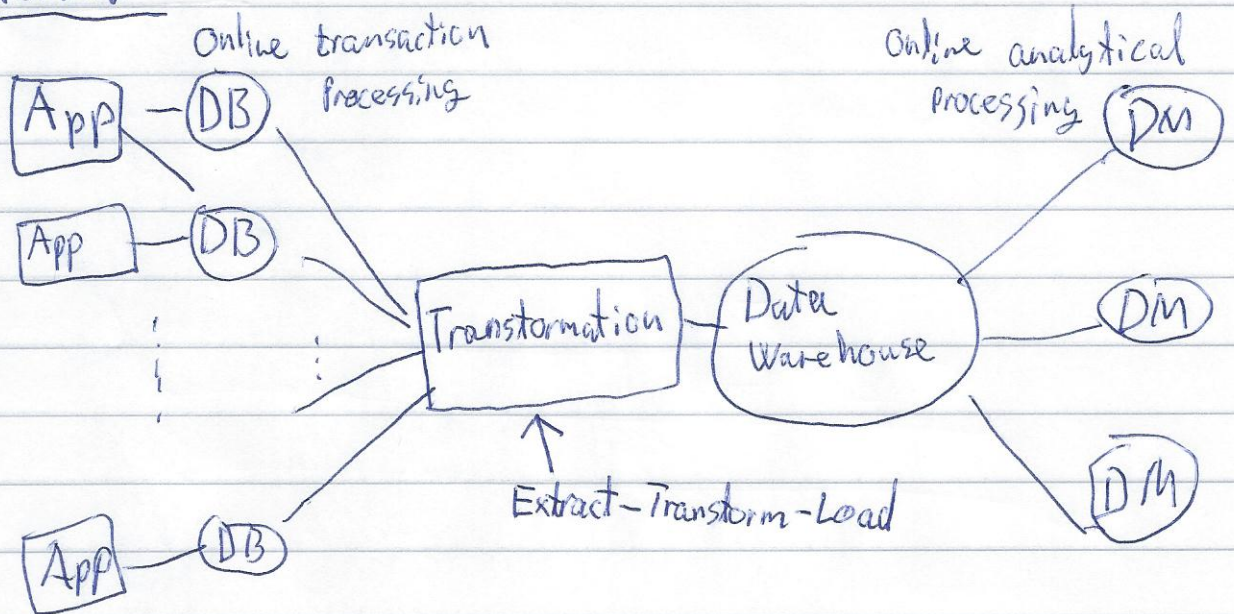


3

Multidimensional databases

Overblik



Hvortor

OLTP:

Ræn mange små queries

Hjppige opdateringer - kan altid løse og opdatere
mindre data volume

OLAP:

Få store opdateringer - enten løse-eller opdateringsstetstand

Få queries - men arbejder på større datasæt.

Model

Facts: what you want to look at: a sale

Dimensions: Beskrivelse af fact (tid, produkt, butik)

- Har et hierachy dag \rightarrow month - year - T

Measures: Hvad man ønsker at analysere: salgspris

- Kan ofte aggregeres (drill down, roll up, slice)

Cube?

Relational design (star)

Denormalized dimensions
Fast

uses known technology

TimeID	dag	month	year
1	11. november 2011	november 2011	2011

VS snowflake

Snowflake skal bruge joins

snowflake sparer plads i dimensionerne.

Typisk er 5% plads = dimensioner

~~SalesID~~

ProductID	TimeID	shopID	SalesPrice
1	1	1	99
	⋮		

Fact table

Each row is a sale event

Could be snapshot fact or cumulative

Dimension changes

Indtil videre har dimensionerne ikke ændret sig. Hvad hvis de gør?

Raw edit

Løsning 1: Ændring af størrelse på attribut

- Ret værdien, \therefore mister gammel info, + god til rettelse

Versioning

Løsning 2: opret ny række i dimension med nye værdier og henvis til denne ved nye facts.

+ mister ikke information, + queries er nemme \therefore kan ikke se hvornår ny række blev indsat.

Two attr

Løsning 3: Hav to værdier af attributer der ændrer sig.

+ muligt at se på to forskellige værdier, \therefore kun sikte to værdier \therefore ikke muligt at se hvornår værdierne ændrede sig.

special facts

Løsning 2.1: Indsæt speciel fact ved oprettelse af ny dimensions row.

+ kan se hvornår ændringen er sket, \therefore fylder mere.

Time stamping

Løsning 2.2: Timestamping

+ correct info i DW, \therefore større dimension tables.