Design Issues

- *When and how to gather data*

  - **Source driven architecture**: data sources transmit new information to warehouse

    - either continuously or periodically (e.g. at night)

  - **Destination driven architecture:** warehouse periodically requests new information from data sources

  - **Synchronous** vs **asynchronous replication**

    - Keeping warehouse exactly synchronized with data sources (e.g. using two-phase commit) is often too expensive

    - Usually OK to have slightly out-of-date data at warehouse

    - Data/updates are periodically downloaded form online transaction processing (OLTP) systems.

- *What schema to use*

  - Schema integration