Use Case 2
Online book Shopping

Use case

An Online book shopping is a very active e-commerce area. A large amount of customer orders is produced every day, and the Company “your best readings” can manage book sales across different stores, book stock on several warehouses, and books delivery.

The CEO wants to know what would be better: to open a new book store, to provide a better service to his customers or to open a new book data warehouse to reduce time book delivery. In addition, the book shop manager may also integrate discount book information from his competitors, he wants to compare pricing schemes, analyze market trends and make new business plans.

Suppose the already mentioned applications have all the required information to make decisions, but they were independently developed as transactional information systems within relational databases.

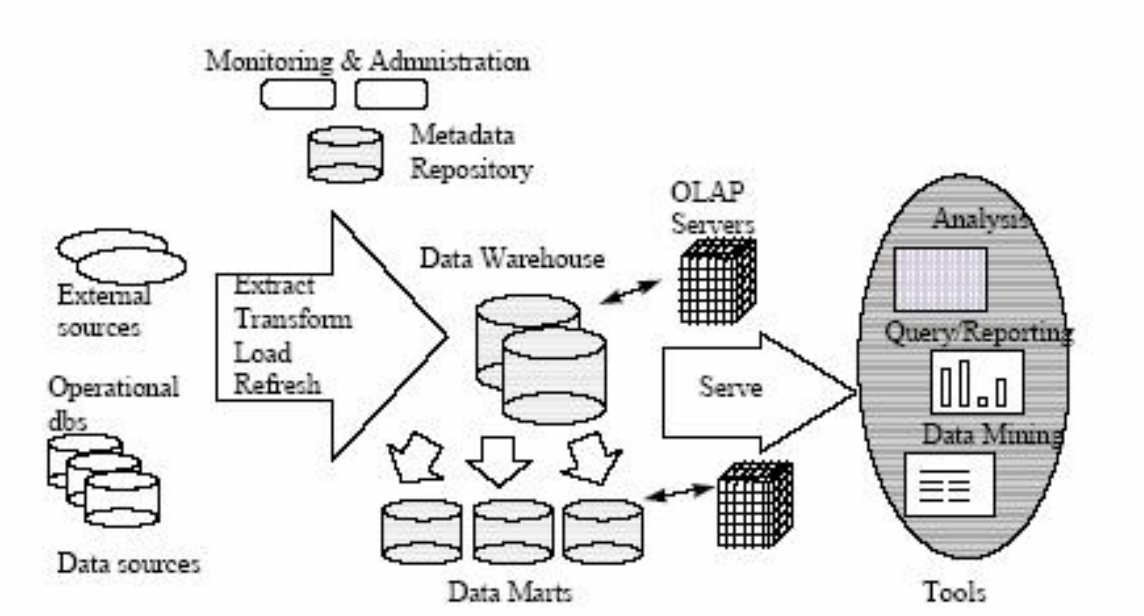
The CEO and the book shops manager would like to ask the following queries:

1.- Obtain the total daily sales by book and subregion
2.- Obtain annual sales per book
3.- Obtain which region had the most transactions during October 2017
4.- Obtain the average of the sale per ticket (invoice) per month in each region
5.- Get which books are bought together more frequently

Solution

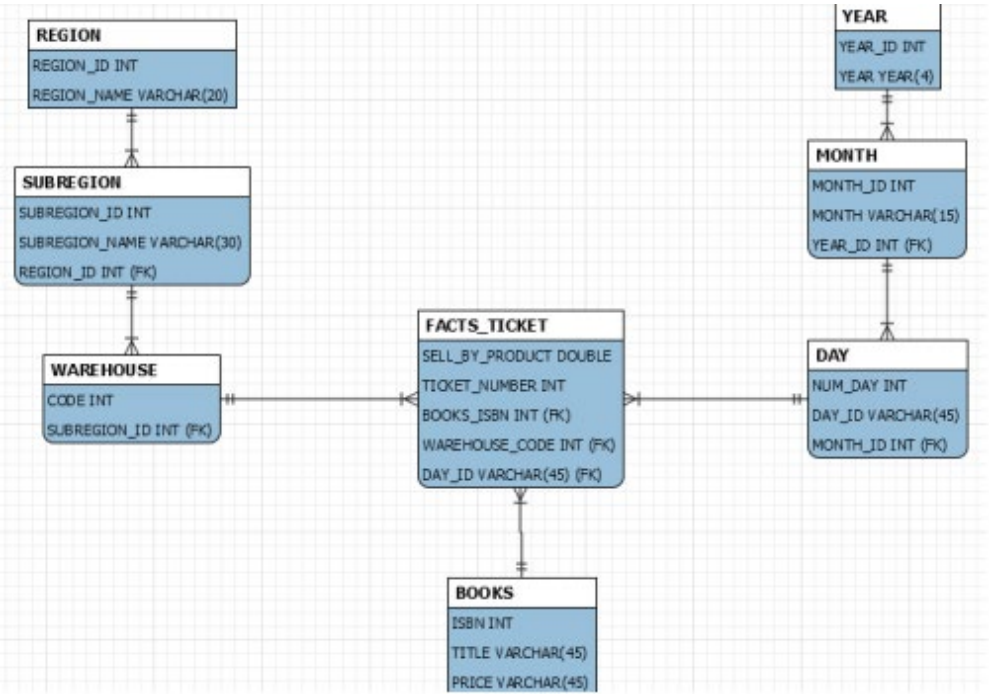
1. Typical architecture according to the type of information system

Since we must analyze the data, so OLAP system would be an ideal system for this use case.



2. Kind of database to implement and its design

The normalized snowflake schema with a columnar binary relational database. Here is the database schema for the same:



3. Answers to the queries required:

a) Obtain the total daily sales by book and subregion

```
select distinct s.subregion_name as "subregion",  d.day as "day", d.num_day as "num_day",  p.isbn
as "ISBN", p.title as "title",  sum(t.sell_by_product) as "total_sales"
from olap_bookshop.subregion s, olap_bookshop.day d, olap_bookshop.books p,
olap_bookshop.facts_ticket t, olap_bookshop.warehouse
where t.warehouse_code = warehouse.code
and warehouse.subregion_id = s.subregion_id
and t.day_id = d.day_id
and t.books_isbn = p.isbn
group by s.subregion_name, d.day, d.num_day, p.isbn
order by num_day asc;
```

a) Obtain annual sales per book

```
Select distinct  an.year as "year", p.isbn as "isbn", p.title as "title", sum(t.sell_by_product) as
"total_sales"
from olap_bookshop.year an, olap_bookshop.books p, olap_bookshop.facts_ticket t
where t.books_isbn = p.isbn
group by an.year, p.isbn
order by  p.isbn asc;
```

a) Obtain which region had the most transactions during October 2017

```
select region_name, max(count)
from (
        select r.region_name,  count(f.ticket_number) as count
        from olap_bookshop.region r, olap_bookshop.facts_ticket f,  olap_bookshop.day d,
        olap_bookshop.month m, olap_bookshop.warehouse t, olap_bookshop.subregion s
        where m.month = 'october'
        and f.day_id = d.day_id
        and d.month_id=m.month_id
        and f.warehouse_code = t.code
        and t.subregion_id = s.subregion_id
        and s.region_id = r.region_id
        group by r.region_name)
group by region_name;
```

a) Obtain the average of the sale per ticket (invoice) per month in each region

```
select r.region_name, m.month, avg(f.sell_by_product)
from olap_bookshop.region r,  olap_bookshop.facts_ticket f,  olap_bookshop.day d,
olap_bookshop.month m, olap_bookshop.warehouse t, olap_bookshop.subregion s
where f.day_id = d.day_id
and d.month_id=m.month_id
and f.warehouse_code= t.code
and t.subregion_id = s.subregion_id
and s.region_id = r.region_id
```

```
group by r.region_name, m.month
order by r.region_name;
```

a) Get which books are bought together more frequently

```
Select prod1.title, pr.title
from olap_bookshop.books pr, olap_bookshop.facts_ticket t,
        (
        select f.ticket_number, p.title,
        from olap_bookshop.books p, olap_bookshop.facts_ticket f
        where p.isbn = f.books_isbn
        group by f.ticket_number, p.title
        ) as prod1
where t.ticket_number = prod1.ticket_number
```

4. According to the proposed solution, how to implement that as books are sold the inventory is deregistered and the amount of the sale corresponds to the merchandise sold?

The student proposed descriptive queries using pivot, drilldown, rollup a transaction AND the last question was answered by SQL or by any data mining technique such as association, clustering, etc.

5. How does it support scalability

The workload depends on the Number of requests per second, the ratio of reads to writes in a database, the number of simultaneously active users, the hit rate on a cache, etc.

As system grows, we can improve hardware capabilities - more RAM, more CPU, faster HDD. However, if hardware reaches its physical limits. There are other options: For instance, creating a clustering, swinging, and dispatching infrastructure to support Multi Parallel.

6. How does it support maintainability

Maintainability is given by a multidimensional or datawarehousing process to obtain a normalized data model snowflake or a denormalized data model such as star schema that can make possible future changes.

7. How does it support security and reliability

Security and reliability can be implemented with database tools.