

ORACLE 的存储体系结构及其对象空间分配的研究^{*}

杨华千¹ 刘勇国¹ 杨德刚^{1,2} 廖晓峰¹

(重庆大学计算机科学与工程学院 重庆400044)¹ (重庆师范学院数学与计算机科学系 重庆400047)²

Research on ORACLE Storage Architecture and its Object Space Allocation

YANG Hua-Qian¹ LIU Yong-Guo¹ YANG De-Gang^{1,2} LIAO Xiao-Feng¹

(Department of Computer Science and Engineering, Chongqing University, Chongqing 400044)¹

(Department of Math & Computer, Chongqing Normal University, Chongqing 400047)²

Abstract With fast development in information times, the database, as the kernel unit of information storage, has been playing an increasingly role in the field of the modern information technology. In this paper, ORACLE storage architecture is analysed from physical and logical aspects. Meanwhile, its object space allocation is also discussed and explained by some examples.

Keywords ORACLE database, Storage architecture, Object space allocation

1 引言

信息时代与网络计算向数据库系统提出了新的挑战,要求系统支持更多的用户和更庞大的数据,并且具备更高的运算速度以及更低的计算成本。ORACLE 数据库作为业界领先的数据库系统,代表了此领域发展的最新进展。它于1997年推出了 ORACLE 8.0 版本,首次满足了 OLTP 与数据仓库应用的需求,该系统能够支持数万个用户同时连接,存储容量达到100TB,并且数据仓库查询速度较以往提高了数倍。由于采用了先进的存储体系结构,因此它具备强大的计算能力,然而对此结构的深入理解通常比较困难。同时,在进行数据库系统设计时,为了最大限度地提高系统性能,充分利用存储空间,要求必须对数据库对象空间做出合适的分配,这是数据库系统设计的重要步骤。因此,本文将主要针对 ORACLE 8i 数据库的存储体系结构及其对象空间分配进行深入探讨。

2 数据库存储体系结构

为了高效地组织、管理数据信息,ORACLE 数据库采用相互关联的储存体系结构表示方法,即逻辑数据存储结构和物理数据存储结构。逻辑存储结构表示数据在逻辑概念上的组织方式,例如数据库、表;物理存储结构则表示现实的数据存储单元,如文件或数据块。

2.1 逻辑存储结构

如图1所示的数据库逻辑存储体系结构,其中表空间为数据库内部数据的逻辑组织结构,每个表空间可拥有一个或多个物理数据文件。通常所建立的数据库对象(如:表、索引、视图、触发器等)逻辑上均存放在表空间中,而物理上存放于表空间所在的数据文件中。对于任何 ORACLE 数据库系统,至少需要一个表空间,即系统表空间,用于存储数据库的数据字典、PL/SQL 程序源代码和编译后的代码以及某些数据库对象(如视图、对象类型声明、同义词、序列等)的定义。但在实际应用中,一般都有多个表空间从逻辑和物理两个方面来组织

库中数据。比如可分别为数据字典信息、表数据、索引数据、系统事务回滚数据、内部系统处理过程中使用的临时数据建立独立表空间。通过为不同的应用数据集分别建立表空间,形成独立的应用数据管理模式。例如,应经常地备份活动频繁的应用数据,而对使用较少的应用数据应降低备份频率。

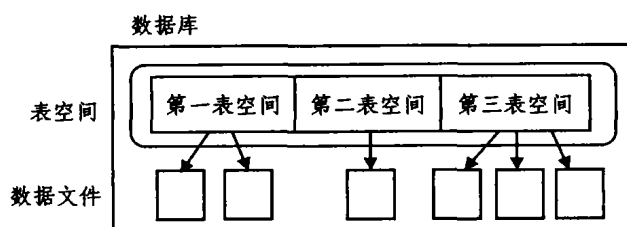


图1 数据库逻辑存储结构图

2.2 物理存储结构

数据文件是数据库系统中数据的物理存储场所。表空间可将数据映射到单一数据文件,或将其映射到多个数据文件上实现数据存储。通过表空间与数据文件间映射方式的变化,可充分利用操作系统和 ORACLE 数据库系统本身的并发 I/O 操作来减少数据访问的响应时间,提高系统性能。此外,ORACLE 数据库要能正常地运行还需要一个控制文件。控制文件包含了数据库物理结构的相关信息(如数据库名称、数据库相关文件的名称和位置信息),它还追踪系统内部信息,记录系统的当前物理状态,包含关于表空间、数据文件和系统备份的信息。为了确保数据库在孤立磁盘发生错误时仍正常工作,ORACLE 允许将控制文件镜像复制到多个地方,此时数据库系统应同时更新控制文件的每个副本。

此外,具体的数据库对象(例如表、索引等)的存储过程如下所述。与数据库系统预分配数据文件作为数据库中表空间的物理存储区相似,ORACLE 预分配数据块的段作为数据库对象的物理存储区。ORACLE 为数据库对象分配一组连续的称为区间(即成组的连续数据块)的数据块,其中段是数据库

^{*}国家自然科学基金(项目编号:60271019)和重庆市应用基础研究资助项目(项目编号:7370)。杨华千 硕士生,主要研究方向为数据库,网络安全。刘勇国 博士生。杨德刚 硕士生。廖晓峰 教授,博导。

对象使用的全部区间的集合。

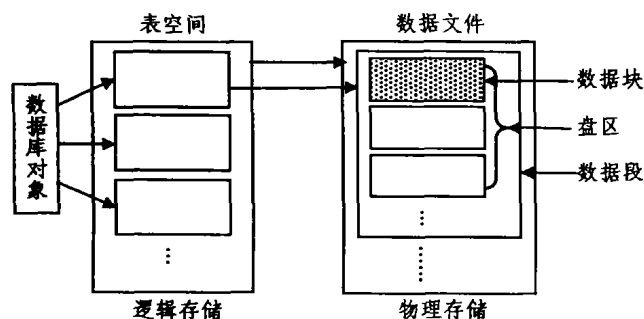


图2 数据库物理存储结构、逻辑存储结构对照示意图

当创建新数据存储对象时,用户可以在指定的表空间创建相应的段(常见段有数据段、索引段、回滚段、临时段等),然后系统从目标表空间使用的数据文件中分配数据块。特别要提到,在几乎所有的大型关系数据库(RDBMS)系统中都有事务处理这个概念。ORACLE 使用称为回滚段的特殊类型段记录事务的回滚数据,如果事务发生回滚,系统从回滚段读取必要的数据库完成数据重建,返回到数据改变前的状态。ORACLE 中回滚段记录信息的方式与其它类型段的操作不同,其操作方式如图3所示。

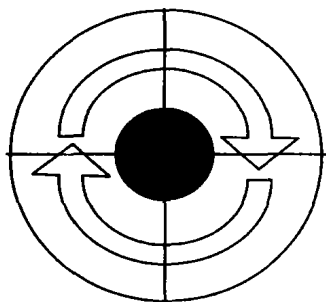


图3 回滚段信息记录示意图

回滚段在系统中表示为一个区间环,事务将回滚信息写向区间环的当前区间并最终将其填满,接着绕到该环的下一区间继续记录回滚信息。

数据块是数据库系统的磁盘存取单元,当进行数据库操作时,系统使用数据块存储和提取磁盘上的数据。当创建数据库时,可以指定数据库中数据块的大小,但数据块的大小必须等于服务器操作系统块的大小或为其整数倍。

3 数据库对象空间的分配

3.1 对象空间分配原则

数据库对象空间的合理分配至关重要,它不仅可以节省存储空间,而且还能够显著地改善数据库的 I/O 性能(因为 ORACLE 在全表扫描读取数据时,不能跨盘区进行操作,因此盘区的大小对其性能有明显的影响),因而在第一次建立数据库对象之前必须估算所需空间的大小,然后在实际使用过程中用统计的方法不断地对空间大小进行调整。在计算数据库对象的空间尺寸时,应了解 ORACLE 系统空间分配规则:

- 1) 只分配整个块,而不分配块的部分。
- 2) 可分配块组,通常是5块的倍数。
- 3) 根据表空间的可用自由空间大小,可对块组分配进行调整。

据上述空间分配规则,设数据库块大小为4kB(此数据可在 initorcl.ora 文件中查看 DB_BLOCK_SIZE 项),创建两个表(具体创建语法可查阅 SQL 相关资料):

表名	初始盘区	下一盘区(Next)	增长百分比(pctincrease)
Exam1	3K	3K	0
Exam2	103K	103K	20

对于 Exam1 表,数据库管理员(DataBase Administrator, DBA)指定初始区间大小为3k,但系统数据库块的大小为4k,根据上述规则 ORACLE 把初始区间大小舍入为4K(1块),然后又把初始区间大小舍入为20k(5块),分配下一盘区时系统使用指定的 Next 值,执行类似计算。对于 Exam2 表,ORACLE 分配第一、二两个区间的算法与 Exam1 表的分配方法相同,大小为120k,在继续为 Exam2 表分配区间时,由于 Exam2 中增长百分比为20,因此从第三个区间开始后一盘区应比前一盘区大20%,在为 Exam2 分配第三个区间时系统回到 Next 的指定值而不是实际分配的值,所以 Exam2 的第三盘区的大小是140k($103 + 103 * 20\% = 123.6K = 30.9$ 块,依据 Oracle 的空间分配规则,实际分配35块),而不是160k(40块)。通过上述计算可得两表的指定空间及实际分配空间分配结果如下:

表	初始盘区	下一盘区	增长百分比	第一盘区	第二盘区	第三盘区
Exam1	3K	3K	0	20K	20K	20K
Exam2	103K	103K	20	120K	120K	140K

综上所述,ORACLE 实际分配的空间与 DBA 预计分配的空间有较大出入。因此,在进行空间分配计算时,应充分考虑数据库的对象空间分配方案,减少 ORACLE 对分配空间的改动。

3.2 数据库区间尺寸的选择问题

在确定区间尺寸时,应首先考虑其对性能的影响,一味地减少表中的区间数量并不能直接提高系统的性能。在某些情况下,为一个表分配多个区间能明显减少 I/O 冲突,提高系统性能。因此,应正确设置区间尺寸,不应仅以表中盘区数量为衡量标准。ORACLE 通过两种方式从表中读取数据: RowID 方式和全表扫描方式。在 RowID 方式中,区间尺寸对读性能影响不大,在此不讨论。通过全表扫描方式读取数据,盘区大小对性能影响较大,在设置盘区尺寸时应充分考虑操作系统 I/O 缓冲区大小的限制,利用 ORACLE 的能力在全表扫描时执行多块读取。举例说明,已知一个尺寸为640k 的表,假定操作系统的 I/O 缓冲区尺寸为64k,如果将此表平均分配到10个盘区上,则每个盘区大小为64k,欲执行全表扫描,ORACLE 须执行10次读操作;如果将640k 全部放在一个盘区上,Oracle 仍需10次读取扫描全表,可见减少盘区并未提高读取性能。当盘区大小不是 I/O 缓冲区大小的整数倍时,读取次数可能增加,仍以上表为例,如将其分配在8个80k 大小的盘区上,则读取一个盘区需要执行两次读操作(因为读操作不能跨盘区进行),全表共需16次读操作。可见在此情况下,减少盘区数还降低了系统的性能。因此为了确定合适盘区尺寸应遵循如下原则:

- 1) 创建明显大于 I/O 容量的区间。
- 2) 创建大小是操作系统的 I/O 缓冲区尺寸倍数的区间。

另一方面,对所有的区间均使用自定义的尺寸,则需更多的时间用于自由空间的管理,为了避免自由空间的浪费,在选择区间尺寸时还应满足一个准则,即每个区间尺寸必须是较小区间尺寸的整数倍。

3.3 对象空间尺寸估算举例

由于数据库的对象类型较多,本文只对常见的非族表空间进行估算。非族表空间的估算只需下列四个值即可确定所需空间尺寸:①数据库的块大小;②表的 pctfree 值;③行的平均长度;④表中的期望行数。

数据库块的大小可查看 init.ora 文件中 DB_BLOCK_SIZE 的设置。一旦设定了数据库块尺寸,其值不能改变;欲进行更改须重建数据库。此外,数据库中数据块并不是全都用于存储记录,还有部分空间用于内部开销,约为90字节。pctfree 用于确定 Insert 期间未使用的自由空间尺寸,以更新先前插入块中的行。已知一个表,列数为5,平均行长度600字节,期望行数为25000,pctfree 为10%,块的大小为4k,则:

每块中用于行记录的字节 = $4096 - 90 - (4096 - 90) * 10\% = 3605.4$ 字节,舍入为3605字节。

每块存储的行数 = $3605 / 600 \approx 6$ 行。

总共所需的块数 = $25000 / 6 \approx 4166$ 块。

至此,在指定盘区时,可先创建一个16MB(4096块)的初始盘区和一个512k的 Next 盘区。这样实际分配给该表的空间为 $4096 + 128 = 4224$ 块,超过估算值1%。对于索引空间的估算,与上述方法类似,只是在索引中块内开销约161字节。

总结 本文通过对 ORACLE 数据库系统的物理和逻辑存储体系结构以及数据库对象空间分配方法进行深入分析,探讨了 ORACLE 数据库在上述方面的技术原理和特点。由于数据库系统和 ORACLE 本身相当复杂,我们将会进一步深入研究其体系结构的运行机制和原理。

参考文献

- 1 Loney K, Theriault M 著. 李纪松,等译. Oracle 8i 数据库管理员手册. 北京:机械工业出版社,2000,7
- 2 Bobrowski S,王焱,等译. Oracle 8体系结构. 北京:机械工业出版社,2000,1
- 3 Stuns D, Thomas B. OCP: Oracle8i DBA 架构与管理及备份与恢复. 北京:电子工业出版社,2001,3

(上接第126页)

SOAP 协议部署服务、传递消息,有别于 CORBA、COM 等基于 TCP/IP 协议的组件构架;G-WebService 中的服务分类已经在信息视角中给出,WSDL 是 ISO 批准的服务描述语言,同服务流描述语言 WSFL 一样都是基于 XML 的服务描述标准格式,UDDI 注册中心将是 G-WebService 各种服务登记、发布的中心,存储了已注册服务的详细访问信息,其角色类似于北美的商业黄页。

5) 技术视角 它是在前述四个视角的基础上,抽取并确定一种系统的实现方案,包括硬件配置和软件的实现,这里仅就“数字城市信息应用服务平台”软件实现,即 G-WebService 的工程体现,举例说明。

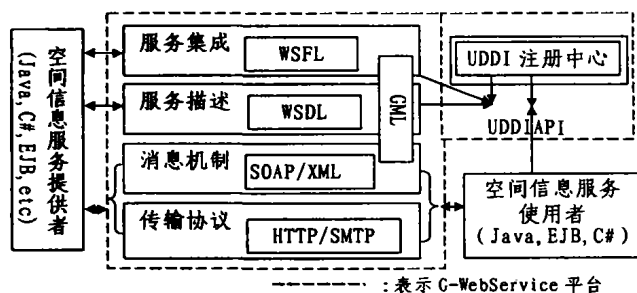


图5 G-WebService 具体的技术实现

结论 WebService 本身是针对商业应用整合而提出的新的概念框架,所以本文在 RM-ODP 模型的基础上提出 G-WebServcie 即空间 Web 服务的框架设计与实现也是一种新

的尝试。不久的将来,许多传统的 GIS、RS 应用以及 LBS(基于位置的服务)将会以网络服务的形式提供给公众,政府和企业,成为数字城市应用服务平台^[3]的核心架构。本文所提出的框架设计以及相关研究成果,已应用到数字北京示范工程的建设,和国家863攻关项目“基于 SIG 框架的数字城市服务系统与示范”中,关键技术研究 and 系统开发也正分阶段深入开展。

参考文献

- 1 Oracle9i Application Server Web Services Developer's Guide, 2002. <http://www.oracle.com>
- 2 Web Services concepts — a technical overview, HP whitepapers, 2000. <http://www.bluestone.com>
- 3 Web Services Architecture Requirements, Editor's Draft 05 June 2002. <http://www.w3.org/2002/ws/arch/2/06/wd-wsa-reqs-20020605.html>
- 4 Blanc X, Gervais M-P. Enterprise Distributed Object Computing Conference, 2000. EDOC 2000. Proceedings. Fourth International, 2000. 86~90
- 5 ISO/IEC 10746-1 ODP Reference Model Part 1. Overview. <http://www.iso.org>
- 6 Doyle A, Reed C. Introduction to OGC Web Services. OGC Interoperability Program White Paper, May 2001
- 7 OGC Interoperability Program Service Model. March 14, 2002. <http://www.opengis.org>
- 8 林绍福,李琦. 数字城市:创建21世纪的智能服务平台. 计算机科学, 2002(7)