

1. Noțiuni introductive

Cuprins

1.1. Definiții ale depozitului de date.....	2
1.2 Evoluția sistemelor informatice destinate asistării deciziilor și pentru inteligența afacerii.	3
1.3. Caracteristicile depozitelor de date.....	4
Orientarea pe subiect	4
Integrarea	5
Persistența datelor.....	5
Caracterul istoric	5
1.4. OLPT versus Data Warehouse	5
Informațiile stocate	5
Operațiile	6
Modificările datelor	6
Proiectarea.....	6
Cantitatea de date citită.....	7
Datele istorice.....	7
Diferențele dintre OLTP și Data Warehouse	7
Bibliografie	9

1. Noțiuni introductive

1.1. Definiții ale depozitului de date

- Depozitele de date (*Data Warehouse*) au apărut din nevoia de:
 - a analiza datele acumulate de-a lungul timpului de marile companii;
 - a obține informații importante, necesare în luarea deciziilor.
 - Exploatarea inteligentă a acestor depozite de date:
 - mărește capacitatea de acomodare la tendințele pieței,
 - determină diminuarea costurilor și creșterea profitului.
 - Datele istorice integrate într-o singură structură care constituie baza pentru procesul de luare a deciziilor a devenit o prioritate a noilor tehnologii informaționale.
 - **W. H. Inmon** - *Building the Data Warehouse, 4th Edition, Wiley, 2005/1996*
 - Un depozit de date este o "colecție de date orientate pe subiecte, integrate, istorice și nevolatile destinată sprijinirii procesului de luare a deciziilor manageriale".
 - **R. Kimball** - *The Data Warehouse Lifecycle Toolkit, 2nd Edition Wiley, 2008/1998*
 - Depozitul oferă acces la datele organizaționale.
 - Datele conținute sunt consistente.
 - Datele pot fi separate și combinate în funcție de fiecare dimensiune sau aspect al afacerii.
 - Depozitul de date include un set de instrumente pentru interogare, analiză și prezentare a informațiilor.
 - **Compania Oracle**
 - Depozitul de date este o colecție de informații strategice derivate direct din sistemele *OLTP* și alte surse externe.
 - Scopul specific al depozitului de date este de a oferi suport de decizie și nu suport pentru activitatea operativă.
-
- ❖ Un depozit de date este un mediu pentru suportul deciziei care preia datele stocate în diferite surse operaționale, le organizează și le face disponibile pentru analize decizionale și rapoarte.
 - ❖ Depozitele de date sunt structuri create pentru stocarea unor volume mari de



date organizate pe domenii, ce constituie subiecte de interes decizional în activitatea companiei.

- ❖ Un aspect important este redundanța datelor care este permisă (de exemplu, data calendaristică se poate exprima și în luni și în semestre și în ani). Deci, informațiile care se pot calcula din datele primare sunt stocate explicit în depozit pentru a fi gata calculate la o eventuală solicitare.
- ❖ Principiul pe care îl urmează este cunoscut sub numele de procesare analitică (*OLAP – On Line Analytical Processing*).
 - Spre deosebire de acesta, principiul pe care se bazează sistemele tranzacționale este numit procesare tranzacțională (*OLTP – On Line Transactional Processing*).
 - *Data Warehouse* este o bază de date relațională, care este creată în special pentru analiză și interogări și mai puțin pentru procesarea tranzacțiilor.

1.2 Evoluția sistemelor informatice destinate asistării deciziilor și pentru inteligența afacerii



- *EIS - Executive Information Systems*
 - Erau aplicații scrise în limbaje de generație a treia (3GL), în limbaje de generație a patra (4GL), C++ sau alte limbaje de programare structurate; erau restrictive, permițând interogări predefinite, concretizate în rapoarte tabelare sau grafice; informațiile furnizate erau limitate la totalul vânzărilor, număr unități produse etc.
 - Permiteau gestiunea informațiilor necesare la nivelul executiv al companiei și luarea unor decizii nestructurate.
 - Ofereau acces la informații relevante pentru managementul la nivel înalt al companiei.
 - Sunt considerate o formă specializată a *DSS* (sisteme suport de decizie).
- *DSS - Decision Support Systems*

- Reprezintă prima generație de aplicații care oferă SQL generat dinamic, permițând utilizatorilor să extragă date din bazele de date relaționale.
- Ofereau managerilor modele complexe și aprofundate de analiză pentru fundamentarea deciziilor.
- *DW&BI - Data Warehousing & Business Intelligence*
 - Reprezintă noua generație de *DSS* care oferă capabilitatea de a crea rapoarte mult mai ușor și mai rapid, folosind multiple surse de date și multiple subiecte de interes, ce pot fi utilizate simultan pentru a oferi o evaluare exactă a afacerii.
 - Tehnologia *BI* permite analiza indicatorilor cheie de performanță, ajutând la atingerea obiectivelor strategice ale companiei, o evoluție rapidă la nivelul managementului relației cu clienții (*CRM*), analiza lanțului de aprovizionare, creșterea vânzărilor, diminuarea pierderilor și, implicit, creșterea profiturilor.

1.3. Caracteristicile depozitelor de date

- Definirea unui depozit de date pornește de 4 caracteristici:
 - orientare pe subiect;
 - integrare;
 - persistența datelor;
 - caracterul istoric.
- Un depozit de date este un mediu pentru suportul deciziei care preia datele stocate în diferite surse operaționale, le organizează și le face disponibile pentru analize decizionale și rapoarte.

Orientarea pe subiect

- Depozitele de date sunt create cu scopul de a permite analiza datelor.
 - De exemplu, ca să se poată obține mai multe informații despre vânzările unei companii, se va construi un depozit de date care va conține toate datele despre vânzări.
 - Utilizând acest depozit vom putea răspunde, de exemplu, la întrebări de tipul
 - Care a fost cel mai bun client al articolului x în ultimul an?
 - Care este evoluția anuală a vânzărilor articolului x ?
 - Care este raportul vânzărilor articolului x în primul trimestru al anului curent relativ la primul trimestru al anului trecut?
 - Posibilitatea de a defini un depozit de date în funcție de un subiect de interes (de exemplu, în funcție de vânzări), evidențiază orientarea pe subiect a sa.

Integrarea

- Integrarea este strâns legată de orientarea pe subiect.
- Depozitul de date trebuie să transpună datele provenite din surse diferite, într-un format consistent.
 - De-a lungul timpului compania a evoluat și a utilizat mai multe surse de date sau utilizează chiar în prezent mai multe surse de date
 - De exemplu, *Fox*, *SQL Server*, *Oracle*, *DB2* etc
- Un depozit de baze de date este considerat integrat atunci când ajunge într-un format consistent
 - Sunt rezolvate conflictele de
 - nume, tipuri de date și dimensiuni
 - inconsistență a unităților de măsură etc
 - Datele sunt transformate, curățate și validate.

Persistența datelor

- Această caracteristică presupune că datele care au fost integrate într-un depozit nu mai trebuie modificate.
- Scopul depozitului de date este de a permite analiza datelor, nu modificarea acestora.

Caracterul istoric

- Cu scopul de a descoperi tendințe în afaceri, analiștii au nevoie de volume mari de date. Acest fapt este în contrast cu sistemele *OLTP*, în care cerințele de performanță fac ca datele istorice să fie ținute în arhive.
- Datele integrate în depozit furnizează informații din perspectivă istorică (pe luni, trimestre, ani etc).
- Managerii companiei pot consulta valorile succesive ale indicatorilor pentru a determina evoluția în timp a acestora.

1.4. OLTP versus Data Warehouse

Informațiile stocate

- Baza de date *OLTP* stochează date:
 - operaționale, referitoare la un anumit proces sau funcție a companiei,
 - de detaliu cu un anumit grad de volatilitate (pot suferi schimbări),
 - care prezintă interes, în primul rând, pentru cei ce le introduc.

- Baza de date *Data Warehouse* stochează date:
 - pentru asistarea deciziei,
 - referitoare la subiecte de interes decizional,
 - centralizate sau derivate din datele operaționale,
 - nu se schimbă în timp,
 - sunt orientate către utilizatorii finali (managerii) de nivel strategic.

Operațiile

- Baza de date *OLTP*
 - Suportă numai operații predefinite.
 - Aplicațiile ar putea fi optimizate sau create special numai pentru acele operații.
- Baza de date *Data Warehouse*
 - Este creată pentru a permite interogări *ad hoc*.
 - Este posibil să nu fie cunoscută dinainte linia pe care o va urma depozitul de date, dar acesta trebuie să fie optimizat, astfel încât să răspundă la o varietate mare de operații de interogare.
 - Trebuie să fie suficient de flexibilă pentru a putea răspunde interogărilor spontane ale utilizatorilor.

Modificările datelor

- Baza de date *OLTP*
 - Utilizatorii realizează, de obicei, în mod individual procedurile de modificare a bazei de date.
 - Este întotdeauna actualizată și reflectă starea curentă a fiecărei tranzacții.
- Baza de date *Data Warehouse*
 - Este actualizată pe baza unui proces *ETL* (*Extract, Transform, Load*).
 - Acesta poate fi rulat noaptea sau săptămânal, folosind tehnici speciale care permit modificarea masivă a datelor.
 - Utilizatorii finali ai unui depozit de date nu fac în mod direct actualizări ale depozitului.

Proiectarea

- Baza de date *OLTP*
 - Sunt utilizate de obicei scheme total normalizate pentru a optimiza performanțele operațiilor *LMD* și pentru a garanta consistența datelor.
- Baza de date *Data Warehouse*

- Sunt utilizate de obicei scheme denormalizate sau scheme denormalizate parțial (cum ar fi de exemplu schema stea) pentru a optimiza performanța interogărilor.

Cantitatea de date citită

- Baza de date *OLTP*
 - O operație tipică *OLTP* accesează doar o mică parte din înregistrări.
 - De exemplu, cererea „Afișează informații despre comanda curentă a clientului x”.
- Baza de date *Data Warehouse*
 - O simplă interogare a depozitului de date poate determina scanarea a mii sau chiar milioane de înregistrări.
 - De exemplu, cererea „Afișează valoarea totală a vânzărilor realizate trimestrul trecut pentru toți clienții persoane juridice”.

Datele istorice

- Baza de date *OLTP*
 - Sistemele *OLTP*, de obicei, înmagazinează date numai din câteva săptămâni sau luni.
 - Acestea rețin datele istorice atâta timp cât este necesar pentru a îndeplini cu succes cerințele tranzacției curente.
- Baza de date *Data Warehouse*
 - Depozitele de date stochează datele pe o perioadă lungă de timp (luni sau ani).
 - Acest lucru oferă suport pentru analiza istorică a datelor.

Diferențele dintre OLTP și Data Warehouse

CARACTERISTICĂ	<i>OLTP</i> Structuri complexe de date	<i>DATA WAREHOUSE</i> Structuri multidimensionale de date
normalizare	mare (de obicei FN3)	mică (denormalizare)
mărimea tabelor	mică	mare
număr de înregistrări/tabele	mic	mare
durată tranzație	mică	mare
număr de utilizatori <i>on line</i>	mare (1000)	mic(<100)

actualizări	frecvent	rar
scanare <i>full-table</i>	rar	frecvent
date istorice	< 90 zile	> 1 an

Bibliografie

1. Connolly T.M., Begg C.E., *Database Systems: A Practical Approach to Design, Implementation and Management*, 5th edition, Pearson Education, 2005
2. Dollinger R., Andron L., *Baze de date și gestiunea tranzacțiilor*, Editura Albastră, Cluj-Napoca, 2004
3. Inmon W.H., *Building the Data Warehouse*, 4th Edition, Wiley, 2005
4. Kimball R., *The Data Warehouse Toolkit*, 3rd Edition, Wiley, 2013
5. Kimball R., Ross M., Thornthwaite W., Mundy J., Becker B., *The Data Warehouse Lifecycle Toolkit*, 2nd Edition Wiley, 2008
6. Oracle and/or its affiliates, *Oracle Database Concepts*, 1993, 2024
7. Oracle and/or its affiliates, *Oracle Database Administrator's Guide*, 2001, 2025
8. Oracle and/or its affiliates, *Oracle Database Performance Tuning Guide*, 2013, 2025
9. Oracle and/or its affiliates, *Oracle Database SQL Language Reference*, 1996, 2025
10. Oracle and/or its affiliates, *Oracle Database PL/SQL Language Reference*, 1996, 2025
11. Oracle and/or its affiliates, *Oracle Database: SQL Tuning Workshop*, 2010, 2025
12. Oracle and/or its affiliates, *Oracle OLAP Customizing Analytic Workspace Manager*, 2006, 2019
13. Oracle and/or its affiliates, *Oracle OLAP DML Reference*, 1994, 2019
14. Oracle and/or its affiliates, *Oracle OLAP User's Guide*, 2003, 2019
15. Oracle and/or its affiliates, *Oracle Warehouse Builder Concepts*, 2000, 2021
16. Oracle and/or its affiliates, *Oracle Warehouse Builder Data Modeling, ETL, and Data Quality Guide*, 2000, 2021
17. Oracle and/or its affiliates, *Oracle Database Data Warehousing Guide*, 2001, 2021
18. Oracle University, *Oracle Database: PL/SQL Fundamentals, Student Guide*, 2009, 2025
19. Poe V., Klauer P., Brobst S., *Building A Data Warehouse for Decision Support*, 2nd Edition, Prentice Hall; 1997
20. Popescu I., Alecu A., Velcescu L., Florea (Mihai) G., *Programare avansată în Oracle9i*, Ed. Tehnică, 2004