What is Distributed Data Processing?
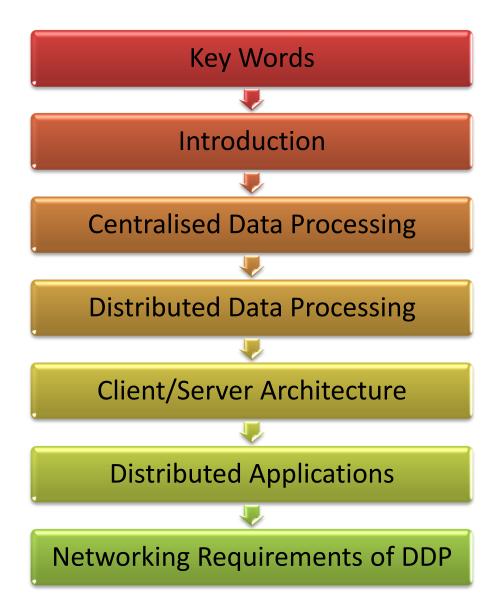
Eva Bezděková

Content

Source of information

- **Encyclopedia of Computer Science**
  - ISBN 0-333-77879-0
- **PDF presentation**
  - web.iku.edu.tr/courses/ee/ee102/ee102/lecture%20notes/network_add.pdf
- **Dictionary of Computing**
  - ISBN 0-747-56622-4
- **Others**
  - www.thefreedictionary.com/distributed+data+processing
  - www.wikipedia.org

Key words

- Network
- Distance
- Data
- Master computer
- Slave computer
- Communication

- Cluster of computers
- Central location
- Dedicated resources
- Processing approaches
- Horizontal Partitioning
- Vertical Partitioning

Introduction

- Distributed data processing allows multiple computers to be working among multiple geographically separate sites where local computers handle local processing needs

Intoduction

- One computer is designed as the master computer

- Up to 99 slave computers can be used to communicate with the central computer and perhaps between each other

- Linked by communication network

Introduction

- Term Distributed Data Processing was firstly used to describe distribution of multiple computers throughout an organization in contrast to a centralised system

Centralized Data Processing

- Centralised computers, processing, data, control, support

Centralized Data Processing

- support is provided by one cluster of computers, generally large computers, located in a central data processing facility

- 1970's - organizations implemented centralized systems, with a mainframe computer doing all the processing at a central location for the whole company

Centralized Data Processing

- **Centralised data:** Most data is stored, accessible and controlled at a central facility

- **Centralised control:** Control of data, applications, processes and infrastructure maintained centrally

- **Centralised support:** Technical support and development staff located in a central facility

Advantages of CDP

- Economy for equipment and personal

- Lack of duplications

- Ease in enforcing standards, security

Distributed Data Processing

- The Distributed Systems is the opposite to the centralized system:

1. computers installed at different sites

2. each of them performing independent data processing

3. each computer is specialized to perform a range of activities (marketing, promotion….)

Distributed Data Processing

- Allow greated flexiblity in structure

- More redundancy

- More autonomy

Why is DDP Increasing?

- Dramatically reduced hardware costs

- Increased desktop power

- Improved user interfaces

- Ability to share data across multiple servers

Reasons for DDP

- **Need for new applications**
  - On large centralised systems, development can take years
  - On small distributed systems, development can be component-based and very fast
- **Need for short response time**
  - Centralised system result is contention among users and processes
  - Distributed systems provide dedicated resources

Key issues

- How does it affect end-users?

- How does it affect management?

- How does it affect productivity?

Benefit of DDP

- Distance and local independence

- End-user productivity

- Increased user involvement and control

- Privacy and security

- Organisational patterns

- Resource sharing

- Incremental growth

- Availability

- Responsiveness

- Vendor independence

Drawbacks of DDP

- More components and dependence on communication means more points of failure

- Difficulties in failure diagnosis

- Incompatibility of components

- Incompability of data

- More complex management and conrol

- Difficulty controlling information resources

- Duplication of effort

Client/Server Architecture

- One of data processing approaches

- Combines advantages of distributed and centralized computing

- Cost-effective, achieves economic of scale

- Flexible, scalable approach

Intranet

Internet

- A specialized form of client/server architecture

- Often use Internet standards

- Content is accesible only to internal users

- Provides access to outside users
  - Customers, suppliers

Distributed applications

- **Horizontal partitioning**
  - Different application on different systems
  - One application replicated on systems
  - Example: Office automation
- **Vertical partitioning**
  - One application dispersed among system
  - Example: retail chain

Networking Requirements of DDP

- **Connectivity requirements**
  - What link between components are necessary?
- **Availability requirements**
  - Percentage of time application or data is available to users
- **Performance requirements**
  - Response time requirements

Thank you for your attention!