

大连理工大学 数据库恢复技术

单世民



大连理工大学

事务的基本概念

- 事务(Transaction)是用户定义的一个数据库操作序列，这些操作要么全做，要么全不做，是一个不可分割的工作单位
- 事务和程序是两个概念
 - ✧ 在关系数据库中，一个事务可以是一条SQL语句，一组SQL语句或整个程序
 - ✧ 一个应用程序通常包含多个事务
- 事务是恢复和并发控制的基本单位

如何定义事务

- 显式定义方式

```
BEGIN TRANSACTION
SQL 语句1
SQL 语句2
SQL 语句3
.....
COMMIT
```

```
BEGIN TRANSACTION
SQL 语句1
SQL 语句2
SQL 语句3
.....
ROLLBACK
```

- 隐式方式

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

事务结束语句

- COMMIT

- ✧ 事务正常结束
- ✧ 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

- ROLLBACK

- ✧ 事务异常终止
- ✧ 事务运行的过程中发生了故障，不能继续执行
- ✧ 回滚事务的所有更新操作
- ✧ 事务滚回到开始时的状态

事务的特性

- 事务的ACID特性:
 - ✧ 原子性 (Atomicity)
 - ✧ 一致性 (Consistency)
 - ✧ 隔离性 (Isolation)
 - ✧ 持续性 (Durability)

一致性与原子性

- 原子性 (Atomicity)

事务是数据库的逻辑工作单位

事务中包括的诸操作要么都做，要么都不做

- 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

- ✧ 一致性状态：数据库中只包含成功事务提交的结果
- ✧ 不一致状态：数据库中包含失败事务的结果

一致性与原子性

- 银行转帐：从帐号A中取出一万元，存入帐号B。
- 定义一个事务，该事务包括两个操作

A	B
$A=A-1$	$B=B+1$

- 这两个操作要么全做，要么全不做
- 全做或者全不做，数据库都处于一致性状态。
- 如果只做一个操作，数据库就处于不一致性状态。

隔离性

- 隔离性
- 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

隔离性

T_1	T_2
① 读A=16	
②	读A=16
③ $A \leftarrow A-1$ 写回A=15	
④	$A \leftarrow A-3$ 写回A=13

T1的修改被T2覆盖了！

持续性

- 持续性也称永久性（Permanence）
 - ✧ 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。

事务的特性

- 保证事务ACID特性是事务处理的任务
- 破坏事务ACID特性的因素
 - 多个事务并行运行时，不同事务的操作交叉执行
 - 事务在运行过程中被强行停止

数据库恢复

- 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
- 故障的影响
 - ✧ 运行事务非正常中断
 - ✧ 破坏数据库

数据库恢复

- 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，能把数据库中的数据从错误状态恢复到某种逻辑一致的状态
 - ✧ 保证事务ACID
- 恢复技术是衡量系统优劣的重要指标

故障-事务故障

- 事务故障

某个事务在运行过程中由于种种原因未运行至正常终止点就夭折了

- 事务故障的常见原因

- ✧ 输入数据有误
- ✧ 运算溢出
- ✧ 违反了某些完整性限制
- ✧ 某些应用程序出错
- ✧ 并行事务发生死锁
- ✧ 。 。 。 。

事务故障的恢复

- 发生事务故障时，夭折的事务可能已把对数据库的部分修改写回磁盘
- 事务故障的恢复：撤消事务（UNDO）
- 强行回滚（ROLLBACK）该事务
- 清除该事务对数据库的所有修改，使得这个事务象根本没有启动过一样

故障-系统故障

- 系统故障
 - ✧ 整个系统的正常运行突然被破坏
 - ✧ 所有正在运行的事务都非正常终止
 - ✧ 内存中数据库缓冲区的信息全部丢失
 - ✧ 外部存储设备上的数据未受影响
- 系统故障的常见原因
 - ✧ 操作系统或DBMS代码错误
 - ✧ 操作员操作失误
 - ✧ 特定类型的硬件错误（如CPU故障）
 - ✧ 突然停电

系统故障的恢复

- 清除尚未完成的事务对数据库的所有修改
 - ✧ 系统**重新启动时**，恢复程序要强行撤消（UNDO）所有未完成事务
- 将缓冲区中已完成事务提交的结果写入数据库
 - ✧ 系统**重新启动时**，恢复程序需要重做（REDO）所有已提交的事务

故障-介质故障

- 介质故障
 - ✧ 硬件故障使存储在外存中的数据部分丢失或全部丢失
 - ✧ 介质故障比前两类故障的可能性小得多，但破坏性大得多
- 常见原因
 - ✧ 硬件故障
 - ✧ 磁盘损坏
 - ✧ 磁头碰撞
 - ✧ 操作系统的某种潜在错误
 - ✧ 瞬时强磁场干扰

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

恢复的实现技术

- 恢复操作的基本原理：冗余
- 利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
- 恢复的实现技术：
复杂，一个大型数据库产品，恢复子系统的代码要占全部代码的10%以上

未雨绸缪

恢复的实现技术

- 恢复机制涉及的关键问题

- α 1. 如何建立冗余数据

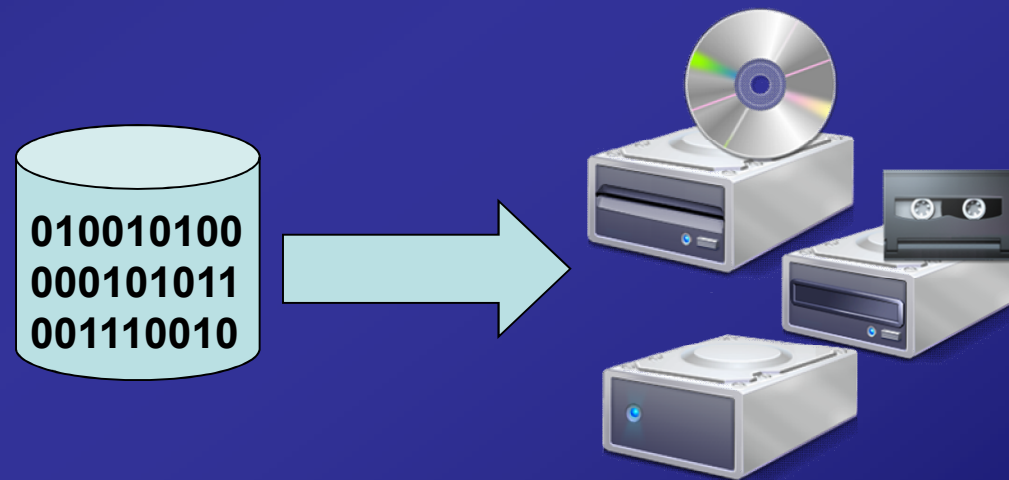
- 数据转储 (backup)
 - 登录日志文件 (logging)

- α 2. 如何利用这些冗余数据实施数据库恢复

恢复的实现技术

- 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。



转储方法-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

动态转储

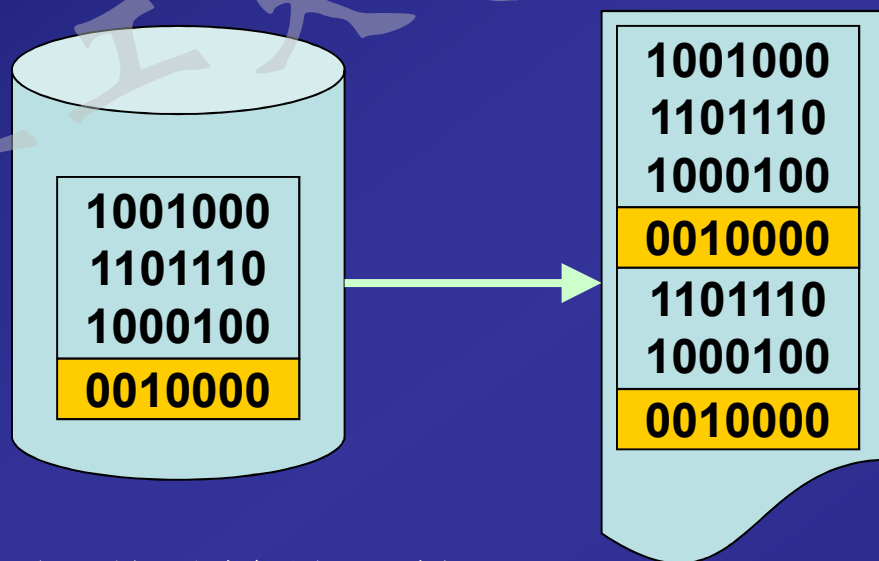
- 转储操作与用户事务并发进行
- 转储期间允许对数据库进行存取或修改
- 优点
 - ✎ 不用等待正在运行的用户事务结束
 - ✎ 不会影响新事务的运行
- 缺点
 - ✎ 不能保证副本中的数据正确有效

动态转储

- 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

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

恢复的实现技术

- 建立冗余数据
 - 数据转储

		转储状态	
		动态转储	静态转储
转储方式	海量转储	动态海量转储	静态海量转储
	增量转储	动态增量转储	静态增量转储

转储策略

- 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：
 - 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

日志文件

- 日志文件
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
- 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
 - ✧ 与事务有关的内部更新操作
 - ✧ 日志文件中的一个日志记录 (log record)

基于记录的日志文件

- 每条日志记录的内容
 - ✧ 事务标识
 - ✧ 操作类型（插入、删除或修改）
 - ✧ 操作对象（记录ID、Block NO.）
 - ✧ 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）

基于数据块的日志文件

- 每条日志记录的内容
 - ▣ 事务标识（标明是那个事务）
 - ▣ 操作对象（记录ID、Block NO.）
 - ▣ 更新前数据所在的整个数据块的值（对插入操作而言，此项为空值）
 - ▣ 更新后整个数据块的值（对删除操作而言，此项为空值）

日志文件的用途

- 用途

- ▣ 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

- 与静态转储后备副本配合进行介质故障恢复

- ▣ 静态转储的数据已是一致性的数据

- ▣ 如果静态转储完成后，仍能定期转储日志文件，则在出现介质故障重装数据副本后，可以利用这些日志文件副本对已完成的事务进行重做处理

- ▣ 这样不必重新运行那些已完成的事务程序就可把数据库恢复到故障前某一时刻的正确状态

日志文件的用途

- 介质故障恢复：日志文件 + 动态转储后备副本
 - ▣ 动态转储数据库：同时转储同一时点的日志文件
 - ▣ 后备副本与该日志文件结合起来才能将数据库恢复到一致性状态。
 - ▣ 利用这些日志文件副本进一步恢复事务，避免重新运行事务程序。

登记日志文件的原则

- 为保证数据库是可恢复的，登记日志文件时必须遵循两条原则
 - ▣ 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

登记日志文件的原则

- 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

恢复策略-事务故障

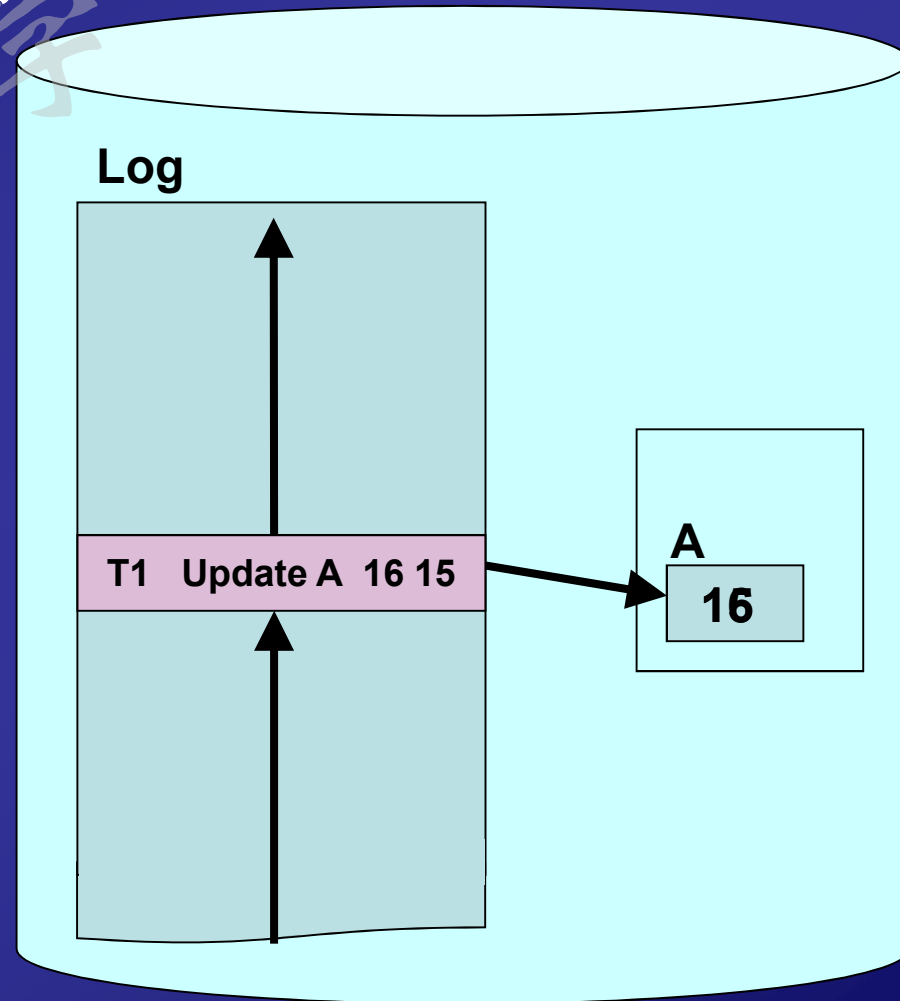
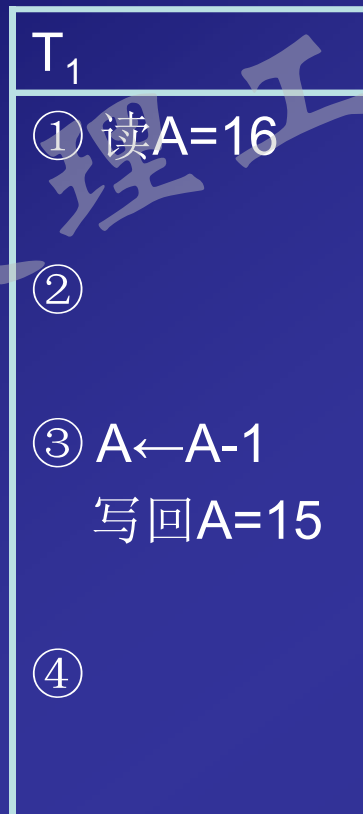
- 事务故障：事务在运行至正常终止点前被中止
- 恢复方法
由恢复子系统应利用日志文件撤消（UNDO）
此事务已对数据库进行的修改
- 事务故障的恢复由系统自动完成，不需要用户干预

事务故障的恢复步骤

- 1. **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。
- 2. **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”（Before 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）
- 3. **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**。
- 4. 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。

事务故障的恢复步骤

- 以Update操作为例



系统故障的恢复

- 系统故障造成数据库不一致状态的原因
 - ✎ 一些未完成事务对数据库的更新已写入数据库
 - ✎ 一些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
 - ✎ 1. Undo 故障发生时未完成的事务
 - ✎ 2. 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

T1 Begin
T2 Begin
.....
T3 Begin
.....
T4 Begin
T5 Begin
T6 Begin
.....
T3 COMMIT
.....
.....
T6 COMMIT
T1 COMMIT
.....

REDO-LIST

T1,T3,T6

UNDO-LIST

T2,T4,T5

系统故障的恢复步骤

- 1. **正向扫描日志文件**（即从头扫描日志文件）
 - ✎ Redo队列: 在故障发生前已经提交的事务
T1, T3, T6.....
 - ✎ Undo队列: 故障发生时尚未完成的事务
T2, T4, T5.....
- 2. **对Undo队列事务进行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
T2, T4, T5.....
- 3. **对Redo队列事务进行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
T1, T3, T6.....

介质故障的恢复

- 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
- 整个恢复过程由两部分组成：
 - ✧ 1. 重装数据库，使数据库恢复到一致性状态
 - ✧ 2. 重做已完成的事务

介质故障的恢复

- 恢复步骤

- ▣ 1. 装入最新的后备数据库副本，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相同的方法（即 REDO+UNDO），才能将数据库恢复到一致性状态。

介质故障的恢复

- 2. 装入有关的日志文件副本，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。

具有检查点的恢复技术

- 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

- 具有检查点（checkpoint）的恢复技术

- ✧ 在日志文件中增加检查点记录（checkpoint）
- ✧ 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

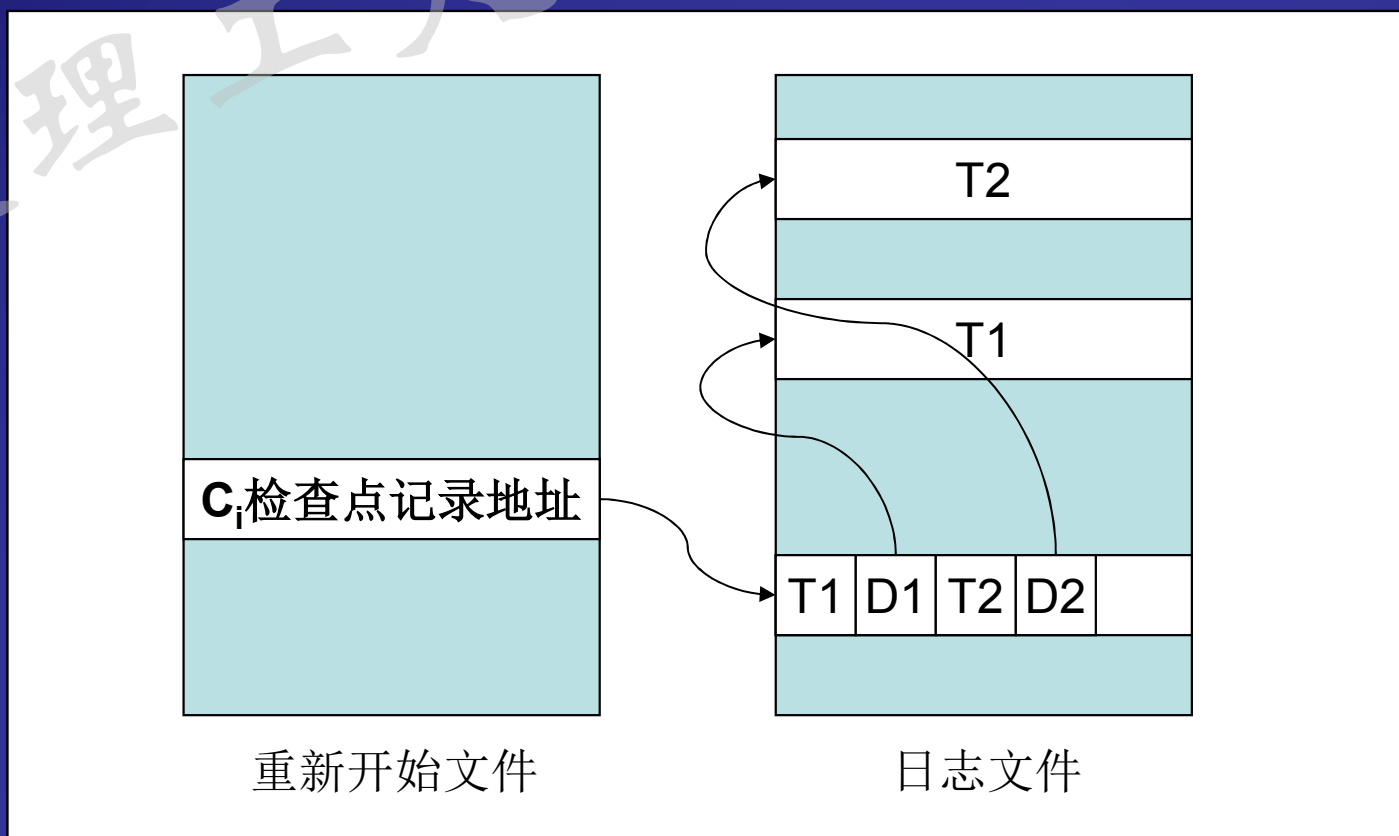
具有检查点的恢复技术

- 检查点的思想



具有检查点的恢复技术

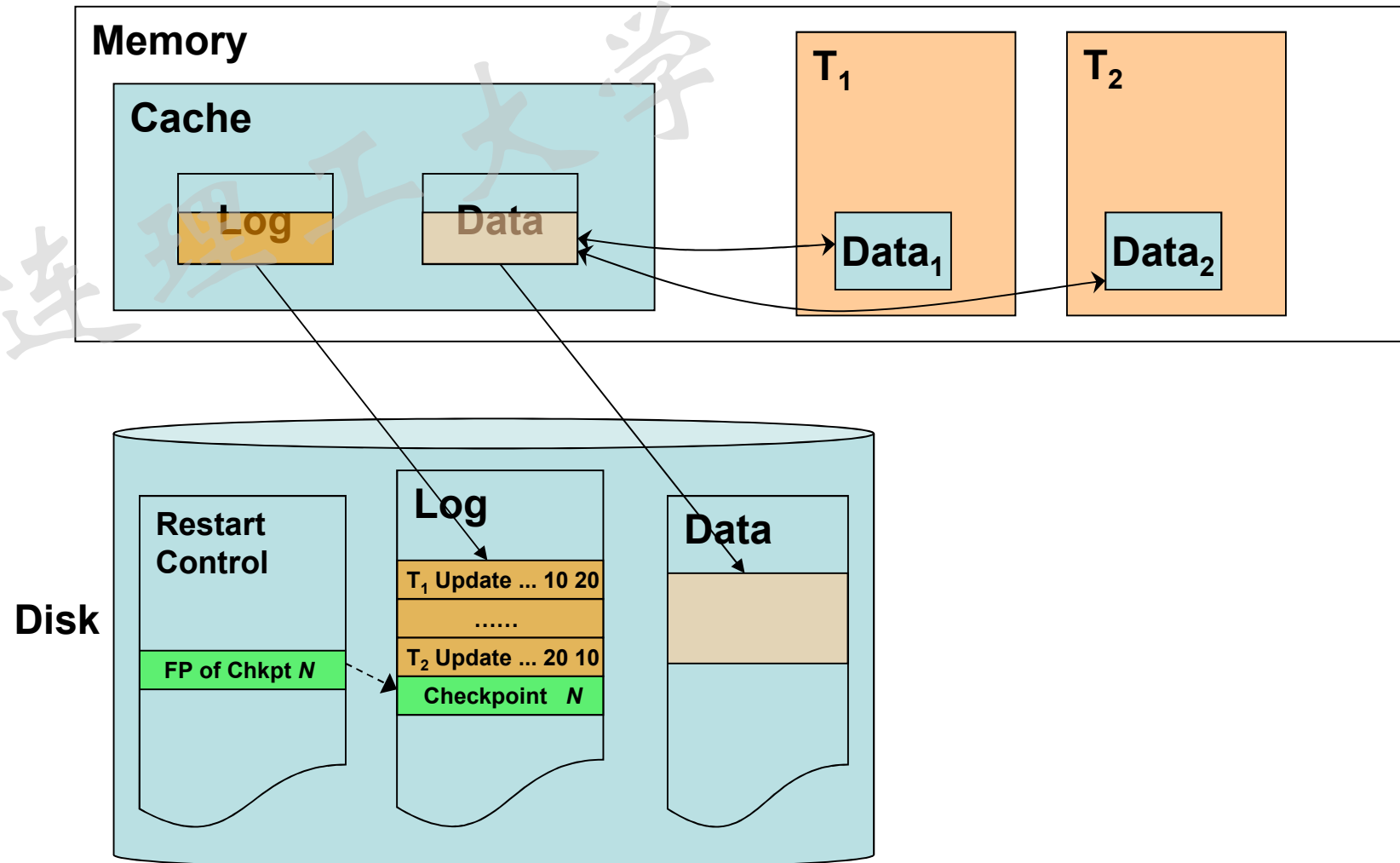
- 具有检查点的恢复技术
 - 具有检查点的日志文件和重新开始文件



检查点技术

- 检查点记录的内容
 - ✧ 1. 建立检查点时刻所有正在执行的事务清单
 - ✧ 2. 这些事务最近一个日志记录的地址
- 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

如何创建检查点



如何创建检查点

- 1. 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。
- 2. 在日志文件中写入一个检查点记录。
- 3. 将当前**数据**缓冲区的所有数据记录写入磁盘的数据库中。
- 4. 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。

建立检查点

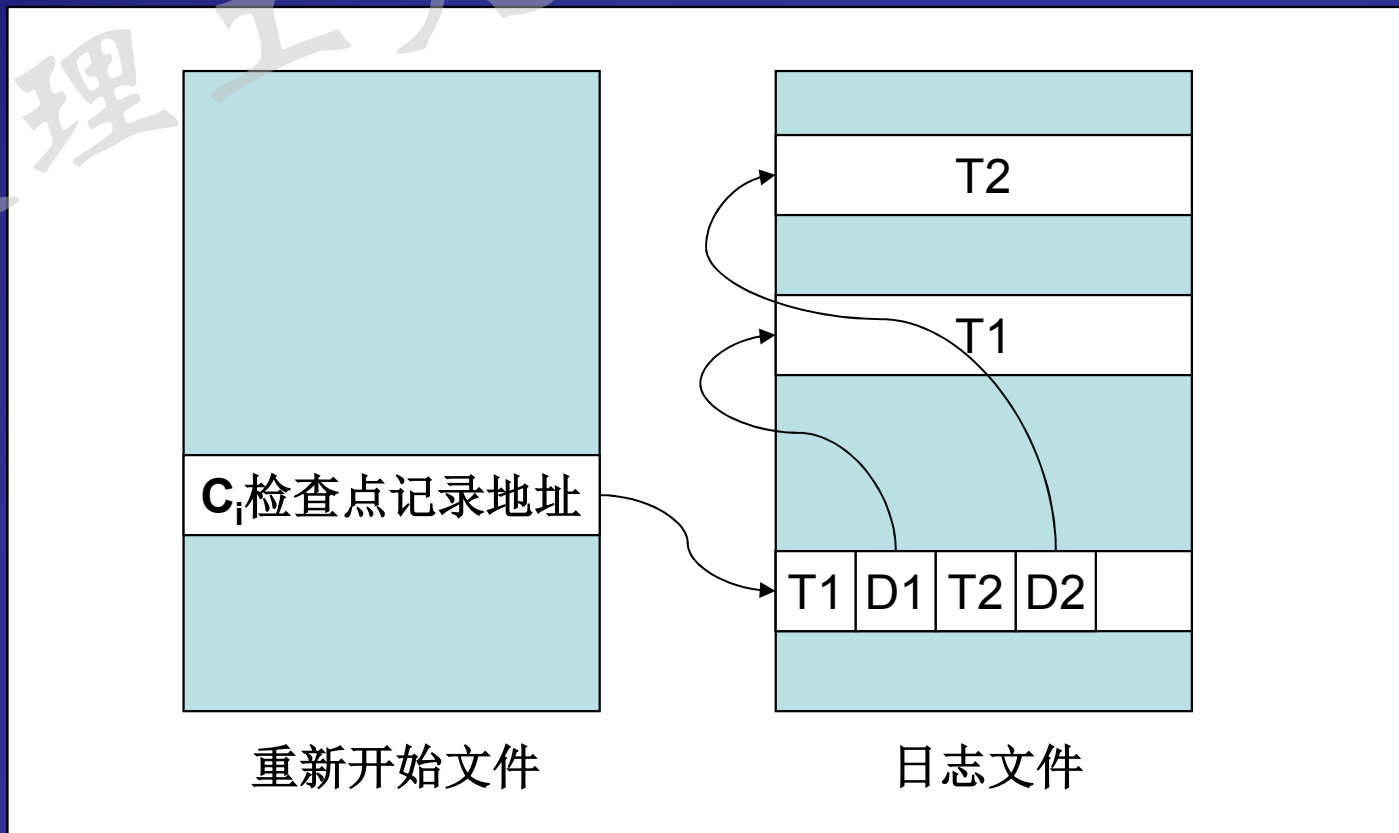
- 定期
 - 按照预定的一个时间间隔
- 不定期
 - 按照某种规则，如日志文件已写满一半建立一个检查点

利用检查点的恢复策略

- 当事务T在一个检查点之前提交，那么T对数据库所做的修改已写入数据库。这样，在进行恢复处理时，就没有必要对事务T执行REDO操作

利用检查点的恢复步骤

- 具有检查点的恢复技术
 - 具有检查点的日志文件和重新开始文件



利用检查点的恢复步骤

Checkpoint N (T1,T2)
T1 Insert
T2 Insert
T2 Delete
T1 Insert
T3 Begin
T3 Update
T1 COMMIT
T3 Update
T3 Update
T3 Insert
T4 Begin
T4 Insert
T4 COMMIT
T2 Delete

ACTIVE-LIST

T1	T2
----	----

UNDO-LIST

T1	T2	T3	T4
----	----	----	----

REDO-LIST

T1	T4
----	----

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

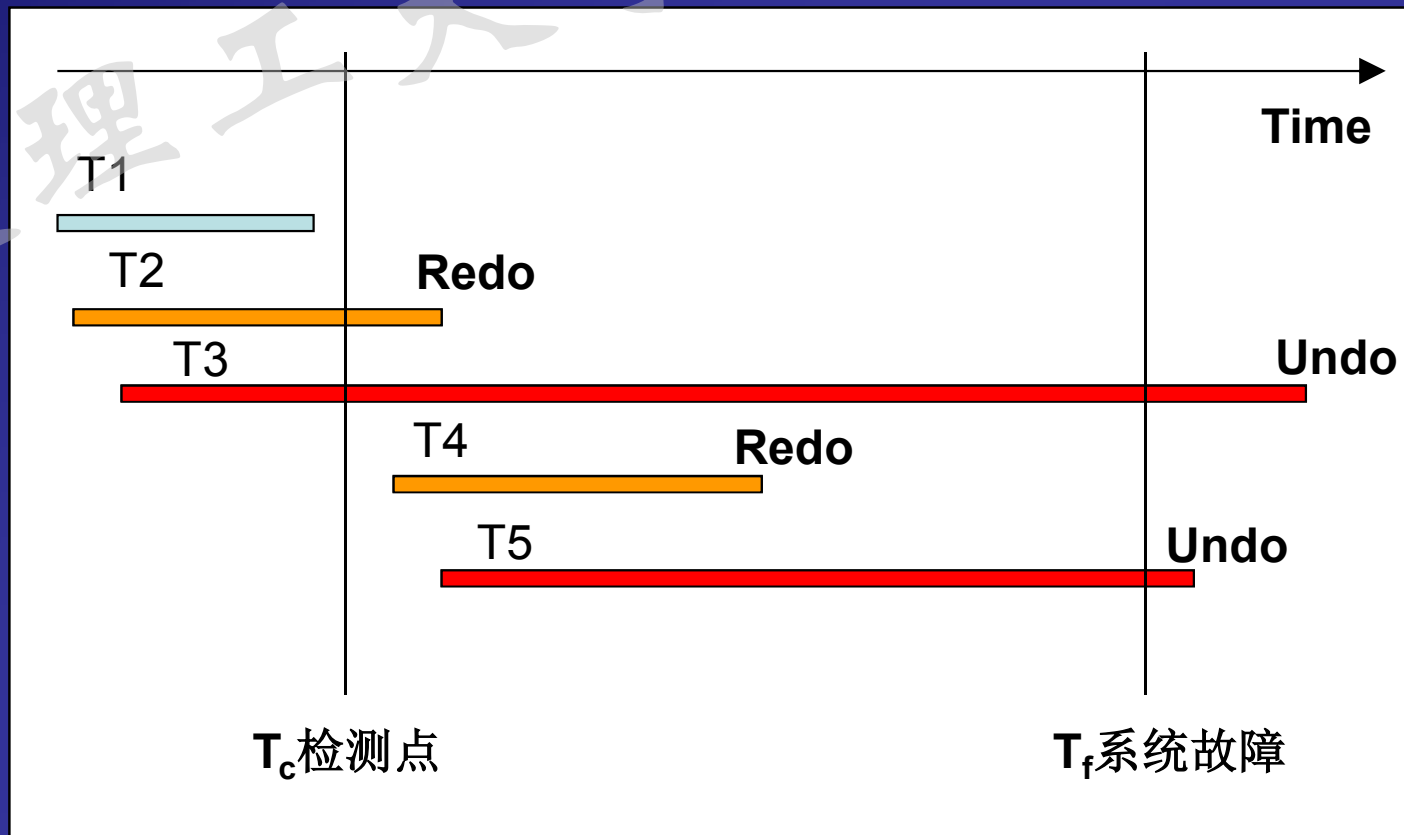


利用检查点的恢复步骤

- 1.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
- 2.由该地址在日志文件中找到最后一个检查点记录
- 3.由该检查点记录得到检查点建立时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。
- 4.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
- 5.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

利用检查点的恢复技术

- 例:



数据库镜像

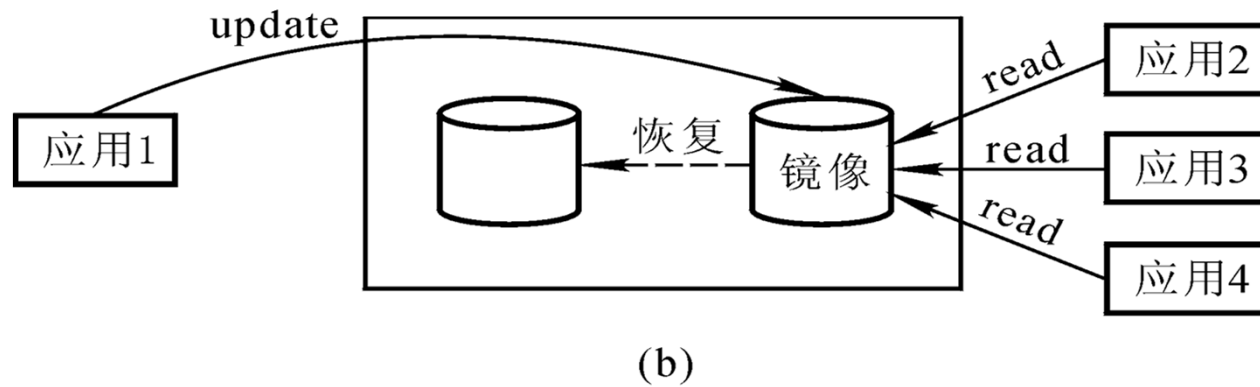
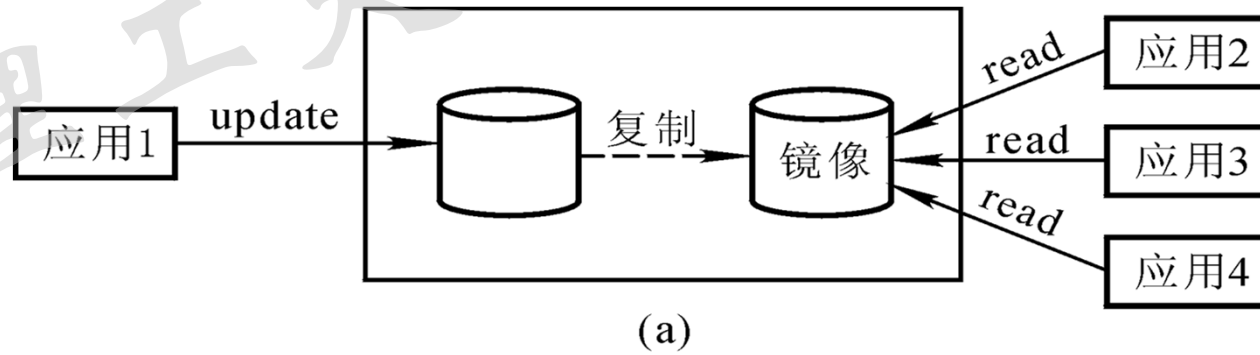
- 介质故障是对系统影响最为严重的一种故障，严重影响数据库的可用性
 - ✧ 介质故障恢复比较费时
 - ✧ 为预防介质故障，DBA必须周期性地转储数据库
- 提高数据库可用性的解决方案
 - ✧ 数据库镜像（Mirror）

数据库镜像

- 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的一致性

数据库镜像



数据库镜像的用途

- 出现介质故障时
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
- 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

小结

- 如果数据库只包含成功事务提交的结果，就说数据库处于一致性状态。保证数据一致性是对数据库的最基本的要求。
- 事务是数据库的逻辑工作单位
 - ✧ DBMS保证系统中一切事务的原子性、一致性、隔离性和持续性

小结

- DBMS必须对事务故障、系统故障和介质故障进行恢复
- 恢复中最经常使用的技术：数据库转储和登记日志文件
- 恢复的基本原理：利用存储在后备副本、日志文件和数据库镜像中的冗余数据来重建数据库

小结

- 常用恢复技术
 - ▣ 事务故障的恢复
 - UNDO
 - ▣ 系统故障的恢复
 - UNDO + REDO
 - ▣ 介质故障的恢复
 - 重装备份并恢复到一致性状态 + REDO

小结

- 提高恢复效率的技术

- ▣ 检查点技术

- 可以提高系统故障的恢复效率
 - 可以在一定程度上提高利用动态转储备份进行介质故障恢复的效率

- ▣ 镜像技术

- 镜像技术可以改善介质故障的恢复效率

大连理工大学

The End

ism



大连理工大学