CHAPTER 11

Data Analytics

Solutions for the Practice Exercises of Chapter 11

Practice Exercises

11.1

Answer:

In a destination-driven architecture for gathering data, data transfers from the data sources to the data warehouse are based on demand from the warehouse, whereas in a source-driven architecture, the transfers are initiated by each source.

The benefits of a source-driven architecture are

- Data can be propagated to the destination as soon as they become available. For a destination-driven architecture to collect data as soon as they are available, the warehouse would have to probe the sources frequently, leading to a high overhead.
- The source does not have to keep historical information. As soon as data are updated, the source can send an update message to the destination and forget the history of the updates. In contrast, in a destination-driven architecture, each source has to maintain a history of data which have not yet been collected by the data warehouse. Thus storage requirements at the source are lower for a source-driven architecture.

On the other hand, a destination-driven architecture has the following advantages.

- In a source-driven architecture, the source has to be active and must handle error conditions such as not being able to contact the warehouse for some time. It is easier to implement passive sources, and a single active

warehouse. In a destination-driven architecture, each source is required to provide only a basic functionality of executing queries.

- The warehouse has more control on when to carry out data gathering activities and when to process user queries; it is not a good idea to perform both simultaneously, since they may conflict on locks.

11.2

Answer:
The relation would be stored in three files, one per attribute, as shown below. We assume that the row number can be inferred implicitly from position, by using fixed-size space for each attribute. Otherwise, the row number would also have to be stored explicitly.

building
Packard
Painter
Taylor
Watson
Watson

room_number
101
514
3128
100
120

capacity
500
10
70
30
50

11.3

Answer:
FILL IN

11.4

Answer:
query:

```
select store-id, city, state, country,
        date, month, quarter, year,
        sum(number), sum(price)
from sales, store, date
where sales.store-id = store.store-id and
        sales.date = date.date
groupby rollup(country, state, city, store-id),
        rollup(year, quarter, month, date)
```

11.5

Answer:
FILL IN

11.6

Answer:
FILL IN