

目 录

第1章 绪论 (1)

- 1.1 分布式系统简述 (1)
 - 1.1.1 分布式系统的定义 (1)
 - 1.1.2 分布式系统中的主要特征 (5)
 - 1.1.3 分布式系统目标 (8)
 - 1.1.4 分布式系统的类型 (12)
- 1.2 分布式系统的拓扑结构 (14)
- 1.3 分布式与计算机网络 (14)
- 1.4 分布式体系结构 (16)
 - 1.4.1 体系结构样式 (16)
 - 1.4.2 分布式系统体系结构 (16)
- 1.5 分布式和云计算系统模型 (19)
 - 1.5.1 协同计算机集群 (19)
 - 1.5.2 计算网格 (21)
 - 1.5.3 对等网络家族 (21)
 - 1.5.4 互联网上的云计算 (23)
- 1.6 分布式系统和云计算软件环境 (24)
 - 1.6.1 用以支持分布式的技术 (24)
 - 1.6.2 分布式对象计算中间件 (26)
 - 1.6.3 组件中间件 (26)
 - 1.6.4 发布/订阅中间件和面向消息的中间件 (27)
 - 1.6.5 面向服务的体系结构(SOA) (27)
 - 1.6.6 分布计算系统的设计问题 (30)
- 1.7 分布式系统的趋势 (32)
 - 1.7.1 泛在联网和现代互联网 (32)
 - 1.7.2 移动和无处不在计算 (33)
 - 1.7.3 分布式多媒体系统 (33)
 - 1.7.4 把分布式计算作为一个公共设施 (34)
- 1.8 分布式系统的挑战 (35)
- 1.9 分布式系统实例 (36)

- 1.9.1 电子商务类 (36)
- 1.9.2 搜索类 (36)
- 1.9.3 社交类 (37)
- 1.9.4 邮箱类 (37)
- 1.9.5 图片及视频类 (38)
- 1.9.6 数据仓库类 (38)
- 1.9.7 云服务类 (38)

习题 (39)

第2章 分布式计算范型 (40)

- 2.1 分布式通信范型 (40)
 - 2.1.1 进程间通信 (41)
 - 2.1.2 客户/服务器范型及变体 (47)
 - 2.1.3 远程过程调用 (54)
 - 2.1.4 远程方法调用 (60)
 - 2.1.5 间接通信范型 (76)
 - 2.1.6 P2P 范型与覆盖网络 (93)

2.2 网络服务范型 (94)

2.3 移动代理范型 (95)

2.4 云服务范型 (99)

习题 (102)

第3章 分布式系统的同步 (104)

- 3.1 分布计算系统中的同步机构 (104)
 - 3.1.1 物理时钟 (105)
 - 3.1.2 逻辑时钟 (111)
- 3.2 系统的全局状态 (117)
 - 3.2.1 全局状态的形式定义 (117)
 - 3.2.2 全局状态的获取 (118)
 - 3.2.3 一致全局状态的充要条件 (119)
- 3.3 互斥 (120)
 - 3.3.1 集中式算法 (121)
 - 3.3.2 非基于令牌的解决方案 (122)
 - 3.3.3 基于令牌的互斥算法 (127)

3.4 选举算法	(133)
3.4.1 Chang 和 Roberts 的算法	(134)
3.4.2 Bully 选举算法	(136)
3.4.3 环选举算法	(137)
3.4.4 非基于比较的算法	(138)
3.5 投标	(139)
3.6 自稳定算法	(141)
习题	(143)
第4章 分布式资源管理	(145)
4.1 资源共享	(145)
4.1.1 数据迁移	(146)
4.1.2 计算迁移	(146)
4.1.3 作业迁移	(147)
4.2 分布式资源管理策略	(147)
4.2.1 局部集中管理方式	(149)
4.2.2 分散式管理方式	(149)
4.2.3 分级式管理方式	(150)
4.2.4 一个分散式资源管理算法	(150)
4.2.5 招标算法	(151)
4.2.6 由近及远算法	(152)
4.2.7 回声算法	(153)
4.3 分布式系统中的死锁问题	(155)
4.3.1 死锁发生的条件	(155)
4.3.2 死锁的图论模型	(155)
4.3.3 处理死锁的策略	(157)
4.3.4 死锁的 AND 条件和 OR 条件	(157)
4.4 分布式死锁预防	(159)
4.4.1 预防死锁的一般方法	(159)
4.4.2 基于时间戳的预防死锁方法	(159)
4.5 分布式死锁检测	(161)
4.5.1 集中式的死锁检测	(162)
4.5.2 分布式死锁检测	(163)

4.5.3 等级式死锁检测	(165)	
4.5.4 死锁检测的实例	(166)	
习题	(169)	
第5章 分布式调度与负载均衡	(171)	
5.1 分布式进程与处理机管理	(171)	
5.1.1 进程与线程	(171)	
5.1.2 分布式进程	(174)	
5.1.3 分布式进程的同步与互斥	(174)	
5.1.4 处理机管理	(174)	
5.2 调度算法概述	(176)	
5.2.1 调度算法的分类	(177)	
5.2.2 调度算法的目标和有效性评价	(179)	
5.3 静态调度	(179)	
5.3.1 任务划分与分配	(180)	
5.3.2 基于任务优先图的任务调度	(183)	
5.3.3 两种最优调度算法	(185)	
5.3.4 基于任务相互关系图的任务调度	(187)	
5.3.5 基于遗传算法和模拟退火算法的任务分配策略	(189)	5.4 动态
调度	(192)	
5.4.1 动态调度的组成要素	(192)	
5.4.2 动态负载平衡算法的分类、设计决策和使用的参数	(194)	
5.4.3 其他相关因素	(197)	
5.5 负载平衡	(198)	
5.5.1 概述	(198)	
5.5.2 负载平衡算法分类	(198)	
5.5.3 负载平衡算法的组成	(199)	
5.5.4 发送者主动算法	(200)	
5.5.5 接收者主动算法	(201)	
5.5.6 双向主动算法	(202)	
5.5.7 梯度模型	(202)	
5.5.8 接收者主动的渗透算法	(203)	
5.5.9 预约策略	(203)	

5.5.10 投标策略	(203)
5.5.11 广播策略	(204)
5.6 智能型任务调度算法	(204)
5.6.1 任务调度中的知识及其表示	(204)
5.6.2 任务调度程序的结构	(205)
5.6.3 任务调度算法的实现	(206)
习题	(206)

第6章 分布式存储与文件系统 (208)

6.1 分布式存储概述	(208)
6.1.1 分布式存储系统的定义与实现问题	(208)
6.1.2 分布式存储系统分类	(209)
6.1.3 存储引擎	(211)
6.1.4 数据模型	(211)
6.2 分布式文件系统	(213)
6.2.1 分布式文件系统的特点和基本要求	(214)
6.2.2 分布式系统的命名要求	(215)
6.2.3 分布式文件系统中的命名方法	(216)
6.2.4 名字的结构	(217)
6.2.5 名字空间	(218)
6.2.6 名字解析	(219)
6.2.7 分布式系统中的名字空间的实现	(221)
6.2.8 分布式文件系统设计	(221)
6.3 同步机制	(230)
6.4 缓存	(230)
6.4.1 文件的远程访问方法	(230)
6.4.2 缓存的粒度和地点	(231)
6.4.3 更新策略、缓存有效性检验和一致性	(231)
6.4.4 缓存和远程服务的比较	(232)
6.4.5 高速缓存(caching, 超缓存)	(233)
6.5 容错和可扩充性	(238)
6.5.1 无状态服务和有状态服务	(238)
6.5.2 可用性与文件复制	(239)

6.5.3 可扩充性	(239)
6.5.4 用线程实现文件服务器	(239)
6.6 安全性	(240)
6.7 GFS	(240)
6.8 Hadoop HDFS	(248)
6.9 其他的分布式文件系统及其比较	(248)
习题	(257)
第7章 分布式数据库技术与应用	(259)
7.1 分布式数据库简介	(259)
7.1.1 分布式数据库系统概述	(259)
7.1.2 分布式数据库特点与分类	(260)
7.1.3 分布式数据库系统的体系结构和组成成分	(264)
7.1.4 分布式数据库管理系统	(265)
7.2 数据分布	(266)
7.2.1 数据分布概念	(266)
7.2.2 数据划分原则及分片方法	(268)
7.2.3 数据分配的原则和方法	(272)
7.3 数据分布方式	(274)
7.3.1 哈希方式	(275)
7.3.2 按数据范围分布	(275)
7.3.3 按数据量分布	(276)
7.3.4 副本与数据分布	(276)
7.3.5 本地化计算	(278)
7.3.6 数据分布方式的选择	(278)
7.4 一致性模型	(279)
7.4.1 严格一致性	(279)
7.4.2 顺序一致性和可线性化一致性	(280)
7.4.3 相关一致性	(281)
7.4.4 FIFO 一致性	(282)
7.4.5 弱一致性	(284)
7.4.6 释放一致性	(285)
7.4.7 数据项一致性	(286)

7.4.8	基本副本协议	(288)
7.5	分布式数据库系统的设计	(290)
7.5.1	分布式数据库系统设计概述	(290)
7.5.2	自顶向下设计分布式数据库	(293)
7.5.3	自底向上设计分布式数据库	(296)
7.6	分布式数据库中的查询处理和优化	(298)
7.6.1	分布式查询优化概述	(298)
7.6.2	全局查询及局部查询处理	(299)
7.6.3	全局查询转换基础知识	(300)
7.6.4	全局查询到逻辑查询的转换	(302)
7.6.5	逻辑查询到物理查询的转换	(309)
7.6.6	基于关系代数等价变换的优化算法	(310)
7.6.7	基于半连接操作的查询优化算法	(310)
7.6.8	基于直接联接操作的查询优化算法	(311)
7.7	分布式数据库中的事务管理和恢复	(311)
7.7.1	原子事务与性质	(311)
7.7.2	事务处理的分类	(311)
7.7.3	基于原子事务处理的局部恢复	(313)
7.7.4	CAP 和 BASE 理论	(313)
7.7.5	分布式事务架构的演进	(315)
7.7.6	并发控制	(322)
7.8	分布式提交协议	(327)
7.9	分布式数据库实例简介	(329)
7.9.1	HBase 分布式数据库	(329)
7.9.2	Google Spanner——全球级的分布式数据库	(333)
7.9.3	SequoiaDB 分布式存储原理	(335)
习题		(336)
第 8 章	分布式系统中容错技术	(338)
8.1	分布式系统中的故障模型	(338)
8.1.1	基本概念	(338)
8.1.2	基本的故障模型	(339)
8.2	容错系统的基本构件	(341)

8.2.1 坚固存储器	(341)
8.2.2 故障—停止处理器	(341)
8.2.3 原子操作	(342)
8.3 节点故障的处理	(342)
8.3.1 向后式恢复	(342)
8.3.2 向前式恢复	(343)
8.4 向后恢复中的问题	(345)
8.4.1 检查点的存储	(345)
8.4.2 检查点方法	(346)
8.4.3 异步检查点	(347)
8.4.4 同步检查点	(348)
8.4.5 混合检查点	(348)
8.4.6 报文日志	(348)
8.5 拜占庭故障的恢复	(350)
8.5.1 恢复中的设计问题	(352)
8.5.2 错误屏蔽和进程复制	(353)
8.5.3 容错系统中的一致性协议	(354)
8.5.4 带有可信部件的拜占庭系统	(358)
8.6 可靠的组通信	(359)
8.6.1 基本的可靠组播方案	(359)
8.6.2 可靠的组播通信中的可扩充性	(360)
8.6.3 原子组播 (atomic multicast)	(361)
8.6.4 处理软件故障	(363)
8.7 分层故障屏蔽和成组故障屏蔽	(364)
8.7.1 分层屏蔽	(364)
8.7.2 成组故障屏蔽	(364)
8.7.3 坚固存储器	(366)
8.7.4 主服务器与备份服务器	(366)
习题	(367)
参考文献	(369)