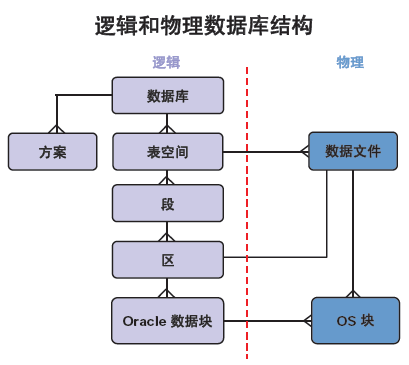
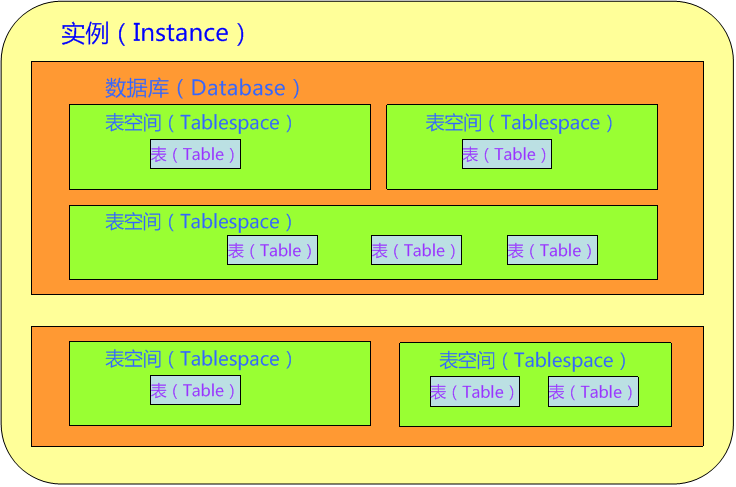
Atitit.数据库的物理存储结构

[](http://photo.blog.sina.com.cn/showpic.html#blogid=9f607ab70100yxyu&url=http://s8.sinaimg.cn/orignal/9f607ab7tbd9daa01ef97)

Oracle和DB2数据库的存储模型如图：



可以看出，数据库中的数据都存储在表空间中。表空间即是管理将逻辑数据库设计映射到操作系统物理存储中的一个数据库对象，用于指明数据的物理位置。关于表空间，以后再讨论。

Oracle数据库磁盘存储的逻辑结构为：一个数据库（Database）对应多个表空间（Tablespace），一个表空间对应多个段（Segment），一个段对应多个区（Extent），一个区对应多个数据块（Data Block），真正的数据就保存在数据块中

对应多个数据块（Data Block），真正的数据就保存在数据块中。这里有以下几点需要说明：

1.Oracle中一个数据块的大小默认是2KB（支持2KB,4KB,8KB，16KB,32KB），而DB2中则默认是4KB（支持4KB,8KB,16KB,32KB）；

2.Oracle中有段（Segment）的概念，而DB2中没有这一概念，表空间直接是各个容器（数据文件）中的区（Extent）组成，不过也还是有一个很弱化的Extent组概念。下面提到的关于Segment的内容则全部是针对Oracle的；

3.Oracle中的数据块称为Oracle Block，而DB2中则直接称为Data Page（数据页）。

4.Oracle中的Extent称为区，而DB2中则称为扩展数据块。为方便阅读，本文中统称为区。

二、 段（segment）（这块儿没有详细记录）

        段用于存储表空间中某一种特定的、具有独立存储结构的数据库对象的数据，它有一个或多个连续的区组成。当创建表、索引、簇等对象时，oracle就会为这些对象分配存储空间（即段），以便存储它们的数据。按照段中所存储的数据的特征、用途不同，可以分为以下几种：

|  |  |  |  |
| --- | --- | --- | --- |
| **段类型** | **名称** | **段类型** | **名称** |
| **Index partition** | 索引分区段 | **Deferred rollback** | 延迟回退段 |
| **Table partition** | 表分区段 | **Undo** | 撤销段（还原段） |
| **Table** | 表段（数据段） | **Temporary** | 临时段 |
| **Cluster** | 簇段 | **Cache** | 高速缓存段 |
| **Index** | 索引段 | **Lob** | 二进制大对象段 |
| **Rollback** | 回退段 | **Lobindex** | 二进制大对象索引段 |

# 2. 表数据在块中的存储以及RowId信息

表的字段除了设计者设计的字段外，还有个伪列Rowid；RowId是oracle每个表记录的位置信息。当表中插入一行记录时，此记录的块就有一个唯一的物理位置，这个位置信息在查询记录时显示为RowId伪列；

### 3. 索引

普通表中记录的存储顺序是没有规律的，在一个数据量很大的表中，如果不引入其他的手段，每次查找小部分记录都是从第一条扫描到最后一条，这样，系统将慢得不能使用；使用索引可以有效解决问题；

索引的本质用途是通过它使读取进程在扫描源表时的数据块范围大大减小了，因此性能大幅度提高；

## 三、索引

数据库以B树格式储存，只解决了按照”主键”查找数据的问题。如果想查找其他字段，就需要建立索引（index）。

所谓索引，就是以某个字段为关键字的B树文件。假定有一张”雇员表”，包含了员工号（主键）和姓名两个字段。可以对姓名建立索引文件，该文件以B树格式对姓名进行储存，每个姓名后面是其在数据库中的位置（即第几条记录）。查找姓名的时候，先从索引中找到对应第几条记录，然后再从表格中读取。

这种索引查找方法，叫做“索引顺序存取方法”（Indexed Sequential Access Method），缩写为ISAM。它已经有多种实现（比如C-ISAM库和D-ISAM库），只要使用这些代码库，就能自己写一个最简单的数据库。

问题，数据到底是按照先来后到存储，还是按照主键的btree。。按道理是无规律，否则没增加一个字段都要重新排序文件？？Yeyou可能prikey自增，会自动排在后面》》？？

Oracle中具有如下几类数据结构：

数据文件：存放数据库中的数据。表和索引是数据库用来存储数据的逻辑结构，数据文件即用来存放表和索引数据。

日志文件：存放数据库操作日志的文件。整个数据库的运行过程被日志化，日志集中存放在日志文件中，用于数据库系统的数据恢复操作。

控制文件：存放数据库物理结构和运行状态信息的文件，存储了数据库名、数据文件和日志文件名、数据库日志归档状态和系统备份状态等信息。

Oracle数据库由下列逻辑结构组成：

系统表空间，用来集中存储Oracle的系统数据。

临时表空间，用于数据库事务处理过程中的临时数据处理。

用户表空间，用于存储用户的数据库对象和最终数据。

还原表空间（回滚表空间），用于存储事务操作过程的数据前映像，以应对用户可能的回退操作。

三、 区（extent）

        区是由物理上连续存放的块所组成的。区 是oracle存储分配的最小单位。一个或多个块组成区，一个或多个区组成段。一个区只能属于一个数据文件。

        当在数据库中创建带有实际存储结构的方案对象时（如表、索引），oracle将为方案对象分配若干个区（视该方案的初始大小而定），以便组成一个对应的段，来为该方案对象提供初始的存储空间。当段中已分配的区都写满，oracle就要在该方案对象所在的表空间中为该段分配下一个新的空白区，以便容纳更多的数据。

四、 块（block）

        块是最小的数据管理单位，即oracle的所有存储的i/o操作都是以块为单位的。相对应的，操作系统执行i/o操作的最小单位是操作系统块。块得大小是操作系统块得大小的整数倍。

​    ​    块得大小是一个表空间的属性。SYSTEM和SYSAUX表空间具有相同的、标准的块大小，这个大小是在创建数据库时（会同时创建这两个表空间）由DB\_BLOCK\_SIZE初始化参数指定的，但在创建数据库之后这个参数值不能改变。

一、 控制文件

        控制文件是一个很小的（通常是数据库中最小的）文件，大小一般在1~5M左右，为二进制文件。但它是数据库中的关键性文件，它对数据库的成功启动和正常运行都是至关重要的，因为它存储了在其他地方无法获得的关键信息，这些信息包括：

        1、数据库的名称

        2、数据文件和重做日志文件的名称、位置、联机\脱机状态和大小

        3、发生磁盘故障或用户错误时，用于恢复数据库的信息（日志序列号，检查点）

        在数据库的运行过程中，每当出现数据库检查点或修改数据库的结构之后，oracle（只能有oracle本身）就会修改控制文件的内容。DBA可以通过ＯＥＭ工具修改控制文件中的部分内容，但DBA和用户都不应该认为的修改控制文件中的内容，否则会破坏控制文件。

三、 重做日志文件

        当用户对数据库进行修改的时候，实际上是先修改内存中的数据，过一段时间后，再集中将被村中的修改结果成批的写入到上面的数据文件中。oracle采取这样的做法，主要是出于性能上的考虑，因为针对数据操作而言，内存的速度比硬盘的速度要快成千上万倍。

        oracle利用“（联机）重做日志文件”随时保存修改结果，即oracle随时将内存中的修改结果保存到“重做日志文件”中。“随时”表示在将修改结果写入到数据文件之前，可能已经分几次写入到“重做日志文件”。因此，即使发生故障导致数据库崩溃，oracle也可以利用重做日志文件中的信息来恢复丢失的数据。只要某项操作的重做信息没有丢失，就可以利用这些重做信息来重现该操作。

        oracle是以循环方式来使用重做日志文件的，所以每个数据库至少需要两个重做日志文件。当第一个重做日志文件被写满后，后台进程LGWR（日志写进程）开始写入第二个重做日志文件；当第二个重做日志文件被写满后，又始写入第二个重做日志文件，以此类推。

自己实现一个最简单的数据库\_电脑教程学习网.htm