

* 第一章

- * 分布式系统定义: **硬件+控制+数据**
- * 基本属性
- * 必要属性
- * 主要特征 1-37
- * 目标 1-38
 - * 可访问性
 - * 开放性
 - * 可扩展性
 - * 可用性
 - * 可靠性
 - * **透明性 1-43、 1-94**
- * 分布式支持技术
 - * Ad hoc: IPC
 - * **中间件** 1-66、1-82 中间件要解决的问题 1-18 中间件的类别、特性 1-20
 - * RPC
- * 安全性: 机密性、完整性、可靠性
- * 可伸缩性 1-90
- * 第二章: 计算范型

- * 范型: 模式、例子、模型
- * 系统模型
 - * 物理模型
 - * 体系结构模型
 - * 基础模型
- * **通信范型 2-14**
 - * 进程间通信 IPC
 - * 远程调用 2-102
 - * 请求-应答协议
 - * RPC: 远程过程调用
 - * RMI: 远程方法调用 2-142
 - * **RMI 与 RPC 的共性与特殊性 2-142**
 - * **RMI 的实现 2-165**
 - * 伺服器 2-164
- * **分布式系统的基本目标: 资源共享**
- * **4+1 视图模型**: 逻辑、开发、进程、物理、场景 2-22
- * 移动代理 2-28
- * 交互范型 2-33 各种计算模型抽象层次
- * 组播通信 2-79
- * IP 组播 2-81
 - * IP 层
 - * UDP
- * **覆盖网络** 2-89
- * RPC**的调用语义 2-132**
 - * 或许
 - * 至少一次

- * 至多一次
- * 组播排序 2-199
 - * FIFO
 - * 因果序
 - * 全序
- * **发布订阅 2-109**
 - * 基于事件
 - * **特征：异构性、异步性、为通知提供不同的传递性保障 2-213**
 - * 基于渠道 2-152
 - * 基于主题
 - * 基于内容
 - * 基于类型
- * 洪泛
- * 过滤
- * 广告
- * 分布式并行计算
 - * 数据并行
 - * 模型并行
 - * 张量并行
 - * 流水线并行
- * 故障模型
 - * 故障分类 2-318
 - * 进程遗漏性故障中最常见的是进程崩溃，利用超时机制解决
 - * 随机故障（拜占庭故障）
- * 第三章：分布式系统的同步与互斥
 - * 物理时钟同步
 - * 外部同步
 - * 通过通信延时时钟进行校准
 - * **Christian 算法计算 3-12**
 - * $\text{客户机校准后时间} = \text{服务器返回 CUTC} + (\text{UTC 到达客户机时间} - \text{客户端发起时钟请求时间} - \text{服务器中断处理时间}) / 2$
 - * 此过程是主动的
 - * Berkeley 算法 3-14
 - * 被动的
 - * 时间服务器定期询问每台机器的时间，并将其取均值，并高速所有机器这个新值。只能拨快和减缓，不能拨慢
 - * 逻辑时钟同步 3-22
 - * 本质：在先发生关系；所有进程在时间的发生顺序上达成一致
 - * 系统的全局状态
 - * 快照算法
 - * 互斥控制
 - * 基于令牌的

- * 令牌环互斥算法 3-73
 - * 基于非令牌的
 - * **Lamport 3-63**
 - * Ricart-Agrawala 3-68
 - * 选举算法 3-78
 - * def: 选出一个进程执行特别的任务
 - * **bully 算法 3-68**
- * 第四章：资源管理
 - * 资源共享：数据迁移；计算迁移；作业迁移
 - * 资源管理
 - * 工作站模型
 - * 分布式资源管理策略 4-18
 - * 招标算法
 - * 由近及远算法
 - * 回声算法
 - * 分布式系统的死锁问题 4-38
 - * 死锁的图论模型
 - * 处理死锁的策略 4-46
 - * 死锁的 and & or 条件 4-48
 - * 分布式死锁预防
 - * 基于时间戳
 - * wait-die
 - * wound-wait
 - * 集中式死锁检测
 - * 假死锁问题
- * 第五章：任务调度与负载均衡
 - * 进程与线程
 - * 分布式进程
 - * 调度单位是任务队列
 - * 处理器管理
 - * 空闲
 - * 等待
 - * 运行
 - * 处理器的通信
 - * 点对点
 - * 广播
 - * 资源调度：静态和动态
 - * 分布式调度的基本目标
 - * 调度算法的有效评价
 - * 静态调度 5-11
 - * 基于任务优先图
 - * 基于任务相互关系图
 - *

* 第六章：存储与文件系统

* 数据库管理系统：层次，网状，关系

* **ACID 6-8**

* A 原子性

* C 一致性

* I 隔离性

* D 持久性

* **No-sql**

* 最佳拍档

* **CAP**

* **传统分布式注重 C，云计算平台注重 A 和 P**

* C：一致性

* 强一致性

* 弱一致性

* 最终一致性（弱一致性的特例）

* 因果一致性

* 读写一致性

* 会话一致性

* 单调读一致性

* 单调写一致性

* Quorum NRW

* A：可用性 6-31

* P：分区容忍性

* CA 满足强一致性、可用性的系统，传统数据库

* AP 满足可用性、分区容忍性的系统，大多数网站架构

* CP 满足一致性、分区容忍性的系统，性能不高

* **BASE：牺牲强一致性来获得可用性和分区容忍性（可伸缩性），云计算就是**

* B：基本可用

* S：软状态

* E：最终一致性

* 分布式存储系统分类 40

* **存储类型 6-42**

* DAS

* SAN

* NAS

* 云存储关键技术

* 容灾备份

* 数据删除

* 存储引擎

* 哈希表

* B 树

* LSM 树

* 数据模型

* 文件模型

* 关系模型

- * 键值模型
- * 表格模型
- * 分布式文件系统 DFS【特点】 50
 - * 高可用，高负载，高性能
 - * 统一命名空间
 - * 扩展性强
 - * 高性能
 - * 高度负载均衡
- * 分布式文件系统的需求 55
 - * 透明性
 - * 一致性
 - * 安全性
 - * 效率
- * **分布式文件系统的透明性 51**
 - * 访问透明性
 - * 位置透明性
 - * 移动透明性
 - * 性能透明性
 - * 伸缩透明性
- * 有状态和无状态服务器
- * 文件系统的三个组成部分：
 - * 平面文件服务器
 - * 目录服务器
 - * 客户端
- * 层次文件系统
 - * 目录树
 - * 目录图
 - * 文件组
- * 分布式文件系统的命名方法 62
 - * 无层次命名
 - * 结构化命名
 - * 基于属性的命名
- * 迭代名称解析
- * 递归名称解析
- * DNS 递归解
- * 同步机制 73
 - * DFS 锁机制
 - * 分布式锁机制
 - * 租赁方式
 - * 基于时间的机制
- * 文件远程访问方法 74
- * 缓存的额粒度和地点
- * 更新策略
 - * 写直达
 - * 写回
 - * 关闭时写回
- * 高速缓存

- * 高速缓存一致性

- * 第七章：分布式数据库

- * 体系结构

- * 分布式数据库定义 7-10

- * Top-down【设计新的】和 Bottom-Up【集成旧的】

- * 分布式数据库的基本特点 7-15

- * 结构特点：物理分布，逻辑相关

- * 应用特点：站点自治

- * 数据特点：数据分布透明性

- * 控制特点

- * 冗余特点

- * 事务特点：维护原子性、一致性、隔离性、持久性（ACID）

- * 数据独立性

- * 逻辑独立性

- * 物理独立性

- * 数据分布独立性

- * nosql: 6-17

- * 查询处理和优化

- * 分布式事务管理

- * **与集中式 DB 相比，分布式 DB 有（数据分布性）特点；与分散式 DB 相比，又具有（逻辑整体性）特点；**

- * **区别系统是分散式还是分布式就是判定系统是否支持（全局应用）**

- * **分布式 DB 中，用户看到的系统如图一个集中式 DBS，因为（位置透明性）（复制透明性）分片透明性、复制透明性和位置透明性**

- * **同类型数据模型，但是 DBMS 不同：同构异质型 DDBS**

- * 分布式数据库系统增加的模式级别 7-26

- * 全局外观模式

- * 全局概念模式

- * 分片模式

- * 分布模式

- * 数据分配与分片 7-30

- * 全局数据库（GDB）通过分片模式->片段数据库（FDB）

- * FDB 通过 分配模式 -> 物理数据库 PDB

- * 分片：从全局关系到片段模式的映射

- * 分配：从片段关系到物理关系的映射

- * 数据分片 7-32

- * 水平

- * 垂直

- * 混合

- * 导出

- * 分片原则

- * 完备性

- * 可重构性

- * 不相交性

- * **分布透明性 7-43**

- * 分片透明性（最高层次）
- * 位置透明性
- * 局部数据模型透明性
- * 数据本地化：全局查询到片段查询的变换 7-99

* 第九章：云计算

- * 云计算定义 1: 9-13

- * 什么是云计算 9-17

* 通过集中式远程计算资源池，以按序分配方式为终端用户提供强大而廉价的计算服务能力

* 云计算分类

* **按照服务层面分 9-23**

* **SaaS：提供各种应用软件服务**

- * 关键技术 9-193

- * 呈现技术
- * 多租户技术

* 架构 9-205

- * 呈现层
- * 调度层
- * 业务层
- * 数据层

* **PaaS：提供软件支撑平台服务**

* **IaaS：提供接近于裸机的计算资源和基础设施服务**

- * 硬件资源作为服务提供给用户

* 主要技术

- * 虚拟化技术
- * 资源动态管理和调度技术
- * EC2

- * I：计算资源；存储资源

- * S：技术模式；商业模式

- * 特征与优势 9-104

* 基础设施云；平台云；应用云

* **按照计算系统类型分类 9-25**

* **公有云**

* **私有云**

* **社区云**

* **混合云**

* 云计算的发展目标

- * 云计算的特点 9-29

- * 按序分配
- * 无限资源
- * 资源共享
- * ...

* **云计算的特征：9-29**

- * **低成本，大规模**
- * **平滑扩展**
- * **资源共享**
- * **动态分配**
- * **跨地域**

* 云计算与并行计算、网格计算、效用计算

* 云计算的优势 9-37

- * 优化产业布局
- * 推进专业分工
- * 提升资源利用力
- * 减少初期投资

* **机遇与挑战，发展动力，趋势 9-45**

* 云计算的关键 9-83

- * 资源虚拟化和弹性调度解决小粒度应用资源共享
- * 大数据存储处理和并行计算服务提供大粒度应用计算能力

* **云计算的关键技术 9-85**

- * **虚拟化：虚拟机**
- * **云计算架构**
- * **资源调度技术**
- * **并行计算技术**
- * ...

* 云计算中新的挑战 9-94

- * 足够稳定
- * 能够伸缩
- * 保证安全
- * 高效率

- * -----
- * 安全性
- * 可用性
- * 可伸缩性
- * 高性能
- * 标准化
- * 基础设施层的基本功能 9-106
- * 系统虚拟化技术 9-116
 - * 硬件仿真
 - * 全虚拟化（最成熟）
 - * 半虚拟化
 - * 硬件辅助虚拟化
 - * 操作系统级虚拟化
- * **传统分布式注重 C，云计算平台注重 A 和 P**
- * MapReduce
- * 分布式文件系统 9-170
 - * 基本特征
 - * 透明性
 - * 并发访问
 - * 高可用性
 - * 基本需求
 - * 数据冗余
 - * 异构性
 - * 一致性
 - * 高效性
 - * 安全性
- * 分布式数据库
- * 分布式协同管理
 - * 常用并发控制方法 9-173
 - * 基于锁
 - * 基于时间戳
 - * 乐观并发控制

* 基于版本的并发控制