IF5014 – Perkembangan Basis Data Terdistribusi

Wikan Danar
Departemen Teknik Informatika
Institut Teknologi Bandung

Current Issues

- Data Delivery Alternatives
- Data Warehousing
- World Wide Web
- Push-Based Technologies
- Mobile Databases

Data Delivery Alternatives

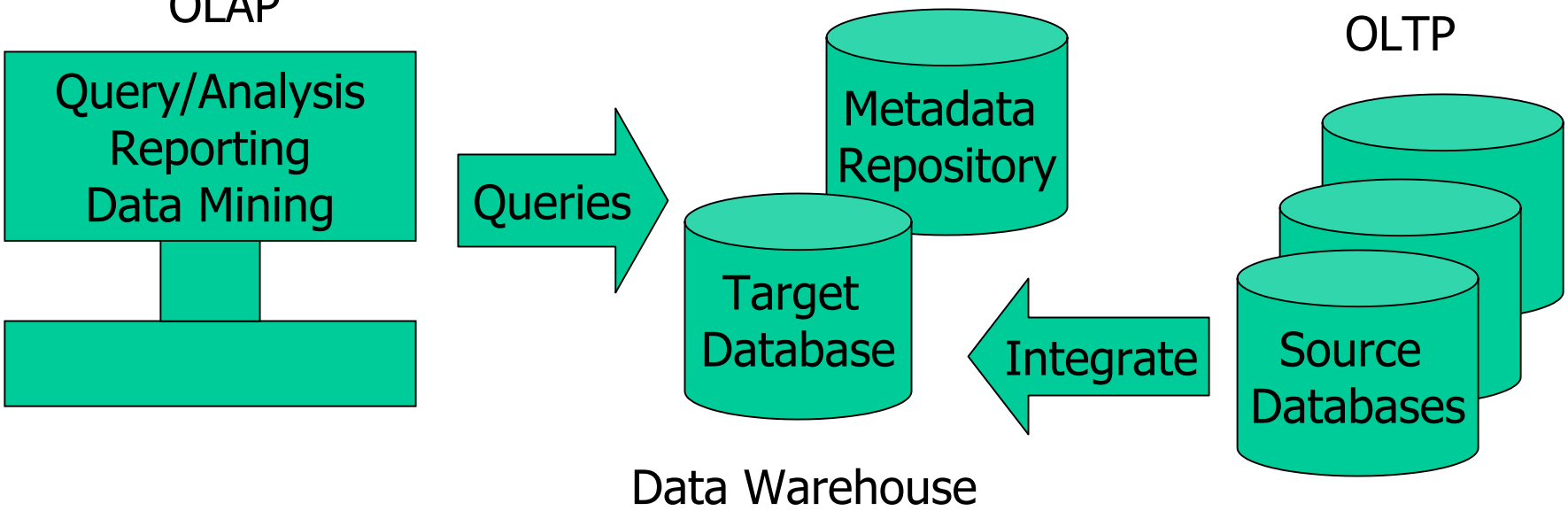
Three orthogonal dimensions:

- Delivery modes
  - Pull-only: data transfer initiated by a client pull
  - Push-only: data transfer initiated by a server push
  - Hybrid: combines client-pull and server-push
- Frequency
  - Periodic: regular & pre-specified repeating schedule
  - Conditional: certain conditions satisfied
  - Ad-hoc / irregular: most performed in pull-based
- Communication Methods
  - Unicast: one-to-one
  - One-to-many: server send to a number of clients

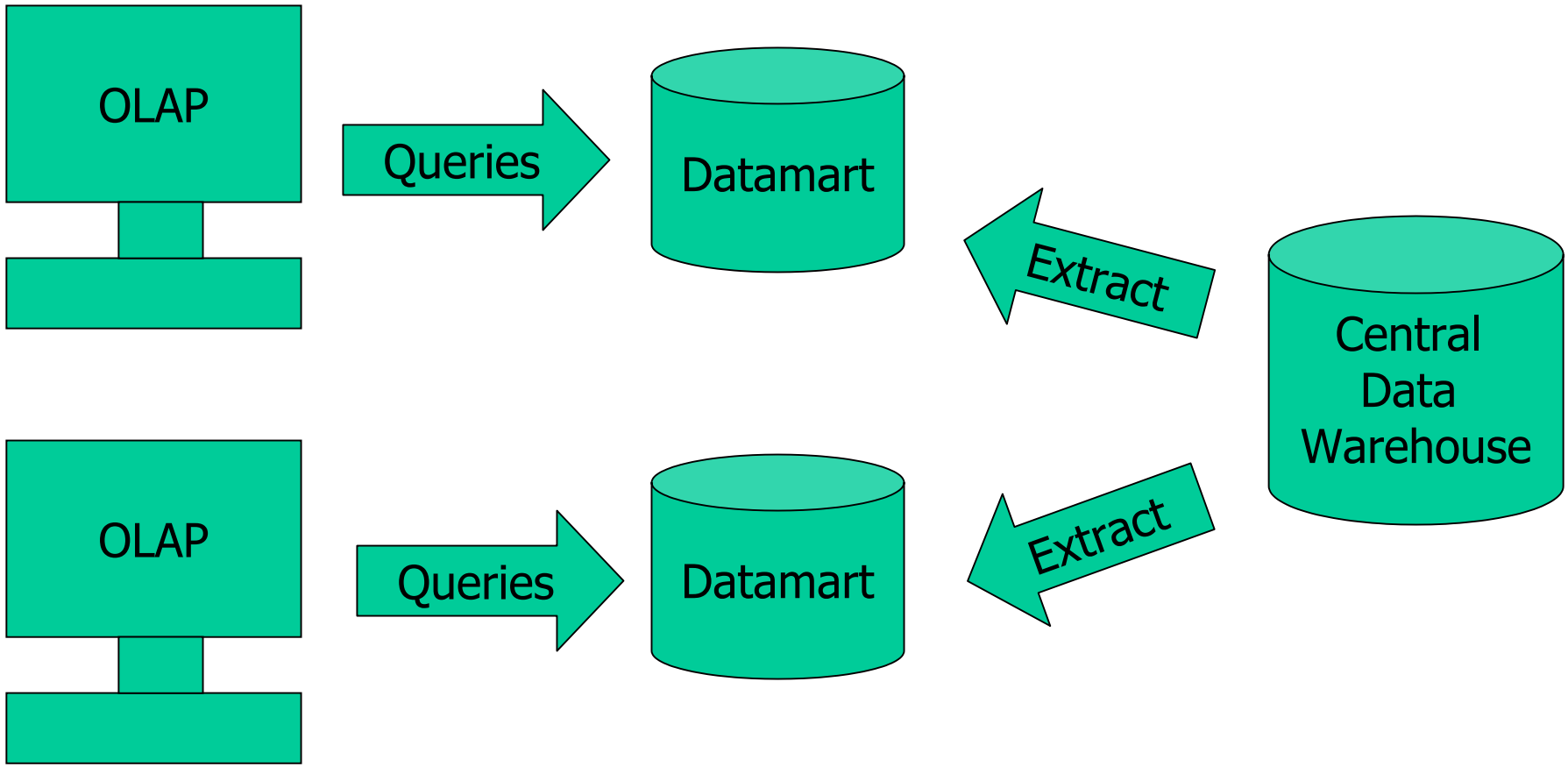
Data Warehousing

- Data warehousing refers to *a collection of technologies aimed at improving decision making*
- Data Warehouse: a subject-oriented collection of data integrated from various operational databases
- Information is classified by subjects of interest to business analysts, such as customers, products, and accounts

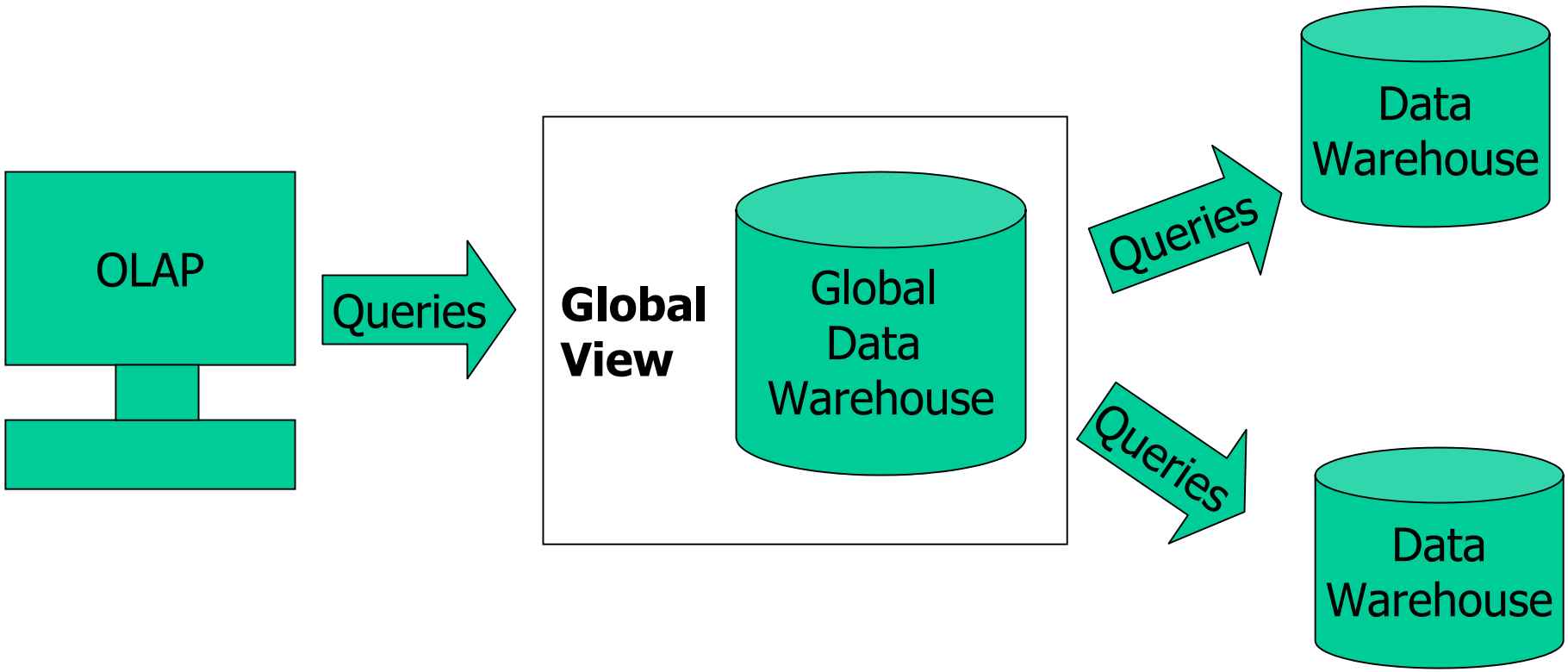
Architecture of a Data Warehouse



Centralized Data Warehouse



Decentralized Data Warehouse



OLAP Data Model

- OLAP data model is *multidimensional*
- Data is represented by a multidimensional array of *numeric measures*, such as sales or revenue, which is useful for analysis
- Aggregation operators:
  - *Roll-up*: increases the level of aggregation by moving up the dimension hierarchy
  - *Drill-down*: increase the level of detail
  - *Slice-and-dice*: corresponds to select-project on a subset of the dimensions

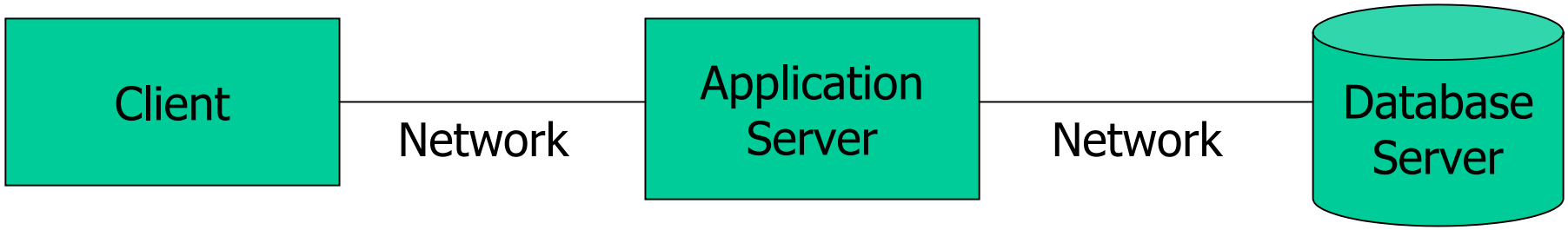
OLAP Servers

- *Multidimensional OLAP* (MOLAP) servers: directly support OLAP operations on multidimensional data structures
- *Relational OLAP* (ROLAP) servers: extend relational databases to support OLAP operations

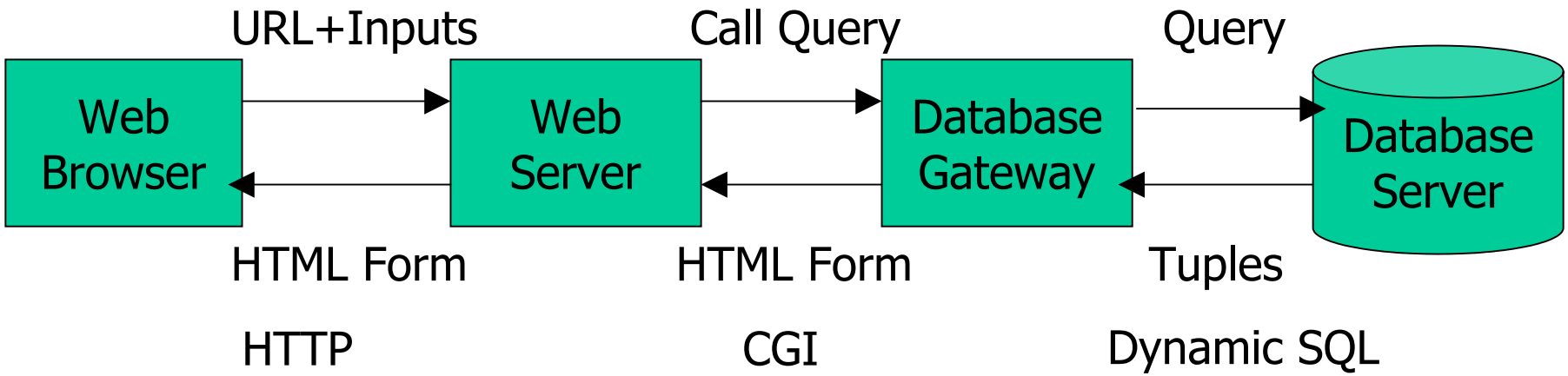
World Wide Web

- Web architecture: client-server
- Standard communication protocol: HTTP (HyperText Transfer Protocol) implemented on top of TCP/IP
- With HTTP, any client browser can request a document on a Webserver using its URL (*uniform resource locator*)
- Home page: the entry point to a Website identified by URL
- HTML (HyperText Markup Language): the standard page description language for the Web
- XML (eXtended Markup Language): a subset of SGML (Standard Generalized Markup Language) and provides a clean separation between content structure and presentation

Database Access in WWW



Three-Tier Client/Server Architecture



Database Access from a Web Browser

Data Structure in Internet

- Fully unstructured data: raw text and image
- Fully structured data: relational or object databases
- Semistructured data: HTML and SGML data

Object Exchange Model (OEM)

An OEM object consists of

1. A label which is the name of the object class
2. A type which is either atomic (integer, string, etc) or set
3. A value which is either atomic or a set of objects
4. An optional object identifier

Architecture for Information Integration

Issues for information integration:

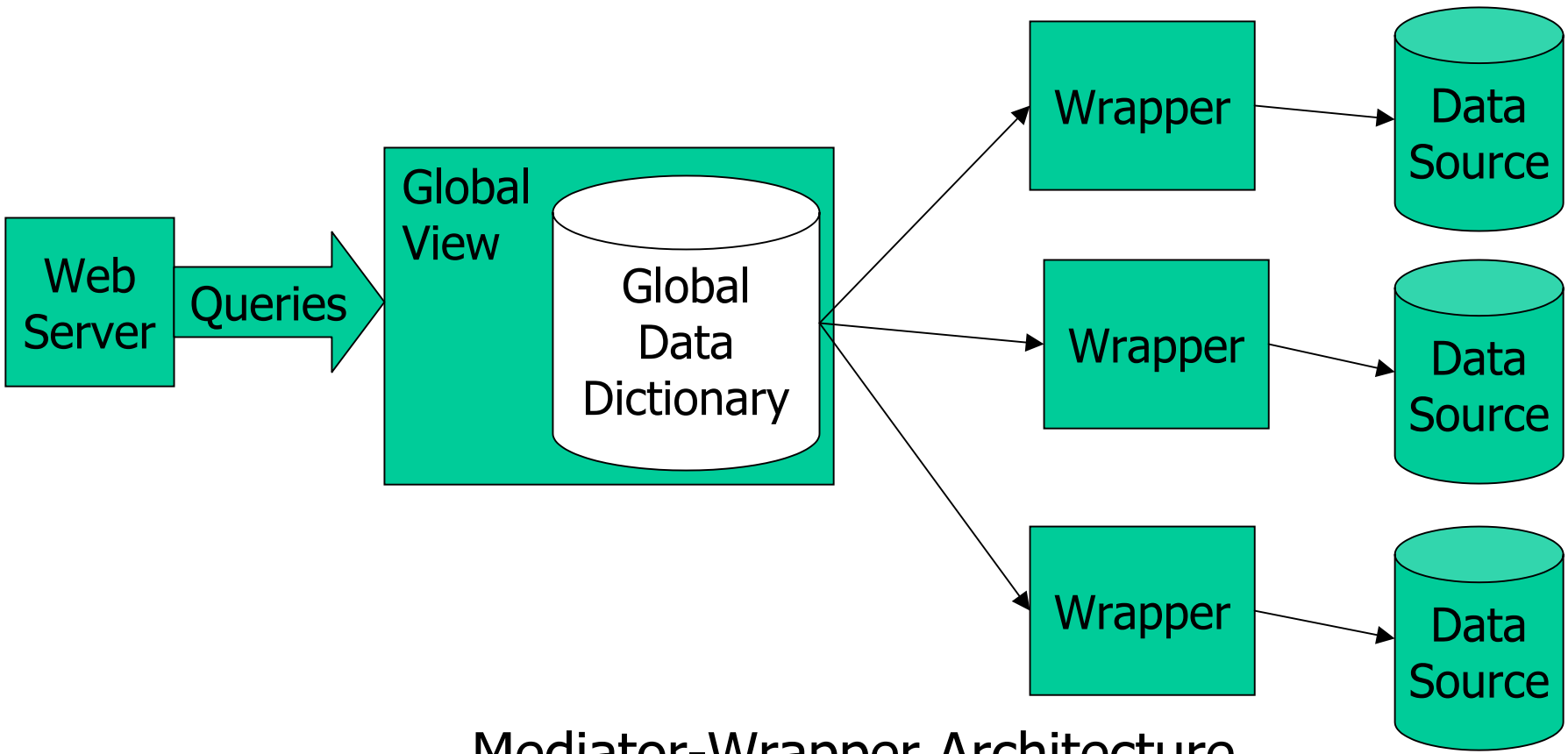
- Number of data sources may be very high
- The space of data sources is very dynamic
- Data sources may have different computing capabilities
- Data sources may be unstructured or semistructured

Wrappers and Mediators

Multidatabase architecture with data source *wrappers* and *mediators* are used to address previous problem

- For each data source, a *wrapper* exports some information about its source schema, data, and query processing capabilities
- A *mediator* centralizes the information provided by the wrappers in a unified view of all available data, decomposes the user query in smaller queries, gathers the partial results and computes the answer to the user query

Mediator-Wrapper Achitecture



Mediator-Wrapper Architecture

Mobile Databases

- A wireless network consists of a "wireline" (fixed) backbone network on which a number of control stations are located
- Each control station coordinates the communication from a mobile computer in its respective cell to another mobile computer in the same cell, or in another cell, or to a stationary computer on the wireline network

Characteristics of Mobile Environment

1. Communication: conducted over wireless networks which are prone to disconnections, noise, echo, and low bandwidth
2. Mobility: address migration, maintenance of directories, and difficulty in locating stations
3. Portability: places restrictions on the type of equipment that can be used in these environment

Characteristics of Mobile Environment (2)

1. The wireless network have restricted bandwidth
2. The power supplies (i.e. batteries) in mobile stations have limited life-times.
3. Because of power restrictions, mobile stations are not available as widely as stationary ones
4. As the name suggests, mobile stations move in these systems