**全国计算机等级考试-计算机三级数据库技术**

## 第十七张试卷

1. 如果数据库文件中的每个查找码值在索引文件中能对应一个索引记录,则成该索引为稠密索引
2. 在三层浏览器/服务器体系结构的数据库应用系统中,触发器是在数据库服务器上执行的.

解析：在三层浏览器/服务器体系结构的数据库应用程序程序中,数据层位于数据库服务器,通过DBMS完成具体的数据储存和数据存取等数据管理功能.触发器是一种特殊测存储过程,它是存储在数据库中供所有用户程序调用的子程序.

1. UML图中,用于表达系统中不同的包,命名空间或不同的项目间彼此关系的图称为包图.

解析：UML图中,用于表达系统中不同的包,命名空间或不同的项目间彼此关系的图称为包图.

4．在SELECT语句中,用于分组的子句是GROUP BY,ORDER BY是排序的子句.

1. SQL Server 2018提供了4个排名函数,其中RANK ROW\_NUMBER函数返回结果集中每行数据在每个分区内的排名,并且每个分区内的排名从1开始,但排名中间数值可能有间断.
2. 在存储过程中输入参数不需要注明,输出参数需要用关键字OUT或OUTPUT注明该参数是输出参数.
3. 在SQL Server 2008中,要在T1表的C1列上建立一个非聚集的唯一索引,请补全下列语句.

CREATE UNIQUE NONCLUSTERED INDEX idx\_t1 ON T1(C1)

关键字UNIQUE使用来注明该索引是唯一索引,关键字NONCLUSTERED是用来注明该索引是非聚集索引.

1. 在SQL Server 2008中,删除数据库用户用DROP关键字.
2. 派生性冗(rong)余列是指表中增加的列由表中的一些数据项经过计算生成.以减少查询时连接操作,避免使用聚合函数.

在数据库调用过程中,在

10．将经常一起使用的具有公共列值的多个表中的数据行存储在一起的数据库存储发方法被称为聚集(Cluster).

11．检查点技术的最大特点是最大限度地减少了数据库完成恢复时所必须执行的日志部分.

12．分布式数据库中,位置独立性,数据分片独立性和数据复制独立性和数据复制独立性形成了分布式数据系统中的分布透明性,使用户完全感觉不到数据是分布,与使用集中式数据完全一样.

13．对一个数据集的数据进行分组,使得每一组内的数据尽可能相似而不同间的数据尽可能的不同,这样数据挖掘方法称为聚类算法/Clustering.

14．数据仓库是一个**面向主题的**,**集成的**,**非易失的**,**且随时变化**的数据集合.

## 第十六张试卷

1. 数据库完整性约束的作用对象分为：**列,元组,关系**三种级别.元组对应于行,关系对应表.
2. 在数据文件的主码属性集上建立的索引称为主索引,在数据文件的非主属性建立的索引称为辅索引.
3. 在UML图中,描述系统中硬件和软件的物理配置情况和系统体系结构的图称为部署图.

注：部署图又称配置图,描述系统中硬件和软件的物理配置情况和系统体系结构.

1. SELECT基本语法：SELECT 查询内容 FROM 表名 WHERE 条件表达式

GROUP BY 待分组的列名 HAVING 列名 ORDER BY 待排序的列明.

1. 使用FOR或者AFTER选项定义的触发器为后触发器.
2. 在对游标数据进行提取的过程中,可以使用@@FETCH\_STATUS全局变量判断数据提取的状态.由于@@FETCH\_STATUS对于在一个连接上的所有游标是全局性的,不管是对哪一个游标,只要执行一次FETCH语句,系统都会对@@FETCH\_STATUS赋一次值,以表明该FETCH语句的执行情况.
3. 在SQL Server 2008中用BACKUP关键字进行数据库备份.
4. 在SQL Server 2008中系统管理员的角色是：sysadmin
5. 垂直分割是根据列的特点进行分割,分割之后所得的所有表中除了都含有主码列外其余列都不同,其查询时会减少I/O次数,但其缺点是查询所有数据时需要Join(连接)操作.
6. 存储方式中的聚集就是将经常一起使用的多个表中的数据行按照其公共列值存储在一起.
7. 在备份语句(BACKUP)中默认情况下使用完整数据库备份,而关键字DIFFERENTIAL专门改变备份：差异备份.
8. 水平片是在关系中从行的角度(元组)依据一定条件划分为不同的片段.而垂直片面是在关系中以列的角(属性)依据一定条件划分为不同对的片段.
9. 在数据仓库中,元数据一般分成技术元数据和业务元数据.
10. 在数据仓库的导出数据或物化视图(实视图)的维护策略中,只在用户查询时发现数据已经过期才进行更新对的策略称为: 延时维护策略.

## 第十五张试卷

1.IDEFO需求建模方法由箭头和矩形框两种元素构成,矩形框代表功能活动,写在矩形框内的动词短语描述功能活动的名称.

2.防火墙技术是用来保证数据库应用系统的网络坏境安全的.

3.在UML的状态机图中,状态之间的转移是由事件驱动的.

4.SELECT INTO 语句表示从一个表中选取数据,然后把该数据插入到另一个表中,该语句结构通常用于创建表的备份复件或者用于对记录进行存档.该语句使用时要求目标表不存在,因为在插入时会自动创建.

5.通过使用EXISTS谓词进行存在性测试,其形式是：WHERE [NOT] EXISTS(子查询)带EXISTS谓语的子查询不返回查询的结果,只产生逻辑真值和逻辑假值.

6.游标指针各个参数含义:

NEXT:返回紧跟在当前行之后的数据行,并且当前行递增为结果行.

PRIOR:返回紧跟当前行前面的数据行,并且当前递减为结果行.

FIRST:返回游标中的第一行并将其作为当前行.

LAST:返回游标中的最后一行并将其作为当前行.

例如：若要提取游标cur1当前行指针所指的前一行数据,使用的选项是:FETCH FRIOR FROM cur1

7.删除架构用DROP关键字.如删除U1结构,使用的语句是：DROP SCHEMA U1.

8.在SQL Server 2008中,有两种身份验证模式,一种是:混合身份验证模式,另一种是:Windows身份验证模式.

9.**物化**视图是一种包含查询结果的数据库对象.其数据不是在视图使用时读取的.而是预先计算的.

10.在数据库动态转储机制中,一般需要使用日志文件才能将数据库恢复到某一时刻的正确状态.

扩:在动态转储中,利用转储文件只能间数据库恢复到转储过程中的某个状态,且转储文件中的数据可能处于不一致状态,只有和日志文件综合起来使用,才能将数据库恢复到一致的状态,或将数据库恢复到故障发生前的状态.

11.分布式数据库系统的恢复控制采用的是最典型的策略是基于: 两阶段的提交协议.

两阶段的提交协议将场地的事物管理器分为协调者和参与者.

A通过协调者在第一阶段询问所有参与者事物是否可以提交,参与者做出应答.

B在第二阶段协调者根据参与者的回答决定事物是否提交.

12.时间序列分析可以称为数据演变分析,它能描述行为随时间变化的对象规律或趋势,并对其进行建模.

13.在关联规则挖掘中,关联规则的成立与否一般用支持度和置信度两个指标进行描述.

## 第十四张试卷

1. 在IDEF1X数据建模方法中,直角矩形用于表示独立实体集,圆角矩形用于表示从属实体集.
2. 三层B/S结构的数据库应用系统中,业务规则是在应用服务器端完成.解析

解析:三层B/S结构系统结构将数据处理功能分解并分布在表示层,应用层,和数据层三个层次上面.

* 表示层:位于客户端,由Web浏览器实现.
* 应用层:位于Web应用服务器,实现面向具体应用领域的业务规则.
* 数据层位于数据库服务器,通过DBMS完成具体的数据存储和数据存取等数据管理功能.

1. 用例模型通过对系统参与者及其重要行为的描述,表达系统的功能需求
2. 在T-SQL语言中,EXCEPT运算符用于实现两个查询结果的差运算.
3. 声明游标的语法是:DECLARE cursor\_name Curosr FOR SQL Statement..
4. SQL Server数据库中创建数据库应用CREATE关键字,修改数据库用ALTER关键字,删除数据库用DROP关键字.

例如:若要修改DB1数据库结构,使用的语句是:ALTER DATABASE DB1…

1. 在SQL Server 2008中,登录账户的来源共有**两种**.

解析:登录账户的来源有两类:

* 第一类:由SQL Server自身负责身份验证的登录账户.
* 第二类:登录到SQL Server的Windows网络账户.可以是组账户或者使用户账户.

1. **增加冗余列的调优方法是指在多个表中增加具有相同语义的列,它常用来在查询时避免连接操作.**
2. **在分割表的调用方法中,水平分割会给应用增加复杂度,特别是在查询所有数据时需要union操作.**

**垂直分割使得数据行变小,一个数据也能存放太多的数据,在查询时就会减少I/O次数,从而提高系统性能.但其缺点是:查询所有数据时需要Join(连接)操作.**

1. SQL Server 2008中,事务日志备份仅可用在完整恢复模式和大容量日志恢复模式中.
2. 每个处理机拥有独立的主存储器和磁盘,不共享任何资源和并行的数据库系统结构称为:无共享结构.
3. 在数据仓库的数据组织中,描述数据的综合或细节程度的指标称为粒度,**粒度越大,表示综合程序越高,粒度越小,表示综合程序越低.**
4. 在分类预测任务中,需要用到的数据一般包括:训练集,测试集和验证集.

## 第十三张试卷

1. 数据库逻辑结构设计阶段:把数据库概念设计的结果(ER模型),转换为具体的数据库关系系统的数据模型.
2. 数据库的实现与部署(实施)阶段主要包括:

* 建立数据库结构
* 数据加载事务
* 事务和应用程序的编码及测试
* 系统集成
* 测试和试运行
* 系统部署

1. **在UML提供的图中,活动图主要用来描述系统,用例,和程序模块中逻辑流程的先后执行次序,并行次序.**
2. 在使用TOP限定结果集中:

* WITH TIES: 表示包括最后一行取值并列的结果.
* TOP n: 表示取查询结果的前n行数据.
* TOP n percent: 表示取查询结果的前n%行数据.

1. 将查询语句产生的结果集指定一个临时命名的名字,这些命名的结果集就称为:公用表表达式.命名后的公共表表达式可以被多次引用.其语法格式为:**WITH 公共表表达式标识符(列表1——>列表n) AS (SELECT statement..)**
2. 在SQL Server 2008中,可以使用DROP FUNCTIN语句从当前数据库中删除一个或多个用户自定义函数.
3. 在SQL Server 2008的触发器中,用于存储INSERT语句所影响的行的副本的临时表为: INSERTED

解析:INSERTED表用于存储INSERT和UPDATE语句所影响的行的副本.

1. 在SQL Server 2008中,用户自定义函数主要包括:标量函数和表值函数两类.
2. 如果要使某用户可以将授予他的权限再转授给其他用户,那么在授权语句中需要加行**WITH GRANT OPTION**选项.
3. 物化(或索引)视图不是在使用时才读取基本表,而是预先计算并保存表连接或聚集等耗时较多的操作结果,这样在查询时大大提高了读取速度.
4. **非预期的事务内部故障** 是不能由事物程序处理的,如运算溢出故障,并发事物死锁故障,违反了某些完整性限制而导致的故障等.
5. **分布式数据库**最基本的三个特征是:

* 本地自治.
* 非集中式管理.
* **高可用**性.

1. 给定一个销售交易数据库,从中找出这些交易中的某些物品和其他物品之间的关系,这种数据挖掘一般称为:关联挖掘(或关联分析,关联规则).
2. **数据仓库**共有四个特性,分别是:

* 面向主题
* 集成
* 非易失
* 随时间变化

## 第十二张试卷

1. 从功能角度数据库应用系统可以划分为如下部分:

* 表示层: 负责所有与用户交互的功能.
* 业务逻辑层: 负责根据业务逻辑需要将表示层获取的数据进行组织后,传递给数据访问层,或将数据访问层的数据进行相应的加工处理.传递给表示层用于展示.
* 数据访问层: 数据访问层负责与DBMS系统进行交互,提取或存入应用系统所需的数据.
* 数据持久层: 负责保存和管理应用系统数据.

1. 三层浏览器/服务器结构的数据库应用系统组成成分:

* 浏览器
* 数据库服务器
* 应用服务器

1. **在UML中,通信图也称协作图,显示了在交互过程中各个对象之间的组织交互关系及对象彼此之间的链接.通信图显示的是对象之间的关系,并不侧重交互的顺序,它没有将时间作为一个单独的维度,而是使用序列号来确定消息及并发线程的顺序.**
2. 在SQL Server中,一组行被称为一个窗口,开窗函数是指可以用于’分区’或’分组’计算的函数,这些函数结合OVER子句对组内的数据进行编号,并进行求和,计算平均值等统计.因此,从这个角度来说:SUM,AVG以及ROW\_NUMBER等都可以被称为开窗函数.开窗函数可以分别应用于每个分区,把每个分区看成一个窗口,并为每个分区进行计算,开窗函数必须放在***OVER***子句前边.
3. 使用AFTER/FOR选项定义的触发器为后触发器,使用INSTEAD OF选项定义的触发器为前触发型触发器.
4. 在T-SQL中,提取cur1游标中倒数第五行数据的语句时:

FETCH ABSOLUTE **-5** FROM cur1..

解析:游标**ABSOLUTE{n|@nvar}**参数意义:

* N或@nvar为正: 返回从游标第一行开始的第n行,并将返回行变成新的当前行.
* N或@nvar为负: 返回从游标最后一行开始之前的第n行,并将返回行变成新的当前行.
* N或@nvar为零: 不返回行,必须是整数常量,并且@nvar的数据类型必须是smallint,tinyint/int.

1. 在SQL Server 2008中,对视图创建唯一聚集索引后,视图的结果集存储在数据库中,就像带有聚集索引的表一样.建有唯一聚集索引的视图被称为**索引视图**,也称为物化视图.
2. SQL Server固定数据库角色有:

* db\_owner: 具有在数据库中进行全部操作的权限,包括配置,维护数据库及删除数据库.
* db\_accessadmin: 具有添加或删除数据库角色的权限.
* db\_securityadmin: 具有管理数据库角色,角色成员以及数据库汇总的语句和对象权限.
* db\_ddladmin: 具有执行数据定义语言(DDL)的权限.
* db\_backupoperator: 具有备份数据库,备份日志的权限.
* db\_datareader: 具有查询数据库中所有用户数据的权限.
* db\_datawrite: 具有插入,删除和更新数据库中所有用户数据的权限.
* db\_denydatareader: 不允许具有查询数据库中所有用户数据的权限,等同于对所有的表和视图授予了DENY SELECT权限.
* db\_denydatawrite: 不允许具有INSERT,DELETE和UPDATE数据库中所有用户数据的权限.

1. 在SQL Server 2008中,可以在查询语句里加入:FOR XML子句实现以XML格式返回查询结果.
2. **在T-SQL中恢复数据库使用RESTORE DATABASE语句,恢复事物日志使用RESTORE LOG语句.**

例如:*RESTORE DATABASE students FROM MyBack*

1. 互联网上的应用服务一直被称作软件既服务,它是一种软件分配模式.**平台既服务是指:通过网络提供操作系统和相关服务.而无需下载或安装.**基础设施既服务是指将用于支持运作的设备对外提供服务,这些设备包括**存储**,**硬件**,**服务器**和**网络组件**.
2. **局部数据模型透明性** 处于分配模式与局部概念模式之间,它使用户在编写应用程序时不但要了解全局数据的分片情况,还要了解各片段的副本复制情况及各片段和它们副本的场地位置分配情况,但是不需要了解各场地上数据库的数据模型.
3. **分布式数据库**的数据分配有:

* 集中式: 所有数据片段都安排在一个场地上.
* 分割式: 所有全局数据有且只有一份,它们被分割成若干片段,每个片段被分配在一个特定场地上.
* 全复制式: 全局数据有多个副本,每个场地上都有一个完整的数据副本.
* 混合式: 全局数据被分为若干个数据子集,每个自己被安排在一个或者多个不同的场地上,但是每个场地未必保存所有数据.

1. 表分割有水平分割和垂直分割两种:

* 水平分割:根据数据行的使用特点进行分割,分割之后所得的所有表的结构都相同,而存储的数据不同.
* 垂直分割:根据数据列的特点进行分割,分割之后所得的所有表中除了都含有主码列外其余列都不相同.

## 第十一张试卷

1. 在三层B/S结构的数据库系统中,***数据库服务器*** 接收应用服务器提出的数据操作请求.按照**事务逻辑**对数据库进行查询和修改,并将数据访问结果返回给应用服务器.

所以可知: 事务处理是在数据库服务器端完成的.

1. ***用例模型*** 通过对系统参与者及其重要行为的描述,表达系统的***功能***需求,既用户的期望.
2. T-SQL 的*SELECT*语句中,使用***DISTINCT***选项可以消除查询结果返回重复行数据.
3. FETCH的功能是通过Transact-SQL服务器游标检索特定行,具体语法:

FETCH

[[NEXT|PRIOR|FIRST|LAST|ABSOLUTE{n|@nvar}|RELATIVE{n|@nvar}] FROM ..]..

1. **建立索引的语法:**

CREATE [UNIQUE][NONCLUSTERED|CLUSTERED] INDEX 索引名 ON 表名(要求建立索引的列名).

* UNIQUE: 表明该索引是唯一索引.
* NONCLUSTERED: 关键字表明该索引是非聚集索引.
* CLUSERED: 关键字表明该索引是聚集索引.

1. 数据库性能优化中,存储优化包括:**建立物化视图**和**聚集**.
2. ***数据库重组***是指DBA按照系统设计要求对数据库存储空间进行全面调整,如调整磁盘分区方法和存储空间,重新安排时刻的存储,整理回收碎块,重建索引等,以提高数据库性能.
3. 当数据库发生系统故障时,***容错对策***是在重新启动数据库后,撤销(UNDO)所有未提交事务,重做(REDO)所有已提交事务.
4. SaaS是指:软件既服务.

解析: Saas: Software as a Serive(软件既服务)的简称.

1. ***聚类***是一类常见的机器学习方法,聚类算法所处理的数据对象一般都没有标记,因此聚类也被称为***无监督***的学习方法.
2. **粒度概念:粒度是用于描述数据坏境中数据综合程度的概念.**

* 粒度越大:综合程序越高,细节程序越低,能回答的查询就越少.
* 粒度越小:综合程序越低,细节程序越高,能回答的查询就越多.

## 第十张试卷

1. 在数据库事务**概要设计阶段**,使用read和write原语来表达存取操作.
2. 如果数据库中的一个基本表中的数据量很小,查询以整表扫描为主,并且插入,删除,更新等操作非常频繁,该基本表可以采用**堆文件**组织方式,因为堆文件无需建立索引,维护代价非常低.
3. 在UML提供的图中,**用例图**用于描述系统与外部系统及用户之间的交互.
4. 使用TOP限定结果集:

* WITH TIES: 表示包括最后一行取值并列的结果.
* TOP n : 表示取查询结果的前n行数据.
* TOP n percent: 表示取查询结果的前**n%**行数据.

1. 删除存储过程的语句: DROP PROC(PROCEDURE) 存储过程名.
2. 在SQL Server 2008的触发器中,用于存储DELETE语句所影响的行的复本的临时表为: DELETED

解析: DELETE表用于存储DELETE和UPDATE语句所影响的行对的复本.

1. SQL Server 2008中,用户自定义函数主要包括:**标量函数**和**表值函数**.
2. 使用REVOKE命令回收用户权限时,如果要将该用户转授给其他用户的权限一并回收,则该命令中需要加上:**CASCADE**选项.CASCADE对应于GRANT语句中的WITH GRANT OPTION,表示要撤销的权限也会从此主体授权或拒绝该权限的其他主体中撤销.
3. 建有唯一聚集索引的视图称为索引视图,定义索引视图的语句如下:

CREATE **UNIQUE** **CLUSTERED** INDEX 索引名 ON 视图名

1. **系统故障**又称软故障,是指数据库在运行过程中,由于硬件故障,数据库软件及操作系统的漏洞,突然停电等情况.导致系统停止运转,所有正在运行的事务以非正常方式终止,需要系统重新启动的一类故障.
2. **分布式数据系统**的组成成分:

* 局部数据库管理系统
* 全局数据库管理系统
* 全局数据字典
* 通信管理

1. ETL工具通常需要完成的处理操作包括抽取,转换和装载.

* 解析:ETL是实现数据集成的主要技术,既填充更新数据仓库的数据抽取,转换,装载的数据采集过程.

## 第九张试卷

1. **运行管理与维护阶段** 包括:

* **日常维护**
* **监控与分析**
* **性能优化调整**
* **系统进化**

1. 三层浏览器/服务器结构的数据库应用系统是由**浏览器**,**应用服务器**和**数据库服务器**三部分组成.
2. 用例之间的关系包含:**扩展**,**使用**,**组合**三种.
3. **聚合开窗函数**只能使用**PARTITION BY**子句或都不带任何语句.
4. 在T-SQL查询语句中使用**CASE**函数,可以达到分情况显示不同类型的数据的目的.
5. 视图又称虚拟表,标准视图的结果集并不永远存储在数据库中.每次通过标准视图查询数据时,数据库管理系统都会在内部将视图的定义替换成对基本表的查询语句,然后再对基本表执行查询,系统会产生很大的开销.***解决的方法就是对视图创建唯一聚集索引来提高查询的性能.对视图创建唯一聚集索引后,视图的结果将存储在数据库中.建有唯一聚集索引的视图又称索引视图,也称物化视图.***
6. 在SQL Server 2008,具有系统管理员权限的 服务器角色是:sysadmin

* bulkadmin: 具有执行BULK,INSERT 语句的权限.
* dbcreater: 具有创建,修改,删除和还原数据库的权限.
* diskadmin: 具有管理磁盘文件的权限.
* processadmin: 具有管理运行在SQL Server中的进程的权限.
* securityadmin: 能够处理大多数的日常事务,但没有系统管理员的超级权限.
* serveradmin: 具有设置服务器级别的而配置选项和关闭服务器的权限.
* setupadmin: 该角色的权限仅局限于添加和删除连接服务器.
* sysadmin: 具有在服务器及数据库上执行任何操作的权限.

1. **派生性冗余列**是指表中增加的列有表中的一些数据项经过计算生成.它的作用是在查询时减少连接操作,避免使用聚合函数,提高查询统计的速度,**由于增加了列,会增加存储空间,也减少了查询时间,因此:增加派生性冗余列的调优方法是典型的以空间换时间的调优方法.**
2. 对视图创建唯一**聚集索引**后,视图的结果集将存储在**数据库中**.建有唯一聚集索引的视图称为索引视图,也称物化视图.
3. 在分布式数据库中,**垂直分片**从列的角度将关系分为不同的片断.

* **水平分片**: 是在关系中从**行**的角度(元组)依据一定条件分为不同的片段.关系中的每一行必须至少属于一个片段.以便在需要时可以重构关系.
* **垂直分片**: 是在关系中从**列**的角度(属性)依据一定条件分为不同的片段,各片段中应该包含关系的主码属性,以便通过连接方法恢复关系.

1. **无共享结构**每个处理机拥有独立的主存储器和磁盘,不共享任何资源,它被认为是支持**并行数据库系统**的最好并行结构,比较适用于银行,出纳,民航售票等OLTP类应用.
2. **数据挖掘**作为知识发现的过程,一般由三个主要阶段组成:

* **数据准备**
* **数据挖掘**
* **结果的解释和评论**

## 第八张试卷

1. 每**TPS**代价值是用于衡量**系统性价比**的指标.
2. **DFD既数据流图**,它是由**数据流**,(对数据的加工)**处理**,**数据存储**和**外部项**组成.
3. RELATIVE{n|@nvar}:使用规则:

* n/@nvar为正:返回从**当前行**开始向后的第n行,并将返回行变成新的当前行.
* n/@nvar为负:返回从当前行开始向前的第n行,并将返回行变成新的当前行.
* n/@nvar为0:返回当前行.
* N必须是整数常量,@nvar的数据类型必须是smallint,tinyint/int.

1. ***CRATE [UNIQUE|CLUSTERED|NONCLUSTERED] INDEX INDEX\_NAME ON TABLE\_NAME(PROPERTY\_NAME)***

其中UNIQUE和CLUSTERED为可选项,分别是建立唯一索引和聚簇索引.

* UNIQUE:表示此索引的每一个索引值只对应唯一的数据.
* CLUSTERED:表示要建立的索引是聚簇索引.既索引项的顺序与表中记录的物理顺序一致的索引组织.
* NONCLUSTERED:是非聚集索引.

1. **新增汇总表**的调优方法是一种以空间换时间的调优方法.
2. 备份日志的语法:

* *BACKUP LOG database\_name (数据库名) TO backup\_device(备份的地方).*

1. **分布式数据库分片类型**:

* 水平分片: 按一定的条件把全局关系的所有分组划分成若不相交的子集,每个子集都是关系的一个片段.
* 垂直分片: 把一个全局关系的属性集分成若干子集,并在这些子集上作投影运算,每个投影称为垂直分片.
* 导出分片: 又称为导出水平分片,既水平分片的条件不是本关系属性的条件,而是其他关系属性的条件.
* 混合分片: 以上三种方法的混合.可以先水平分片再垂直分片,或垂直分片再水平分片,或其他形式的分片,但他们的结果是不相同的.

1. 数据仓库的数据模型一般被划分为:

* 概念模型
* 逻辑模型
* 物理模型

1. 数据挖掘方法中,**关联规则** 挖掘就是用于发现数据库中数据间的关联性的.

## 第七张试卷

1. SQL Server中根据索引的实现方式,索引技术可以被分成两大类:

* **有序索引**
* **散列索引**

1. 在分布式数据库中,使用**分配模式**来描述各片段到物理存放场地的映像.
2. SQL Server中有五种约束类型:

* **主键约束** (PRIMARY KEY)
* **外键约束** (FOREIGN KEY)
* **唯一约束** (UNIQUE)
* **缺省约束** (DEFAULT)
* **检查约束** (CHECK)

1. 当一个事务发生故障时需要终止并取消所有已执行的数据修改时应执行(日志文件撤销)**UNDO**操作.
2. 对数据库的访问是通过SQL Server 2008的权限层次结构进行管理的.可以通过GRANT,DENY,REVOKE语句来管理这个权限层次结构.

* **GRANT:** 允许一个数据库用户或角色执行所授权限指定的操作.
* **DENY:** 拒绝一个数据库用户或角色的特定权限,并阻止它们从其他角色中继承这个权限.
* **REVOKE:** 收回已经授权的权限.

1. 为了保证事务操作的完整性,需要利用到事务性质中的**原子性**.
2. **OLAP**的实现技术主要分为以下三类:

* 基于关系数据库的OLAP(POLAP,Relational OLAP).
* 基于多维数据库的OLAP(MOLAP,Multi-Dimensional OLAP).
* 混合型OLAP(HOLAP,Hybrid OLAP).

1. **数据库备份** 分为以下几种:

* 完全备份:备份所有数据和日志.
* 差异备份:备份自上一次数据库完全备份以来产生的数据改变.

1. UNION操作符用于合并两个或多个SELECT语句的结果集.

## 第六张试卷

1. 对数据库系统的监控分为**手动监控机制**和**自动监控机制**两种.
2. 在数据库系统中,**外模式/模式**映像关系用于提供数据与应用程序间的逻辑独立性.

* 当数据库模式发生改变时,通过调整外模式/模式间的映像关系,使得应用程序不必随之修改,从而保证数据与应用程序间的逻辑独立性,简称数据的逻辑独立性.

1. 在SQL中,使用DISTINCT关键字消除重复的元组.
2. 数据库的一致性指的是事物执行成功则全部提交,必须使所有数据都具有一致的状态,如果一个事务提交失败,所有做过的更新全部撤销.
3. 如果数据库中的一个基本表的数量很少,且插入,删除,更新等操作频繁,该基本表最佳采用的文件结构是**堆文件.**

* 堆文件也称为无序文件,在堆文件中,记录随机地存储在文件物理空间中,新插入的记录存储在文件的末尾.

1. 在UML模型中,通信图是交互图的一种,也称作协作图,它是表达对象之间的联系以及对象间发送和接收信息的图.

## 第五张试卷

1. 数据库关系系统为用户和应用程序提供了DDL语言,用此语言可以定义数据库的模式结构,但此语言执行后的结果不能回滚.

* 数据库模式定义语言DDL(Data Description Language)是用来描述数据库中要存储的现实世界实体的语言.

1. 当用户在一个关系表的某一列上建立一个非聚集索引(该表没有聚集索引)时,数据库关系系统会自动为该索引维护一个索引结构,该索引结构中的记录是由**索引列列值**和它相对对应的**指针**构成的.
2. 在UML中,当要描述状态之间的转换时,可通过时间图来体现时间因子的作用.

* 时间图适用于当状态的转移与时间密切相关时,同时强调时间因子在状态转移过程中的重要作用.

1. SQL Server 2008支持三种用户自定义函数:

* 标量函数:可以出现在SELECT语句目标列中.
* 内嵌表值函数:必须放在SELECT语句的FROM子句中.
* 多语句表值函数:必须放在SELECT语句的FROM子句中.

1. INSTEAD OF 创建前触发器,FOR/AFTER创建后触发器.
2. SQL Server 2008游标声明的语句格式为:

* *DECLARE vendor\_cursor CURSOR FOR SELECT ..*

1. SQL Server 2008中,用户数据库数据文件的主要扩展名为:MDF
2. SQL Server 2008中设置不允许用户获得对表的某种操作权限的命令为:

* *DENY 对象权限名 ON {表名|视图} TO {数据库用户名|用户角色名}*

1. 优化SQL语句执行效率属于数据库应用程序运行管理与维护阶段的任务.
2. 为了避免活锁的现象发生,DBMS一般采用先来先服务(**FIFS**)策略处理事务的数据操作请求.
3. 分布式数据库的分片透明性是最高级别的透明性.它使得在编写程序时用户只需要对全局关系进行操作,这样简化了应用程序的维护.其位于全局概念与分片模式之间.
4. MOLAP称为基于多维库的OLAP,这种OLAP的核心是多维数据库技术,MOLAP工具以多维数据库的形式将元数据,基础事实数据和导出数据存储在以**多维数组为基本存储结构的多维数据库中.**

## 第四张试卷

1. 在UML中,用例模型由用例,系统和角色三部分组成.
2. 在SQL Server 2008中,用于判断游标数据提取状态的全局变量是: **@@FETCH\_STATUS**
3. 计算两个日期只差的函数是: **dateiff();**

* DATEDIFF(datepart,startdate,enddate)函数可以返回两个日期之间的天数.

1. 数据库系统中的故障的种类可以分为:

* 事物内部故障(预期事务内部故障和非预期事务内故障)
* 系统故障
* 介质故障
* 计算机病毒

1. 两阶段加锁协议可以保证事务调度的可串行性.
2. 差量备份只记录数据库上一次完全转储后的变化部分,这样可以提高效率,同时保证了备份数据的完整性.
3. 数据在网络中传输时,则是以整个关系(也可以是片段)传输,显然这是一种冗余的方法.在一个关系传输到另一场地后,并非每个数据都参加与连接操作,因此,不参与连接的数据或无用的数据不必在网络中来回传输.采用半连接操作即可在网络中只传输参与连接的数据.继而减少场地之间的数据传输量.
4. 常用的OLAP多维分析操作有切片,切块,旋转,钻取和卷起.
   * **卷起**:是用户能够从多个角度多侧面观察数据.卷起是在数据立方体中执行聚集操作.
   * **钻取**:是通过在维级别中下降或通过引入某个/某些维来细致的观察数据.
   * **切片/切块**:实现局部数据的显示,帮助用户从众多混杂的数据中进行选择,选装就是改变维的方向.
   * **旋转**:改变维的方向.
5. 元数据是关于数据对的数据,在数据仓库系统中,元数据可以帮助数据仓库管理员和数据仓库的开发人员非常方便地找到他们所关心的数据.元数据是描述数据仓库内数据的结构和建立的数据,可将其按用途的不同分为两类:

* **技术元数据**
* **业务元数据**

## 第二张试卷

1. 数据库系统的三级模式结构是由外模式,模式,内模式三级构成.

* **外模式/模式映像**:保证了数据库系统中的数据具有较高的逻辑独立性.
* **模式/内模式映像**:保证了数据库系统中的数据具有较高的物理独立性.

1. 系统设计阶段:

* 概念设计:整个数据库设计的关键,他通过对用户需求进行综合,归纳与抽象,形成一个独立于具体的DBMS的概念模型
* 逻辑设计:将概念结构转换为某个DBMS所支持的数据模型,并对其进行优化.
* 物理设计:为逻辑数据模型选取一个最合适应用环境的物理结构.

1. 设用UML设计某数据应用系统,设计人员规划了一组应用程序集,该集合由动态链接库和可执行程序构成.为了展现这些应用程序集间的组织和依赖关系,以对源码,可执行程序的发布等进行系统建模,应采用的**UML图是组件图**.
2. 聚类函数*COUNT(DISTINCT|ALL<列明>)*,表示对某列的值的计数.如果指定DISTINCT短语,则表示在计算时要取消指定列中的重复值.
3. 标识列的定义使用关键字:identity(唯一列,避免重复).
4. 在SQL Server 2008中,设U1是SQL Server身份验证模式的数据库服务器登录账户,现希望U1在该数据库服务器上具有系统管理员权限,请补全下列语句:

* *EXEC sp\_addsrvrolemember ‘U1’,’sysadmin’*

其中’sp\_addsrvrolemember’是定义好的存储过程.其作用是为了登录账户赋角色去权限.

1. 数据库管理系统中的加锁协议规定了事务的加锁时间,持锁事件,其中三级加锁协议就可以完全保证并发事物数据的一致性.

* 三级加锁协议也称为三级封锁协议,它保证了正确地调用事务的并发操作,是事务对数据库对象加锁,解锁必须遵循的一种规则.在运用X锁和S锁对数据库对象加锁时,还需要约定一些规则,例如何时申请X锁,持锁时间,何时释放.

1. 通常数据库的转储机制有三种:

* 完全转储:是指对整个数据库中的数据全部重新备份.效率低,时间长.
* 差量转储:指基于上一次转储基点之后变化转储,是对上次转储之后对所有文件中修改或删除的记录的转储.效率高,时间短.
* 增量转储:是对数据库中凡是有记录变化的文件的整个文件进行复制.效率介于前两种转储方式之间.

1. 数据仓库是为了构建新的分析处理环境而出现的一种数据储存和组织技术,基本特征包括:

* 面向主题的
* 集成的
* 非易失的
* 随时不断变化的

数据结合,用来支出管理人员的决策.

## 第一张试卷

1. 数据库管理系统提供了数据定义语言(**DDL**),用于定义各种数据库对象.数据定义语句经DDL编译器编译后,各种对象的描述信息存放在数据库的数据字典中.

* 数据字典是对系统中各类数据描述的集合,是进行详细的数据收集和分析所获得的主要成果,数据字典在数据库设计中占有很重要的地位,通常包括:
  1. 数据项
  2. 数据结构
  3. 数据流
  4. 数据储存
  5. 数据处理

1. UML中的顺序图主要用于描述系统内对象之间的消息发送和接收序列.

* 它有两个坐标轴:纵向表示时间的持续过程.横向表示对象每一个对象用矩形框表示,纵向的虚线表示对象在序列中的指向情况.
* 顺序图的第一个消息一般在左边第一个对昂的生命线顶端,其它消息按时间顺序相继插入顺序图中,后面发生的消息的线应比前发生的消息的线画的低一些,以表示它们之间的时间关系.

1. WITH TIES 一般是和TOP/ORDER BY相结合使用,**表示包括与最后一行ORDER BY后面的参数取值并列的结果.**
2. 数据表中对于某一列数据的唯一性约束有UNIQUE和PRIMARY约束,但两者之间有区别.

* RRIMARY KEY:强调主键不可出现相同元素,但是强制一列或多列组合(不是主键)的唯一性时应使用UNIQUE而不是RPIMARY KEY约束.
* UNIQUE:允许出现空值,PRIMARY KEY约束不允许出现空值.

1. 在SQL Server 2008中,设log1是SQL Server身份验证模式的数据库服务器登录账户,现要授权log1具有创建数据库的权限,请补全下列语句:

* *EXEC sp\_addsrvrolemember ‘1og1’,’dbcreator’*

1. 数据库管理系统一般通过周期性检查事务等待图来实现死锁检测.
2. 系统故障造成数据库不一致的原因有两个:
   1. 未完成的事务对数据库的更新可能已写入数据库.
   2. 已提交的事务对数据库的更新可能还留在缓冲区还没来得及写入数据库.

对于故障发生前已提交的事务(既有BEGIN Transaction又有commit)做**REDO**操作.对于故障发生时尚未完成的事务(只有BEGIN Transaction,而没有commit或rollback)做**undo**操作.

## 新增题库(卷一)

1. 在IDEF1X数据建模方法中,多对多联系又称为非确定联系.

* 在IDEF1X数据建模方法中,所有实体集间的联系都必须用确定联系来描述,不允许出现不确定联系,多对多联系为非确定联系,非确定联系需要分解为若干个一对多的联系.

1. 如果数据库中的一个基本表中的数据量很少,且插入,删除,更新等操作频繁,则该表可以采用堆文件组织方式,因为堆文件无需建立索引,维护代价非常低,虽然堆文件的数据访问效率较低,但是在数据量很少时,定位文件记录的时间非常短.
2. 在使用UML进行系统建模过程中,一个系统中可能存在许多用例,用例之间的关系如下:

* 扩展/extends
* 使用
* 组合

1. 在SQL Server 2008的LIKE子句中,有以下特殊字符:

* 下划线(\_):匹配任意一个字符.
* 百分号(%):匹配0到多个字符.
* 方括号([]):转义.
* 尖号(^):排除一些字符进行匹配.

1. DEFAULT约束用于向列中插入默认值,如果没有规定其他的值,那么会将默认值添加到所有的新纪录.
2. 在SQL Server 2008中,删除数据库中用户的语法为:

* DROP USER 用户名.

1. 水平分割会给应用增加复杂度,它通常在查询时需要多个表名.查询所有数据需要union操作
2. SQL Server 2008中有三种事务日志备份:

* **纯日志备份**:包含一定间隔的事务日志记录而不包含在大容量日志恢复模式下执行的任何大容量更改的备份.
* **大容量备份**:包含日志备份记录以及大容量操作更改的数据页的备份,不允许对大容量操作日志备份进行时点恢复.
* **结尾(尾部)备份**:对可能已损坏的数据库进行的日志备份,用于捕获尚未备份的日志记录.尾日志备份在出现故障时进行,用于防止丢失工作,可以包含纯日志记录或大容量操作日志记录.

1. 在各种并行数据库系统结构中,层次/Hierarchical结构综合了共享内存,共享磁盘,共享结构的特点,将结构分成顶层和底层两层,顶层是无共享结构,底层是共享内存或共享磁盘结构.
2. 在数据仓库中,元数据一般分为**技术型元数据**和**业务型元数据**
3. 在机器学习中,分类算法需要用到的样本数据一般分为:

* 训练集
* 测试集
* 验证集
* 其划分比例一般为 0.6:0.2:0.2,对原数数据进行三个数据集的划分,也是为了防止模型过拟合

## 新增题库(卷二)

1. 性能需求分析主要包括:

* 数据操作响应时间
* 系统吞吐量
* 允许并发访问的最大用户数和每TPS代价值(Price per TPS).

1. **DFD**建模方法的组成:

* 数据流
* 处理
* 数据储存
* 外部项

1. 在UML中,用于描述系统,用例和程序模块中逻辑流程的先后次序并行次序的图称为活 动图.

* 活动图主要用于描述系统,用例和程序模块中逻辑流程的先后次序,并行次序.活动图用于低层次程序模块的作用类似于流程图.但活动图还可以描述并行操作,而流程图只能描述串行操作,每个活动图都有一个起点,但不一定有终点.活动由圆角矩形框标识.

1. 在SELECT语句中:

* ORDER BY:对插叙结果进行排序.
* DISTINCT:消除重复行.

1. 在SQL Server 2008中,SQL查询语句的格式为:

* *SELECT COUNT(DISTINCT A) OVER (PARTITION BY B) FROM C.*

1. 如果希望以降序索引某个列的值可以在列名称之后添加保留字: DESC

* *CREATE INDEX index\_name ON table\_name(column\_name DESC)*

1. 删除SQL Server登录账户语法为:*DROP LOGIN login\_name*
2. 物化视图是一种保存了查询结果的数据库对象,其中的数据是预先计算并保存在对象中的.
3. 空间换取查询时间常用的几个方法:

* 增加冗余表(计算汇总表).
* 增加冗余列(包括计算字段)
* 增加索引(包括计算索引)
* 增加索引视图(物化视图)
* 数据缓存.

1. 非预期的事物内部故障指不能由事物程序处理的,如:运算溢出故障,并发事物死锁故障,违反了某些完整性限制而导致的故障等,事物故障的恢复由系统自动完成,对用户透明的.
2. 在数据仓库环境中,数据的粒度设计时一种重要的设计问题,**它会影响到数据仓库中数据量以及系统能回答的查询的类型.同时影响存放在数据仓库中的数据量的大小,在数据仓库中的数据量大小与查询的详细程序之间要做出权衡.**
3. 在企业数据仓库架构中,操作型数据存储ODS一般情况下主要用来支持OLAP和全局型OLTP两类应用.

在ODS上课实行的全局应用大致可分为:

* 实现企业全局的OLTP操作
* 实现即时的OLAP操作.