1. 事务的基本概念
2. 事务的定义

一个数据库操作序列

一个不可分割的工作单位

恢复和并发控制的基本单位

1. 事务和程序比较

在关系数据库中，一个事务可以是一条或多条SQL语句，也可以包含一个或多个程序

一个程序通常包含多个事务

1. 定义事务
2. 显式定义方式
   1. COMMIT

BEGIN TRANSACTION

SQL 语句1

SQL 语句2

……

COMMIT

事务正常结束，操作序列全部执行完成

提交事务的所有操作（读+更新）

事务中所有对数据库的更新永久生效

* 1. ROLLBACK

BEGIN TRANSACTION

SQL 语句1

SQL 语句2

……

ROLLBACK

事务异常终止，操作序列全都**不执行**

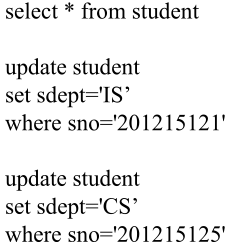
**回滚**事务的所有更新操作，事务回到开始时的状态

1. 隐式定义

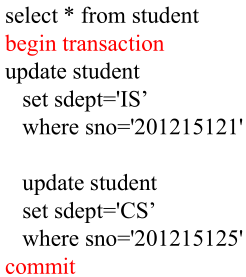
当用户没有显式地定义事务时，DBMS按缺省规定自动划分事务

在默认情况下，SQL Server2008采用自动提交方式，即**如果没有显示定义事务，则一个SQL语句为一个事务**

示例1：在查询分析器中执行下面的一组SQL语句，DBMS将理解为 3 个事务



示例2：DBMS将理解为 2 个事务



1. 事务的ACID特性
2. 原子性（Atomicity）

事务是数据库的逻辑工作单位

事务中包括的所有操作**要么都做，要么都不做**

1. 一致性（Consistency）

事务执行的结果必须是使数据库从一个一致性状态变到另一个一致性状态

只要事务的原子性不遭到破坏，就能保证一致性。

1. 隔离性（Isolation）

**对并发执行而言一个事务的执行不能被其他事务干扰**

一个事务内部的操作及使用的数据对其他并发事务是隔离的

并发执行的各个事务之间不能互相干扰

1. 持续性（Durability ）

持续性也称永久性

一个事务一旦提交，它对数据库中数据的改变就应该是永久性的

接下来的其他操作或故障不应该对其执行结果有任何影响

1. 数据库恢复概述
2. 故障是不可避免的

计算机硬件故障

系统软件和应用软件的错误

操作员的失误

恶意的破坏

1. 故障的影响

运行事务非正常中断

破坏数据库

1. 数据库管理系统对故障的对策

DBMS提供恢复子系统

保证故障发生后，能把数据库从**错误状态**恢复到**某一已知的正确状态**，这就是**数据库的恢复**

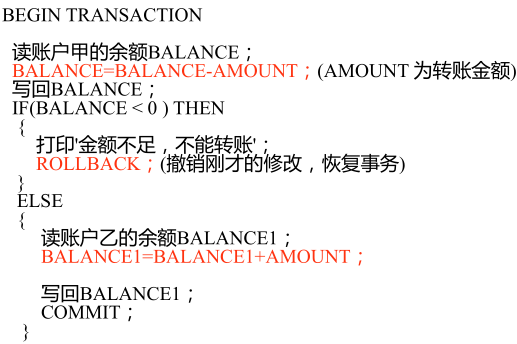
保证事务ACID

1. 恢复技术是衡量系统优劣的重要指标

数据库恢复机制是数据库管理系统的重要组成部分，占整个系统代码的百分之十以上

1. 故障种类
2. 事务内部的故障
3. 有的是可以通过事务程序本身发现的

示例：银行转账事务，这个事务把一笔金额从一个账户甲转给另一个账户乙



这个例子所包括的两个更新操作要么全部完成要么全部不做。否则就会使数据库处于不一致状态，例如只把账户甲的余额减少了而没有把账户乙的余额增加

在这段程序中若产生账户甲余额不足的情况，应用程序可以发现并让事务回滚，撤销已作的修改，恢复数据库到正确状态

1. 有的是非预期的

事务内部更多的故障是非预期的，是不能由应用程序处理的

* 1. 运算溢出
  2. 并发事务发生死锁而被选中撤销该事务
  3. 违反了某些完整性限制等

以后，事务故障仅指这类**非预期的故障**

事务故障的恢复：**撤消事务（UNDO）**

1. 系统故障
2. 概述

系统故障：称为软故障，是指造成系统停止运转的任何事件，使得系统要重新启动

* 1. 整个系统的正常运行突然被破坏
  2. 所有正在运行的事务都非正常终止
  3. 不破坏数据库
  4. **内存中数据库缓冲区**的信息全部丢失

1. 系统故障的常见原因

特定类型的硬件错误（如CPU故障）

操作系统故障

DBMS代码错误

系统断电

1. 系统故障的恢复
   1. 发生系统故障时，**事务未提交**

**恢复策略：强行撤消（UNDO）所有未完成事务**

* 1. 发生系统故障时，**事务已提交，但缓冲区中的信息尚未完全写回到磁盘上**

**恢复策略：重做（REDO）所有已提交的事务**

1. 介质故障
2. 概述

介质故障：称为硬故障，指**外存故障**

* 1. 磁盘损坏
  2. 磁头碰撞
  3. 操作系统的某种潜在错误
  4. 瞬时强磁场干扰

1. 介质故障的恢复

装入数据库发生介质故障前某个时刻的数据副本

重做自此时开始的所有成功事务，将这些事务已提交的结果重新记入数据库

1. 计算机病毒

一种人为的故障或破坏，是一些恶作剧者研制的一种计算机程序

可以繁殖和传播

破坏、盗窃系统中的数据

破坏系统文件

1. 故障小结

各类故障，对数据库的影响有两种可能性

一是**数据库本身被破坏**

二是**数据库没有被破坏**，但数据可能不正确，这是由于**事务的运行被非正常终止**造成的



1. 恢复实现技术
   * 1. 概述

恢复操作的基本原理：**冗余**，用**存储在系统其它地方的冗余数据**来**重建**数据库中已被破坏或不正确的那部分数据

恢复机制涉及的关键问题

如何建立冗余数据（数据转储（backup）、登录日志文件（logging））

如何利用这些冗余数据实施数据库恢复

* + 1. 数据转储

1. 转储

转储是指DBA将**整个数据库**复制到磁带或另一个磁盘上保存起来的过程

这些备用的数据文本称为后备副本或后援副本

如何使用：数据库遭到破坏后可以将后备副本重新装入，重装后备副本只能将数据库恢复到**转储时**的状态

按转存**状态**，可分为静态转储，动态转储

按转存**方式**，可分为海量转储，增量转储

1. 静态转储

在系统中**无运行事务时**进行转储

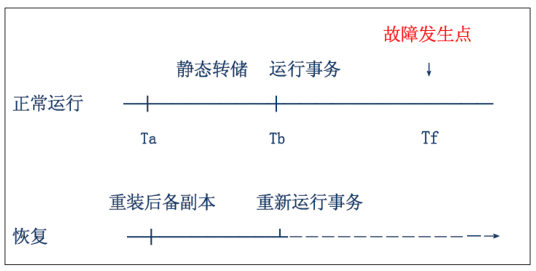
转储开始时**数据库处于一致性状态**

转储期间**不允许对数据库的任何存取、修改活动**

优点：实现简单

缺点：降低了数据库的可用性

**转储必须等用户事务结束，新的事务必须等转储结束**



1. 动态转储

**转储操作**与**用户事务**并发进行

转储期间**允许对数据库进行存取或修改**

优点：不用等待正在运行的用户事务结束，不会影响新事务的运行

缺点：不能保证副本中的数据正确有效

利用动态转储得到的副本进行故障恢复**需要把动态转储期间各事务对数据库的修改活动登记下来，建立日志文件，后备副本加上日志文件才能把数据库恢复到某一时刻的正确状态**

1. 海量转储与增量转储

海量转储: 每次转储**全部**数据库

增量转储: 只转储上次转储后**更新过**的数据

海量转储与增量转储比较，从恢复角度看，使用海量转储得到的后备副本进行恢复往往更方便，但如果数据库很大，事务处理又十分频繁，则增量转储方式更实用更有效

1. 小结



应**定期**进行数据转储，制作后备副本，但转储又是十分耗费时间和资源的，不能频繁进行

DBA应该根据数据库使用情况确定适当的转储周期和转储方法

例如：每天晚上进行动态增量转储；每周进行一次动态海量转储；每月进行一次静态海量转储

1. 登陆日志文件
2. 日志文件的格式和内容

日志文件(log)是用来**记录事务对数据库的更新操作**的文件

日志文件的格式：以**记录**为单位的日志文件、以**数据块**为单位的日志文件

* 1. 以**记录为单位**的日志文件内容

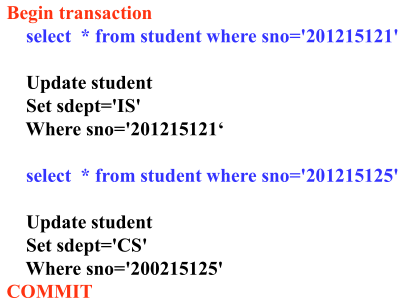
各个事务的开始标记(BEGIN TRANSACTION)

各个事务的结束标记(COMMIT或ROLLBACK)

各个事务的所有**更新操作**

以上均作为日志文件中的一个日志记录 (log record)

下面的事务会产生4条日志记录



以记录为单位的日志文件，每条日志记录的内容

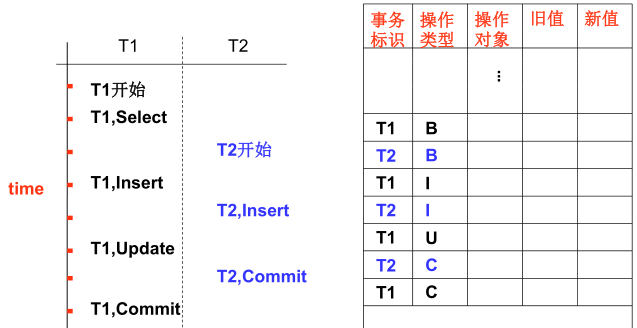
事务标识（标明是哪个事务）

操作类型（插入、删除或修改）

操作对象（记录内部标识）

更新前数据的旧值（对插入操作而言，此项为空值）

更新后数据的新值（对删除操作而言, 此项为空值）



* 1. 以数据块为单位的日志文件内容

以数据块为单位的日志文件，每条日志记录的内容

事务标识（标明是那个事务）

被更新的数据块

1. 日志文件的作用

进行事务故障恢复

进行系统故障恢复

协助后备副本进行介质故障恢复

1. 登记日志文件
   1. 基本原则

登记的次序严格按并行事务执行的时间次序

**必须先写日志文件，后写数据库**

写日志文件操作：把表示这个修改的日志记录写到日志文件

写数据库操作：把对数据的修改写到数据库中

* 1. 为什么要先写日志文件

写数据库和写日志文件是两个不同的操作

在这两个操作之间可能发生故障

如果先写了数据库修改，而在日志文件中没有登记下这个修改，则以后就无法恢复这个修改了

如果先写日志，但没有修改数据库，按日志文件恢复时只不过是多执行一次不必要的UNDO操作，并不会影响数据库的正确性

1. 恢复策略
2. 事务故障的恢复

事务故障：事务在运行至正常终止点前被终止

恢复方法：由恢复子系统应利用日志文件撤消（UNDO）此事务已对数据库进行的修改

事务故障的恢复由系统自动完成，对用户是透明的，不需要用户干预

事务故障的恢复步骤：

·反向扫描文件日志（即从最后向前扫描日志文件），查找该事务的更新操作

·对该事务的更新操作执行逆操作。即将日志记录中“更新前的值” （Befor Image, BI）写入数据库

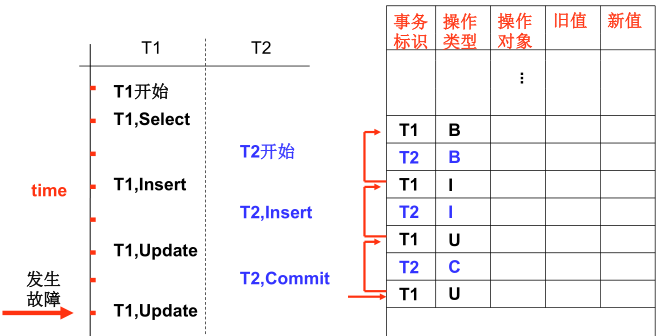
插入操作，“更新前的值”为空，则相当于做删除操作

删除操作，“更新后的值”为空，则相当于做插入操作

若是修改操作，则用BI 代替 AI（After Image）

·继续反向扫描日志文件，查找该事务的其他更新操作，并做同样处理

·如此处理下去，直至读到此事务的开始标记，事务故障恢复就完成了



1. 系统故障的恢复

系统故障造成数据库不一致状态的原因

未完成事务对数据库的更新已写入数据库

已提交事务对数据库的更新还留在缓冲区没来得及写入数据库

恢复方法：Undo故障发生时未完成的事务、Redo 已完成的事务

系统故障的恢复由系统在重新启动时自动完成，不需要用户干预

系统故障的恢复步骤：

·正向扫描日志文件（即从头扫描日志文件）

重做(REDO) 队列：在故障发生前已经提交的事务，这些事务既有BEGIN TRANSACTION记录，也有COMMIT记录

撤销 (Undo)队列：故障发生时尚未完成的事务，这些事务只有BEGIN TRANSACTION记录，无相应的COMMIT记录

·对撤销(Undo)队列事务进行撤销(UNDO)处理

反向扫描日志文件，对每个UNDO事务的更新操作执行逆操作，即将日志记录中“更新前的值”写入数据库

·对重做(Redo)队列事务进行重做(REDO)处理

正向扫描日志文件，对每个REDO事务重新执行登记的操作，即将日志记录中“更新后的值”写入数据库

1. 介质故障的恢复

重装数据库，使数据库恢复到一致性状态

重做已完成的事务

介质故障的恢复需要DBA介入，DBA的工作：重装最近转储的数据库副本和有关的各日志文件副本；执行系统提供的恢复命令

具体的恢复操作仍由DBMS完成

介质故障恢复步骤：

·装入最新的后备数据库副本(离故障发生时刻最近的转储副本) ，使数据库恢复到最近一次转储时的一致性状态

对于静态转储的数据库副本，装入后数据库即处于一致性状态

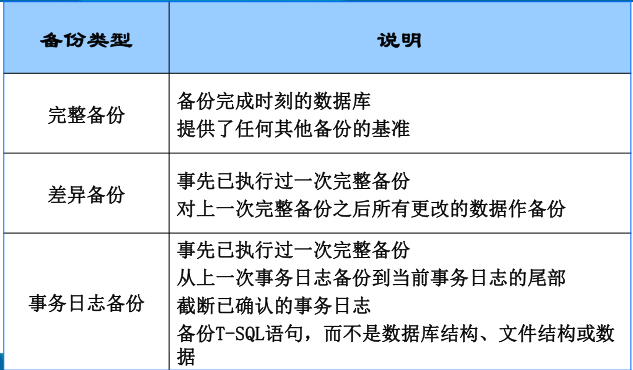
对于动态转储的数据库副本，还须同时装入转储时刻的日志文件副本，利用与恢复系统故障的方法（即REDO+UNDO），才能将数据库恢复到一致性状态

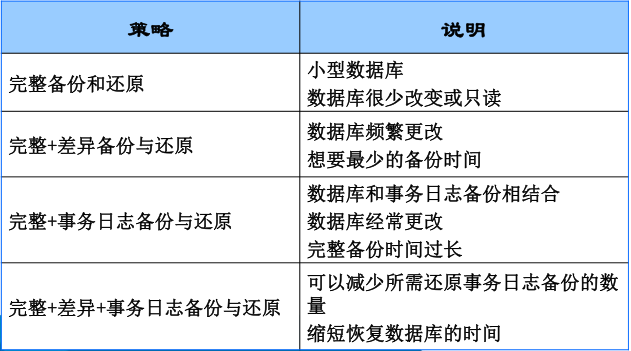
·装入有关的日志文件副本(转储结束时刻的日志文件副本) ，重做已完成的事务

首先扫描日志文件，找出故障发生时已提交的事务的标识，将其记入重做队列

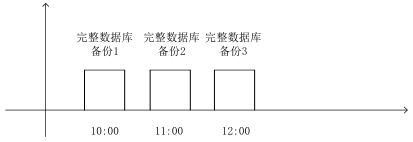
然后正向扫描日志文件，对重做队列中的所有事务进行重做处理。即将日志记录中“更新后的值”写入数据库

1. 备份和还原策略





**示例1：完整备份和还原**



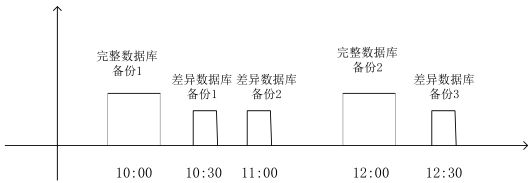
只能选择任意的一个完整数据库备份进行还原

还原到10:00

还原到11:00

还原到12:00

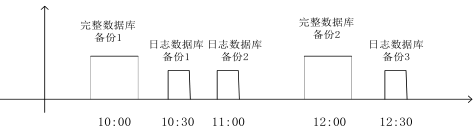
**示例2：完整备份+差异备份与还原**



如果需要还原到11:00时的数据库状态：完整数据库备份1+差异数据库备份2

如果需要还原到12:30时的数据库状态：完整数据库备份2+差异数据库备份3

示例3：完整备份+日志备份与还原



如果需要还原到11:00时的数据库状态：完整备份1+日志备份1+日志备份2

如果需要还原到12:30时的数据库状态：完整备份2+日志备份3 或 完整备份1+日志备份1+日志备份2+日志备份3

如果需要恢复到10:45时的状态：完整备份1+日志备份1+日志备份2，指定到10:45的恢复即时点

1. 带有检查点的恢复技术
2. 问题的提出
3. 两个问题

搜索整个日志将耗费大量的时间

REDO处理：重新执行，浪费了大量时间

1. 解决方案

在日志文件中增加检查点记录

增加重新开始文件

恢复子系统在登录日志文件期间动态地维护日志

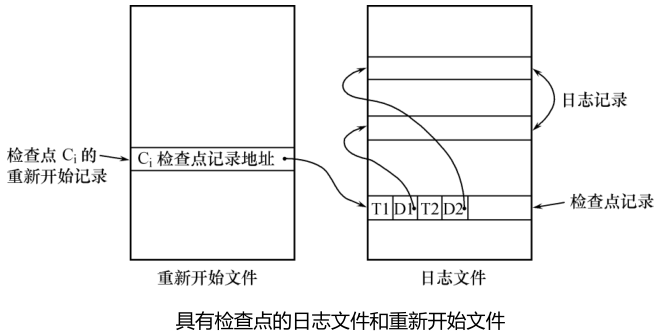
1. 检查点技术
2. 检查点记录的内容

建立检查点时刻所有正在执行的事务清单

这些事务最近一个日志记录的地址

1. 重新开始文件的内容

记录各个检查点记录在日志文件中的地址



1. 动态维护日志文件的方法

周期性地执行建立检查点，保存数据库状态的操作

具体步骤：

·将当前日志缓冲区中的所有日志记录写入磁盘的日志文件上

·在日志文件中写入一个检查点记录

·将当前数据缓冲区的所有数据记录写入磁盘的数据库中

·把检查点记录在日志文件中的地址写入一个重新开始文件

1. 建立检查点

恢复子系统可以定期或不定期地建立检查点，保存数据库状态

定期：按照预定的一个时间间隔，如每隔一小时建立一个检查点

不定期：按照某种规则，如日志文件已写满一半建立一个检查点

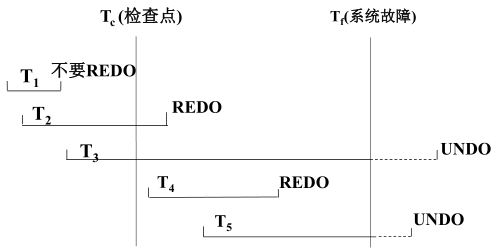
1. 利用检查点的恢复策略

使用检查点方法可以改善恢复效率

当事务T在一个检查点之前提交，T对数据库所做的修改已写入数据库

写入时间是在这个检查点建立之前或在这个检查点建立之时，在进行恢复处理时，没有必要对事务T执行REDO操作

系统出现故障时，恢复子系统将根据事务的不同状态采取不同的恢复策略



T1：在检查点之前提交

T2：在检查点之前开始执行，在检查点之后故障点之前提交

T3：在检查点之前开始执行，在故障点时还未完成

T4：在检查点之后开始执行，在故障点之前提交

T5：在检查点之后开始执行，在故障点时还未完成

恢复策略：

T3和T5在故障发生时还未完成，所以予以撤销

T2和T4在检查点之后才提交，它们对数据库所做的修改在故障发生时可能还在缓冲区中，尚未写入数据库，所以要REDO

T1在检查点之前已提交，所以不必执行REDO操作

1. 利用检查点的恢复步骤

·从重新开始文件中找到最后一个检查点记录在日志文件中的地址，由该地址在日志文件中找到最后一个检查点记录

·由该检查点记录得到检查点建立时刻所有正在执行的事务清单ACTIVE-LIST

建立两个事务队列UNDO-LIST，REDO-LIST，把ACTIVE-LIST暂时放入UNDO-LIST队列，REDO队列暂为空

·从检查点开始正向扫描日志文件，直到日志文件结束

如有新开始的事务T i ，把T i 暂时放入UNDO-LIST队列

如有提交的事务T j ，把T j 从UNDO-LIST队列移到REDO-LIST队列

·对UNDO-LIST中的每个事务执行UNDO操作，对REDO-LIST中的每个事务执行REDO操作

1. 数据库镜像
2. 概述

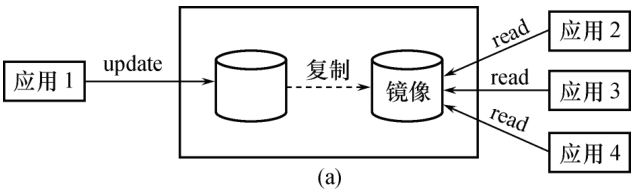
介质故障是对系统影响最为严重的一种故障，严重影响数据库的可用性

介质故障恢复比较费时，为预防介质故障，DBA必须周期性地转储数据库

提高数据库可用性的解决方案：数据库镜像（Mirror）

数据库镜像：DBMS自动把整个数据库或其中的关键数据复制到另一个磁盘上

DBMS自动保证镜像数据与主数据库的一致性，每当主数据库更新时，DBMS自动把更新后的数据复制过去（如图所示）

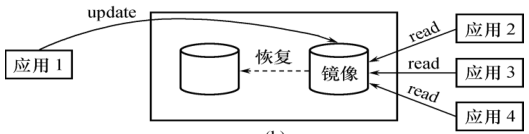


1. 数据库镜像的用途
2. 出现介质故障时

可由镜像磁盘继续提供使用

同时DBMS自动利用镜像磁盘数据进行数据库的恢复

不需要关闭系统和重装数据库副本(如图所示)



1. 没有出现故障时

可用于并发操作

一个用户对数据加排他锁修改数据，其他用户可以读镜像数据库上的数据，而不必等待该用户释放锁

1. 频繁地复制数据自然会降低系统运行效率，在实际应用中用户往往只选择对**关键数据**和**日志文件**镜像，而不是对整个数据库进行镜像