

- 1、分布式事务除了要考虑对数据的存取操作序列之外，还需要考虑（ ）
A、通信原语 B、控制报文 C、通信原语和控制报文 D、无
- 2、分布式事务的实现模型包括（ ）
A、进程模型和服务器模型 B、进程模型和主从控制模型
C、服务器模型和三角控制模型 D、主从控制模型和三角控制模型
- 3、对于进程模型和服务器模型，描述正确的是（ ）
A、进程模型中进程可以为多个事务进行服务
B、服务器模型中进程可以为多个事务进行服务
C、进程模型只需要少量的进程
D、服务器模型需要大量进程
- 4、分布式事务管理的目标不包括（ ）
A、高效的事务执行效率 B、系统可靠性 C、系统可用性 D、数据依赖性
- 5、关于分布式事务执行的控制模型，描述不正确的是（ ）
A、主要分为三种控制模型
B、主从控制模型不允许任意两个从节点之间进行通讯
C、三角控制模型就是 P2P 模型
D、层次控制模型扩展性好，但是实现上更加复杂
- 6、关于 CAP、ACID、BASE 描述，不正确的是（ ）
A、ACID 和 BASE 是 CAP 理论的两个分支
B、CAP 不能同时满足一致性、可用性和分区容错性
C、ACID 要求同时满足原子性、一致性、隔离性和持久性
D、BASE 要求同时满足基本可用、软状态、强一致性
- 7、关于弱事务型数据库和强事务型数据库，描述不正确的是（ ）
A、前者强调可用性和分区容错性，后者强调一致性
B、前者复杂性高扩展性低
C、后者数据完整性和实效性较高
D、前者基本不支持跨行跨表事务
- 8、关于故障描述正确的是（ ）
A、软故障是可预期故障 B、硬故障也可以自行恢复
C、分布式数据库系统中存在通信故障 D、通信故障即是网络分割故障
- 9、故障恢复模型不需要（ ）
A、数据 B、日志文件 C、Redo 恢复策略 D、Undo 恢复策略
- 10、日志文件中，不包括（ ）
A、数据日志 B、事务日志 C、命令日志 D、检查点日志
- 11、关于缓冲区更新策略描述不正确的是（ ）
A、fix/flush 更新策略需要 redo 处理
B、fix/no_flush 更新策略 redo 处理
C、non_fix/flush 更新策略需要 undo 处理
D、non_fix/no_flush 更新策略需要 redo 处理和 undo 处理

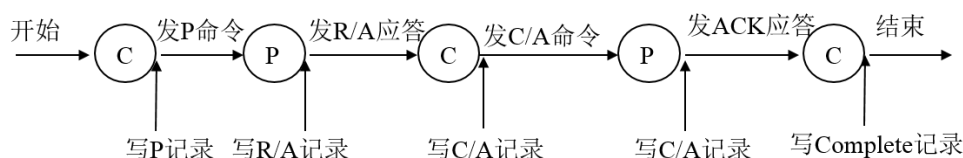
- 12、关于两阶段提交协议的终结协议（场地故障时使用）描述正确的是（ ）
- A、参与者等待协调者的最终决定 B、协调者重新发消息给参与者
C、参与者会访问其他参与者的状态再做最终决定 D、终结即全局废弃
- 13、关于大数据数据库系统的故障检测技术描述，不正确的是（ ）
- A、自适应故障检测有利于提高系统容错性能
B、租约机制故障检测中监控进程会定期想被监控进程发放租约
C、Gossip 协议是强一致性协议
D、大数据数据库系统故障恢复策略会与多副本管理策略相结合
- 14、并发控制是不能保证（ ）
- A、事务的一致性 B、事务的隔离性 C、数据的一致性 D、数据的持久性
- 15、有如下四个历程 H1、H2、H3：
- H1：R1(x)R1(y)W1(x)W1(y)R2(x)R2(y)W2(x)W2(y)
H2：R1(x)R1(y)W1(x)R2(x)W1(y)R2(y)W2(x)W2(y)
H3：R1(x)R1(y)W1(x)R2(x)R2(y)W1(y)W2(x)W2(y)
H4：R1(x)W1(x)R2(x)W2(x)R1(y)W1(y)R2(y)W2(y)
- 请问以下错误的是（ ）
- A、H1 是串行历程，其余三个是并行历程 B、H2 与 H1 等价
C、H3 与 H1 等价 D、H4 与 H1 等价
- 16、以下关于加锁描述，不正确的是（ ）
- A、锁粒度可以大到整个数据库，也可以小到属性和字段
B、锁可以分为排他锁和共享锁
C、锁可以分为显示锁和隐式锁
D、锁粒度越小，系统开销越小，但提高了并发度
- 17、以下关于基本的两段封锁协议和严格的两段封锁协议的描述，不正确的是（ ）
- A、基本的两段封锁协议可以避免丢失修改错误
B、基本的两段封锁协议可以避免不能重复读错误
C、两者最大的不同点在于解锁时刻不同
D、两者都不能保证事务更新数据的永久性
- 18、关于基于锁的并发控制方法的实现方法，描述不正确的是（ ）
- A、集中式实现方式简单，方便管理
B、分布式实现方式可以避免中心场地负载重和单点失效
C、集中式和分布式实现方式都需要多个 2PL 调度器
D、对复制数据的封锁可以采用基于特定副本的封锁方法和基于投票的封锁方法
- 19、关于分布式死锁管理描述，不正确的是（ ）
- A、死锁等待图有局部等待图 LWEg 和全局等待图 GWEg
B、无论是局部等待图还是全局等待图，哪一个有回路都代表有死锁
C、集中式死锁检测和层次死锁检测均需要多个死锁检测程序
D、分布式死锁检测由各场地共同完成，每个场地对等
- 20、关于大数据数据库并发控制技术描述，不正确的是（ ）

- A、部分大数据数据库系统为提高系统性能，抛弃了事务的概念
- B、大数据数据库系统中读事务通常可以采用最新读、快照读、非一致性读
- C、大数据数据库系统中读模式三种均可以读到未提交事务的数据
- D、写事务模式采用了 Paxos 一致性协议完成写日志

21、如果分布式数据库系统总共有 5 个节点/场地，其中 1 号场地为协调者，2、3、4、5 号场地为四个参与者，且任意节点之间传输消息平均时间为 3ms。针对两阶段提交协议 2PC（决定阶段和执行阶段），请回答以下问题：

- 1) 请描述在两阶段中第一阶段（决定阶段）中协调者节点 1 需要做哪些动作？
- 2) 如果采用集中式 2PC 实现，问传递消息总数为多少？传输消息总用时为多少？
- 3) 如果采用分布式 2PC 实现，问传递消息总数为多少？传递消息总用时为多少？
- 4) 如果采用分层式 2PC 实现，问传递消息总数为多少？传递消息总用时最短为多少？最长为多少？
- 5) 如果采用线性 2PC 实现，问传递消息总数最少为多少？最多为多少？传递消息总用时最短为多少？最长为多少？

22、利用两阶段提交协议，场地故障发生在不同点，采用不同的恢复策略。



- 1) 如果某部分用户的事务均落在协调者场地 C 上，当协调者场地 C 出现故障，读取日志文件，获取如下日志序列：
 $\{B_1, D_1^1, B_2, D_1^2, D_2^1, K, D_1^3, B_3, C_1, D_2^2, B_4, D_3^1, A_2, C_3, D_4^1\}$
 其中 B 代表事务开始，下角标代表事务编号，D 代表数据操作，上角标代表事务内操作编号，C 代表事务提交，K 代表检查点、A 代表事务废弃。请描述恢复过程（即初始化 redo 和 undo 表，到最终的 redo 和 undo 表的变化过程）。
- 2) 如果故障点发生在第一参与者场地 P，且在参与者 P 写 R/A 记录之前出现错误，会出现什么现象？采取什么样的恢复策略？
- 3) 如果故障点发生在第二个协调者场地 C，协调者在写 P 记录之后，写 C/A 记录之前出错，出错时协调者已将 P 命令发送给参与者，会出现什么现象？采取什么样的恢复策略？
- 4) 如果在协调者和参与者进行通信过程中，丢失了 R/A 应答报文，会做如何处理以实现恢复？

23、假设分布式数据库系统共有三个场地 N1、N2、N3 上，每个场地上存在如下事务以及事务依赖关系：

N1: {T1, T4, T7} {T1→T5, T4→T1, T7→T2}

N2: {T2, T5, T8} {T2→T3, T5→T9, T8→T3}

N3: {T3, T6, T9} {T6→T1, T9→T6}

同时，假设在 T1 事务执行之前 X 值为 3，节点 N1 上的 T1 和 T7 两个事务（其中 B 代表

事务开始，C 代表事务提交，R 代表读 X 操作，W 代表更新 X 操作，下角标代表事务编号，上角标代表事务内操作编号）可能执行情况如下：

时间	t1	t2	t3	t4	t5	t6	t7	t8	t9
事务 T1	B ₁	R ₁ ¹			R ₁ ²		R ₁ ³	C ₁	
事务 T7		B ₇	R ₇ ¹	W ₇ ¹		R ₇ ²			C ₇

请回答以下问题：

- 1) 假如分布式数据库系统采用基于加锁的方式进行并发控制，请画出 N1、N2 和 N3 三个场地的全局死锁等待图，并根据此图判断是否存在死锁？如果存在死锁，可以采用哪两种方法解决？是否可以预防死锁？
- 2) 假如分布式数据库系统采用多版本时间戳算法进行并发控制，针对 N1 场地上的 T1 和 T7 两个事务执行情况如上表所示，事务 T1 和 T7 的时间戳（开始时间）分别是 t1 和 t2，那么 R₁² 和 R₇² 两个操作的时间戳是多少？如果 W₇¹ 操作把 X 更新为 5，那么 R₁¹、R₇² 和 R₁³ 结果分别是多少？除了多版本时间戳算法之外，还有哪些基于时间戳的并发控制方法？
- 3) 基于锁的并发控制方法和基于时间戳的并发控制方法分别是通过事务的什么来维护 and 实现并发事务的可串行化的？