

# 分布式数据库系统

---

讲解人：李鸿岐

# 分布式数据库管理和相关概念

---

## 发展起源

□ 分布数据库管理系统兴起于70年代中期，推动其发展来自两方面：

一方面是应用需求，另一方面是硬件环境的发展。

如：全球及全国范围内的航空/铁路/旅游订票系统、银行通存通兑系统、水陆空联运系统、跨国公司管理系统、连锁配送管理系统等。

□ 这些应用都涉及地理上分布的团体、组织的局部业务管理和系统全局管理，采用成熟的集中式数据库管理系统已无法实现应用需求（分散在不同地理位置上数据的一致性、完整性及有效性）。

□ 在硬件环境上提供了功能强大的计算机和成熟的广域范围公用数据网及局域范围局域网的硬件环境支持。

## 基本概念

---

- ❑ 分布式数据库系统（Distributed Database System, DDBS）是地理上分散而逻辑上集中的数据库系统。即通过计算机网络将地理上分散的各局域结点连接起来共同组成一个逻辑上统一的数据库系统。因此可以说：分布式数据库系统是**计算机网络**技术和**数据库**技术相结合的产物。
- ❑ **计算机网络**允许对数据进行分布式处理，传统**数据库**则注重对数据的集中管理和受控访问。
- ❑ 分布式数据库系统与集中数据库系统一样，包含两个重要部分：分布式数据库（DDB）和分布式数据库管理系统（DDBMS）。

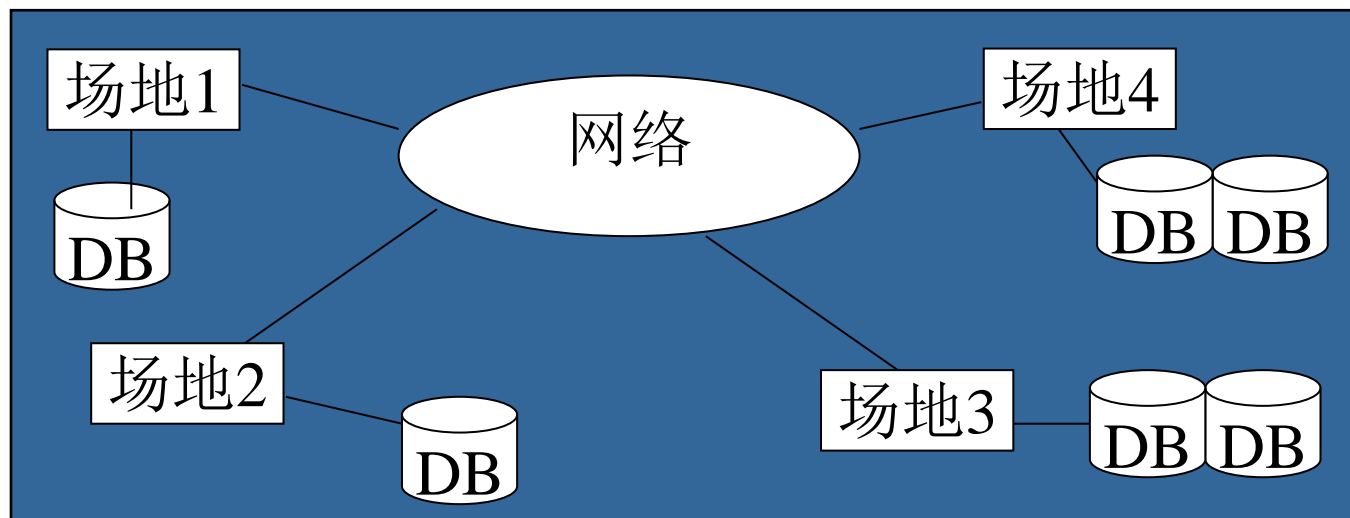
# 基本概念

## □ 节点/场地 (node/site)

被连接的逻辑单位称为节点 (node) 或场地 (site)。节点/场地可以是单独一台计算机，也可以是由局域网组成。

## □ 分布式数据库 (DDB) :

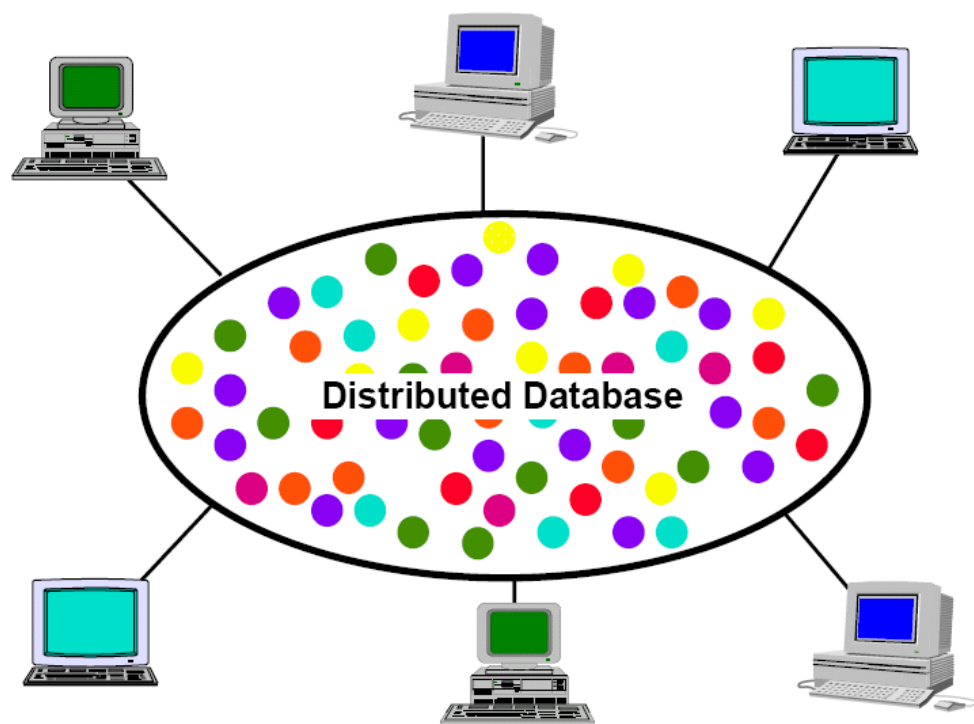
分布在一个计算机网络上的多个逻辑相关的数据库的集合。也就是说，分布式数据库是一组结构化的数据集合，逻辑上属于同一系统，物理上分布在计算机网络的各个不同节点上。



## □ 分布式数据库管理系统 (DDBMS) :

管理分布式数据库的软件系统。

### 分布式数据库—用户视图



EMP

ENO	ENAME	TITLE
E1	J. Doe	Elect. Eng.
E2	M. Smith	Syst. Anal.
E3	A. Lee	Mech. Eng.
E4	J. Miller	Programmer
E5	B. Casey	Syst. Anal.
E6	L. Chu	Elect. Eng.
E7	R. Davis	Mech. Eng.
E8	J. Jones	Syst. Anal.

ASG

ENO	PNO	RESP	DUR
E1	P1	Manager	12
E2	P1	Analyst	24
E2	P2	Analyst	6
E3	P3	Consultant	10
E3	P4	Engineer	48
E4	P2	Programmer	18
E5	P2	Manager	24
E6	P4	Manager	48
E7	P3	Engineer	36
E7	P5	Engineer	23
E8	P3	Manager	40

PROJ

PNO	PNAME	BUDGET
P1	Instrumentation	150000
P2	Database Develop.	135000
P3	CAD/CAM	250000
P4	Maintenance	310000

PAY

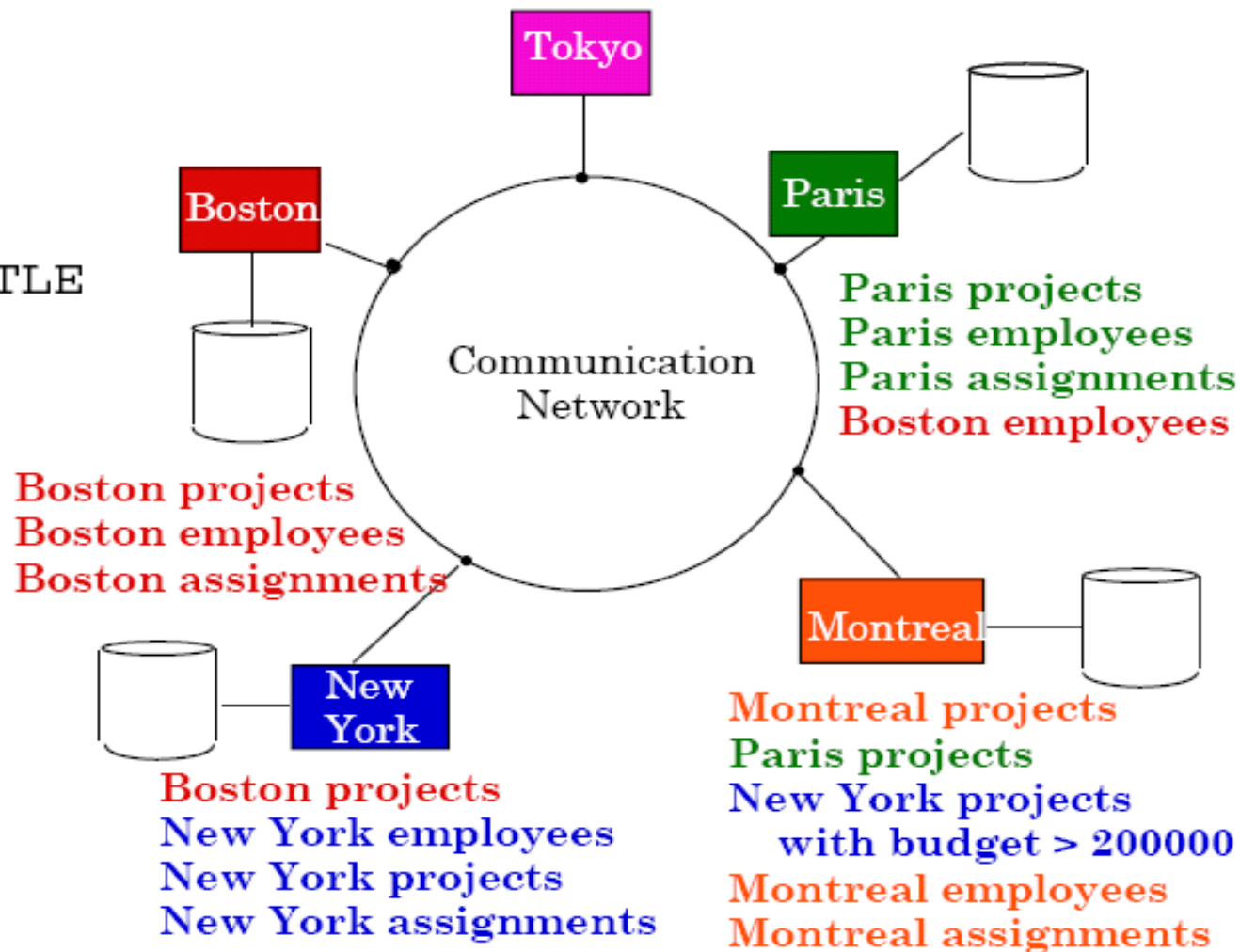
TITLE	SAL
Elect. Eng.	40000
Syst. Anal.	34000
Mech. Eng.	27000
Programmer	24000

# 基本概念

```
SELECT ENAME, SAL
FROM   EMP, ASG, PAY
WHERE  DUR > 12
AND    EMP.ENO = ASG.ENO
AND    PAY.TITLE = EMP.TITLE
```

## □ 分布式数据库的两大特性:

- 数据透明性
- 场地自治性



# 基本概念

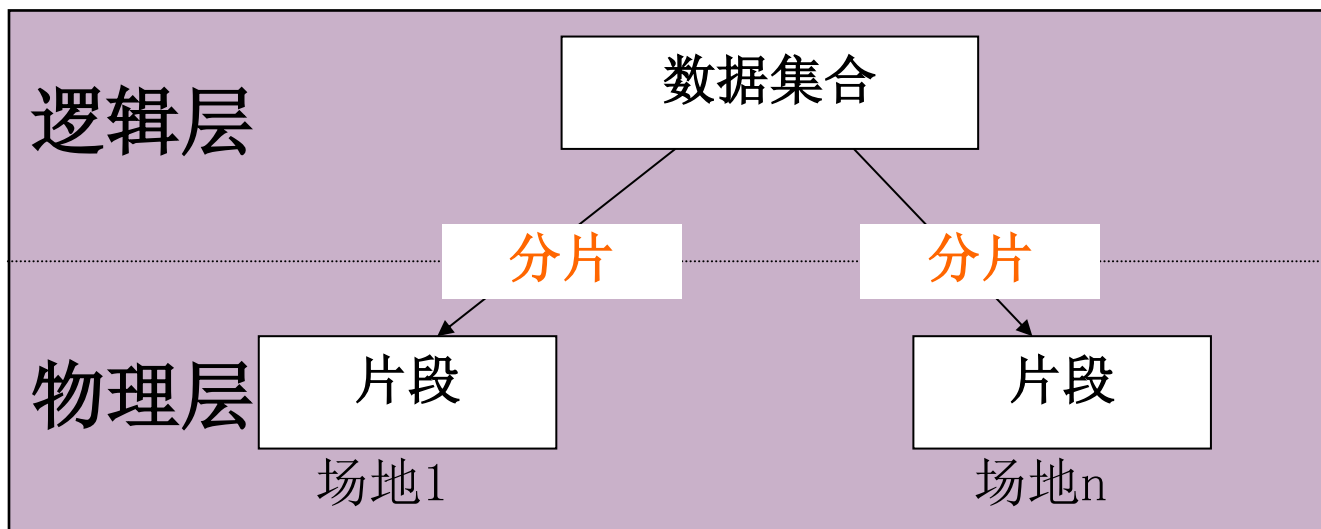
位置透明性

命名透明性

## ■ 数据透明性：

□ 基于分布库的划分层次：全局外层（用户层）、全局概念层、局部概念层和局部内层。

□ 可知：应用程序与系统实际数据组织相分离，即数据具有独立性或透明性。具体体现为分布透明性、复制透明性和分片透明性。



□ 分布透明性/数据组织透明性/网络透明性：全局用户看到的是全局数据模型的描述，用户像使用集中数据库一样，不需考虑数据的存储场地和操作的执行场地。

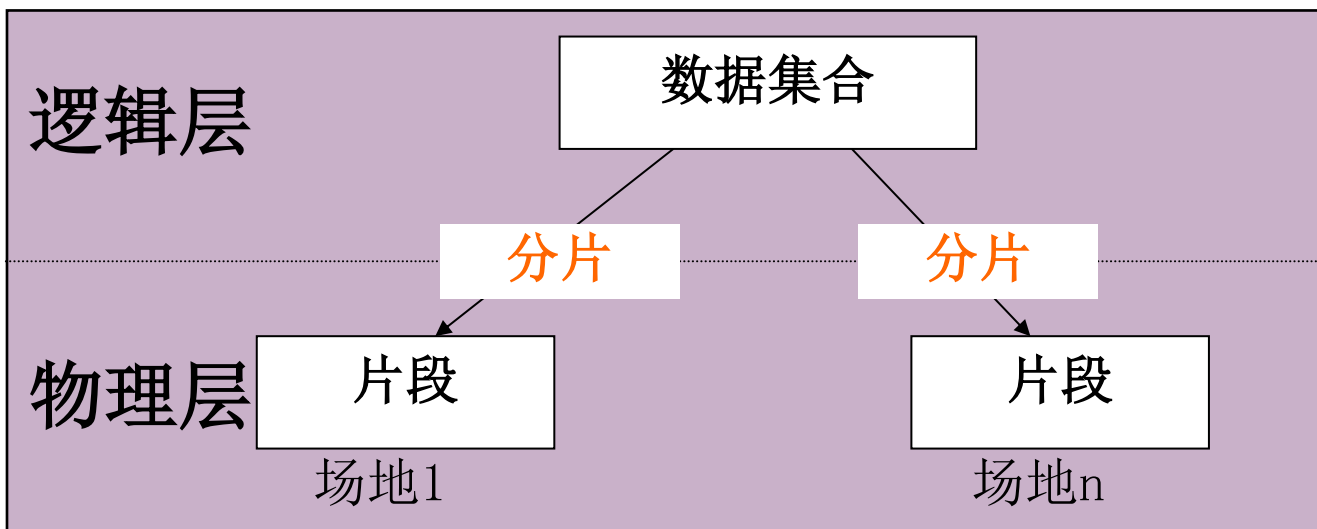


# 基本概念

## ■ 数据透明性：

□ 基于分布库的划分层次：全局外层（用户层）、全局概念层、局部概念层和局部内层。

□ 可知：应用程序与系统实际数据组织相分离，即数据具有独立性或透明性。具体体现为分布透明性、复制透明性和分片透明性。



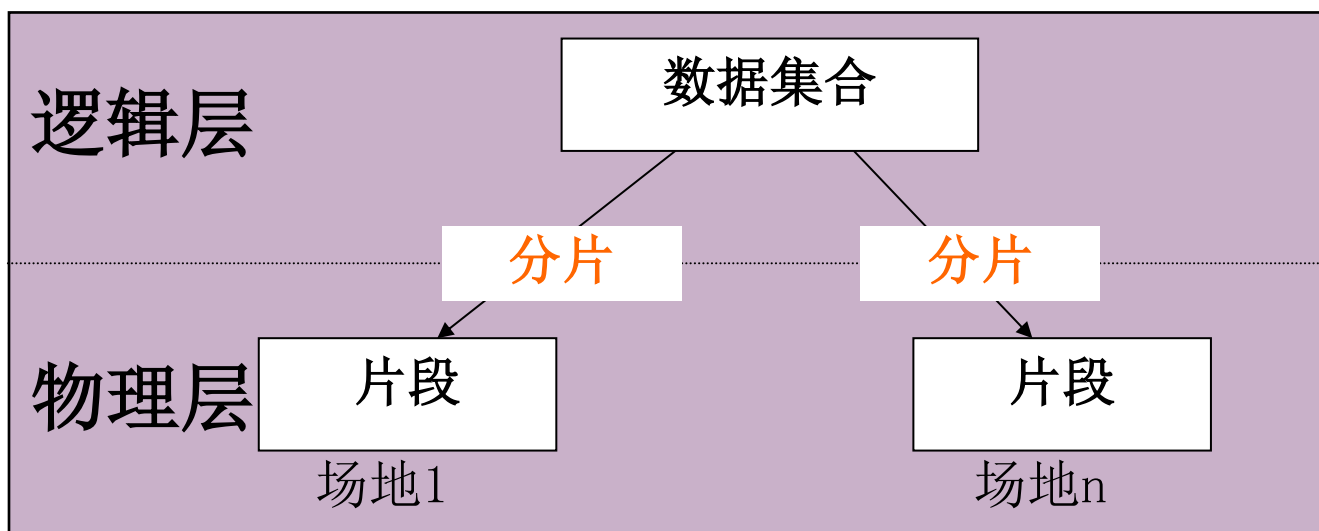
□ 复制透明性：分布式数据库支持有控制的数据冗余，即数据可重复存储在不同的场地上，提高系统的可用性和可靠性，加强系统处理的并行性。

# 基本概念

## ■ 数据透明性：

□ 基于分布库的划分层次：全局外层（用户层）、全局概念层、局部概念层和局部内层。

□ 可知：应用程序与系统实际数据组织相分离，即数据具有独立性或透明性。具体体现为分布透明性、复制透明性和分片透明性。

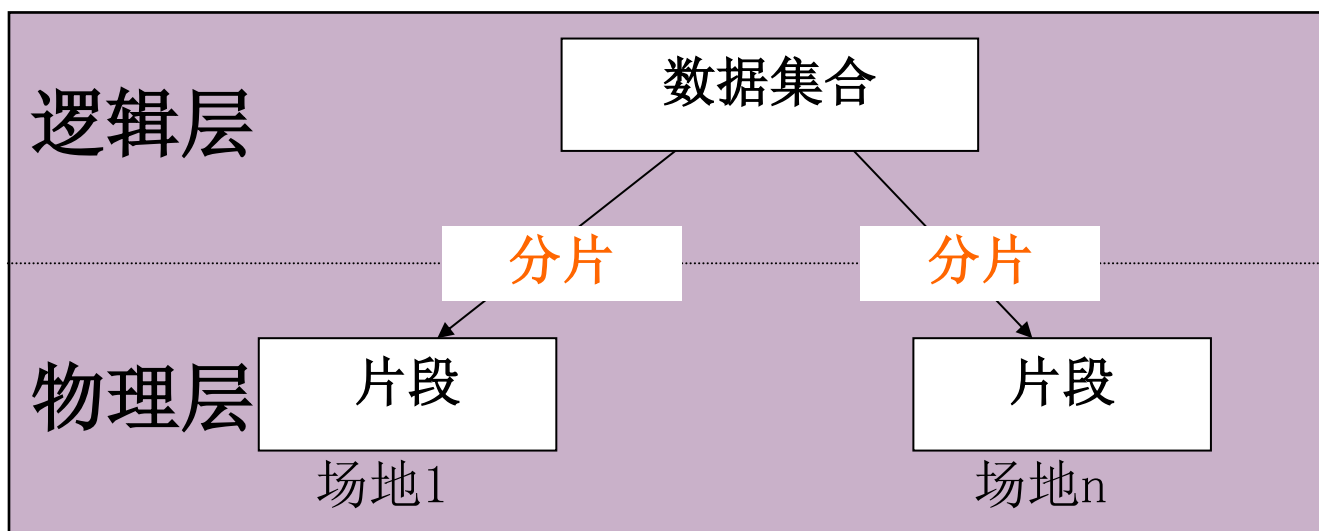


□ 分片透明性：分片是将一个关系分成几个子关系之中的一个子关系。根据实际需求，一个分片可能存储在不同的场地上（实际存储在场地上的分片称分段）。

# 基本概念

## ■ 数据透明性：

□ **逻辑层**是用户语义层，**物理层**是具体实现细节。逻辑层的语义与物理层的实现分离开来，对高层系统和用户隐蔽了实现细节。即**应用程序无须考虑数据分片细节**，对分片的调用（分片到分段的映射）由系统自动完成。



□ **分片透明性**：**分片**是将一个关系分成几个子关系之中的一个子关系。根据实际需求，一个分片可能存储在不同的场地上（实际存储在场地上的分片称分段）。

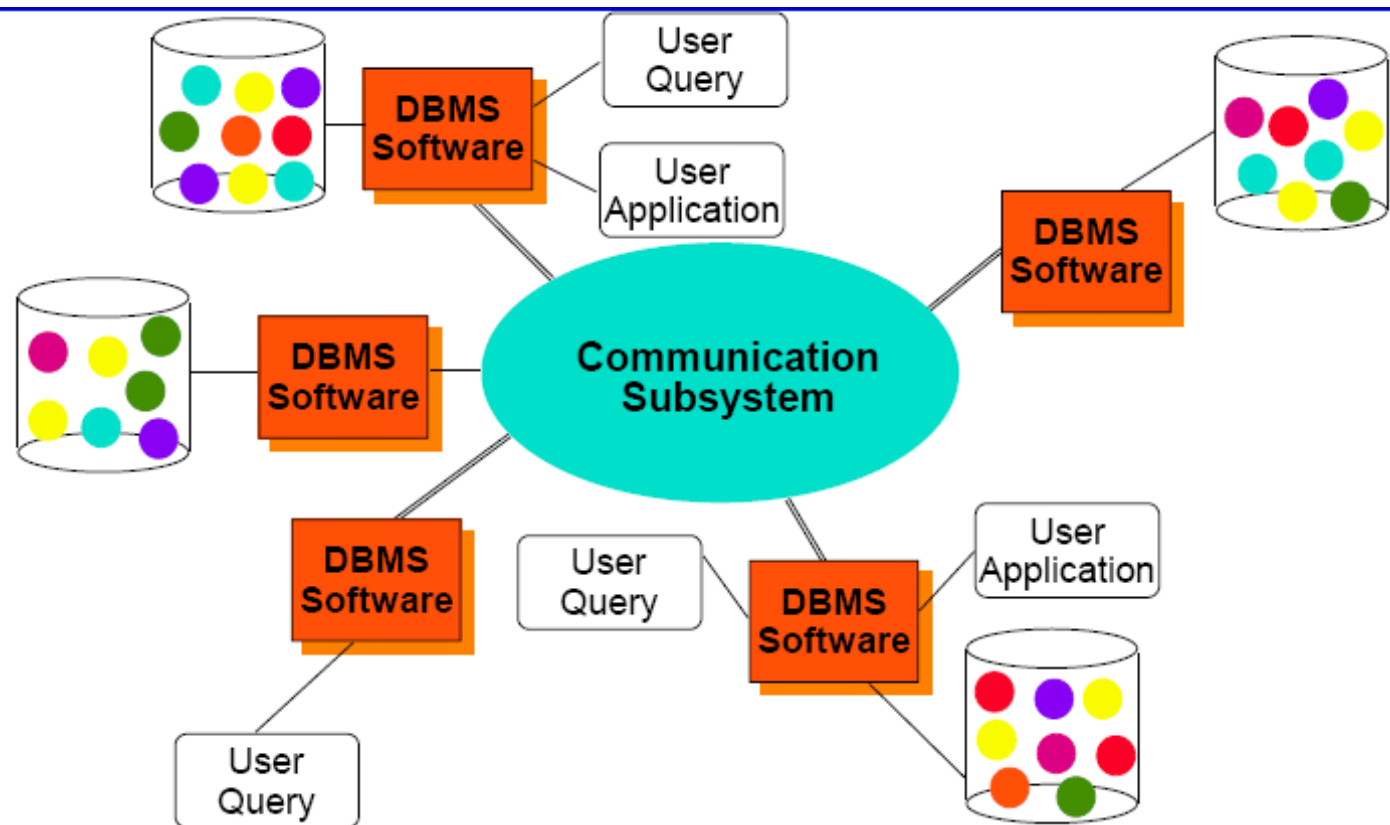
# 基本概念

## ■ 场地自治性：

□ 在分布式数据库系统中，多个场地或节点的局部数据库在逻辑上集成为一个整体，并为分布式数据库系统的所有用户使用，这种应用称**全局应用**，其用户称**全局用户**。

□ 分布式数据库系统也允许用户只使用本地的局部数据库，这种应用称为**局部应用**，其用户为**局部用户**。

□ 这种局部用户独立于全局用户的特性称**局部数据库的自治性**，也称**场地自治性**。



■ 局部用户与全局用户

- **设计自治性：** 局部DBMS能独立决定它自己局部库的设计
- **通信自治性：** 局部DBMS能独立决定是否和如何与其他场地的**DBMS**通信
- **执行自治性：** 局部DBMS能独立决定以何种方式执行局部操作

# 基本概念

## □ DDB和多处理器系统的区别

集中式管理系统存在于一个场地上；

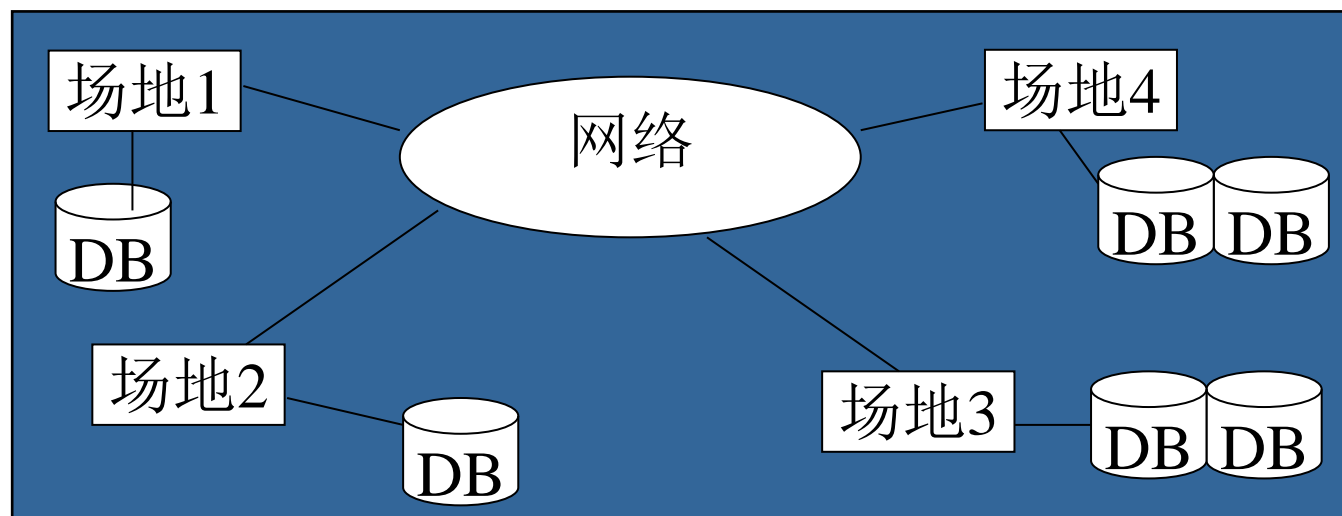
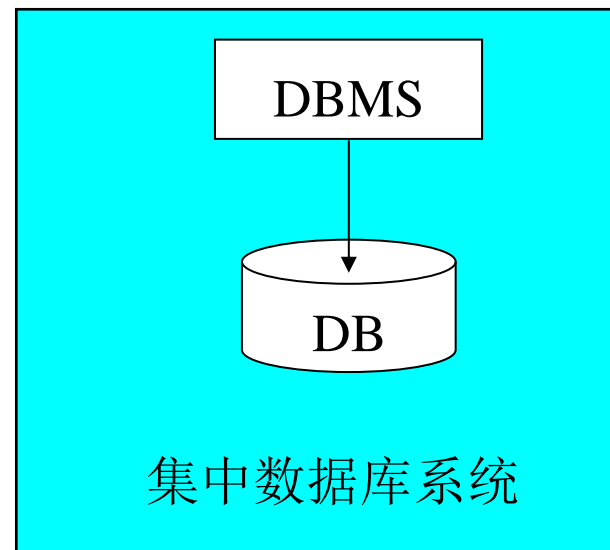
分布式管理系统分布在网络上的多个场地上。

并行系统根据三要素P（处理机）、M（内存）、D（磁盘）组成方式不同，有三种体系结构，其系统

结构如下页图所示。



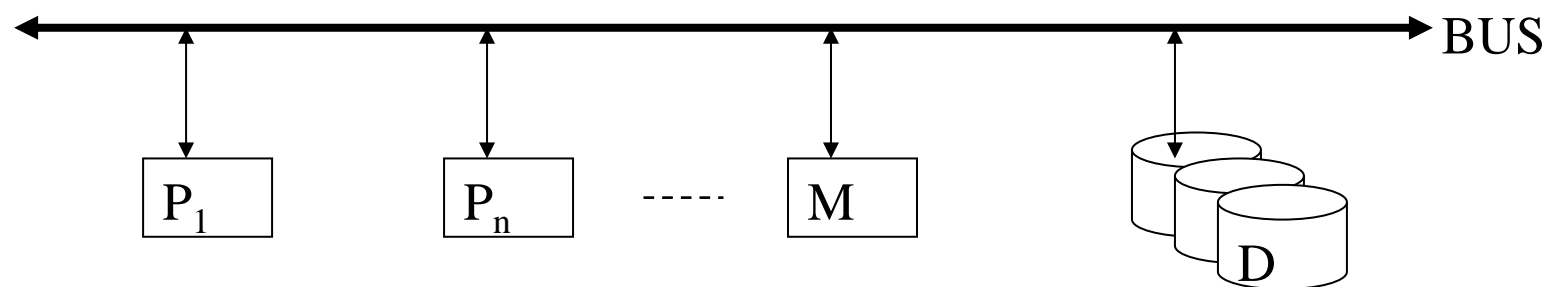
多处理器体系



# 基本概念

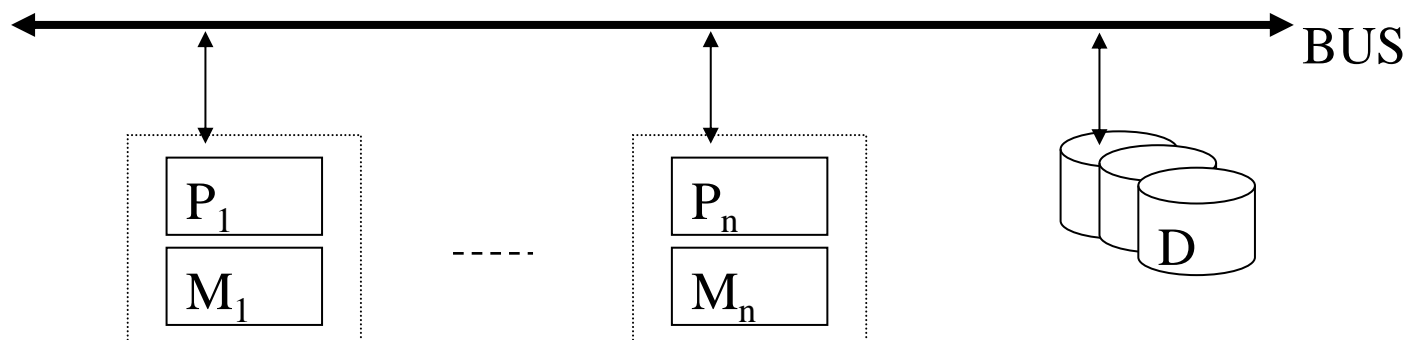
## □ DDB和多处理器系统的区别

## □ 共享内存系统（Shared-Memory Architecture）



多个处理器共享磁盘存储，也共享主存

## □ 共享磁盘系统（Shared-Disk Architecture）

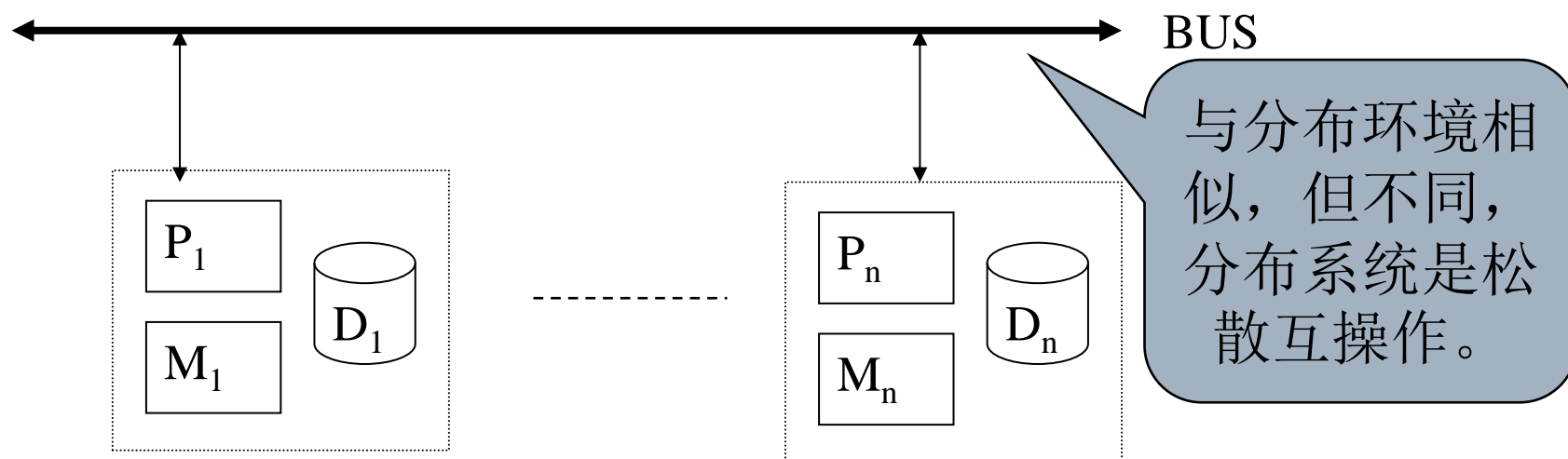


多个处理器共享磁盘存储，每个处理器有各自的主存

## 基本概念

### □ DDB和多处理器系统的区别

#### □ 无共享系统（**Shared-Nothing Architecture**）



每个处理器都有各自的主存和二级内存，不存在公共内存

最基本不同是操作模式。**多处理器系统**是对称的，由一定数量的处理器、内存组件组成，由一个或多个同样的操作系统严格控制每一个处理器的任务。

而**分布计算系统**是**异构**的操作系统以及硬件环境。

# 基本概念

---

## □ DDB和多处理器系统的区别

## □ 分布式数据库系统不是多处理器系统

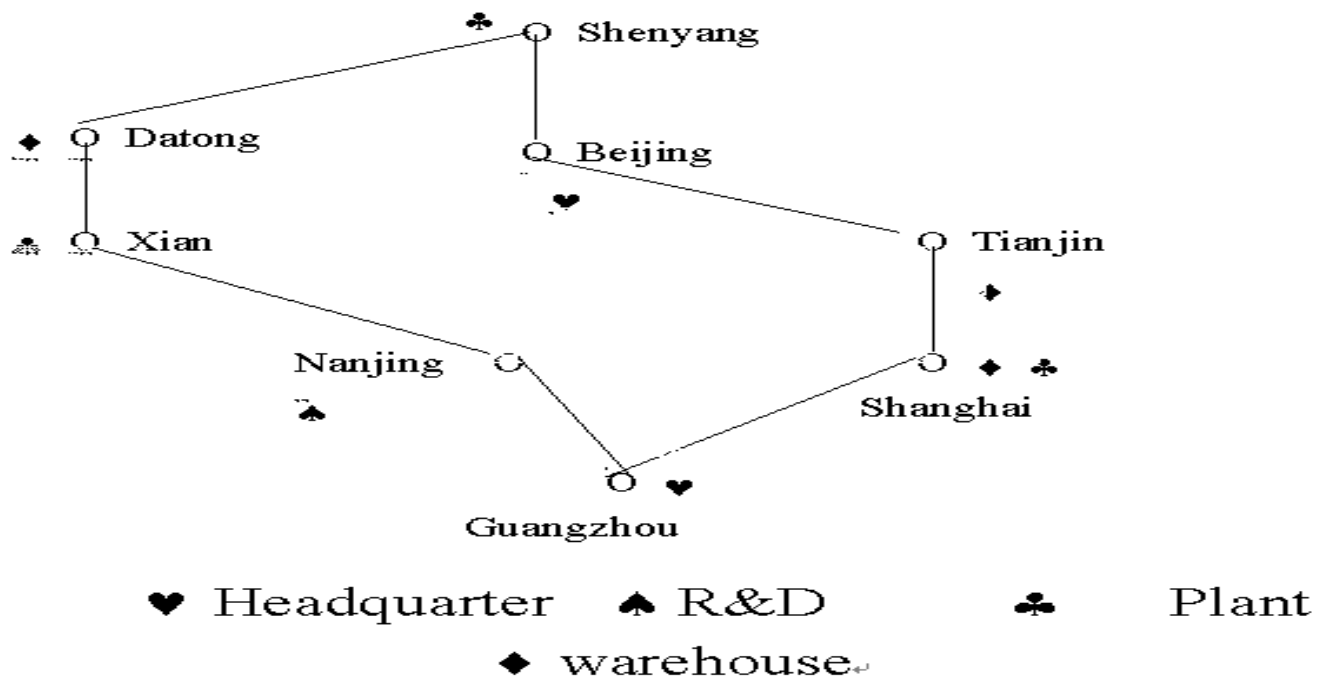
例：一个全国范围的加工制造公司的分布系统

- a) 每一地域的分公司保存自己公司的雇员信息（ **employees** ）；
- b) 研发部门（ **R&D facility** ）维护其研发项目信息；
- c) 加工工厂保存工程信息并可访问研发场地的研发信息和仓库数据（ **warehouse** ）；
- d) 总部保存区域的市场销售信息，并可访问工厂和仓库的帐目数据（ **inventory** ）。



# 基本概念

## □ DDB和多处理器系统的区别



## □ 特点:

✓ 数据分布在各个场地上，如公司的雇员（employee）和库房的帐目信息（inventory）。

## □ 特点:

✓ 查询涉及多个场地，如查询年龄小于45岁的雇员，涉及所有的场地雇员信息。

# 基本概念

---

□ DDB和多处理器系统的区别

□ 分布式数据库系统不是基于服务器的系统

□ DDB 与 C/S DB的区别

■ 分布数据与集中数据（Distributed data versus centralized data）

■ 集中的功能和分布的功能（Centralized function versus distributed function）

■ 基于服务器的系统：以功能集中为核心，适合强一致性、低扩展需求的场景。

■ 分布式数据库：以功能分布为核心，适合高可用性、弹性扩展的大规模场景。

# 基本概念

## □ DDB和多处理器系统的区别

## □ 分布式数据库系统不是基于服务器的系统

## □ DDB 与 C/S DB的区别

- 基于服务器的系统： 银行转账系统：需确保转账操作的原子性和一致性，适合强集中的主从架构。

核心逻辑：所有功能（如事务处理、数据存储、权限管理）由中心节点掌控。

**优势**：强一致性、简单架构、易于维护。 **劣势**：单点故障风险、扩展能力有限。

- 分布式数据库： 全球电商平台：用户分布在不同地区，需就近访问数据并支持高并发，适合分布式架构。

核心逻辑：功能分散到多个节点，各节点具备一定自治能力。

**优势**：高可用性、弹性扩展、容灾能力强。 **劣势**：一致性模型复杂（如最终一致性）、跨节点事务成本高。

## 基本概念

□ 一个数据库系统若被称为是**分布式数据库**，则它至少要满足如下条件：

- ☑ 通过**计算机网络连接数据库节点**。具有多个计算机，称为站点/场地/节点。这些站点必须通过下层通信网络连接起来以传输站点之间的数据和命令。
- ☑ 已连接数据库在**逻辑上是相互关联**的。
- ☑ 已连接节点之间**不存在同构约束**。

所有节点在数据、硬件、软件上不必是相同的。

# 分布库系统的作用和特点

## □ 优点

- 提高系统性能，如：吞吐率和响应速度提高。
  - e. g., 对数据库进行分片，将数据存储在最频繁使用的地方；  
大型数据库分布于多个站点上，每个站点只存放较小的数据库
- 可利用现有设备和系统，降低投资。
- 系统允许存储副本，提高可靠性、可用性和并行执行度。
- 根据实际需要，可增加或减少某一场地。系统具有可扩展性。
- 系统资源和数据分布在物理上不同的场地上，为系统所有用户共享。

# 分布库系统的作用和特点

## □ 缺点

■ 系统不很完善。

■ 系统复杂。

数据复制、场地故障、分布式事务处理

■ 维护开销大。

■ 需分布控制。

■ 安全问题：

用户授权和认证问题、访问权限控制问题、数据完整性约束问题

■ 系统难以改变。

最小代价查询优化

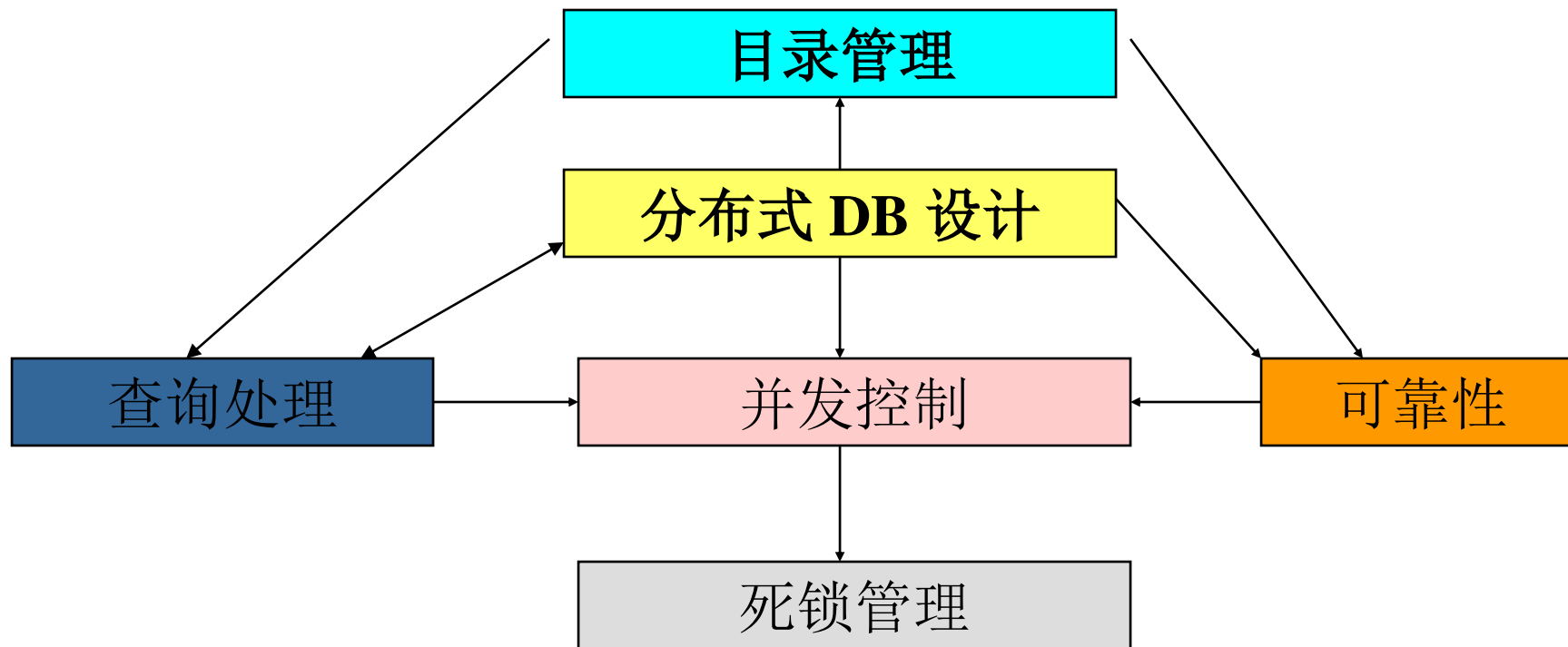
多用户并发访问的同步

事务的一致性和隔离性

死锁的解决与预防

# 分布库系统

## □ 关键技术及相关关系



# 分布式数据库系统类型

---



## 分布式数据库系统类型

□ 分布式数据库系统的共同点是数据和软件分布于多个站点上，这些站点通过某种形式的通信网络连接在一起

分布式数据库系统的分类方法很多。为全面、系统地对分布式数据库系统进行分类，采用分布式数据库的三个特性（**分布性**、**异构性**、**自治性**）组成的三维空间图来描述分布式数据库的类型。

### □ 三个基本特性

■ **分布性**是指系统的各组成单元是否位于同一场地上。 分布式数据库系统是物理上分散、逻辑上统一的系统，即具有分布性。而集中式数据库系统集中在一个场地上，所以不具有分布性。

# 分布式数据库系统类型

- **异构性**是指系统的各组成单元是否相同，不同为异构，相同为同构。异构主要：
  - ✓ **数据异构性**：指数据在格式上、语法和语义上存在不同。
  - ✓ **数据系统异构性**：指各个场地上的局部数据库系统是否相同。如：均采用ORACLE数据库系统的同构数据系统，或某些场地采用ORACLE数据库系统，某些场地采用XML数据库系统的异构系统。
  - ✓ **平台异构性**：指计算机系统是否相同。如：均为微机系统组成的平台同构系统或由VAX或ALPHA系统等异构平台组成的系统。
- **自治性**是指每个场地的独立自主能力。自治性通常由设计自治性、通讯自治性和执行自治性三方面来描述。根据系统的自治性，可分为**集中式系统**、**联邦式系统（Federated DB）**和**多数据库系统（Multi DB）**。

# 分布式数据库系统类型

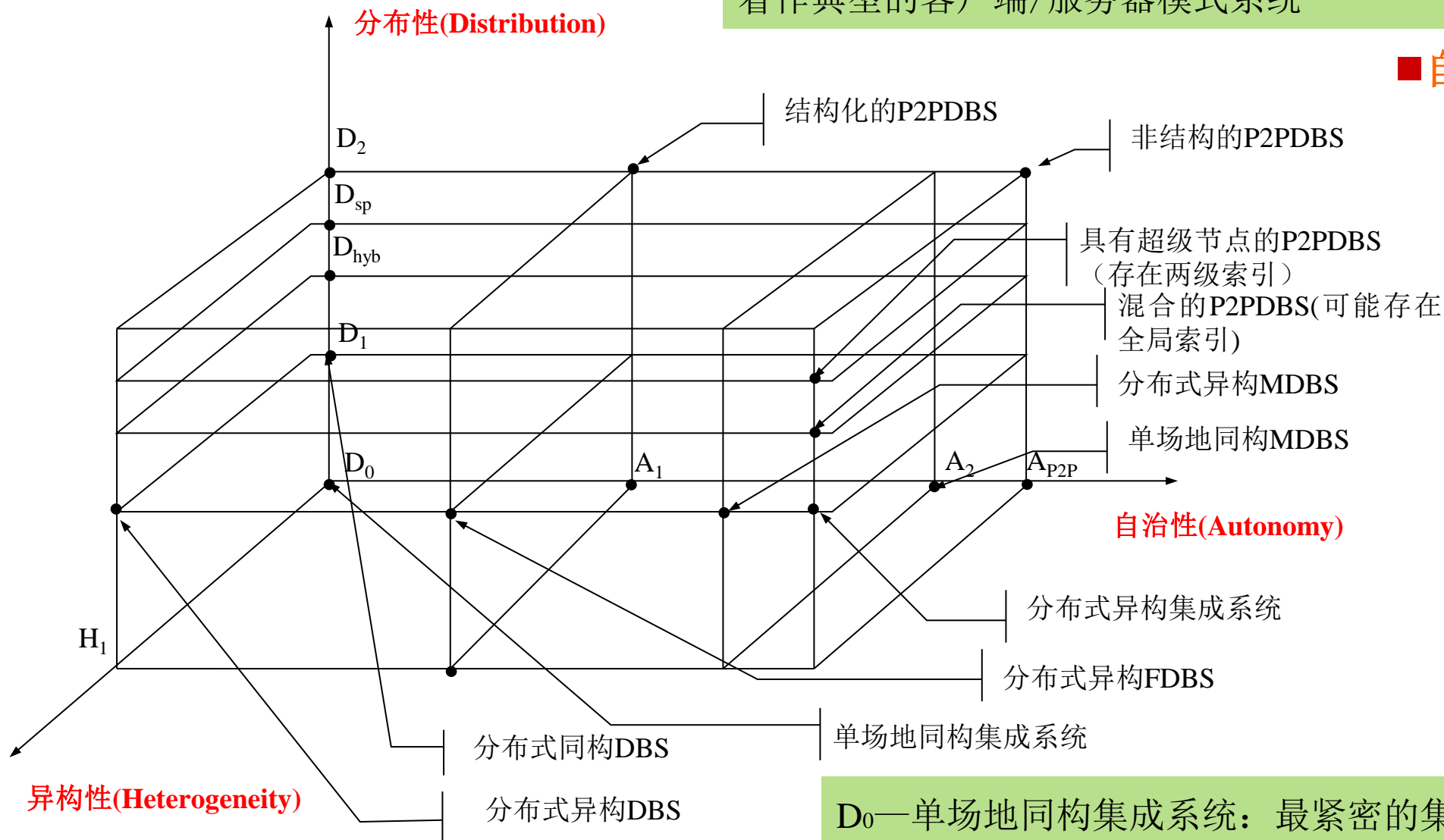
D<sub>1</sub>—分布式同构数据库系统：具有集成视图和分布特性，可看作典型的客户端/服务器模式系统

## ■ 自治性：

✓ **集中式系统**即为传统的数据库系统。

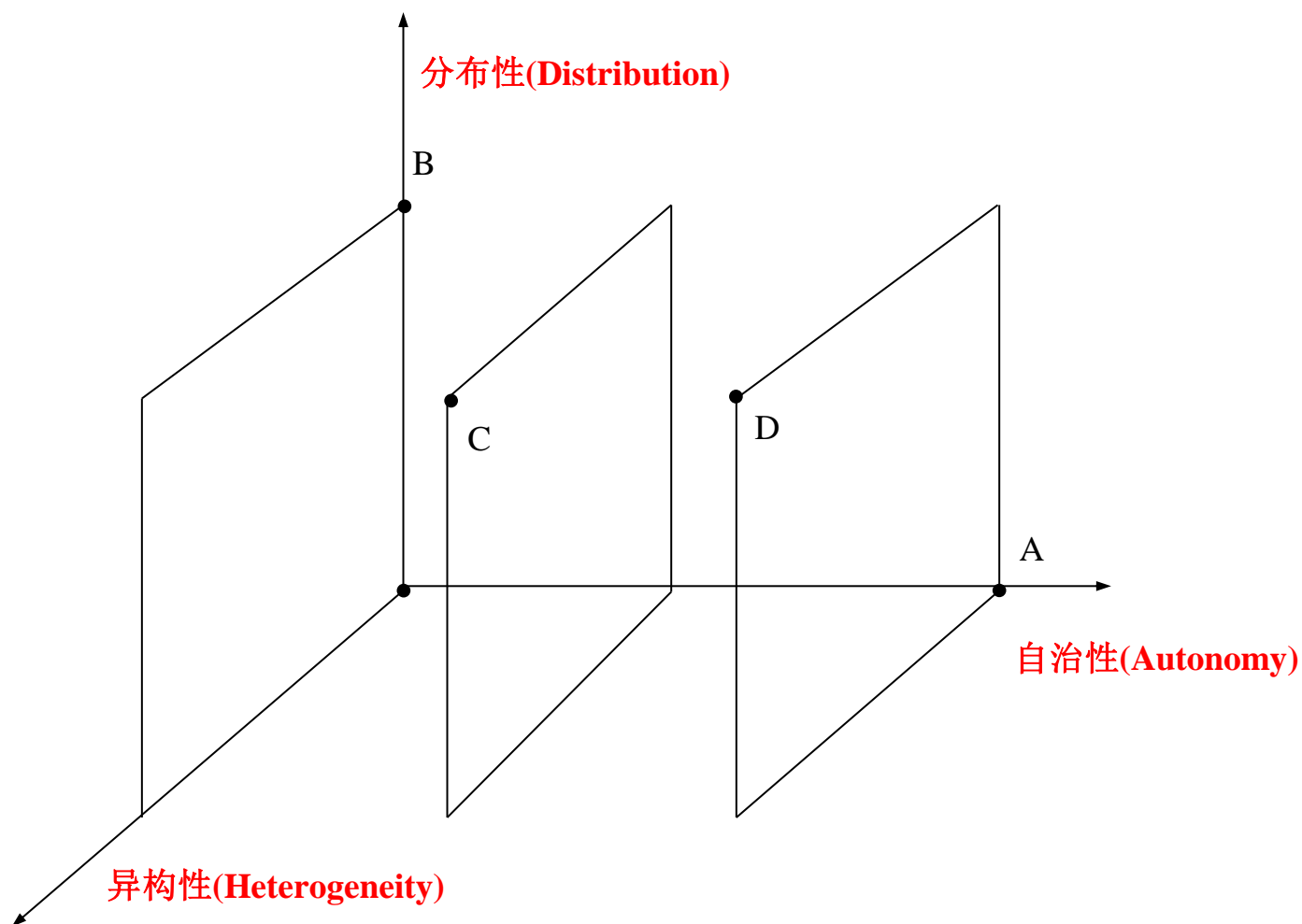
✓ **联邦式系统**：实现需要交互的所有数据库对之间的一对一连接。

✓ **多数据库系统**：是指若干相关数据库的集合。若各个数据库可以存在于多个场地，也可分布在多个系统内进行管理。多数据库管理系统是对多个数据库管理系统进行统一管理，并提供透明访问。



D<sub>0</sub>—单场地同构集成系统：最紧密的集成系统，可看作共享内存和硬盘的多处理机系统

# 分布式数据库系统类型



## ■ A: 传统的集中式数据库系统

具有完全的自治性，但完全不具有分布性和异构性；

## ■ B: 纯粹的分布式数据库系统

概念上的架构，对所有系统的访问都是通过DDBMS的一个站点获得，意味着无本地自治性

## ■ C: 联邦式数据库系统

数据库的联盟在某些全局的视图或模式上为应用提供共享

## ■ D: 多数据库系统或对等网络数据库系统

具有完全的自治性，没有全局模式，在需要时可由应用交互式创建；

# 分布式数据库系统类型

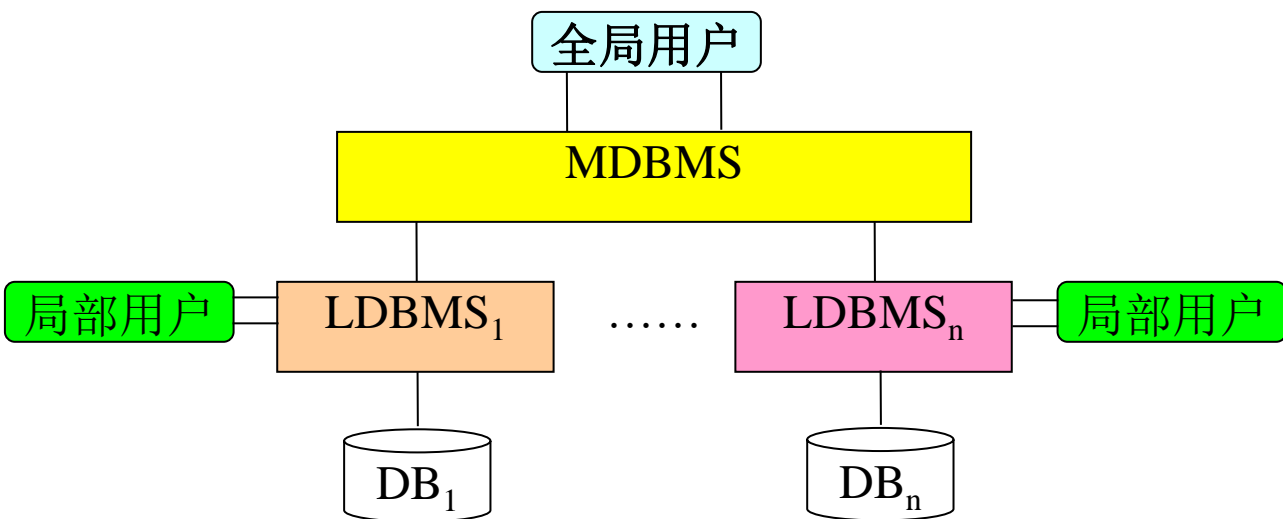
## 多数据库系统（MDB）

■ MDBS 是在已经存在的数据库系统 (称为局部数据库系统LDBs) 之上为用户提供一个统一的存取数据的环境。

## DDB 和 MDB的不同

■ MDB：本地数据管理器（data manager）都是完备的DBMS；每个本地数据管理器都具有自己独立的DDL，DML和TM。

■ DDB：只有一个数据管理器和依赖于该管理器的存在于各局部场地上的组件数据管理器（component-data-manager）。



多数据库系统逻辑结构

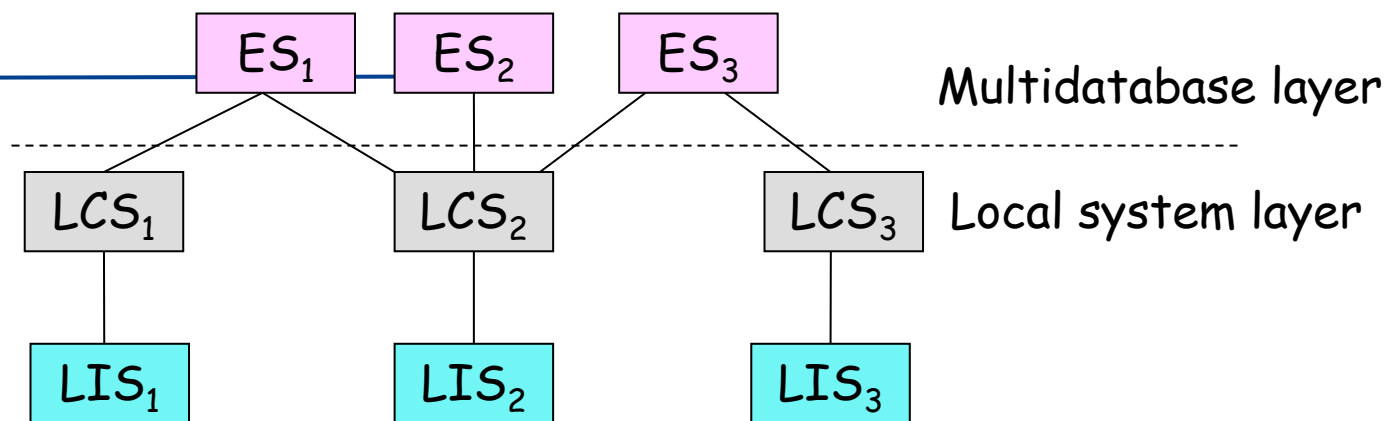
# 分布式数据库系统类型

## 多数据库系统 (MDB)

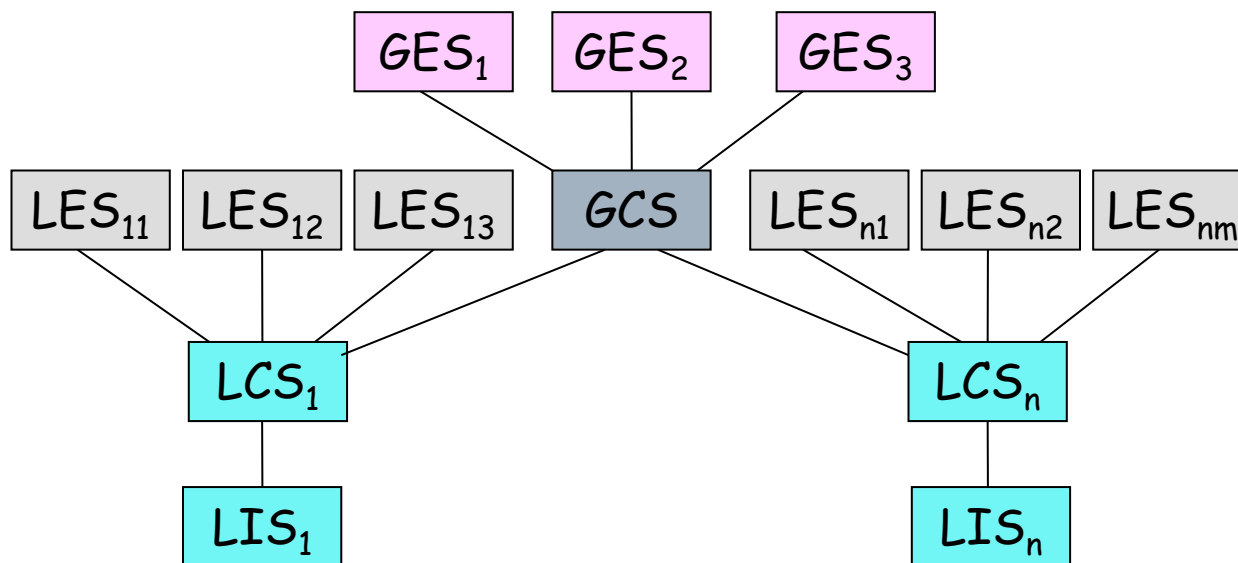
## DDB 和 MDB的不同

■ DDBS: 全局概念模式GCS是整个数据库的概念模式，是所有局部概念模式LCS的集合。

■ MDBS: 全局概念模式是可共享的各局部概念模式的集合，并不是所有局部概念模式的集合，MDBS甚至可以没有全局概念模式。



## 具有全局概念模式的MDBS模型

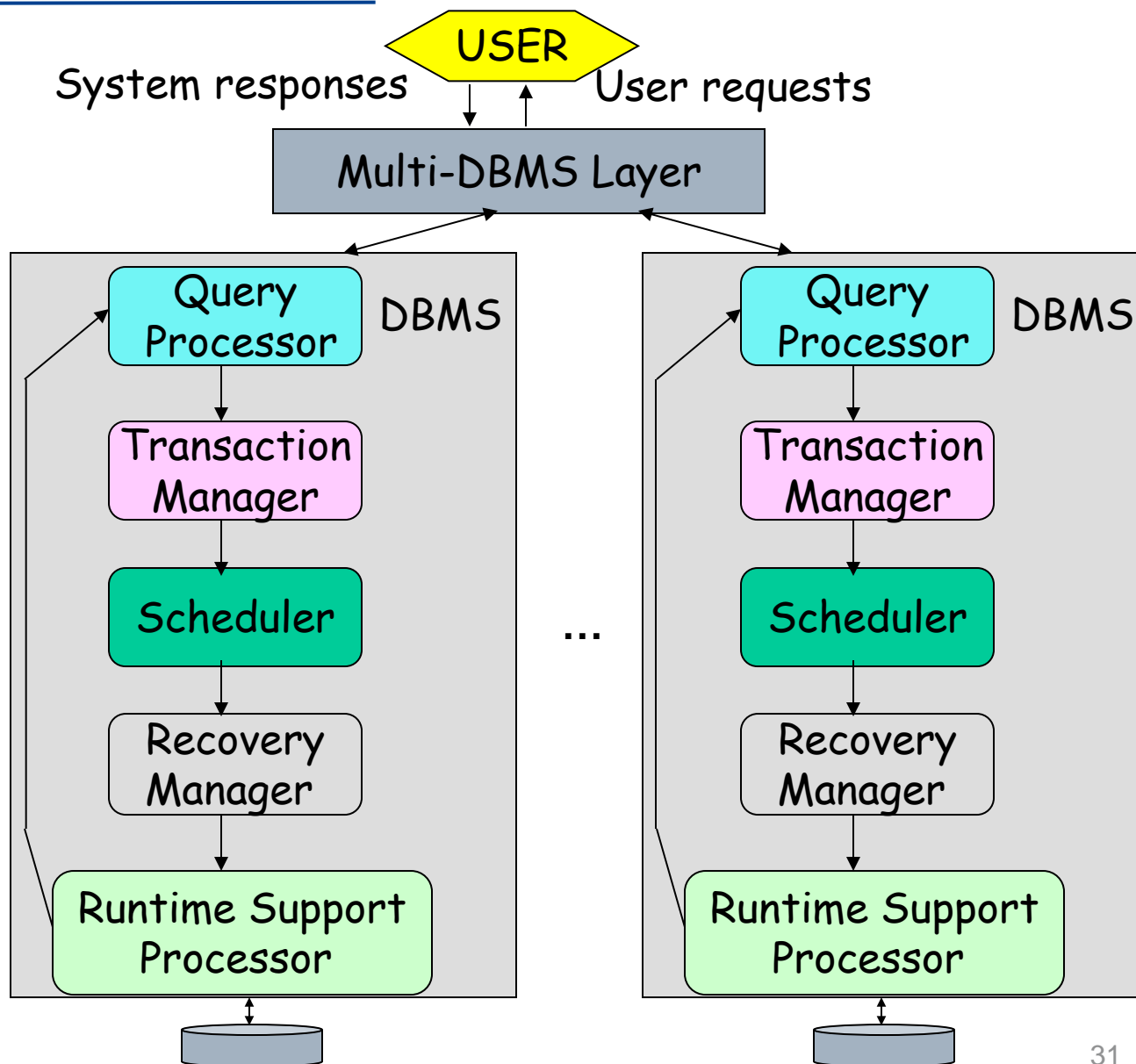


## 没有全局概念模式的MDBS模型

# 分布式数据库系统类型

## 多数据库系统 (MDB)

### 组件结构



# 分布式数据库系统类型

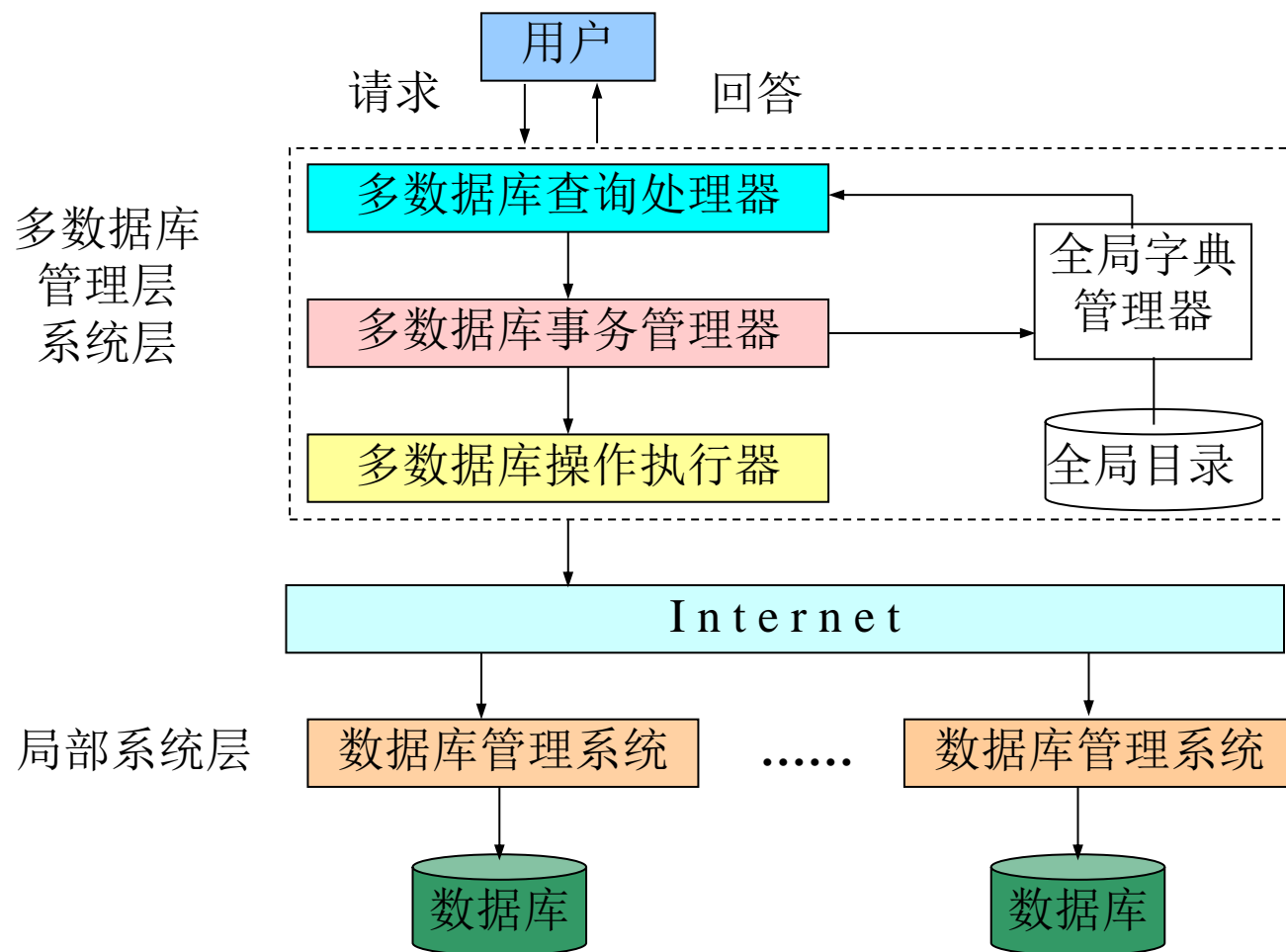
## ❑ 多数据库系统（MDB）

### 软件实现结构

#### ❑ 多数据库管理系统（MDBMS）

由多个完全独立的数据库管理系统组成，各个数据库管理系统各自管理不同的数据库。

❑ 多数据库管理系统通过运行在这些独立的成员数据库管理系统之上的一层软件来支持对各个不同数据库的访问。



一种多数据库管理系统的实现结构



## 分布式数据库系统类型

---

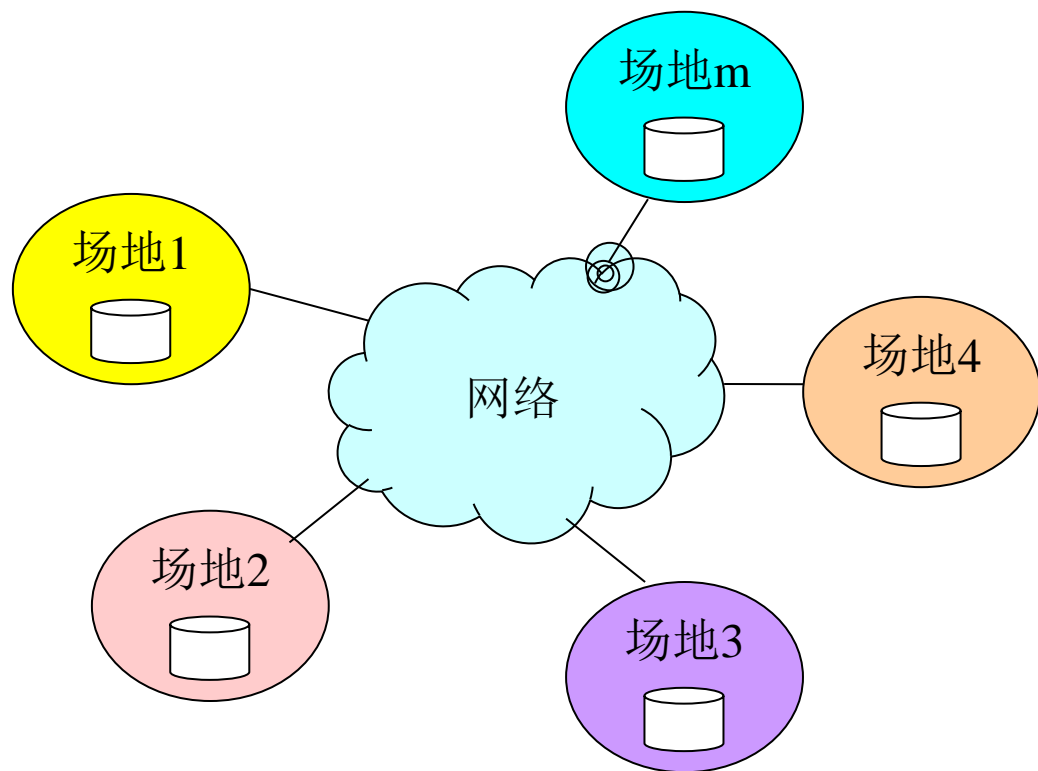
### □ P2PDBS（对等型数据库系统）

# 系统体系结构

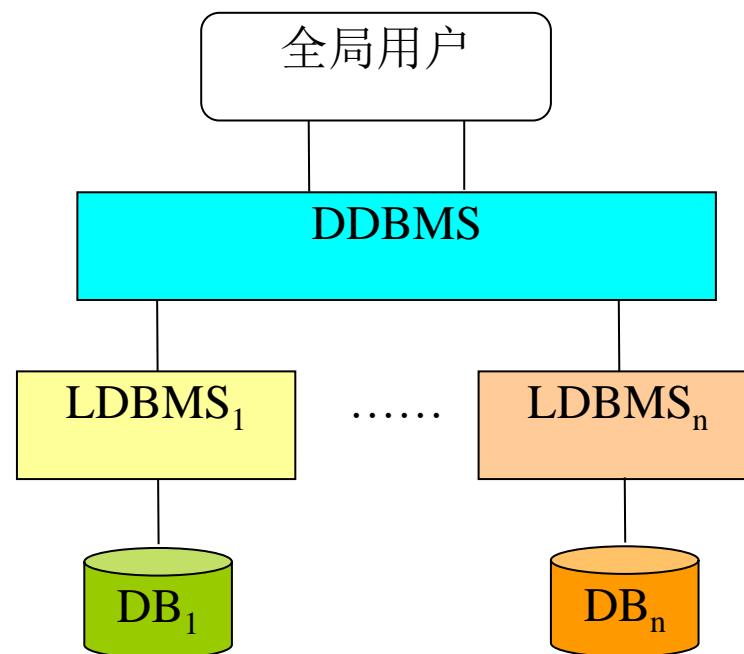
---

# 分布式数据库系统体系结构

## 物理结构和逻辑结构



DDBS的物理结构



DDBS的逻辑结构

# 分布式数据库系统体系结构

□ **系统的体系结构（Architecture）** 指按组件、组件功能以及它们的交互作用定义系统的结构。

□ **参考体系结构的目的**

- **讨论的框架**

- **标准**

□ **DBMS 标准化**

- **基于功能**

- **基于组件**

- **基于数据**

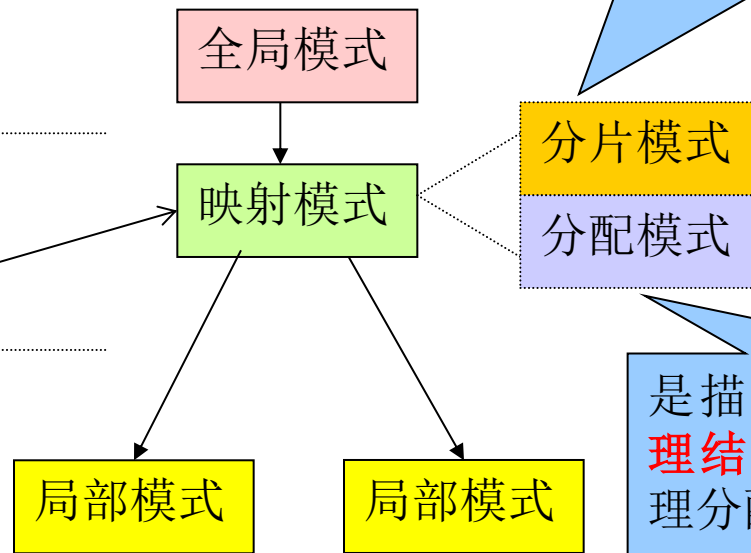
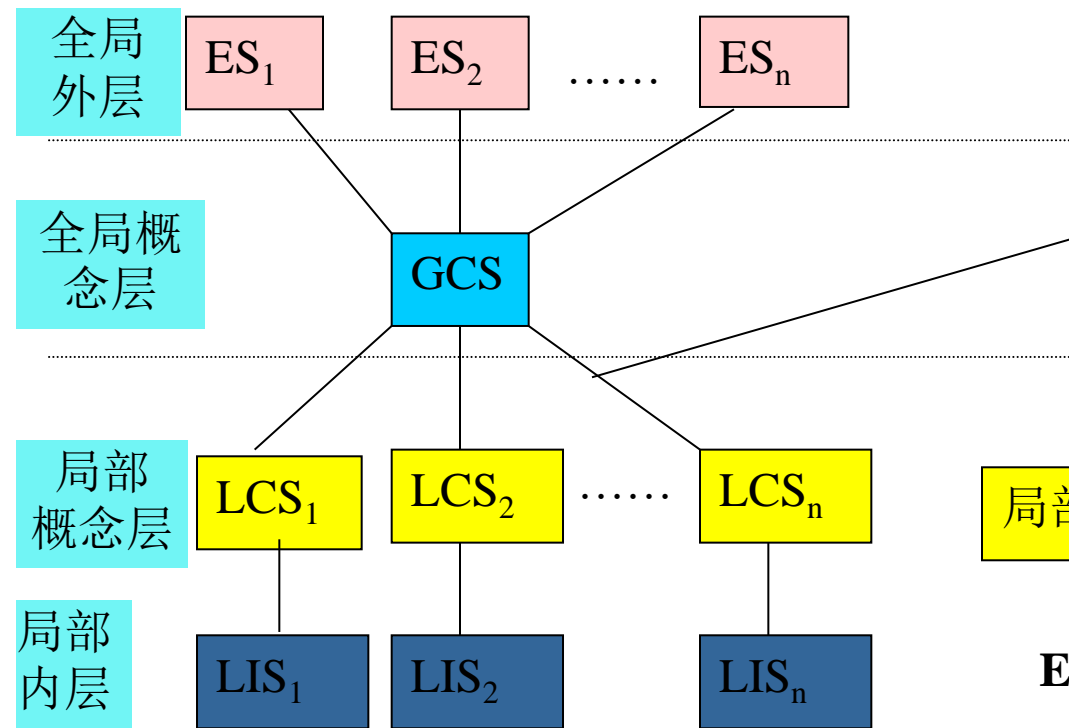
- **最早的是基于数据组织ANSI/SPARC Architecture（external view, conceptual view, internal view）的三级模式结构**

American National Standards Institute, Standards Planning And Requirements Committee: 美国国家标准协会的计算机与信息处理委员会中的标准计划与需求委员会。

# 分布式数据库系统模式结构

□我国分布式数据库系统标准草案中给定的抽象为四层的模式结构：**全局外层**、**全局概念层**、**局部概念层**和**局部内层**。模式与模式之间是映射关系。

描述全局数据的逻辑划分视图。即**全局数据逻辑结构**根据某种条件的划分为**局部数据逻辑结构**。每一个逻辑划分形成一个**分片**。在关系数据库中，一个关系中的一个子关系称该关系的一个片段。



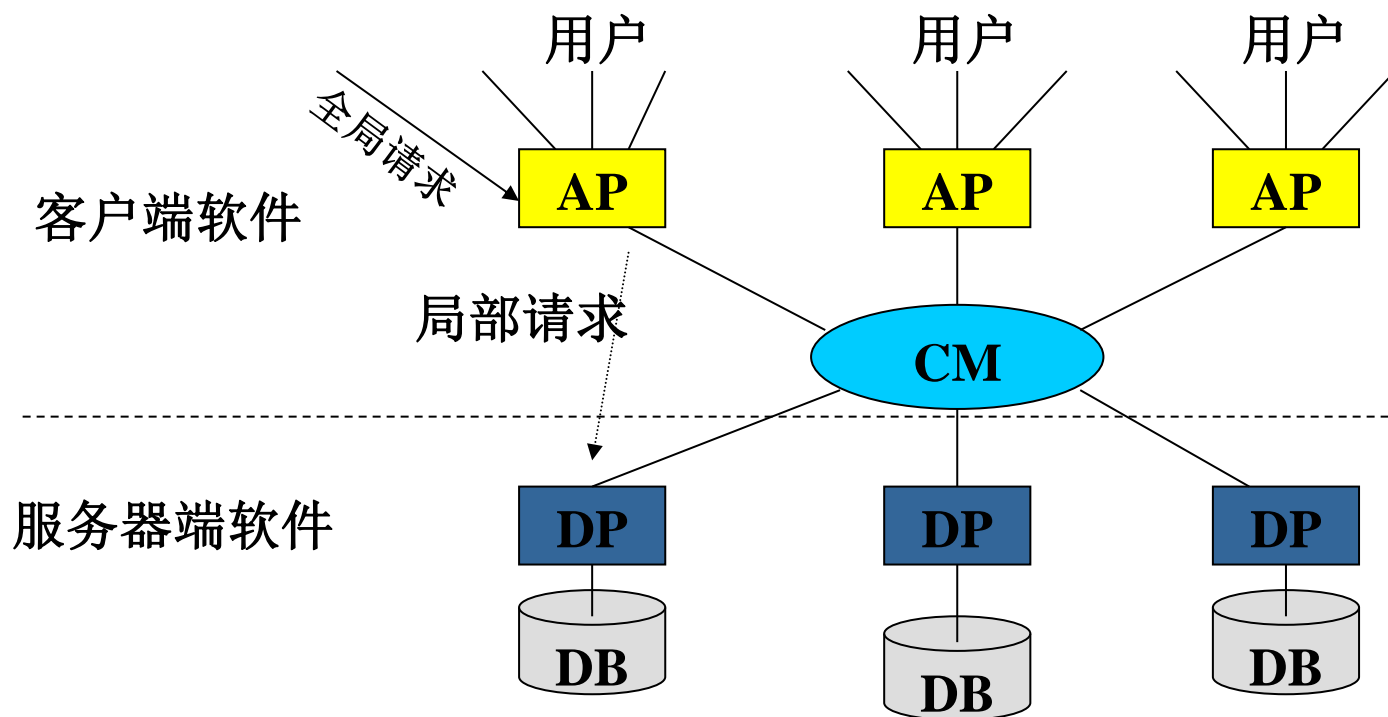
是描述局部数据逻辑的**局部物理结构**，即划分后的分片的物理分配视图。

ES: 全局模式 GCS: 全局概念模式

LCS: 局部概念模式 LIS: 局部内模式

# 分布式数据库系统体系结构

## □ 基于客户端/服务器结构的系统结构



其中，AP：应用处理器 DP：数据处理器 CM：通信处理器

**通信管理器(CM)**负责为AP和DP在多个场地之间传送命令和数据。

### 应用处理器(AP)

用于完成分布数据处理的软件，如：处理访问多个场地的请求，查询全局字典中分布信息等；把用户命令翻译成适合于局部场地执行的规范化命令格式；将查询返回的结果数据从规范化格式转换成用户格式...

### 数据处理器(DP)

负责进行数据管理的软件，类似于一个集中式数据库管理系统（DBMS）；将规范化命令翻译成物理命令，并发的执行物理命令，并返回结果数据...

# 分布式数据库系统体系结构

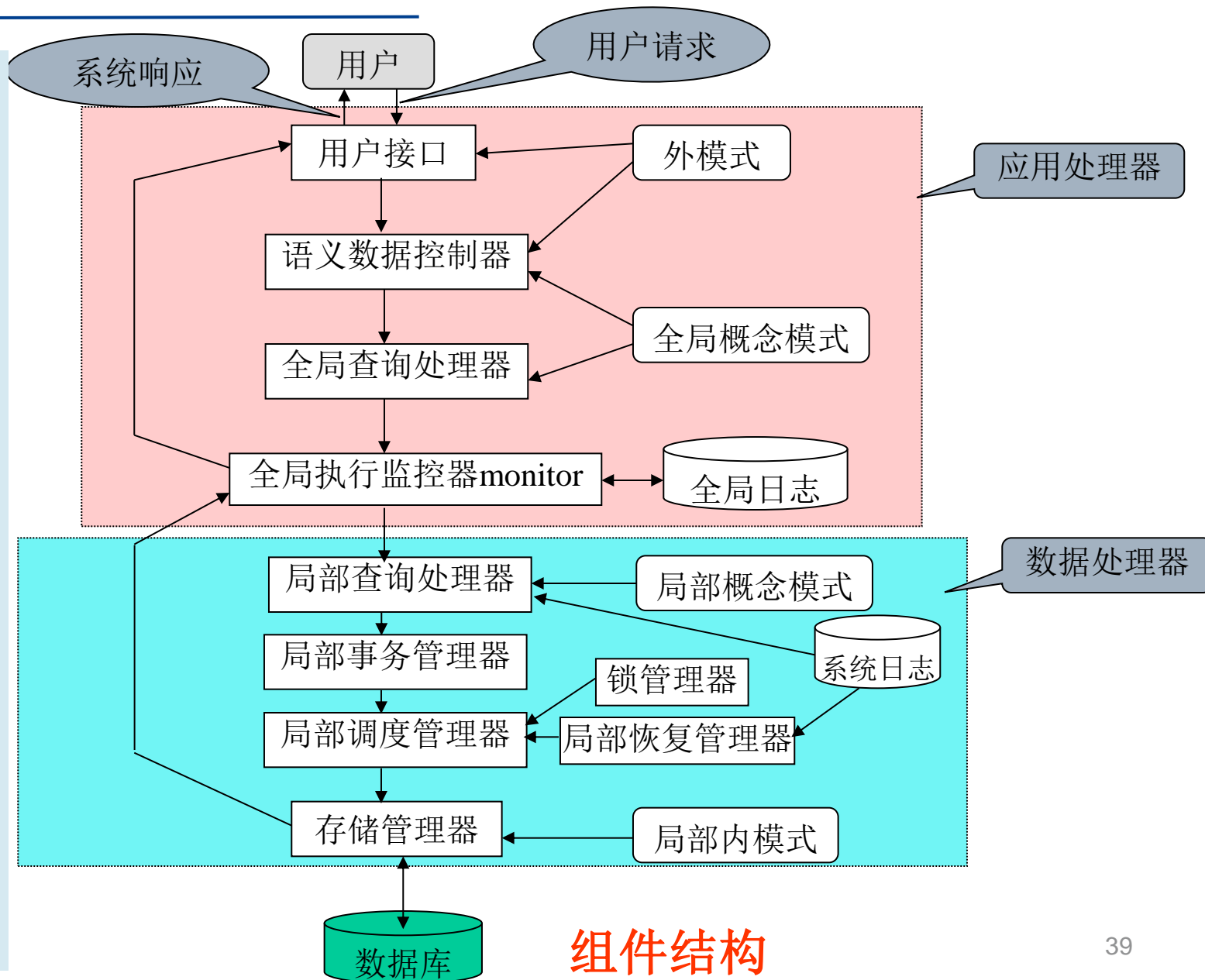
## 应用处理器（AP）功能

■ **用户接口**：检查用户身份，接受用户命令，如：SQL命令。

■ **语义数据控制器**：视图管理、安全控制、语义完整性控制（全局概念模式）。这些约束定义在字典中。

■ **全局查询处理器**：将用户命令翻译成数据库命令；生成全局查询的分布执行计划；收集局部执行结果并返回给用户。

■ **全局执行监控器（全局事务管理器）**：调度协调和监视AP和DP之间的分布执行；保证复制数据的一致性；保证全局事务的原子性。



# 分布式数据库系统体系结构

## ❑ 数据处理器DP功能

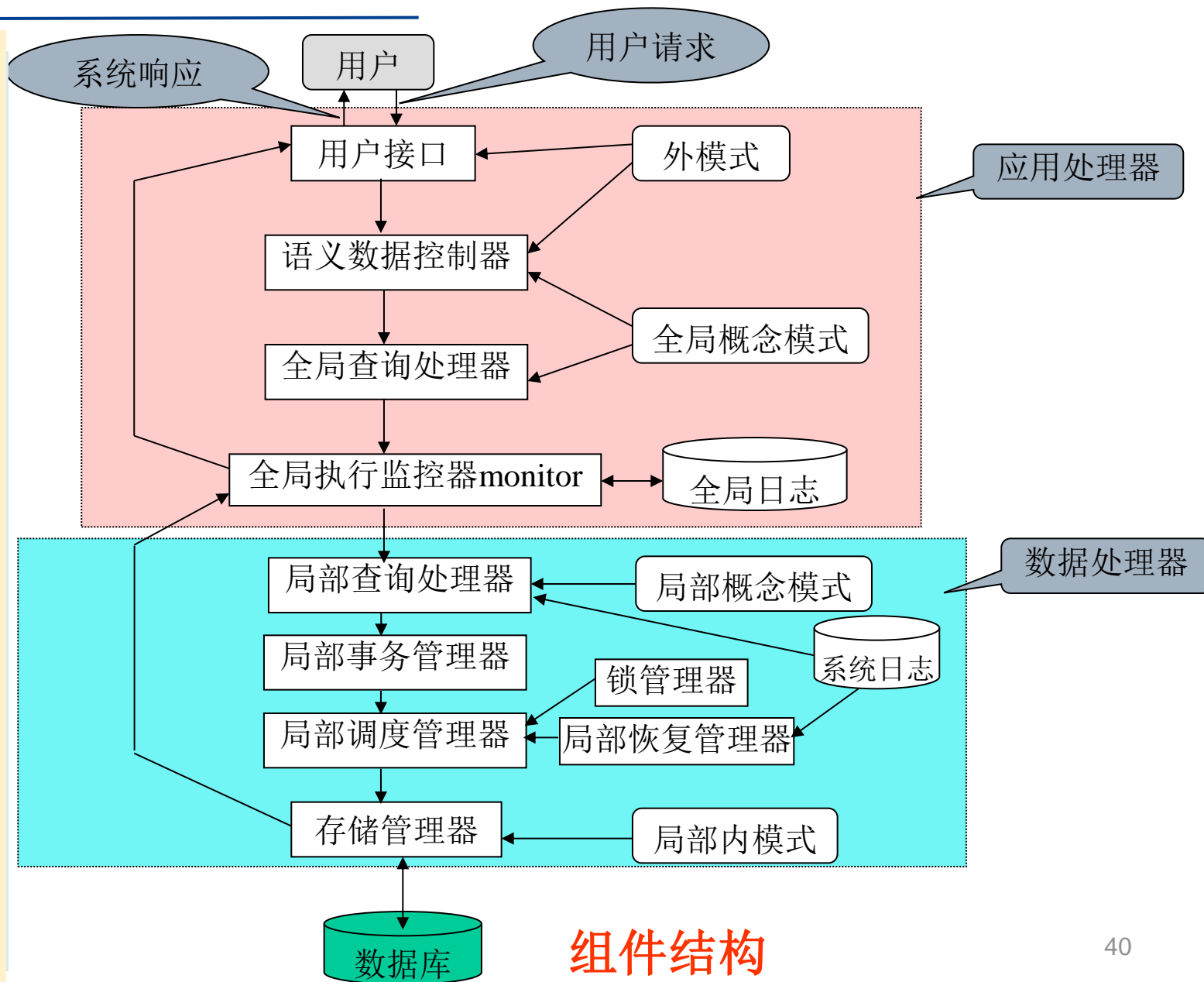
■ **局部查询处理**：实现全局命令到局部命令的转换；访问路径选择器，选择最好的路径执行。

■ **局部恢复管理器**：维护本地数据库一致性的故障恢复。

■ **存储管理器**：按调度命令访问数据库；控制数据库缓存管理器；返回局部执行结果；保证子事务执行的正确性。

■ **局部事务管理器**：以局部子事务为单位调度执行，保证子事务执行的正确性。

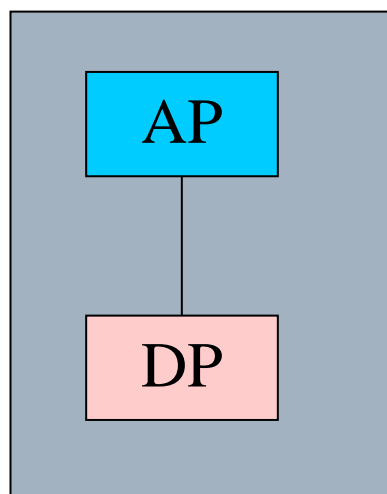
■ **局部调度管理器**：负责局部场地上的并发控制，按可串行化调度和执行数据操作。





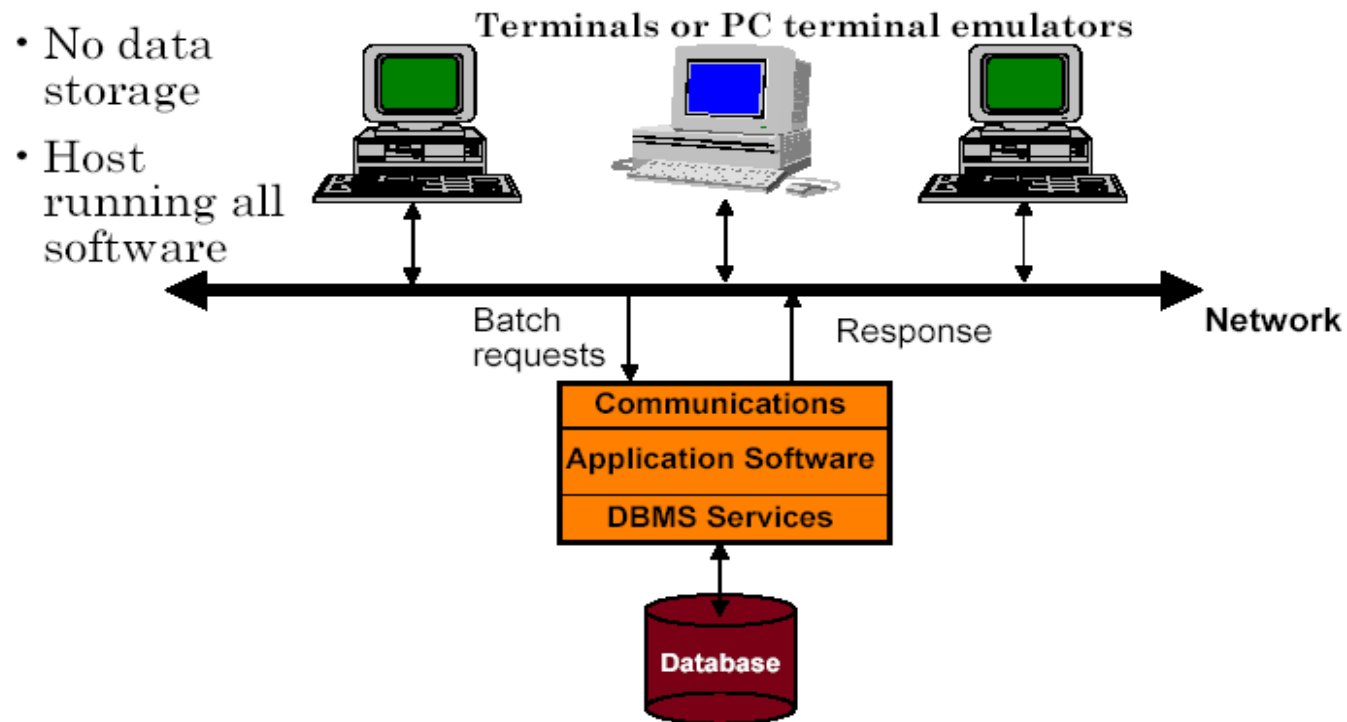
# 分布式数据库系统体系结构

## 不同系统的系统结构图



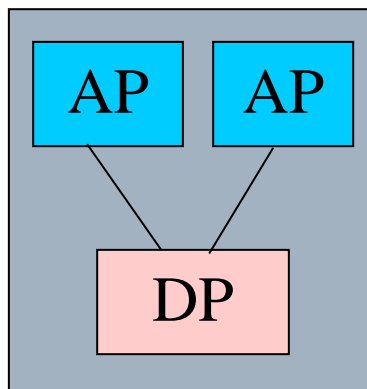
a. 集中式  
数据库系统结构

## Timesharing Access to a Central Database

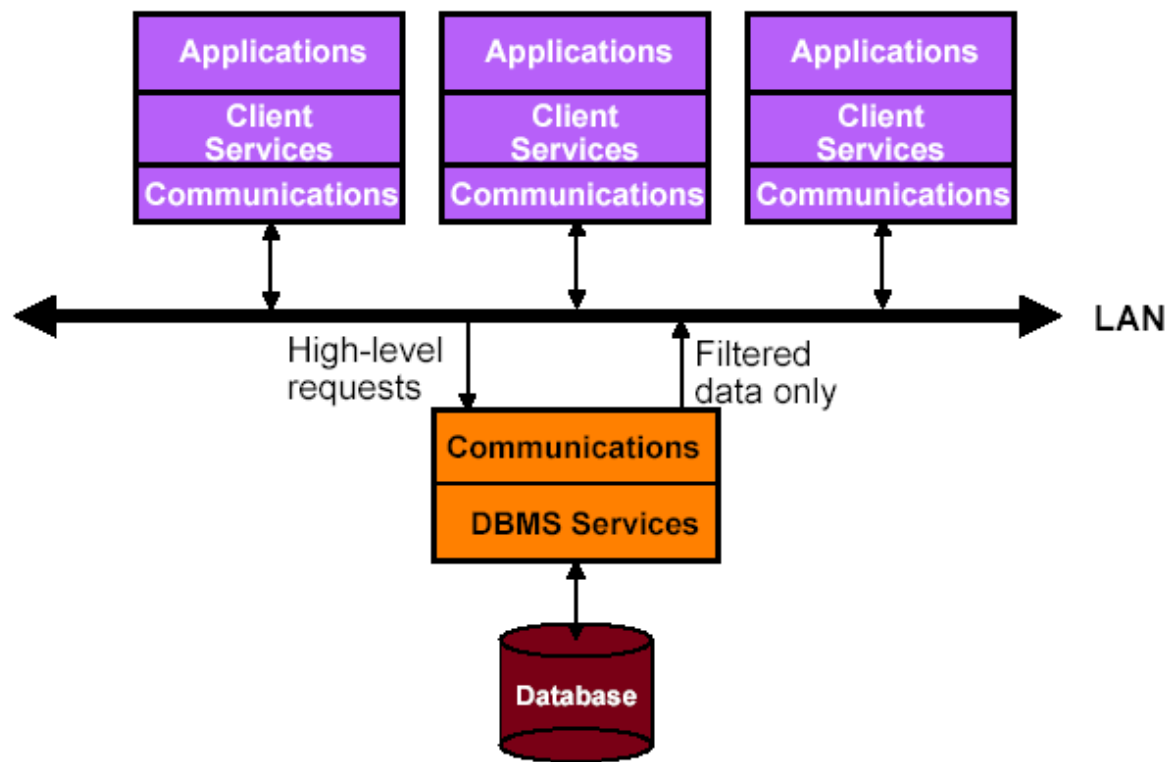


# 分布式数据库系统体系结构

## 不同系统的系统结构图



### Multiple Clients/Single Server

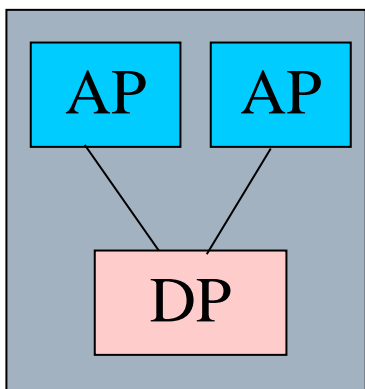


b. 多客户/单服务器(c/s)

多AP，单DP系统结构属于网络数据库服务器系统结构

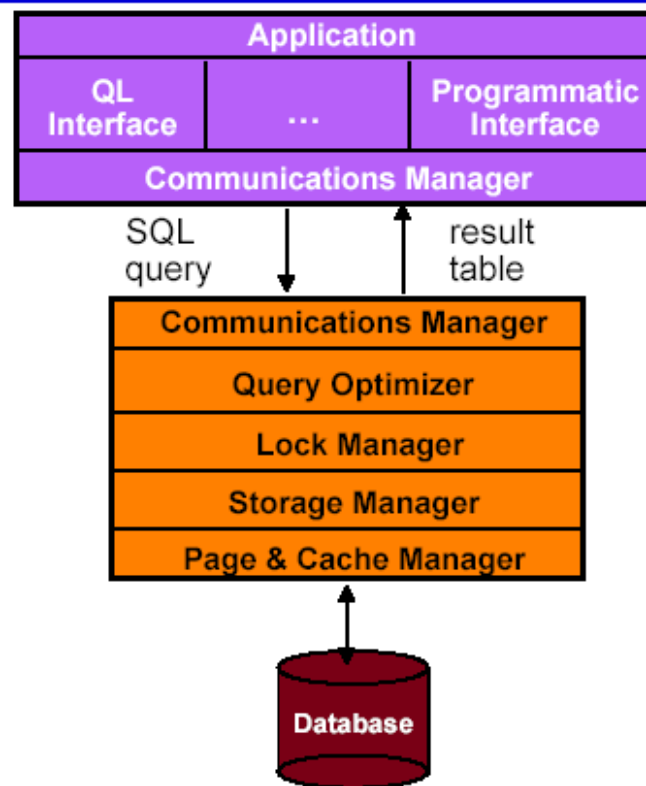
# 分布式数据库系统体系结构

## 不同系统的系统结构图



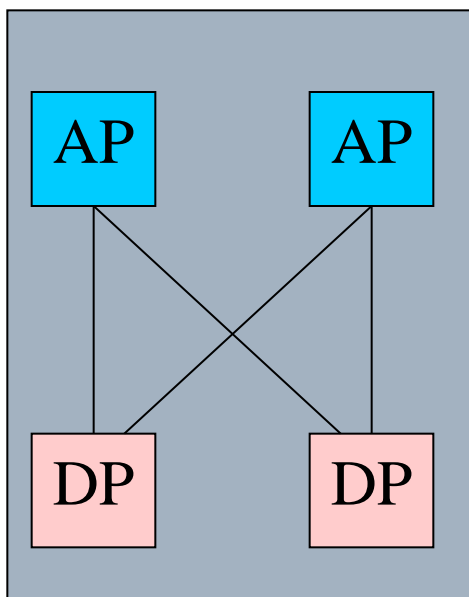
b. 多客户/单服务器(c/s)

## Task Distribution



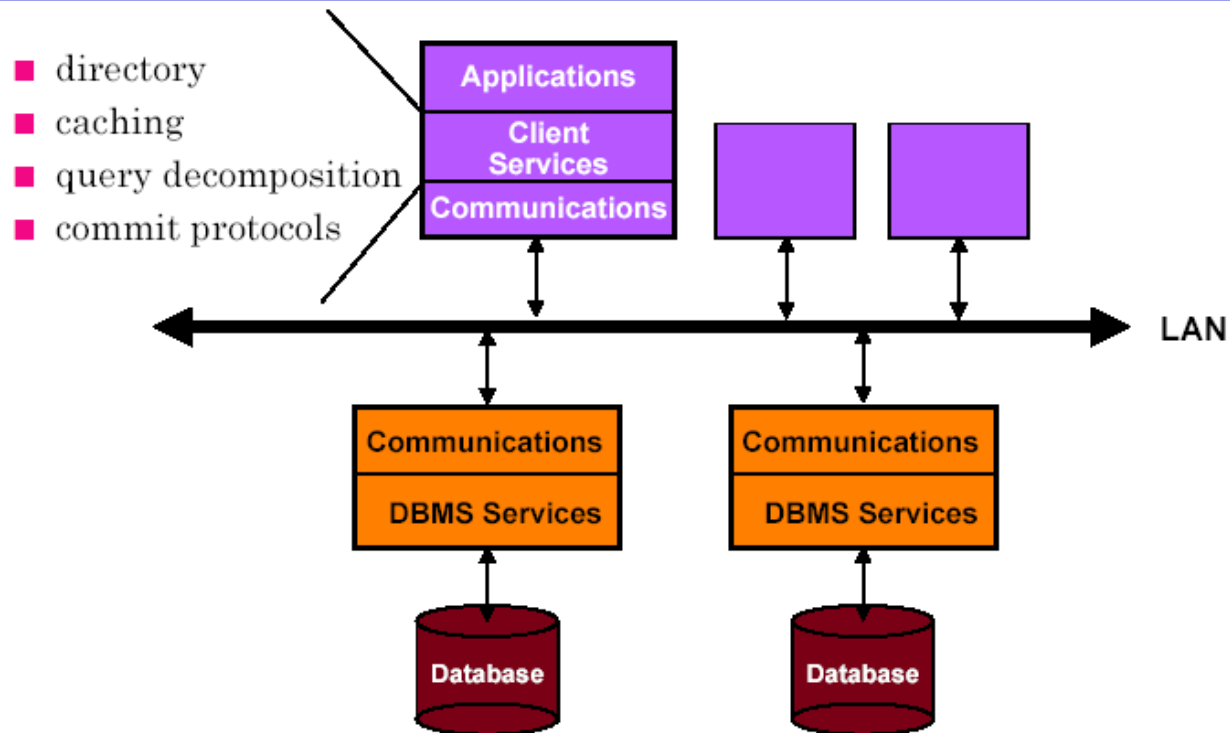
# 分布式数据库系统体系结构

## 不同系统的系统结构图



### c. 多客户/多服务器 (mc/ms)

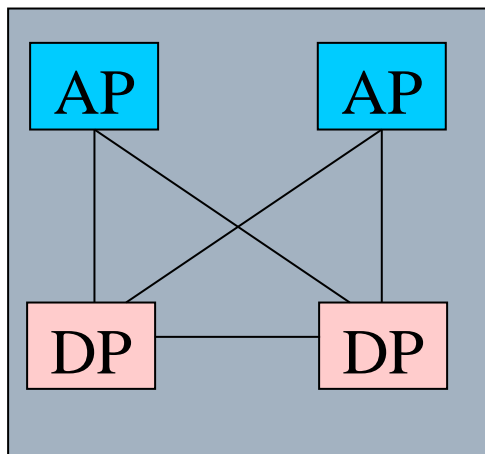
#### Multiple Clients/Multiple Servers



单AP，多DP系统结构属于并行数据库系统

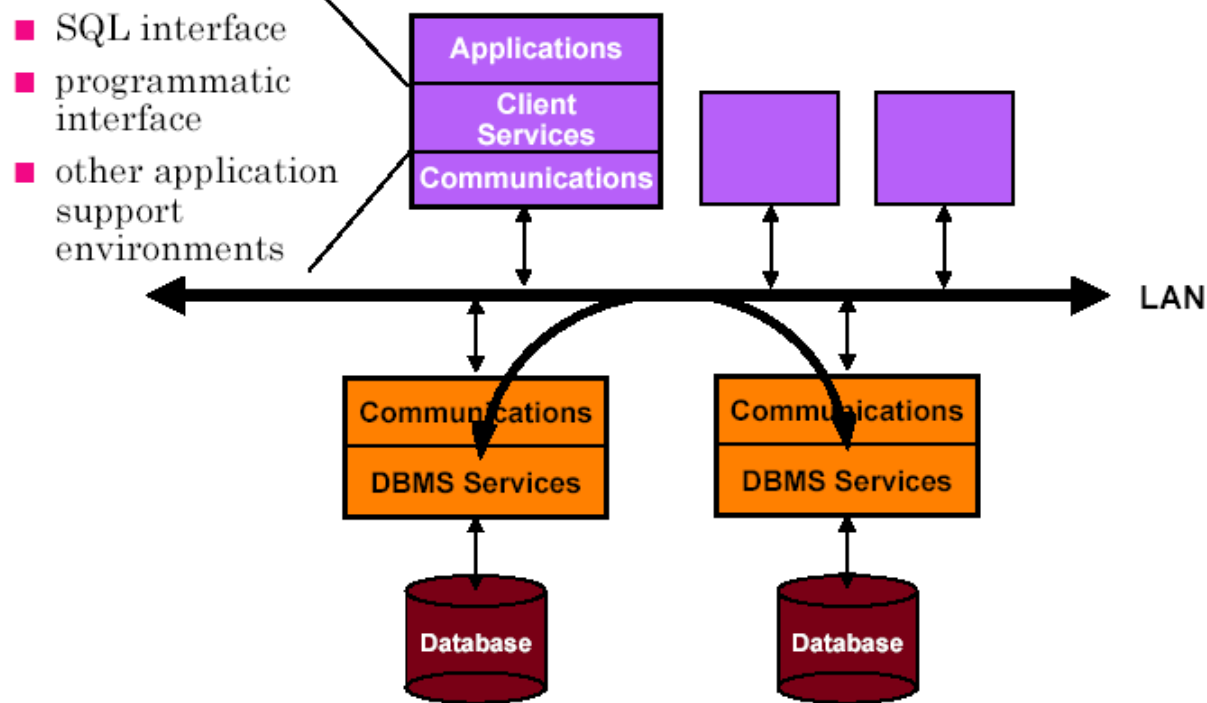
# 分布式数据库系统体系结构

## 不同系统的系统结构图



d. 服务器结构

### Server-to-Server

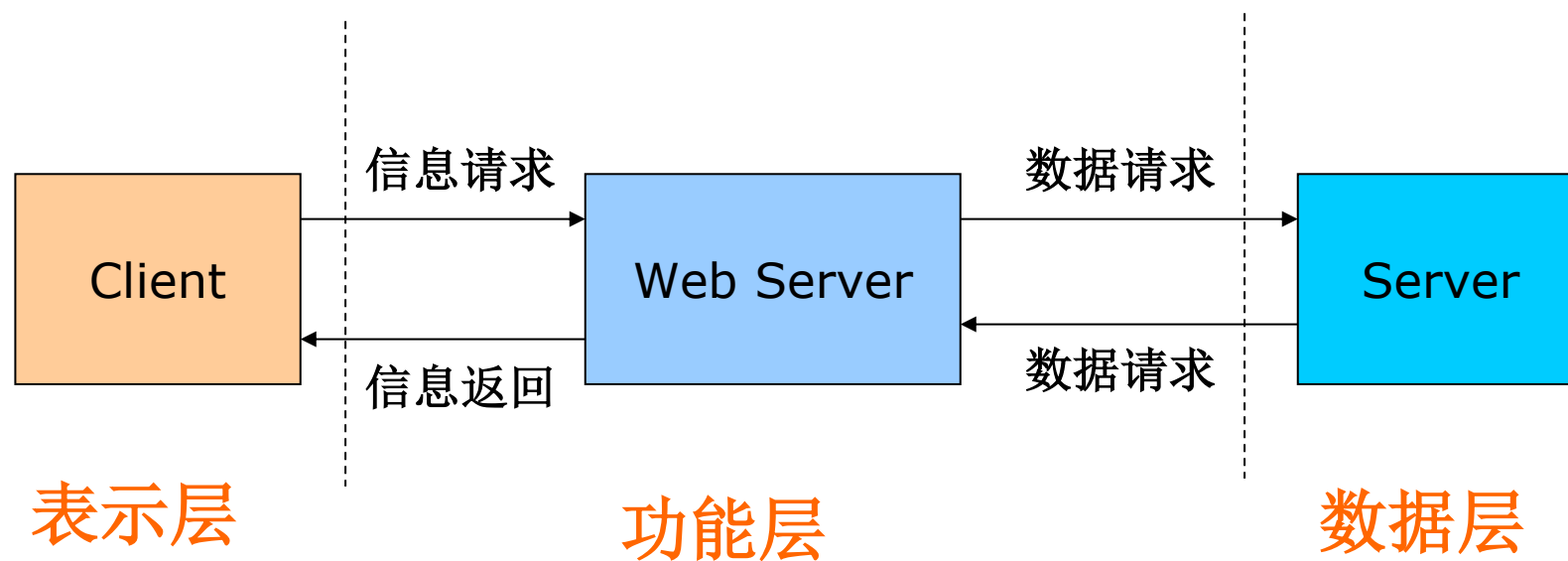


多AP多DP属于典型的分布式数据库系统体系结构

# 分布式数据库系统体系结构

## 不同系统的系统结构图

### Browse/Server

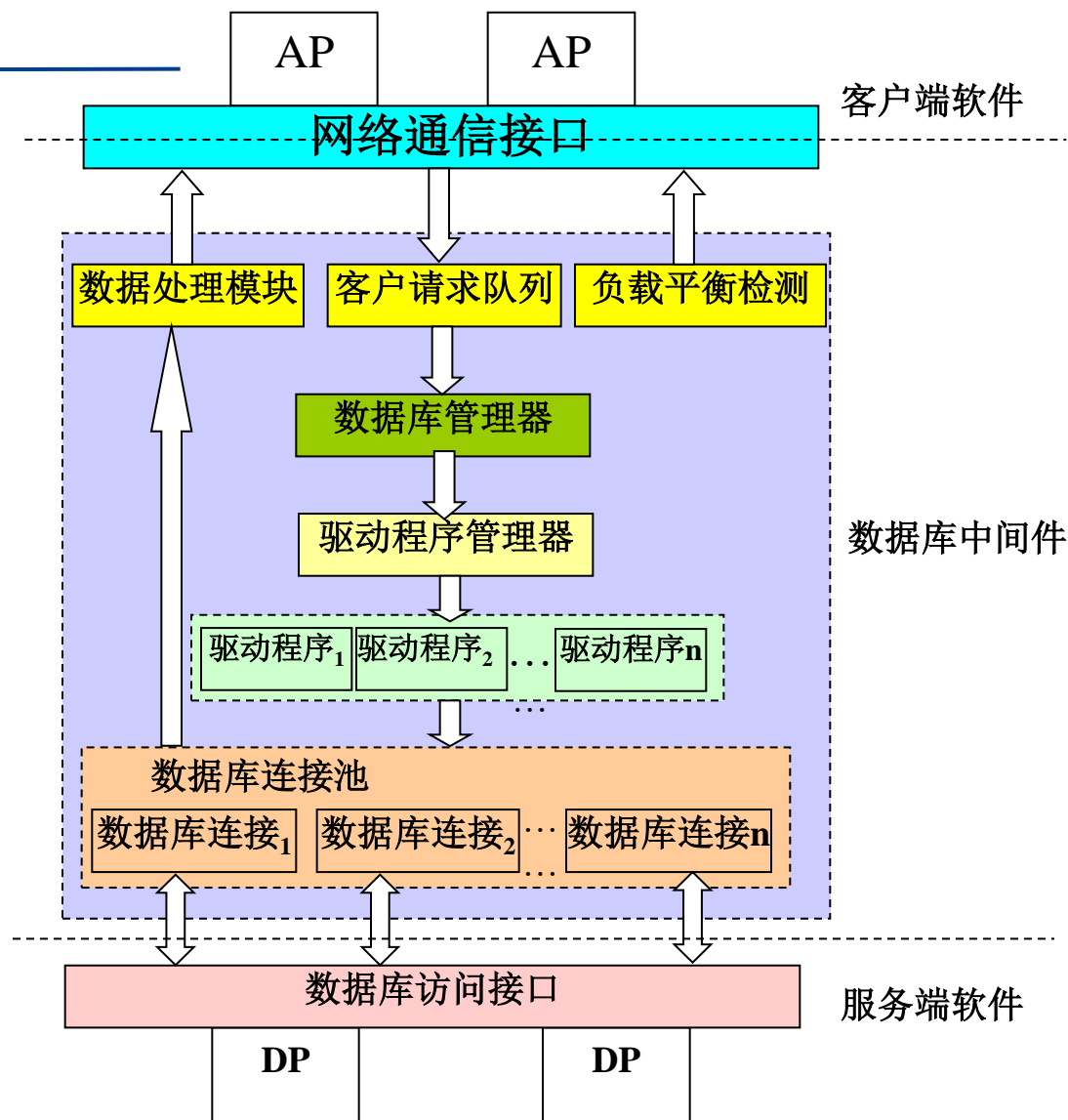


传统的客户端/服务器结构是由**全局事务管理器**统一协调和调度事务的执行，属于**紧耦合**模式，导致系统的复杂度高，资源利用率低。

# 分布式数据库系统体系结构

## 不同系统的系统结构图

- **客户请求队列**：负责存放所有AP发来的数据请求，同时缓存客户的响应结果；
- **负载均衡监测**：负责监控数据库服务器的状态及性能，为中间件提供调度依据；
- **数据处理**：负责处理从数据库返回的数据，按一定的方式将数据传送给AP。
- **数据库管理器**：负责接收客户请求队列中的客户请求，调用相应的驱动程序管理器，完成相应的数据库查询任务。
- **驱动程序管理器**：负责调度相应的数据库驱动程序，实现与数据库的连接；
- **连接池**：实现与物理数据库的连接。



基于“中间件”的客户/服务器的体系结构

## 关于本讲内容

---



**祝各位学习愉快!**



# 感谢观看！

讲解人：李鸿岐