

- 1、分布式数据库管理系统 DDBMS 兴起于 ( )  
A、 60 年代中期      B、 60 年代后期      C、 70 年代前期      D、 70 年代中期
- 2、分布式数据库系统是 ( ) 的数据库系统  
A、 地理和逻辑上分散      B、 地理和逻辑上集中  
C、 地理上分散和逻辑上集中      D、 地理上集中和逻辑上分散
- 3、分布式数据库的主要特点是 ( )  
A、 数据透明性和场地自治性      B、 分片透明性和通信自治性  
C、 分布透明性和设计自治性      D、 复制透明性和执行自治性
- 4、分布式数据库系统 ( )  
A、 不能提高系统性能      B、 不能降低投资成本  
C、 允许存在副本      D、 分布式数据库的安全性和完整性
- 5、基于层次的 DBMS 系统结构不包括 ( )  
A、 应用处理器      B、 通信处理器      C、 数据处理器      D、 数据库系统
- 6、应用处理器包括 ( ) (按照从上到下顺序)  
A、 用户接口、语义数据控制器、全局查询处理器、全局执行监控器  
B、 语义数据控制器、用户接口、全局查询处理器、全局执行监控器  
C、 用户接口、全局查询处理器、语义数据控制器、全局执行监控器  
D、 语义数据控制器、用户接口、全局执行监控器、全局查询处理器
- 7、分布式数据库系统的异构型不包括 ( )  
A、 数据异构性      B、 数据库管理系统异构性  
C、 平台异构性      D、 应用系统异构性
- 8、分片 ( )  
A、 不能减少网络传输      B、 不能增大事务处理的局部性  
C、 提高数据的可用性      D、 不能提高查询效率
- 9、分片原则不包括 ( )  
A、 可重构性      B、 完备性      C、 不相交性      D、 可重复性
- 10、水平分片的关系代数表示中使用了哪些操作符 ( )  
A、 投影      B、 选择      C、 连接      D、 投影和选择
- 11、以下 ( ) 不属于分布库  
A、 全分隔式分布库      B、 全复制分布库  
C、 部分复制分布库      D、 集中库
- 12、复制技术 ( )  
A、 不能减少网络负载      B、 提高系统性能  
C、 能有效解决数据一致性维护问题      D、 避免了更多的数据冲突
- 13、查询优化的内容不包括 ( )  
A、 执行运算的次序      B、 执行运算的方法  
C、 所访问副本的场地      D、 查询请求的场地
- 14、查询处理和查询优化的顺序, 正确的是 ( )  
A、 查询分解, 查询存取优化, 数据局部化, 局部查询优化

- B、 查询分解，数据局部化，查询存取优化，局部查询优化  
C、 全局查询优化，片段查询优化，局部查询优化，存取优化  
D、 全局查询优化，存取优化，片段查询优化，局部查询优化
- 15、查询分解的步骤（ ）  
A、 约简、规范化、分析、查询重构  
B、 分析、规范化、约简、查询重构  
C、 规范化、分析、约简、查询重构  
D、 规范化、约简、分析、查询重构
- 16、关于场地描述不正确的是（ ）  
A、 逻辑上分为查询场地、源数据场地和执行场地  
B、 查询场地可以是源数据场地  
C、 查询场地不可以是执行场地  
D、 执行场地可以使源数据场地
- 17、关于查询总时间和响应时间不可能发生的是（ ）  
A、 总时间大于响应时间  
B、 总时间小于响应时间  
C、 总时间大于等于响应时间  
D、 总时间不等于响应时间
- 18、假设学生选课表 SC (sno: 5B, cno: 5B, grade: 5B) 总共有 10000 条元组，假设平均每个学生选修 50 门课，投影出来学号和成绩两个属性之后 (Sno 不用去重)，问投影结果大小约为（ ）。  
A、 10KB                      B、 100KB                      C、 2KB                      D、 20KB
- 19、假设学生选课表 SC 占用外存 10000 块 (Block)，学生表 S 占用外存 100 块 (Block)，问采用哈希连接方式 (一个桶能全部装入内存缓冲区块) 的代价是多少（ ）次 IO。  
A、 30300                      B、 50500                      C、 20200                      D、 40400
- 20、常见的直接连接算法不包括（ ）  
A、 归并排序连接      B、 嵌套循环连接      C、 索引连接      D、 半连接
- 21、在学生选课系统中，底层数据库系统的数据模型存在四张表格，分别如下：  
(1)、学生表: Stu (Sno: 2B, Sname: 3B, Dept: 3B, Uno: 2B);  
(2)、课程表: Cou (Cno: 2B, Cmane: 3B, Dept: 3B);  
(3)、选课表: SC (Sno: 2B, Cno: 2B, Grade: 2B);  
(4)、学校表: Uni (Uno: 2B, Uname: 3B, Province: 3B)。
- 数据库系统分片需求:
- (1)、学生表 Stu 按照 Dept 为“CS”和“非 CS”分为两个分片 S1 和 S2，其中 S1 分片中有 5000 条元组，大小 50KB; S2 分片中有 15000 条元组，大小为 150KB;
- (2)、选课表 SC 按照课程是给本学院上的还是给其他学院上的分为两个分片 SC1 和 SC2，其中 SC1 分片中有 150000000 条元组，大小 900MB; SC2 分片中有 1500000 条元组，大小为 9MB，其中不及格的选修记录占比 10%，且“CS”学院学生选修“非 CS”学院课程的选修记录占比 10%;
- (3)、课程表 Cou 和学校表 Uni 数据较少，不分片;
- (4)、以上为 1 个学校的数据，共 10 个学校的数据，10 个学校数据均匀分布。

数据库系统各分片的场地分配如下：

(1)、场地 N1: S1、SC1、Cou、Uni;

(2)、场地 N2: S2、SC2、Cou、Uni。

问：

(1)、查询浙江工业大学“CS”学院学生选修“非 CS”学院课程且成绩不及格(<60)的学生姓名、课程名和成绩。画出该查询的片段查询优化树。

(2)、假设以 N1 为查询场地，以 N1 为执行场地，执行过程如下：

- 把 SC2 中满足条件（学生选课中成绩不及格）的元组（Sno, Cno, Grade）选择出来，作为 SC2'；
- 把 SC2'传输到 N1 场地，然后和 S1、Cou 以及 Uni 进行连接得到最终结果。

假设通信单位传输代价为 1，通信启动代价为 0，计算通信代价。

(3)、假设以 N1 为查询场地，以 N2 为执行场地，执行过程如下：

- 把 S1 中满足条件（浙江工业大学“CS”学院学生）的元组（Sno）选择出来，作为 S1'；
- 把 S1'和 SC2 进行连接，同时选择成绩不及格的元组，就找到满足条件（浙江工业大学“CS”学院学生选修“非 CS”学院课程”且成绩不及格(<60)的学生）的选课信息（Sno, Cno, Grade），作为 SC2'；
- 把 SC2'传输到 N1 场地，然后和 S1 以及 Cou 进行连接。

假设通信单位传输代价为 1，通信启动代价为 0，计算通信代价。

(4)、题（3）采用半连接实现查询，基于题（2）和（3）计算该半连接的收益和代价分别是多少？并判断该半连接是否是收益半连接？