the common business model and common staging area.

------------------------------------------------------------------------

What is a Data Mart marketing?

A data mart is a subject-oriented database that is often a partitioned segment of an enterprise data warehouse. The subset of data held in a data mart typically aligns with a particular business unit like sales, finance, or marketing. Data marts accelerate business processes by allowing access to relevant information in a data warehouse or operational data store within days, as opposed to months or longer. Because a data mart only contains the data applicable to a certain business area, it is a cost-effective way to gain actionable insights quickly.

An actionable database developed exclusively for marketers, a Marketing Datamart serves as your repository for real-time insights into the consumers you serve, the effectiveness of your targeting campaigns, and performance trends related to your business. Intalytics will partner with you to design and construct a custom datamart that meets your needs and business requirements.

Each Marketing Datamart that Intalytics develops and maintains consists of the following components: feed maintenance, data hygiene, campaign management, and reporting and analytics.

------------------------------------------------------------------------

SUPPLY CHAIN DATA MANAGEMENT

Supply chain data management is the foundation for smart supply chain decisions, but can be a very complex task. With our in-depth expertise of enabling data for some of the largest supply chains in the world, HAVI can help you:

- Get to your single source of truth for predictive and prescriptive analysis and planning

- Improve supply chain visibility and transparency with consistent, reliable end-to-end data

- Improve collaboration with both internal and external parties

- Reduce administrative labor

- Improve efficiencies with inventory monitoring

-Create an efficient process for leveraging data from external trading partners

- Improve your supply chain risk management capabilities.

------------------------------------------------------------------------

Datamart Management System

The Data mart Management system is a highly programmed software whose task is to maintain and keep track of all the records of the goods that are available in stock or is not available too in an inventory management system.

As we all know, Inventory is basically a place where all the retail things and various goods and items are stored and maintained and is sell to the buyers. When a good is sell, all the details like to which retailer the good is being selling or which item and how many in quantity and at what prices are being it is selling. All the details is being managed in the data mart system.

------------------------------------------------------------------------

What does Data Access?

Data access refers to a user's ability to access or retrieve data stored within a database or other repository. Users who have data access can store, retrieve, move or manipulate stored data, which can be stored on a wide range of hard drives and external devices.

There are two ways to access stored data: random access and sequential access. The sequential method requires information to be moved within the disk using a seek operation until the data is located. Each segment of data has to be read one after another until the requested data is found. Reading data randomly allows users to store or retrieve data anywhere on the disk, and the data is accessed in constant time.

--------------------------------------------------------------------

What is Middleware?

Middleware is the software that connects network-based requests generated by a client to the back-end data the client is requesting. It is a general term for software that serves to "glue together" separate, often complex and already existing programs.

Middleware programs come in on-premises software and cloud services, and they can be used independently or together, depending upon the use case. While cloud providers bundle middleware into cloud services suites, such as middleware as a service (MWaaS) or integration PaaS (iPaaS), industry researchers note that many businesses still choose independent middleware products that fit their specific needs.

Why is it called middleware?

With network-based interactions, a client, or requesting program, can make a request. That client is typically an application that resides on the front end, which is where the user interacts with software. Resources such as databases, message queues, NoSQL data stores and file servers are often referred to as being part of the back end.

------------------------------------------------------------------------

What is decision support system (DSS)?

A decision support system (DSS) is a computer program application that analyzes business data and presents it so that users can make business decisions more easily. It is an "informational application" (to distinguish it from an "operational application" that collects the data in the course of normal business operation).Typical information that a decision support application might gather and present would be:

- Comparative sales figures between one week and the next

- Projected revenue figures based on new product sales assumptions

- The consequences of different decision alternatives, given past experience in a context that is described

A decision support system may present information graphically and may include an expert system or artificial intelligence (AI). It may be aimed at business executives or some other group of knowledge workers.

------------------------------------------------------------------------

What is EIS Reporting?

EiS Technologies provides next-generation, Oracle-platform–based reporting with its

powerful eXpress Reporting for Oracle E-Business Suite. Discover why EiS clients are

able to administer all of their operational and ad hoc reports from a single system that is

securely integrated with your Oracle E-Business Suite application. In addition, eXpress

Reporting offers comprehensive business views, analytics, and dashboards.

--------------------------------------------------------------------

What is Relational database tool?

A relational database organizes data into tables which can be linked—or related—based on data common to each. This capability enables you to retrieve an entirely new table from data in one or more tables with a single query. It also allows you and your business to better understand the relationships among all available data and gain new insights for making better decisions or identifying new opportunities.

For example, imagine your company maintains a customer table that contains company data about each customer account and one or more transaction tables that contain data describing individual transactions.

The columns (or fields) for the customer table might be Customer ID, Company Name, Company Address, etc the columns for a transaction table might be Transaction Date, Customer ID, Transaction Amount, Payment Method, etc. The tables can be related based on the common Customer ID field. You can, therefore, query the table to produce valuable reports, such as a consolidated customer statement.

------------------------------------------------------------------------

**What is Online Analytical Processing(OLAP)?**

**is the technology behind many Business Intelligence (BI) applications. OLAP is a powerful technology for data discovery, including capabilities for limitless report viewing, complex analytical calculations, and predictive “what if” scenario (budget, forecast) planning.**

**How is OLAP Technology Used?**

**OLAP is an acronym for Online Analytical Processing. OLAP performs multidimensional analysis of business data and provides the capability for complex calculations, trend analysis, and sophisticated data modeling. It is the foundation for many kinds of business applications for Business Performance Management, Planning, Budgeting, Forecasting, Financial Reporting, Analysis, Simulation Models, Knowledge Discovery, and Data Warehouse Reporting. OLAP enables end-users to perform ad hoc analysis of data in multiple dimensions, thereby providing the insight and understanding they need for better decision making.**

**------------------------------------------------------------------------**

**What is relational online analytical processing (ROLAP)?**

**Relational online analytical processing (ROLAP) is a form of online analytical processing (OLAP) that performs dynamic multidimensional analysis of data stored in a relational database rather than in a multidimensional database (which is usually considered the OLAP standard).**

------------------------------------------------------------------------

**What does Data Mining?**

**Data mining is the process of analyzing hidden patterns of data according to different perspectives for categorization into useful information, which is collected and assembled in common areas, such as data warehouses, for efficient analysis, data mining algorithms, facilitating business decision making and other information requirements to ultimately cut costs and increase revenue.**

**Data mining is also known as data discovery and knowledge discovery.**

**The major steps involved in a data mining process are:**

**- Extract, transform and load data into a data warehouse**

**- Store and manage data in a multidimensional databases**

**- Provide data access to business analysts using application software**

**- Present analyzed data in easily understandable forms, such as graphs.**

**--------------------------------------------------------------------**

**What is Internet?**

**The Internet, sometimes called simply "the Net," is a worldwide system of computer networks -- a network of networks in which users at any one computer can, if they have permission, get information from any other computer (and sometimes talk directly to users at other computers). It was conceived by the Advanced Research Projects Agency (ARPA) of the U.S. government in 1969 and was first known as the ARPANet. The original aim was to create a network that would allow users of a research computer at one university to "talk to" research computers at other universities.**

**A side benefit of ARPANet's design was that, because messages could be routed or rerouted in more than one direction, the network could continue to function even if parts of it were destroyed in the event of a military attack or other disaster.**

**How the Internet works?**

**Physically, the Internet uses a portion of the total resources of the currently existing public telecommunication networks. Technically, what distinguishes the Internet is its use of a set of protocols called Transmission Control Protocol/Internet Protocol (TCP/IP). Two recent adaptations of Internet technology, the intranet and the extranet, also make use of the TCP/IP protocol.**

**------------------------------------------------------------------------**

**What does Web Browser?**

**A web browser is a software program that allows a user to locate, access, and display web pages. In common usage, a web browser is usually shortened to "browser." Browsers are used primarily for displaying and accessing websites on the internet, as well as other content created using languages such as Hypertext Markup Language (HTML) and Extensible Markup Language (XML).**

**Browsers translate web pages and websites delivered using Hypertext Transfer Protocol (HTTP) into human-readable content. They also have the ability to display other protocols and prefixes, such as secure HTTP (HTTPS), File Transfer Protocol (FTP), email handling.**

**In addition, most browsers also support external plug-ins required to display active content, such as in-page video, audio and game content.**

**--------------------------------------------------------------------**

**Relational Databases For Data Warehousing:**

**A data warehouse is a relational database that is designed for query and analysis rather than for transaction processing. It usually contains historical data derived from transaction data, but it can include data from other sources. It separates analysis workload from transaction workload and enables an organization to consolidate data from several sources.**

**In addition to a relational database, a data warehouse environment includes an extraction, transportation, transformation, and loading (ETL) solution, an online analytical processing (OLAP) engine, client analysis tools, and other applications that manage the process of gathering data and delivering it to business users.**

**The Data Warehouse**

**integrates and stores the enterprise data. It is designed to optimize query and analysis performance, end-user understandability and scalability. The data marts are subsets of the data warehouse focused on a specific subject area. The data warehouse can be used to feed multidimensional databases (OLAP servers) for more specialized analytical applications. The central corporate data warehouse is the cornerstone of the data warehousing environment, but this function also includes other special purpose data stores like the ODS (Operational Data Store), Multidimensional databases (MDDB), and specialized downstream data marts.**