

# Agregacija podataka u Data Warehouse sistemima

---

MATERIJALIZOVANI POGLEDI I  
AGREGACIONE FUNKCIJE

# Sadržaj

## Šema DW BP i agregirani podaci

Upravljanje agregiranim podacima

Pogledi u DW sistemima

Materijalizacija pogleda

Vrste materijalizacije pogleda

Završne napomene

# Šema sa agregiranim podacima

---

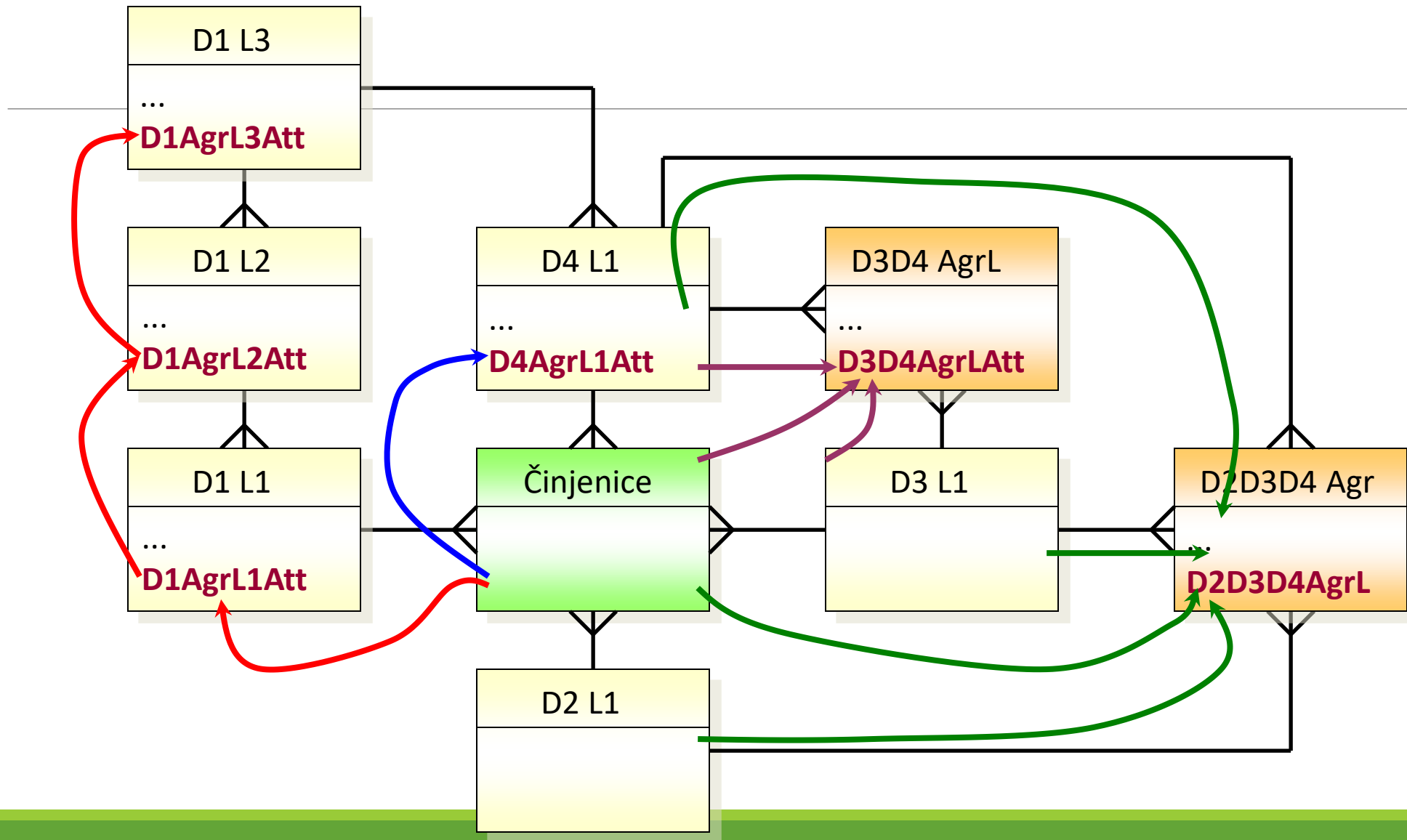
## Agregirani (sumarni) podaci

- sumarni podaci po zadatim dimenzijama
  - redundantni, ali neophodni za efikasnu podršku različitih upita i analiza
  - čuvaju se u DW bazi podataka
  - generišu se (izračunavaju) u ECTL procesu
    - na osnovu formiranih činjeničnih podataka
    - često uz primenu skupovnih funkcija, kao što su: SUM, COUNT, MIN, MAX, AVG, STDDEV, itd.

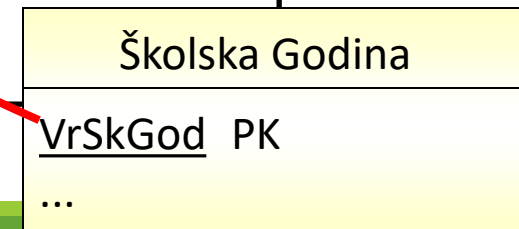
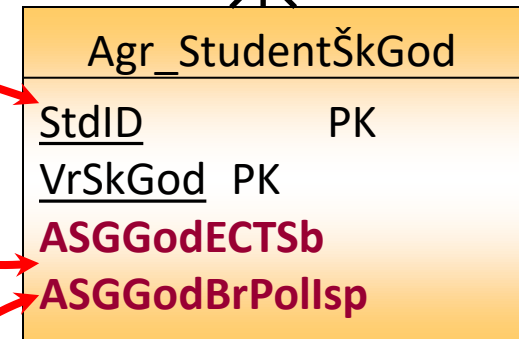
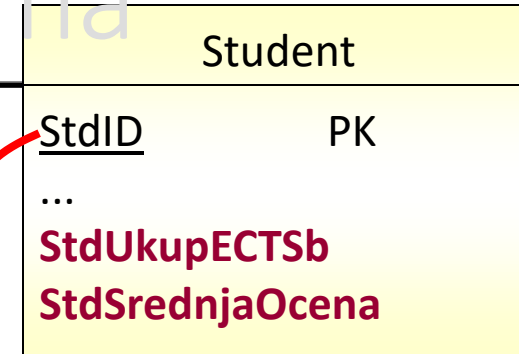
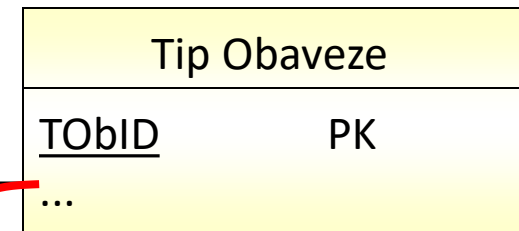
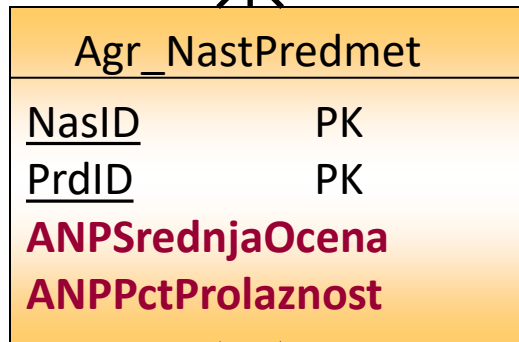
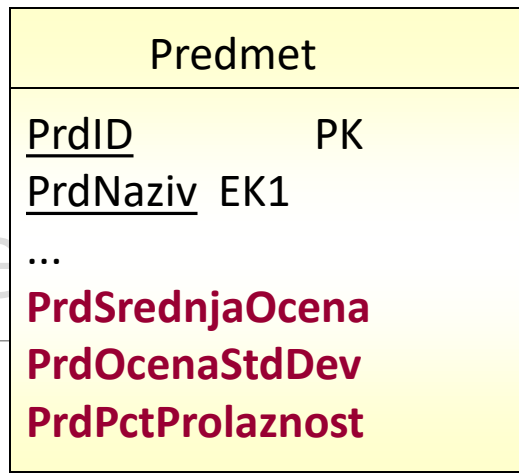
## Nivoi agregacije

- bilo koja kombinacija dimenzija
- bilo koji nivo u hijerarhijskoj strukturi dimenzije

# Šema sa agregiranim podacima



Še  
egir  
acima



# Sadržaj

---

Šema DW BP i agregirani podaci

Upravljanje agregiranim podacima

Pogledi u DW sistemima

Materijalizacija pogleda

Vrste materijalizacije pogleda

Završne napomene

# Upravljanje agregiranim podacima

---

## Summary Management

- projektovanje DW šeme BP sa agregiranim podacima
- implementacija agregacije u DW šemi BP i ECTL
- optimizacija upita nad strukturama agregiranih podataka
- praćenje upotrebe agregiranih podataka

# Upravljanje agregiranim podacima

---

## Summary Management

- projektovanje DW šeme BP sa agregiranim podacima
  - specifikacija atributa agregiranih podataka
  - specifikacija nivoa agregacije
    - izbor odgovarajućih dimenzija
    - izbor odgovarajućih nivoa hijerarhija dimenzija
    - izbor odgovarajućih kombinacija dimenzija
    - izbor granularnosti vremenske dimenzije
      - ne mora se zahtevati ista granularnost za sve periode vremena
      - npr. jedinica dana za poslednju godinu, jedinica meseca za prethodnu godinu i jedinica tromesečja za ranije godine
  - specifikacija algoritama za agregiranje podataka
    - inicijalnog izračunavanja agregiranih vrednosti
    - kasnijih modifikacija (osvežavanja) agregiranih vrednosti

# Upravljanje agregiranim podacima

---

## Summary Management

- projektovanje DW šeme BP sa agregiranim podacima
  - specifikacija načina izračunavanja i memorisanja agregiranih vrednosti
    - agregirane vrednosti se izračunavaju prilikom inicijalnog punjenja i osvežavanja i memorišu u DW BP ili
    - agregirane vrednosti se izračunavaju na zahtev programa i ne memorišu u DW BP
    - da li izračunavanje može koristiti druge (i koje) agregirane vrednosti iz DW BP
    - izračunavanje može biti sprovedeno na nivou
      - servera DW BP
      - aplikativnog programa za upotrebu DW BP

# Upravljanje agregiranim podacima

---

## Summary Management

- implementacija agregacije u DW šemi BP i ECTL
  - kreiranje materijalizovanih pogleda
  - implementacija postupaka inicijalizacije i osvežavanja agregiranih podataka
- optimizacija upita sa agregiranim podacima
  - zasnovana na
    - upotrebi materijalizovanih pogleda
    - prestrukturiranju (preuređivanju) SQL SELECT naredbi
- praćenje upotrebe agregiranih podataka
  - upotreba inteligentnih alata za praćenje upotrebe agregiranih podataka i unapređenja u šemi DW BP
    - summary advisor

# Sadržaj

---

Šema DW BP i agregirani podaci

Upravljanje agregiranim podacima

Pogledi u DW sistemima

Materijalizacija pogleda

Vrste materijalizacije pogleda

Završne napomene

# Pogledi u DW sistemima

---

## SQL pogled

- virtuelna tabela, čiji se sadržaj generiše na zahtev upita
- imenovana SELECT naredba, sačuvana u rečniku podataka SUBP
- koristi se u upitima, ili za definisanje drugih pogleda
- eksplicitno definisani pogled - sintaksa

```
CREATE VIEW view_name [(view_attr_list)]
```

```
AS SELECT ...
```

- implicitno definisani pogled - sintaksa

```
SELECT ...
```

```
FROM (SELECT ...) alias, ...
```

```
WHERE ...
```

# Pogledi u DW sistemima

---

## Upotreba pogleda u DSS

- često potrebna, kako bi se analitičar podataka fokusirao samo na potrebni deo podataka
- upiti u DSS često zahtevaju agregirane podatke
  - izvedene iz velike količine podataka tabele činjenica

## Materijalizacija pogleda

- neophodna da bi se obezbedile dobre performanse takvih upita
- realno moguća, jer tabele činjenica sadrže relativno nepromenljive podatke u dužem periodu vremena

# Pogledi u DW sistemima

---

## **Materijalizovani pogled**

- tabela ili skup tabela sa agregiranim podacima
  - čiji se sadržaj generiše na osnovu sadržaja baznih tabela DW BP
- DW baza podataka se može, takođe, posmatrati kao jedan materijalizovani pogled
  - nad BP OLTP sistema i nad spoljnim izvorima podataka
- koristi se za efikasnu realizaciju upita u DSS

# Pogledi u DW sistemima

---

## **Materijalizovani pogled**

- mogući kriterijumi kada materijalizovati pogled zasnivaju se na proceni parametara
  - koliko vrsta upita može biti pokriveno materijalizovanim pogledom
  - kakva je frekvencija pokretanja upita koji mogu biti pokriveni materijalizovanim pogledom
  - kakvo je očekivano poboljšanje performansi upita koji mogu biti pokriveni materijalizovanim pogledom
  - koliko materijalizovani pogled zahteva prostora u BP
  - uolikoj meri materijalizovani pogled usložnjava osvežavanje DW BP

# Pogledi u DW sistemima

---

## Primer

- data zvezdasta šema
  - Product({ProdId, ProdName, Category, Price}, {ProdId})
  - Location({ShopId, ShopName, City, District}, {ShopId})
  - Time({TimeId, Date, Week, Month, Quarter, Year}, {TimeId})
  - Sales({ProdId, TimeId, ShopId, SalesData}, {ProdId+TimeId+ShopId})

- dat virtuelni pogled

```
CREATE VIEW Prod_Shop_Sales
  (ProdId, Category, ShopId, City, TimeId, SalesData) AS
  SELECT p.ProdId, p.Category, l.ShopId, l.City, s.TimeId,
         s.SalesData
  FROM Product p, Location l, Sales s
  WHERE p.ProdId = s.ProdId AND l.ShopId = s.ShopId
```

# Pogledi u DW sistemima

---

## Primer

- posmatra se upit nad nematerijalizovanim pogledom

```
SELECT Category, City, SUM(SalesData)
FROM Prod_Shop_Sales
GROUP BY Category, City
```

- očekivano je da se upit realizuje tako što se naziv Prod\_Shop\_Sales zamenjuje definicijom pogleda

# Pogledi u DW sistemima

---

## Primer

- upit u "razvijenoj" formi

```
SELECT v.Category, v.City, SUM(v.SalesData)
FROM (
    SELECT p.ProdId, p.Category, l.ShopId, l.City, s.TimeId,
           s.SalesData
    FROM Product p, Location l, Sales s
    WHERE p.ProdId = s.ProdId AND l.ShopId = s.ShopId) v
GROUP BY v.Category, v.City
```

# Pogledi u DW sistemima

---

## Primer

- materijalizacija pogleda Prod\_Shop\_Sales
- sa ciljem unapređenja performansi upita

```
SELECT Category, City, SUM(SalesData)
FROM Prod_Shop_Sales
GROUP BY Category, City
```

- tako što se eliminiše vreme potrebno
  - za spajanja tabela
    - Sales, Product i Location
    - koje se projektuju na skup kolona
      - {ProdId, Category, ShopId, City, TimeId, SalesData}
      - ili na bilo koji podskup ovog skupa kolona
- formiranjem posebne, bazne tabele koja će sadržati podatke iz Prod\_Shop\_Sales

# Pogledi u DW sistemima

---

## Primer

- materijalizacija pogleda Prod\_Shop\_Sales
- zahtev prema optimizatoru upita
  - da je u stanju da izvrši inteligentnu reorganizaciju upita (query rewriting)

```
SELECT Category, City, SUM(SalesData)
FROM Product NATURAL JOIN Location
      NATURAL JOIN Sales
GROUP BY Category, City
```

- u oblik koji obezbeđuje optimizovano izvršavanje

```
SELECT Category, City, SUM(SalesData)
FROM Prod_Shop_Sales
GROUP BY Category, City
```

# Pogledi u DW sistemima

---

## Primer

- posmatra se upit

```
SELECT p.Category, l.City, SUM(s.SalesData)
FROM Product p, Location l, Sales s, Time t
WHERE p.ProdId = s.ProdId
      AND l.ShopId = s.ShopId
      AND t.TimeId = s.TimeId
      AND t.Year = 2000
GROUP BY p.Category, l.City
```

- upit ne može biti realizovan samo uz pomoć materijalizovanog pogleda Prod\_Shop\_Sales
  - nedostaju podaci o godini (Year), u dimenziji Time

# Pogledi u DW sistemima

---

## Primer

- zahtev prema optimizatoru upita
  - da je u stanju da izvrši inteligentnu reorganizaciju upita
  - tako da koristi materijalizovani pogled Prod\_Shop\_Sales i spajanje s tabelom Time

```
SELECT v.Category, v.City, SUM(v.SalesData)
FROM Prod_Shop_Sales v JOIN Time t
ON (v.TimeId = t.TimeId)
WHERE t.Year = 2000
GROUP BY Category, City
```

# Sadržaj

---

Šema DW BP i agregirani podaci

Upravljanje agregiranim podacima

Pogledi u DW sistemima

Materijalizacija pogleda

Vrste materijalizacije pogleda

Završne napomene

# Materijalizacija pogleda

---

## Klasični relacioni SUBP

- kreiranje i odloženo inicijalno punjenje tabele
  - **CREATE TABLE** mat\_view\_name ...
  - **INSERT INTO** mat\_view\_name **AS SELECT** ...
- kreiranje i istovremeno inicijalno punjenje tabele
  - **CREATE TABLE** mat\_view\_name **AS SELECT**...

## Relacioni SUBP sa OLAP mogućnostima (ROLAP Engine)

- upotreba specijalizovanog koncepta
  - **CREATE MATERIALIZED VIEW** mat\_view\_name
  - ...
  - AS SELECT**...

# Materijalizacija pogleda

---

## SQL materijalizovani pogled

- posebna vrsta bazne tabele
- formira se i ažurira preuzimanjem i agregiranjem podataka iz jedne ili više drugih tabela
  - koje se mogu ali ne moraju nalaziti na istom serveru BP
- ažurira se, u principu, posredno - osvežavanje
  - propagacijom operacija ažuriranja nad osnovnim tabelama, od kojih je nastala, ili ponovnom inicijalizacijom sadržaja
  - u zadatim vremenskim intervalima ili na zahtev
- može biti indeksiran i particioniran
- optimizator upita je u stanju da izvrši reorganizaciju upita nad baznim tabelama (query rewriting)
  - tako da se upotrebe odgovarajući materijalizovani pogledi

# Materijalizacija pogleda

---

## SQL materijalizovani pogled

- politika održavanja sadržaja
  - definiše kada i kako osvežavati materijalizovani pogled
  - definiše se u postupku projektovanja DW sistema
- **vreme osvežavanja sadržaja**
  - **trenutno**
    - u momentu izvođenja transakcije nad originalnim podacima
  - **odloženo**
    - u zadatom trenutku vremena
    - u regularnim vremenskim intervalima
    - na zahtev, prilikom izvođenja upita nad materijalizovanim pogledom
    - na eksplicitni zahtev administratora

# Materijalizacija pogleda

---

## SQL materijalizovani pogled

- **način osvežavanja sadržaja**
  - **inkrementalno ("brzo") osvežavanje**
    - sprovode se samo izmene sadržaja, saglasno sprovedenim operacijama nad tabelama sa originalnim podacima
    - zahteva kreiranje dnevnika materijalizovanog pogleda
      - **CREATE MATERIALIZED VIEW LOG ...**
    - pogodnije sa stanovišta performansi, kada je logički moguće
  - **kompletno osvežavanje**
    - reinicijalizuje se kompletan sadržaj materijalizovanog pogleda
      - ponovnim izvršavanjem SELECT naredbe
    - sprovodi se samo kada inkrementalno osvežavanje nije logički moguće
      - u situacijama kada se agregirane vrednosti izračunavaju primenom tzv. holističkih (kompletno neaditivnih) funkcija (posebna tema)

# Materijalizacija pogleda

---

## SQL materijalizovani pogled

- sintaksa za kreiranje

```
CREATE MATERIALIZED VIEW mvw_name  
[TABLESPACE tbs_name PCTFREE nn]  
[STORAGE (INITIAL n[M|K] NEXT n[M|K] PCTINCREASE n)]  
[NOLOGGING]  
[PARALLEL (DEGREE [DEFAULT | value])]  
BUILD [DEFERRED | IMMEDIATE]  
[USING [NO] INDEX [TABLESPACE tbs_name]]  
REFRESH [FORCE | FAST | COMPLETE]  
[ON DEMAND | ON COMMIT | START WITH date NEXT date]  
[ENABLE QUERY REWRITE]  
AS SELECT ...
```

# Materijalizacija pogleda

---

## SUBP Oracle: tehničke mogućnosti

- generisanje statističkih podataka o materijalizovanom pogledu ili tabeli

- SQL

```
ANALYZE TABLE table [COMPUTE STATISTICS | ESTIMATE STATISTICS SAMPLE nn PERCENTS]
```

- PL/SQL

```
BEGIN
```

```
  DBMS_STATS.GATHER_TABLE_STATS (ownname =>'schema', tablename => 'table');
```

```
END;
```

# Materijalizacija pogleda

---

## SUBP Oracle: tehničke mogućnosti

- inicijalizacija / osvežavanje sadržaja materijalizovanog pogleda
  - PL/SQL paket DBMS\_MVIEW sa servisima
    - DBMS\_MVIEW.REFRESH
    - DBMS\_MVIEW.REFRESH\_ALL\_MVIEWS
    - DBMS\_MVIEW.REFRESH\_DEPENDENT
  - primer
    - DBMS\_MVIEW.REFRESH('<list\_of\_mvs>', 'methods')
    - DBMS\_MVIEW.REFRESH('sch1.mv1,sch2.mv2', 'fc')
      - osvežavanje svih pogleda, nabrojanih u okviru prvog argumenta
      - f – FAST način osvežavanja, ako je moguć
      - c – COMPLETE način osvežavanja, alternativno

# Materijalizacija pogleda

---

## SUBP Oracle: tehničke mogućnosti

- Summary Advisor
  - PL/SQL paket DBMS\_MVIEW sa servisima
    - pribavljanje podataka o kreiranim i potencijalnim materijalizovanim pogledima i mogućnostima za njihovu upotrebu
  - DBMS\_MVIEW.ESTIMATE\_MVIEW\_SIZE
  - DBMS\_MVIEW.EXPLAIN\_MVIEW
  - DBMS\_MVIEW.EXPLAIN\_REWRITE

# Materijalizacija pogleda

---

## SUBP Oracle: tehničke mogućnosti

- Summary Advisor
  - PL/SQL paket DBMS\_ADVISOR
    - Deo serverskog savetodavnog okruženja
    - Skup ekspertskih sistema koji prepoznaju i razrešavaju probleme loših performansi svih komponentni na serveru BP
  - DBMS\_ADVISOR.TUNE\_MVIEW
    - praćenje "workload" statističkih podataka, uz pomoć utility programa Oracle Trace
    - o agregiranim podacima za materijalizovane pogleda
      - koliko puta je pogled upotrebljavan u optimizaciji upita
      - "cost / benefit" podaci – procena isplativosti kreiranja materijalizovanog pogleda
    - generisanje preporuka u vezi kreiranja, spajanja, podele ili ukidanja materijalizovanih pogleda

# Materijalizacija pogleda

---

## SUBP Oracle: tehničke mogućnosti

- Summary Advisor
  - PL/SQL paket DBMS\_DIMENSION
    - pribavljanje podataka o OLAP dimenzionim strukturama
    - validacija OLAP dimenzionih struktura
  - PL/SQL paket DBMS\_ADVANCED\_REWRITE
    - zadavanje semantički ekvivalentnih opcija u formiranju naredbi upita
    - za potrebe optimizacije upita, zasnovane na prestrukturiranju naredbe upita

# Sadržaj

---

Šema DW BP i agregirani podaci

Upravljanje agregiranim podacima

Pogledi u DW sistemima

Materijalizacija pogleda

Vrste materijalizacije pogleda

Završne napomene

# Vrste materijalizacije pogleda

---

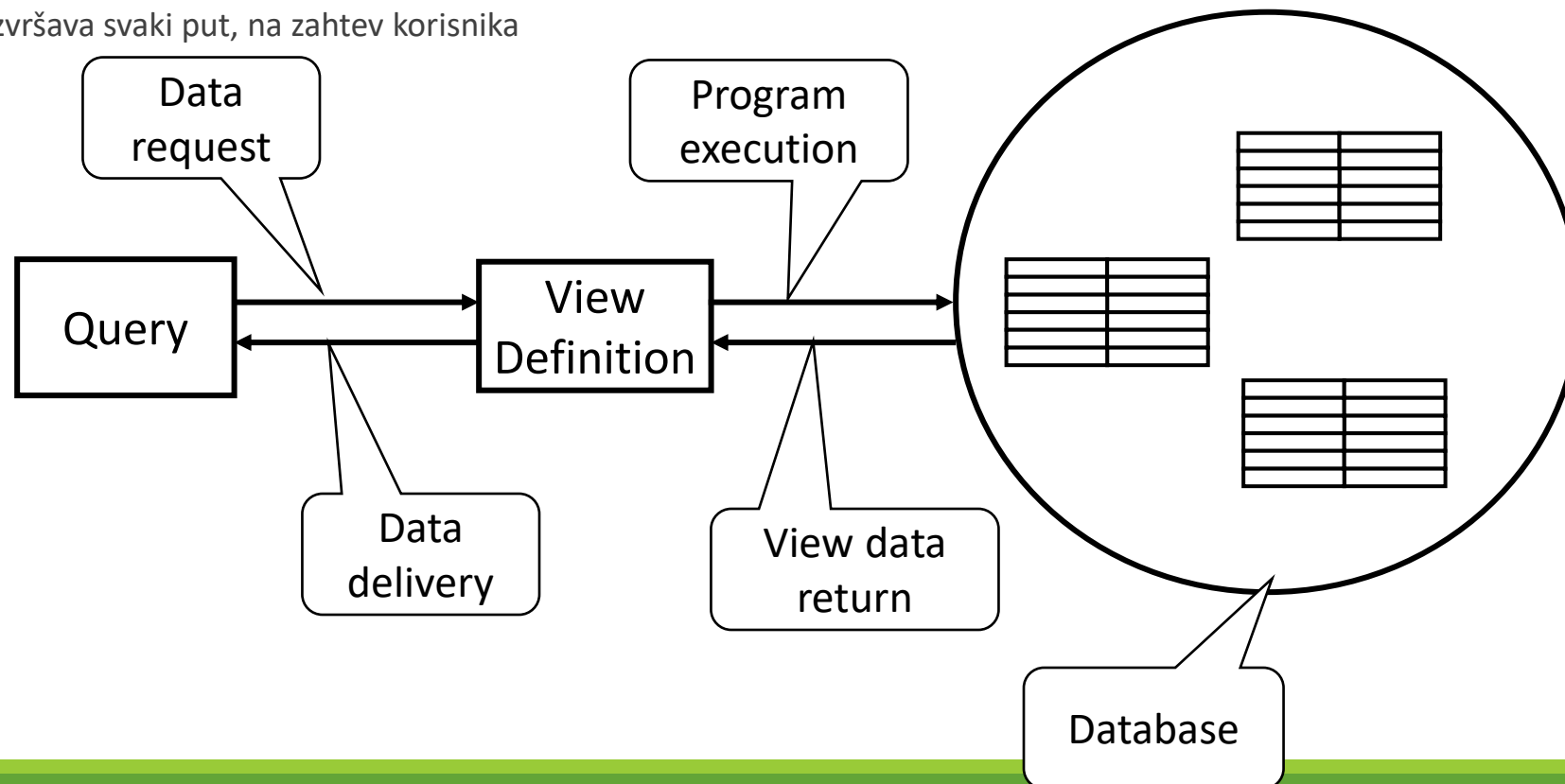
## Mogući oblici materijalizacije pogleda

- kroz program (pure program)
- izvedeni pogled (derived data view)
- kroz podatke (pure data view)
- kroz indeks (pure index)
- kombinacija podaci i indeks (hybrid data and index)
- agregirani pogled (aggregate view)
  - agregirani pogled sa spojem (join aggregate view)
  - pogled nad jednom tabelom (single table aggregate view)
  - pogled samo sa spojem (join only aggregate view)

# Vrste materijalizacije pogleda

## Kroz program (Pure Program)

- SQL definicija virtuelnog pogleda u samom programu
- upit se izvršava svaki put, na zahtev korisnika



# Vrste materijalizacije pogleda

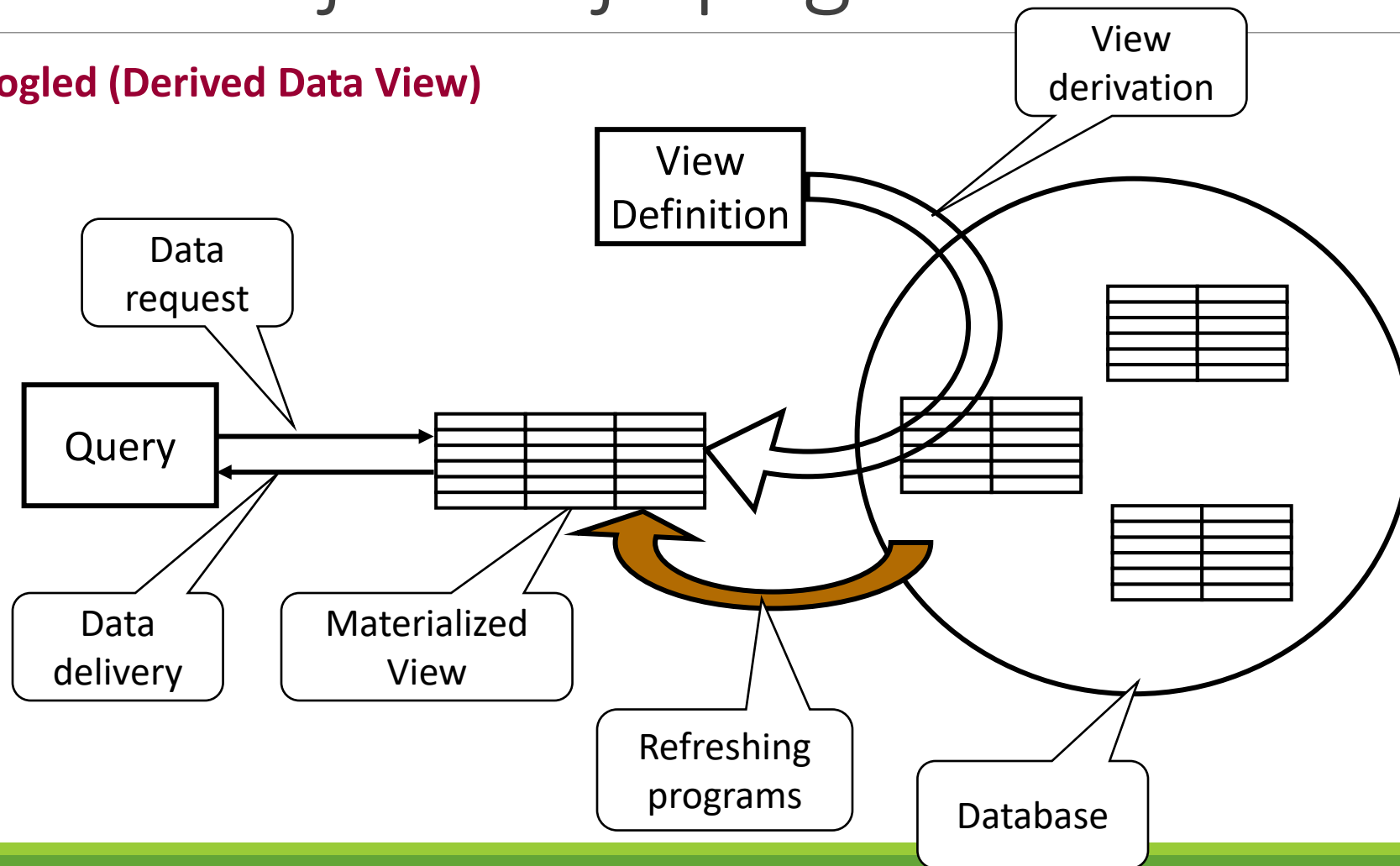
---

## Izvedeni pogled (Derived Data View)

- izvedeni podaci se generišu izvršavanjem upita i kreiranjem materijalizovanog pogleda
- materijalizovani pogled se dalje koristi u upitima
- materijalizovani pogled se (trenutno) osvežava
  - kao posledica ažuriranja baznih tabela, na kojima je zasnovan
- omogućena je optimizacija upita, zasnovana na upotrebi materijalizovanog pogleda
- očuvanje konzistencije podataka materijalizovanog pogleda zahteva posebno procesorsko vreme

# Vrste materijalizacije pogleda

## Izvedeni pogled (Derived Data View)



# Vrste materijalizacije pogleda

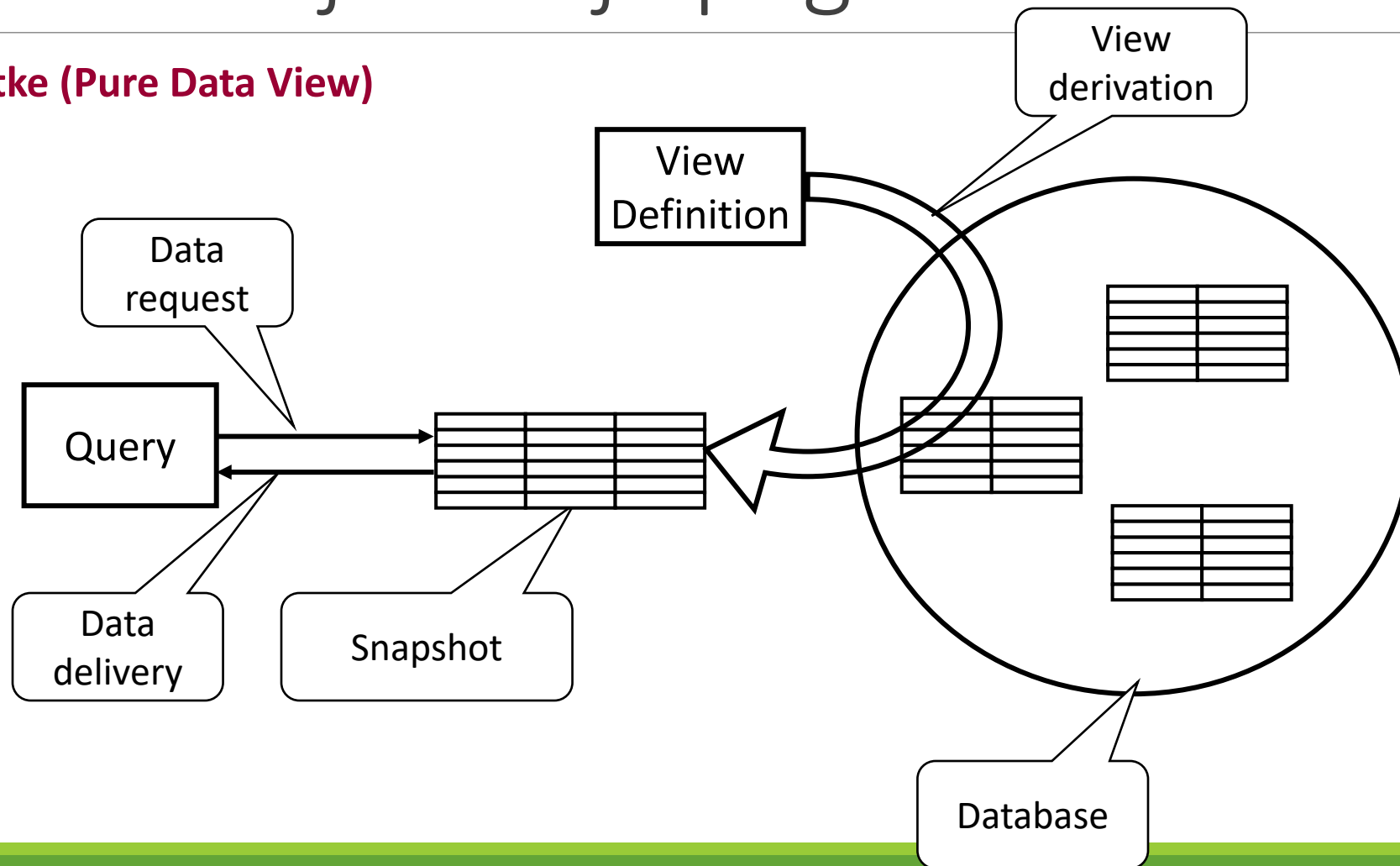
---

## **Kroz podatke (Pure Data View)**

- putem materijalizovanog pogleda, bez uključene procedure osvežavanja
  - naziva se i presek (slika stanja) - snapshot
- sadržaj materijalizovanog pogleda se periodično obnavlja
  - kompletnim izvršavanjem SELECT naredbe, ugrađene u definiciju pogleda
- očuvanje konzistencije podataka materijalizovanog pogleda zahteva posebno procesorsko vreme

# Vrste materijalizacije pogleda

## Kroz podatke (Pure Data View)



# Vrste materijalizacije pogleda

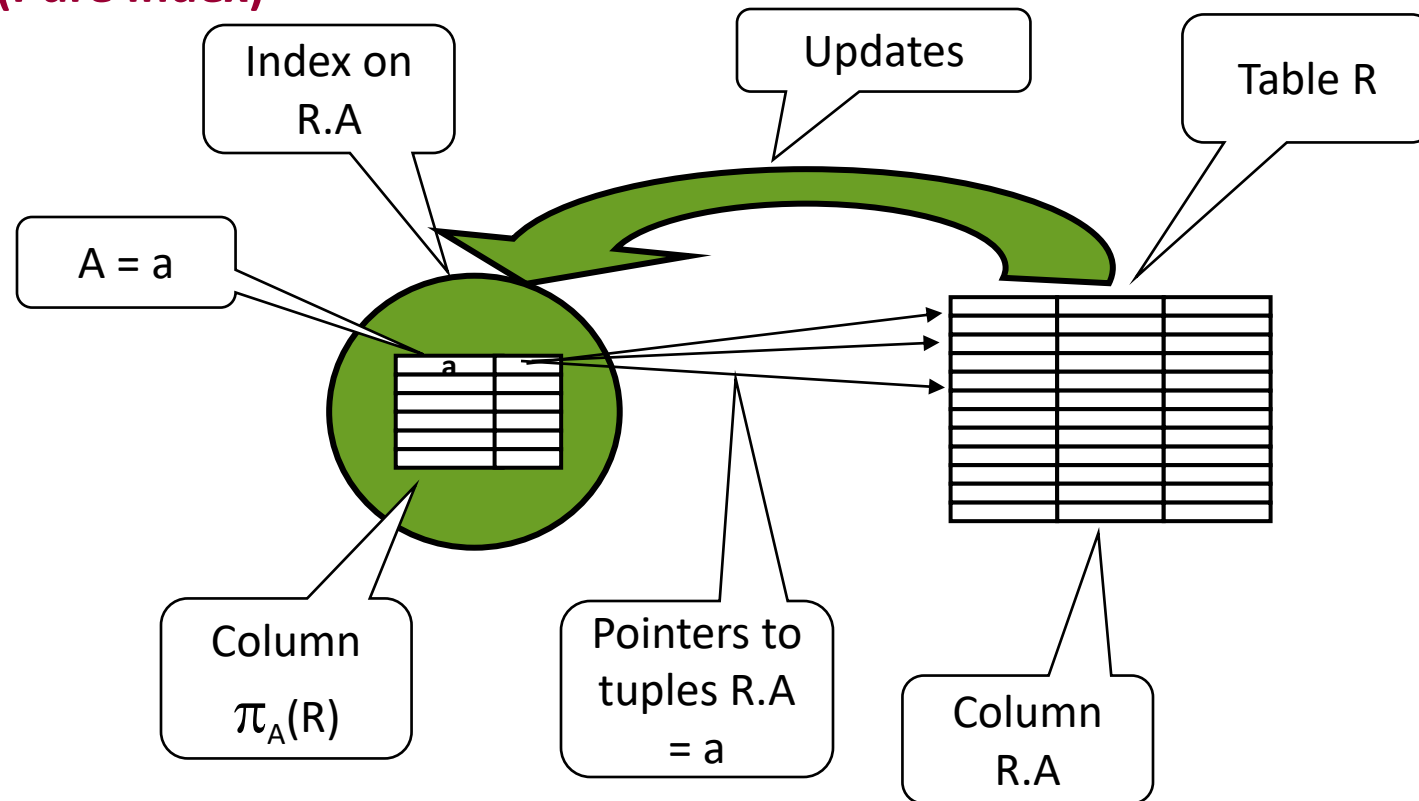
---

## Kroz indeks (Pure Index)

- indeks - jedna vrsta materijalizovanog pogleda
  - sadrži replicirane vrednosti atributa sa adresama odgovarajućih torki u baznoj tabeli
  - osvežava se trenutno u sklopu DML operacija ažuriranja
- moguće namene indeksa
  - poboljšanje performansi upita sa uslovom selekcije koji obuhvata indeksirane attribute
    - selekcija samo neophodnih torki iz indeksirane tabele
  - poboljšanje performansi upita sa operacijom spajanja
    - selekcija samo neophodnih torki iz indeksiranih tabela
    - posebno, ako su torke sa istim indeksiranim vrednostima grupisane u isti ili susedne blokove
  - kompletna realizacija određenih vrsta upita nad indeksom
    - karakteristično za DW sisteme

# Vrste materijalizacije pogleda

## Kroz indeks (Pure Index)



# Vrste materijalizacije pogleda

---

## **Kombinacija podaci i indeks (Hybrid Data and Index)**

- kombinacija tehnika materijalizacije "izvedeni pogled" i "kroz indeks"
- često zahtevane vrednosti atributa se materijalizuju
  - vrši se projekcija sadržaja tabele po izabranim atributima
  - od formirane projekcije pravi se materijalizovani pogled
  - torke u materijalizovanom pogledu povezuju se s originalnim torkama putem pokazivača
- očekuje se postizanje boljih performansi upita nad materijalizovanim atributima
  - pošto projektovani sadržaj tabele zauzima manje prostora
    - torke u materijalizovanom pogledu su manjeg kapaciteta
    - broj torki u materijalizovanom pogledu može biti manji

# Vrste materijalizacije pogleda

## Kombinacija podaci i indeks (Hybrid Data and Index)

Student Index

StudId	S_Name	Pointer

Student Table

StudId	S_Name	Address	Email	Phone

Course Table

CourId	C_Name	NoOfPts	Outline	Lecturer

Course Index

CourId	C_Name	Pointer

# Vrste materijalizacije pogleda

---

## Poređenje materijalizovani pogled / indeks

- sličnosti
  - zauzimaju poseban prostor
  - moraju biti osvežavani
  - obezbeđuju bolje performanse upita koji ih upotrebljavaju
  - upiti su nezavisni od njihove egzistencije
    - ne referenciraju se eksplicitno u tekstu upita
    - ako ih ima, optimizator upita može da ih iskoristi
    - u protivnom, upit će takođe biti korektno realizovan
- razlike
  - materijalizovani pogled može biti eksplicitno referenciran u tekstu upita, a indeks ne
  - materijalizovani pogled može biti dodatno indeksiran, a indeks ne

# Vrste materijalizacije pogleda

---

## **Agregirani pogled (Aggregate View)**

- najšire shvatanje agregacije podataka
- obuhvata sledeće slučajeve
  - celu DW BP nad zvezdastom šemom, ili nekim drugim oblikom šeme DW BP
    - i tabele dimenzija i tabele činjenica predstavljaju poglede nad nekim operativnim podacima, npr. iz OLTP BP
    - mere činjenica predstavljaju agregirane vrednosti operativnih podataka
  - oblik materijalizacije tipa "izvedeni pogled"
  - oblik materijalizacije tipa "kroz podatke"
- klasifikacija
  - agregirani pogled sa spojem (join aggregate view)
  - pogled nad jednom tabelom (single table aggregate view)
  - pogled samo sa spojem (join only aggregate view)

# Vrste materijalizacije pogleda

---

## Agregirani pogled sa spojem (Join Aggregate View)

```
CREATE MATERIALIZED VIEW shop_sales_mv
AS
  SELECT ShopName, SUM(SalesData) AS SumSales
  FROM Location s, Fact f
  WHERE s.ShopId = f.ShopId
  GROUP BY ShopName
```

# Vrste materijalizacije pogleda

---

## Agregirani pogled nad jednom tabelom (Single Table Aggregate View)

- roll-up nad tabelom činjenica

```
CREATE MATERIALIZED VIEW shop_avg_mv AS  
  SELECT ShopId, AVG (SalesData) AS AvgSales,  
          SUM(SalesData) AS SumSales,  
          COUNT(*) AS CountSales  
FROM Fact  
GROUP BY ShopId
```

- Napomena
  - funkcije COUNT i SUM su uključene da bi bilo omogućeno inkrementalno osvežavanje

# Vrste materijalizacije pogleda

---

## Agregirani pogled samo sa spojem (Join Only Aggregate View)

```
CREATE MATERIALIZED VIEW shop_sales_mv
AS
  SELECT s.ShopId, ShopName, t.TimeId, Month, Sales
  FROM Location s NATURAL RIGHT JOIN Fact f
              NATURAL LEFT JOIN Time t
```

- spoljnja spajanja sprečavaju gubljenje torki činjenica
  - u situaciji narušavanja ograničenja referencijalnih integriteta
  - kada toraka činjenice nije spojiva ni sa jednom torkom dimenzije

# Sadržaj

---

Šema DW BP i agregirani podaci

Upravljanje agregiranim podacima

Pogledi u DW sistemima

Materijalizacija pogleda

Vrste materijalizacije pogleda

Završne napomene

# Završne napomene

---

## **Ugrađeni materijalizovani pogled**

- materijalizovani pogled koji je ugrađen u definiciju drugog materijalizovanog pogleda
- materijalizovani pogled može referencirati druge materijalizovane poglede, kao i bazne tabele, u isto vreme

# Završne napomene

---

## Izbor pogleda za materijalizaciju

- često kompleksan problem
  - postoji širok spektar upita i pogleda koji bi se mogli upotrebaljavati u tim upitma
  - s druge strane, uvođenje materijalizovanih pogleda ima znatan uticaj na
    - povećanje potrebnog memorijskog prostora
    - produženje vremena potrebnog za održavanje, tj. osvežavanje podataka u DW BP
- cilj
  - izabrati mali, ali pogodan skup pogleda za materijalizaciju
  - koji može podržati najveći broj upita, označenih kao najznačajniji

# Završne napomene

---

## Primer

- dva međusobno "simetrična" upita

```
SELECT Category, SUM(SalesData)
FROM Product p, Sales s
WHERE p.ProdId = s.ProdId
GROUP BY Category
```

```
SELECT City, SUM(SalesData)
FROM Location l, Sales s
WHERE l.ShopId = s.ShopId
GROUP BY City
```

# Završne napomene

---

## Primer

- postoji više mogućih načina za ubrzavanje navedenih upita, kao na primer
  - formirati dva materijalizovana pogleda samo da pokriju spojeve za: (Sales JOIN Product) i (Sales JOIN Location)
  - formirati dva materijalizovana pogleda, svaki da kompletno pokrije po jedan upit
  - formirati sledeći materijalizovani pogled

```
CREATE MATERIALIZED VIEW ProdLocSales (Category, City, Total) AS  
SELECT Category, City, SUM(SalesData)  
FROM Sales s, Product p, Location l  
WHERE s.ProdId = p.ProdId AND s.ShopId = l.ShopId  
GROUP BY Category, City
```

# Završne napomene

---

## Primer

- inteligentni optimizator upita favorizuje jedan od dva moguća upita nad materijalizovanim pogledom
  - oba dobijena prestrukturiranjem polaznih SELECT naredbi

```
SELECT Category, SUM(Total)
FROM ProdLocSales
GROUP BY Category
```

```
SELECT City, SUM(Total)
FROM ProdLocSales
GROUP BY City
```

- Pitanje: koji od ova dva i zašto?
- Savet: pogledati GROUP BY klauzulu u ProdLocSales

# Sadržaj

Šema DW BP i agregirani podaci

Upravljanje agregiranim podacima

Pogledi u DW sistemima

Materijalizacija pogleda

Vrste materijalizacije pogleda

Završne napomene

# Pitanja i komentari

---



# Aggregacija podataka u Data Warehouse sistemima

---

MATERIJALIZOVANI POGLEDI  
AGREGACIONE FUNKCIJE

*Kraj prezentacije*