

Oracle11gR2数据仓库性能增强Oracle认证考试 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/645/2021_2022_Oracle11gR_c102_645235.htm Oracle 11g R2于2009年9月正式发布，根据业内人士的比较，此次增添的新特性令Oracle数据库的使用更加方便。在这篇文章中，我们将介绍Oracle 11g R2的数据仓库性能增强的特性。数据仓库性能增强 Oracle 11g R1为数据仓库环境提供了一些非常优秀的新特性，最值得一提的就是几个新的分区方法，可以限制只为受影响的分区收集优化统计，并改进了SQL访问顾问，它可以为大表提供分区建议。Oracle 11g R2建立在这些新特性上，改善了并行查询性能，提高了数据仓库抽取、转换和加载操作的效率和性能，以更快的速度刷新物化视图。实例化限制（instance caging），Oracle 8i引入了数据库资源管理器（Database Resource Manager，DRM），通过资源计划指令它提供了对应用程序会话组的CPU资源利用率的限制，后来发行的数据库版本大大改善了DRM限制的范围和粒度，包括对特定资源消耗组限制I/O吞吐量，DRM最大的缺陷就不能针对某个具体实例限制CPU利用率，比如仅限制那些CPU利用率太高的实例，只要一限制CPU利用率，相同服务器上的其它实例也会遭受限制。现在即使遇到一个包含16，32甚至64颗CPU的数据库服务器也不奇怪，这些服务器上面往往运行了几十个数据库实例共享这些CPU资源，但这样做有严重的资源分配不均问题，Oracle 11g R2使用一个叫做实例化限制的新特性来克服这个问题，通过设置CPU_COUNT初始化参数为每个实例配置合适的值。DRM可以限制相同服务器上多个数据库的CPU资源，确保

不会出现一个数据库实例消耗完所有CPU资源。自动化并行度，数据仓库应用程序通常会利用并行处理信息的优势，迅速有效地处理数据，特别是在大表上查询数据时，或者是一个非常复杂的连接查询时经常使用并行查询，在执行并行操作时，优化器应该使用并行度（Degree of Parallelism，DOP），可以在查询本身内指定（通过PARALLEL优化器提示），也可以作为表或索引本身的一个属性（通过它的PARALLEL属性），但要精确地确定一个合适的DOP是有难度的，通常需要对表是如何连接，哪些索引对并行处理有益，同时执行哪些查询都要有一个详细的了解。Oracle 11g R2为我们带来了好消息，现在它可以自动为任何并行语句确定DOP，优化器使用两个新的初始化参数：PARALLEL_DEGREE_POLICY和PARALLEL_MIN_TIME_THRESHOLD来计算自动化并行度（ADOP）。例如，如果PARALLEL_DEGREE_POLICY被设为AUTO，如果查询确实能从并行操作受益，Oracle 11g R2优化器会首先确定一个合适的DOP值，如果查询的预计执行时间超出了PARALLEL_MIN_TIME_THRESHOLD可接受的值（单位：秒），并且有足够资源支持并行处理，它会允许查询执行，否则就会延迟执行，直到有足够的资源释放出来，这样可以防止一个并行查询过度地消耗资源，例如，所有并行执行线程，或集群环境中的所有CPU都被其它非并行操作占用了，值得注意的是ADOP特性无法扩展到并行恢复或并行复制，它们只适用于并行查询。并行数据缓存，在内存中并行执行：许多Oracle数据库用户发现，Oracle真正应用集群（RAC）为OLTP应用程序提供了高可用性，同时也为数据仓库应用程序提供了出色的性能表现，因为它可以在RAC集群

的不同节点上的多个实例间执行并行查询，Oracle 11g R2更好地利用了这一架构，因为它可以智能缓存在RAC集群数据库内并行查询使用的缓冲区。如果新的PARALLEL_DEGREE_POLICY初始化参数被设为AUTO，然后执行一个并行查询，Oracle 11g R2开发一个策略映射所有必需的数据块（如数据或索引块），以便更有效地在RAC数据库实例的多个数据库缓冲区高速缓存中有效地执行，如果在并行查询执行期间，一个RAC实例需要读取相同的缓冲区，它会自动使用这个映射确定哪一个远程实例持有这个缓冲区，这对物理I/O有积极影响，因为它利用了整个数据库所有实例的缓冲区缓存，分配给并行执行的查询。以大块执行并行DML，一个简单的例子就是发票模块，它需要扫描来自多个事务源（如支票、信用卡或ETF）的数以百万计的客户付款，然后更新每张发票的平衡，以及客户应付款的总额的平衡。新的DBMS_PARALLEL_EXECUTE包提供了一套有趣的功能集，在这种情况下最有用，因为它允许大的DML语句并行执行，因此它可以按大块处理数据，大块可以简单地基于选择源表的rowid、主键值或任何程序定义的规则集定义。一旦定义了这些分块规则，DBMS_SCHEDULER用于处理每个数据大块，并提交修改的数据，在处理过程中，如果某个特定的大块失败，可以重新调度处理这个大块。最后，Oracle 11g R2提供了多个新的并行任务元数据视图

（DBA_PARALLEL_EXECUTE_*），它们可以跟踪DBMS_PARALLEL_EXECUTE定义的边界，以及每个大块的处理进度。更快的物化视图刷新，数据仓库应用程序经常使用物化视图，特别是在改写查询时，以便它能够从物化视

图的小数据集获得必要的信息，而不用直接查询基础表，当物化视图的基础表变化频繁时，使用物化视图日志将变化的数据填充到物化视图中。Oracle 11g R2提供了从刷新进程外清洗物化视图日志的功能，性能也因此得到了改善，特别是那些包含了求和，连接和二者皆有的物化视图。同样，在ON COMMIT FAST REFRESH子句上新的WITH COMMIT SCN指令告诉Oracle 11g R2使用基于SCN的物化视图日志快速刷新。

ORACLE_LOADER预处理程序，将旧系统数据加载到数据仓库可能是一个昂贵的操作，因为使用SQL*Loader批量载入工具将数据载入数据库需要大量的系统资源，包括数据仓库表空间第一层存储的成本，以及旧系统的平面文件的存储成本。这些平面文件以压缩格式存储可以降低存储成本，但在SQL*Loader将旧系统数据加载到数据仓库之前，这些文件可能是非经压缩的，意味着需要额外的存储容纳这些未压缩的数据，遗憾的是，这个限制也应用到外部表，它们使用ORACLE_LOADER访问方法直接从原有平面文件检索数据，许多 Oracle用户已经使用外部表从原有的平面文件快速构建了报表，甚至使用INSERT INTO ... SELECT FROM语句进行第一次求和，其它消息型数据就直接加载到数据仓库中。因为ORACLE_LOADER访问方法已经得到了增强，它可以对SQL*Loader操作和外部表进行预处理，新的PREPROCESSOR 子句指定一个存在的DIRECTORY对象，以及该目录下的可执行程序，由这个可执行程序预处理原有的文件数据，例如，使用gunzip解压工具对其解压，然后，ORACLE_LOADER访问方法从标准输出流（stdout）中读取数据，就好像是直接从原有的文件中读取数据一样。

100Test 下载频道开通，各类考试题目直接下载。详细请访问
www.100test.com