

大话 Oracle RAC

集群 高可用性 备份与恢复

张晓明 编著



第1版

大话

Oracle RAC

集群 高可用性 备份与恢复

—— 深入浅出 ——



人民邮电出版社

Oracle RAC

集群 高可用性 备份与恢复

张晓明 编著



人民邮电出版社
北京

图书在版编目 (C I P) 数据

大话Oracle RAC : 集群、高可用性、备份与恢复 /
张晓明编著. -- 2版. -- 北京 : 人民邮电出版社,
2011.8

ISBN 978-7-115-25628-7

I. ①大… II. ①张… III. ①关系数据库—数据库管
理系统, Oracle 10g IV. ①TP311.138

中国版本图书馆CIP数据核字(2011)第101374号

内 容 提 要

本书以 Oracle 10g 为基础,对 Oracle RAC 进行了全面的介绍和分析。全书分为两个部分,共 14 章,第 1 部分是集群理论篇,这部分从集群基础知识入手,通过分析集群环境和单机环境的不同,介绍了集群环境的各个组件及其作用,以及集群环境的一些专有技术,包括 Oracle Clusterware、Oracle Database、ASM、Cache Fusion 等。第 2 部分是实践篇,每一章都针对 RAC 的一个知识点展开讲解,包括 Oracle Clusterware 的维护、HA 与 LB、备份、恢复、Flashback 家族、RAC 和 Data Guard 的结合使用、RAC 和 Stream 的结合使用,最后对 ASM 进行深入介绍,并给出性能调整的指导思想。

本书按照“发现问题→解决问题→实践与理论相结合”的方式进行介绍,首先对现实问题进行分析,然后提供合适的解决方案,最后自然地引出 Oracle 中的理论知识点,这种讲解方法能够有效地降低阅读难度,帮助读者更好地掌握相关技能。

本书可以作为数据库开发人员、数据库管理员、数据库初学者及其他数据库从业人员的工作参考手册,也可以作为大中专院校相关专业师生的参考用书和相关培训机构的培训教材。

大话 Oracle RAC: 集群 高可用性 备份与恢复

◆ 编 著 张晓明

责任编辑 杜洁

◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号

邮编 100061 电子邮件 315@ptpress.com.cn

网址 <http://www.ptpress.com.cn>

北京鑫正大印刷有限公司印刷

◆ 开本: 787×1092 1/16

印张: 36.25

字数: 950 千字

2011 年 8 月第 2 版

印数: 7 521—11 520 册

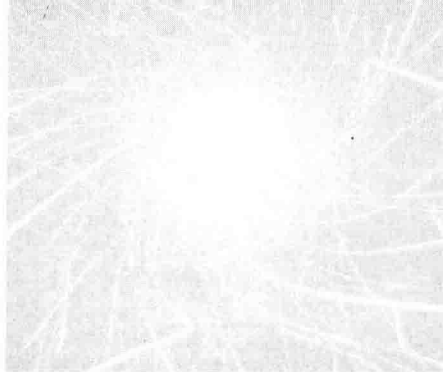
2011 年 8 月北京第 1 次印刷

ISBN 978-7-115-25628-7

定价: 79.00 元

读者服务热线: (010)67132692 印装质量热线: (010)67129223

反盗版热线: (010)67171154



新版序

2006 年的某一天突然动了写书的念头，记不清是在公交车上，还是在地铁上，还是在哪个地方诞生的这个灵感，不记得当时在干什么，也说不清楚究竟是什么原动力促使我有了这个想法，总之这个想法在脑海中反反复复酝酿了许久。更值得庆幸的是，我在忐忑与纠结中迎来了《大话 Oracle RAC——集群 高可用性 备份与恢复》（下简称《大话 Oracle RAC》）的面世，并见证它被广大的读者所接受和认可。

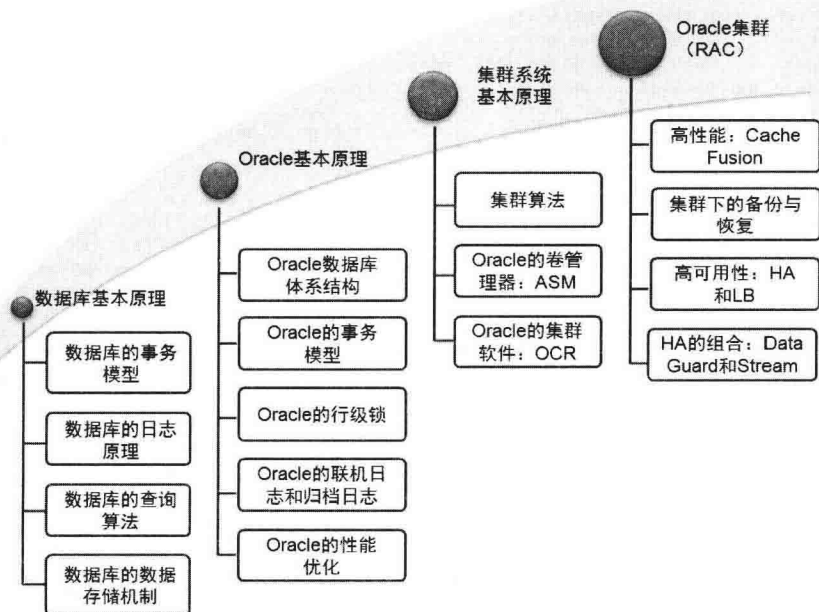
自《大话 Oracle RAC》出版以来，我陆续收到了许多热心读者的反馈，有肯定的、有提出建议的、有指出错误的；有的读者觉得读着吃力、有的读者意犹未尽……总之，热心读者的厚爱和大力支持使我萌生了修订此书的念头。

但是，应该如何进行修订？这个问题着实让我伤透了脑筋。

Oracle 现在已经到了 11.2 版本了，每次 Oracle 新版本一推出，我们这些 DBA 们都要跟着与时俱进。既然 Oracle 11.2 的流行是迟早的事，所以《大话 Oracle RAC》中提到的几个主题，以及 Oracle 11g 中与这些主题相关的改动，自然也是本书应该覆盖的内容。

《大话 Oracle RAC》原本是我自己的学习笔记，也可以说是我对自己这些年 DBA 生涯的一个总结，因此，书中介绍的内容都是我自己觉得很重要的知识点，如下图所示。

修订主要从 4 个方面入手。第一，根据读者的反馈，修正了书中的一些错误，并且对一些有歧义的地方进行说明。第二，理论基础部分删掉了一些过时的内容，同时对于一些重点技术，强调了 Oracle 11g 中的变化。第三，增加了大量的案例和总结，尤其对一些所谓的“高手秘籍”，进行了独家曝光和披露。第四，加入了更多的参考资料，比如 Metalink DOC ID，毕竟，一本书的篇幅是有限的，网络上还有更多的宝藏等你去挖掘。



就我个人的学习经验而言，我从不盲目追求新功能。这也和 DBA 的职责有关，DBA 的职业特点是“稳定压倒一切”。因此，我并不排斥新东西，只是接受新东西的前提比较苛刻：或者是旧功能的升级版，并且旧功能不完善，确实影响到当前的工作；或者它使用一个全新的理念、方法来解决之前无法解决的问题，而且该理论、方法已经在其他的领域得到了验证。

IT 行业里有一个重要的原则“不要重复发明轮子”，Oracle 公司其实是这个理论的最忠实的实践者。我一直建议，学习 Oracle 时一定要跳出 Oracle 产品本身，Oracle 的很多功能其实都是业内已有技术的 Oracle 翻版而已。比如，Oracle 的 Wallet 就是古老的 PKI，而 Oracle RAC 中的心跳、Cache Fusion 和其他集群（比如 JBoss 集群）的思路几乎一样。跳出产品去学习产品，不仅可以触类旁通、融会贯通，而且，你会发现许多所谓的新功能原来不过如此。

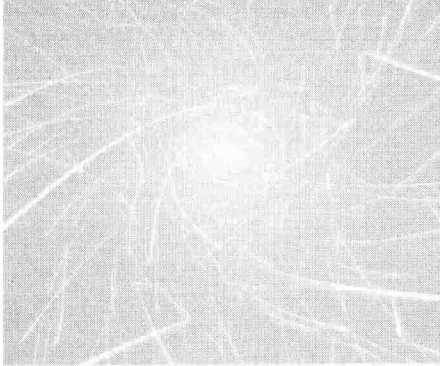
此外，不要迷信所谓的高手和秘籍。当然“他山之石可以攻玉”，前人的经验都是值得借鉴的财富。不过，沉迷于技巧并不是正确的学习态度。我们参考前人的经验时，应该关注的是他们的思想、他们的技术沉淀、他们的成长经历，而不是用了什么隐含参数、用了什么命令。所谓“台上一分钟，台下十年功”，牛人都是熬出来的，没有谁是随随便便成功的。

要重视基本理论、日常养成。到今天为止，我最经常翻看资料还是 Oracle Concept、官方文档，每次阅读都会有新的体会，所谓常读常新。很多人一看 Oracle 官方文档就打不起精神、就犯困，虽然这可以理解，但卓越就是在日复一日的枯燥中磨练出来的。

那么赶紧趁现在，掘一个洞，藏起来，勤力修炼，秘密练兵，待有朝一日，破土而出，非得像十七年的蝉一样，混着桂花香，大鸣大放……

与大家共勉！

张晓明
2011 年 7 月



序

18年前，当我第一次接触 Oracle 数据库时，Oracle 数据库对我来说就像一位带着面纱的美女，神秘而又憧憬。看着我们的 DBA 花了两三天的时间，在键盘上输入几百条甚至上千条命令，经过几次失败，才成功地在一台 HP G30 小型机上安装时，我就想也许还是做一个数据库开发员更适合我。那时的 Oracle 还没有图形化安装界面，所有操作都是通过命令行完成的。作为开发人员的我们不敢也不允许对数据库做数据增、删、改、查之外的其他操作。数据库的安装、配置、建表、建索引、增加用户等也只有 DBA 才能完成。

但在今天，情况已经完全不一样了，以前只有 DBA 能完成的工作现在大部分都可以由开发人员或其他非 DBA 人员来完成。大部分的开发人员都有在自己的 PC 机上安装 Oracle 数据库的经历，运行 Oracle Universal Installer 和运行其他普通的图形界面应用程序已经没什么区别，用户甚至可以不用提供任何参数，只是点击几下鼠标即可完成，整个过程可以不超过 30 分钟。但这并不意味着我们不需要 DBA，就像波音 747 飞机上的驾驶员，虽然 747 的自动导航系统可以控制飞机从起飞、巡航到降落，但我相信没有一个乘客会说不需要飞机驾驶员了，相反我们会提出更高的要求，因为我们需要更安全、更舒适的飞行。同样，DBA 也是如此。

我从事 IT 工作已近 20 年，还从来没像今天这样关心数据库的安全和可靠性，因为数据库已经成为我们一切业务的核心。我们公司为电信企业提供业务服务，这就要求我们的数据库一年 365 天不间断地可靠运行，任何一种中断都是不能接受的。每天数据库要处理上亿次的交易，每天数据量接近 8GB，并还在不断增长中。为提高数据库的可用性、扩展性和可靠性，我们的 DBA 团队尝试了不同的技术，最后采用了 Oracle RAC。经过整个团队的努力，现在我们已经成功地在我们的生产环境部署了 Oracle RAC。张晓明先生，既是

本书的作者，也是带领我们 DBA 团队成功实施 Oracle RAC 的技术领头人。我相信不久，随着 RAC 技术越来越成熟，会有越来越多的公司选择 RAC。作为 DBA，就像作者在书中写的那样：“对于还从没接触过 RAC 的 DBA 来说，现在该是未雨绸缪的时候了。”

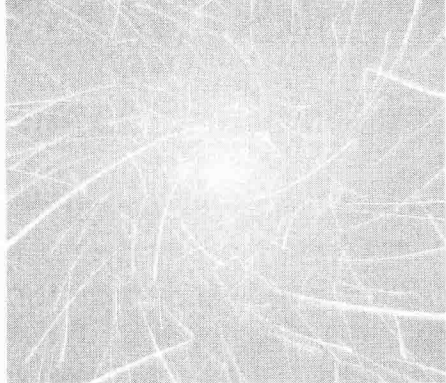
在书中，作者也遵循了著名的 Oracle 大师，Oracle 公司副总裁 Thomas Kyte 提倡的 “*Test thing yourself and base decision on proven examples*”。他从一个简单的实施实例出发，循序渐进、由浅入深地讲解了 Oracle RAC 的实施细节、RAC 的工作原理和实施、维护过程中可能碰到的问题和解决技巧。我相信所有的 DBA 都可受益匪浅。

IBM 创始人 Thomas J. Watson 曾说过：“*Though have been father of every advance since time began, ‘I didn’t think’ has cost the world millions of dollars*”，我想这本书能带给你的不仅是知识和技术，更是一种不同的思维和观点。我希望每位读者能从书中得到更多的思考，并为这个世界节省更多的 “millions of dollars”。



林镇武 副总裁

北京无限新锐网络科技有限公司



前 言

从开始动笔到定稿出版,这本书的编写几乎花费了我 10 个月的时间,大大超出了最初 3 个月的计划。这个过程虽然不能说寝食难安,但其中的酸甜苦辣也是如人饮水,冷暖自知,如果一开始就知道要投入这么大精力,我真怀疑自己是否有勇气开始动笔。

回顾自己学习 RAC 的经历,我对它又爱又恨。因为对技术本身的痴迷加上高“薪”技术的诱惑,所以爱;因为没有能力掌控,所以恨。时至今日,我仍然不敢说自己已完全掌握了 RAC, RAC 的学习过程是对我所有学习过的 IT 技术的大复习,包括 OS、Network、Storage、Database 等,甚至可能还有很多领域我从未涉足。

相信很多对 Oracle 有一定经验的 DBA 和我有同样的感觉, RAC 比普通的 Oracle 更难入门。不仅因为比比皆是的海涩艰深的术语,也不仅因为它覆盖的技术领域太广,更主要的是可用的参考资料太少。我翻遍了所有能够获得的书籍资料,但是这些资料都是从很基本、很琐碎的细节开始,从下到上(Bottom-Up)的学习方式是虽然看到了“树”,但是看不到“林”。虽然无树不成林,技术细节总是必要的,但过于琐碎的分解给我造成更大的迷惑,到底 RAC 的总线是什么? GCS、GES、GRD、PCM 这些东西是如何堆砌在一起的?

这段痛苦经历促使我决定以自己的方式讲述 RAC 的故事,如果大家能在这个旅程中豁然开朗,对我而言将是一件很有趣又备感欣慰的事。

本书特点

本书最大的特点是侧重于理论讲述和实战演练,尤其是对理论的剖析有一定的深度,并通过大量完整的案例来论证这些理论。DBA 和开发人员特点不同,对于开发人员来说,从某种角度上,强调创造力决定高度。但是 DBA 要求的是

稳定压倒一切，DBA 必须去适应自己所使用的产品，而不能期待产品去适应自己，更没有机会去改变产品，因此了解这个产品的工作原理是非常必要的。

现在网络发达，遇到问题时，有论坛、BBS 各种渠道可以利用。相信大家都有类似这样的体会：“哇，这家伙怎么能想到这样做？怎么敢这样做？”。其实这些高手的所谓“特技”都不是偶然，也不是贸然下手，只是他对产品有比你更深刻的认识，知道问题的前因，知道行动的后果，所以看起来他可以“天马行空、为所欲为”，而你只能“亦步亦趋、如履薄冰”。因此，我个人更强调对基础的掌握，而不要迷信有什么技巧或捷径。请相信只要能在一个领域持续上数年，你就在这个领域里能成为专家。

采用这种方式还有一个原因，RAC 覆盖知识面广、资料难觅也是众所周知的。时至今日，我不敢妄谈精通，只是尽可能地把所知中规中矩地记录下来，不敢随意发挥，希望能够抛砖引玉，并对大家有所启发和帮助。

本书适用读者

本书不是一本 Oracle 数据库的基础入门书籍，而是关于 Oracle RAC 的入门指导。

这本书适合于初、中级数据库管理员和数据库开发人员，但是本书不会特别讲述什么是 SGA，什么是数据文件，什么是字典视图，什么是日志，也不会专门介绍如何创建表空间、用户等。如果你对上面这些名词、操作都非常陌生，那么这本书不适合你，请先夯实单实例的基础（可以参考下面的“学习资源推荐”）再来看这本书。

本书假定读者已经脱离了入门阶段。如果你对 Oracle 单实例数据库有一定基础，并做过类似 DBA 工作，但是由于条件所限，一直对 RAC 敬仰如滔滔江水，渴望拥有却不知如何下手，那么这本书就非常适合你。如果你需要在最短时间内充电 RAC 相关知识，并且敢在简历上写下“有一定 RAC 经验”，那么这本书就是为你而准备的。

可能你已经有了有一定的 Oracle 使用经验，比如做过些安装、管理、备份操作，对 Oracle 数据库有了一定感性认识，但是并没有完整的理论框架，那是最好的，你可以通过这本书学习到 Oracle 最核心的理论基础，并看到这些理论在单实例和 RAC 环境下的各种应用，不过还是建议你找一本基础入门的书籍放在手边，方便速查。

学习资源推荐

- http://www.oracle.com/pls/db102/portal.all_books, Oracle 10.2 官方文档，最权威的文档，适合系统学习。
- <http://www.oracle.com/technology/index.html>, OTN，资料新、覆盖面广，尤其适合快速入门、图文并茂；缺点是分类方式不太适合中国用户的习惯。
- <https://metalink.oracle.com>, Metalink，各种深度的文档一应俱全，也是所有补丁的源地，不过用户需要申请账号。
- <http://asktom.oracle.com>, 大名鼎鼎的 AskTom。
- <http://www.databasejournal.com> 和 <http://www.oracle-base.com> 这两个站点程度类似，经常会有针对某专题的系列文章，适合初学者。

本书结构与内容

本书共 14 章，分为集群理论篇、实战篇两大部分。

工欲善其事，必先利其器，学习 RAC，环境是最重要的。所以第 1 章“RAC 初体验”将引导读者在一台 PC 上搭建出一个 2 节点的 RAC 环境，并在建好之后快速感受 RAC 两大亮点——负载均衡、高可用性。这一章内容虽然是介绍安装，但并不是简单的单击“上一步”、“下一步”按钮，其间渗透了许多作者的体会和心得，所以即使有安装经验的读者，也无妨做一次快速翻阅，相信也会有所收获。

从第 2 章至第 5 章主要是基础理论介绍，第 2 章首先概述了集群的概念、分类，并在其基础上对 RAC 架构做一个快速浏览。接下来的 3 章，每一章都讨论了 RAC 的一个组件。

如果对 RAC 环境做一个解剖，可以发现集群软件架构可以分为 3 个组件：Clusterware Manager（集群件）、Distribute Lock Management（DLM，分布式锁管理器）、Oracle 实例。这 3 个组件在 Oracle 的发展历史各不相同。Oracle 在版本 9 之前一直是依赖厂商的集群件产品，比如 Sun 的 Clusterware。直到 Oracle 9 才开始提供自己的集群件产品 Cluster Manager，但是只支持部分平台。到了版本 10，这个产品已经演变成一个功能完整、支持所有平台的独立产品 Oracle Clusterware，这个产品不仅支持 RAC，而且还能为其他非 Oracle 软件提供高可用支持。第 3 章专门介绍 Oracle Clusterware。而第 7 章会演示如何利用这个产品搭建一个 Web 集群。

无论 Oracle 还是 IBM、微软公司的数据库产品，无论单实例数据库还是集群数据库，它们核心功能都是一样的——事务处理，并在保证数据完整的前提下提供最大的并发支持。大家都知道，数据库系统是通过“锁机制”来实现的，而在集群环境下，使用的是“分布式锁机制”。其实不仅仅是数据库，所有支持多用户并发的系统都需要某种“锁机制”，因此无论“锁”还是“分布式锁”都不是 Oracle 所特有的，它是所有这些系统的共性。

Oracle 单实例的锁机制包括两种：Lock 和 Latch，Oracle RAC 同样也离不开这两种锁机制，但是 Oracle RAC 又多了一种新的锁——分布式锁，这种锁的管理机制叫作分布式锁管理器（Distributed Lock Manager，DLM）。Oracle 集群最初使用的是操作系统的 DLM，但是操作系统的 DLM 仅适用于数量很少的资源的分布式管理，比如对文件，对于 Oracle 数据库动辄 G 级的数据量，这种 DLM 就非常不适合。所以 Oracle 开发了自己的 DLM 机制，在 RAC 之前的 OPS 中，这种技术叫作 PCM（Parallel Cache Management），到了 RAC，这个技术改名为 Cache Fusion。可以说，RAC 的学习、管理、优化都是围绕着 Cache Fusion 进行的。不过与 Clusterware 不同，Oracle 并没有把 DLM 独立为一个产品，而是把 DLM 功能集成到数据库内核中，如果安装过程中检测到 Clusterware 的存在，就会激活内核的 DLM 功能。

第 4 章围绕着“锁”在单实例和 RAC 的使用，详细讲述了 RAC 原理。这一章是全书最重要的部分，也是作者花费心血最多的章节，希望读者详细研读。

以“锁”为总线来学习 RAC，这个灵感来自于我在雅虎工作期间主持的一次内部培训，也正是这个灵感造就了本书。

当时雅虎的一个系统要从 SQL Server 转移到 Oracle，大家希望能对 SQL Server 和 Oracle 的区别有所了解，于是我就主持了一次内部培训。当时对培训主题的选择是煞费苦心，如果从两个产品的媒体发布数据、销售数量做比较，对于技术人员显然没有任何意义；如果单从产品功能、用户友好性来比较，也没有任何说服力。在否定了若干方案后，我开始思考“数据库最本质的功能是什么？”，思索的结果就是最终选定“并发和锁”作为比较的主题。通过介绍两种产品的“锁机制”是如何实现的，以及各自的实现方法对并发能力的影响，从而达到比较的目的。

的。这个讲义（PPT 格式）可以在我的 Blog（<http://space.itpub.net/75321>）找到。需要说明的是，这个讲义是以 Oracle 9i 和 SQL Server 2000 为基础，其中关于 Oracle 9i 的部分对 Oracle 10g 仍然有效，但是 SQL Server 2000 部分对其他版本是否适用，就需要根据具体情况来定。

这次内部培训之后，借助这个灵感，我对 RAC 知识重新梳理，终于找到了学习 RAC 最有效的方法（至少我个人这样认为），也最终产生这本书。

第 5 章带领读者认识 ASM，这是 Oracle 10 推出的自动存储管理系统，这个系统同时综合了集群文件系统和卷管理器两方面功能，也是 Oracle 10 RAC 的首选存储方案。

本书的第 6 章至第 14 章是实战部分，每一章都是一个独立的讨论主题。

Oracle 10g 中的管理工具被大大加强，提供了基于 Web 界面的管理方式。这些工具虽然简单好用，但对于初学者来说恰恰也是缺点，首先这些工具隐藏了背后的运作机制，其次对于这些工具本身的学习也会分散读者的精力，何况在某些现场根本没有图形界面可以使用。所以，我更推荐读者在学习阶段还是使用“笨”方法：手工输入命令+观察输出结果。在有一定的经验以后，再去使用各种图形工具，更何况 Oracle 中需要记忆的命令数量也是非常有限的，差不多一章的内容就可以全部覆盖。

因此，第 6 章把 RAC 环境中可能用到的命令进行整理，并按照一种便于记忆的方式进行归类。但这一章并不是命令手册，很多命令会有些较深入的分析。当然在以后的章节学习中，读者可能还会需要反复查阅这一章。

第 7 章介绍 RAC 最主要的亮点“HA 和 LB”，第 8 章介绍 RAC 的备份，第 9 章介绍恢复技术中的完全恢复、不完全恢复，第 10 章介绍其他恢复技术，包括数据块恢复、Flashback 功能家族介绍。

Oracle 的高可用家族除了 RAC 还有两个兄弟：Data Guard 和 Stream Replication，相对于 RAC，这两种方案都比较容易上手，也是目前企业常用的 HA 方案。但是简单好学并不是代表功能会打折扣。这三兄弟每个既能独当一面，也可以互相配合使用，提供更强大的功能。但是，这三个兄弟针对的问题并不一样，适用场合也不尽相同。因此学习 Oracle 的高可用方案时，必须搞清楚这些方案的区别，最贵的并不一定是最适合的。因此，第 11 章比较了 RAC 和 Data Guard 的区别，以及二者如何搭配使用。第 12 章解释了 RAC 和 Stream Replication 的区别和合作。

第 13 章是对 ASM 的深度介绍，其中 ASM Alias 部分很重要，对于用户可能遇到的问题给出了具体的解决办法。

第 14 章主要介绍了调优方面的内容，但只介绍了 SQL 调整策略以及 RAC 环境下要考虑的特殊内容。之所以只选择这些内容，是因为 Oracle 性能调整是一个非常大的主题，经典话题包括主机、内存、网络调整，SQL 语句调整、等待事件调整等，新话题可以讨论 Oracle 10g 新添的若干工具 AWR、ASH、ADDM 等。如果在一章中把这些内容都一网打尽，只能是蜻蜓点水，与其这样不痛不痒地堆砌，还不如有针对性地深入讨论几个重点。需要更多了解性能优化内容的读者，可以参考其他资料和文档。

本书声明

（1）除了特别声明外，本书实例中使用的操作系统是 Red Hat AS 4.0 Update 4，使用的数据库是 Oracle 10 R2 版。

(2) 部分工具软件的下载地址。

Oracle Database、Oracle Clusterware、ASM Lib 驱动

<http://www.oracle.com/technology/global/cn/software/index.html>

Red Hat Advanced Server

<https://www.redhat.com/apps/download>

(3) 作者在编写本书过程中，以“够用就好”为原则，尽量覆盖到 RAC 所有知识点。但所有观点都出自作者的个人见解，疏漏、错误之处在所难免，欢迎大家指正。读者如果有好的建议或者学习本书过程中遇到问题，欢迎发送邮件到 mydba123@gmail.com，希望能够与大家一起交流和进步。

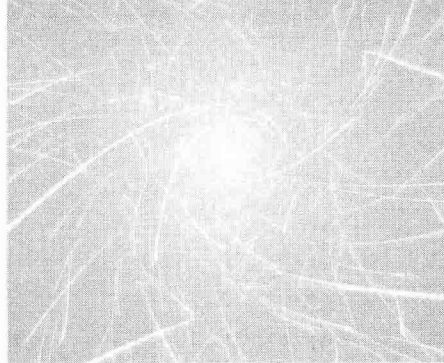
感谢

在此，感谢芮玉奎先生，他在内容结构的编排方面给我很多帮助，感谢 eygle，正是他的推荐，才促成我与人民邮电出版社的这次合作。在本书编写过程中，还得到了很多朋友的支持和帮助，包括清华大学的焦丽静，IT168 的贝贝，阿里巴巴的红星、大鹏，Oracle 公司的 rebecca、bryanxu，爱康网的 cindy、Viky，网友 nitenike、sunadmin、nyan、firebing、Kane、Jacky、Rainbow、Ashily 等，在此一并感谢。

在看得见的地方学习知识，在看不到的地方学习智慧。同时也祝愿大家在 Oracle RAC 的学习道路上顺风顺水。

张晓明

2011 年 7 月



目 录

第 1 部分 集群理论篇

第 1 章 RAC 初体验	2
1.1 本书使用环境	3
1.1.1 硬件环境	3
1.1.2 软件环境	4
1.1.3 本书使用的环境	6
1.2 如何在 PC 机上搭建 RAC 环境	7
1.2.1 需要下载的软件	7
1.2.2 安装过程	8
1.3 任务列表	16
1.4 规划阶段	17
1.4.1 确认主机名和 3 个 IP 地址	17
1.4.2 存储方案选型	17
1.5 实施阶段	19
1.5.1 主机配置	19
1.5.2 安装 Oracle Clusterware	27
1.5.3 安装 Oracle Database	33
1.5.4 配置 Listener	35
1.5.5 创建 ASM	38
1.5.6 创建数据库	40
1.6 客户端测试	43
1.6.1 客户端配置	43
1.6.2 体验 Failover	44
1.6.3 体验 LoadBalance	45
1.6.4 修改归档模式	46
第 2 章 集群、RAC 和 MAA	48
2.1 什么是集群	48
2.1.1 系统可靠性	49

2.1.2	系统性能	49
2.1.3	集群的技术基础	50
2.2	集群术语	51
2.2.1	两大关键特性: 可扩展性和高可用性	51
2.2.2	两大核心能力: 负载均衡和故障转移	52
2.2.3	集群组成元素	52
2.2.4	集群分类	52
2.3	Oracle 的高可用架构 (MAA)	53
2.3.1	MAA 实例	54
2.3.2	双机热备——一个被遗漏的技术	55
2.4	RAC 集群	56
2.4.1	存储层	57
2.4.2	网络层	57
2.4.3	集群件层	58
2.4.4	应用层	58
2.5	RAC 集群环境的特殊问题	59
2.5.1	并发控制	59
2.5.2	健忘症 (Amnesia)	59
2.5.3	脑裂 (Split Brain)	59
2.5.4	IO 隔离 (IO Fencing)	60
2.6	RAC 的前世今生	61
2.6.1	Oracle Database 的历史	61
2.6.2	Oracle RAC 的历史	62
2.6.3	OPS 和 RAC 的区别	64
2.7	RAC 的好处	65
2.7.1	增加可用性	66
2.7.2	扩展性	67
2.7.3	更容易管理	67
2.7.4	其他选择	68
2.8	小结	68
第 3 章	Oracle Clusterware	69
3.1	Clusterware 和 RAC 的关系	69
3.2	Oracle Clusterware 组成	70
3.2.1	磁盘文件	70
3.2.2	Clusterware 后台进程	76
3.2.3	网络组件	83
3.3	Clusterware 的日志体系	87
3.4	Clusterware 和厂商集群的关系	88
3.5	小结	89
第 4 章	RAC 原理	90
4.1	数据库基本原理	90

4.1.1	并发访问和数据一致性	91
4.1.2	事务和隔离级别	92
4.1.3	Oracle 支持的隔离级别	93
4.2	Oracle 单实例的并发控制机制	94
4.2.1	Lock	94
4.2.2	数据记录的行级锁	95
4.2.3	Latch	99
4.2.4	Latch 和 Lock 对比	99
4.2.5	进一步理解	100
4.3	RAC 下的并发控制	102
4.3.1	DLM 中资源和锁	103
4.3.2	Non-Cache Fusion 资源	104
4.3.3	Cache Fusion 资源	105
4.3.4	GRD (Global Resource Directory)	106
4.3.5	PCM Lock	107
4.3.6	Cache Fusion	108
4.3.7	RAC 并发控制总结	113
4.4	RAC 架构	113
4.4.1	SGA 的变化	114
4.4.2	后台进程的变化	114
4.4.3	文件	116
4.4.4	SCN	117
4.4.5	Cache Fusion、GCS、GES	118
4.5	RAC 和 Clusterware 的交互	119
4.5.1	Clusterware 层	119
4.5.2	RAC 层	120
4.6	小结	123
第 5 章	存储方案——ASM 初步	124
5.1	Shared-Disk 和 Shared-Nothing 架构	124
5.2	Oracle 10g RAC 的存储方案介绍	125
5.3	ASM 架构	125
5.3.1	ASM 实例	126
5.3.2	文件	127
5.4	配置 ASM	127
5.4.1	安装位置	127
5.4.2	创建 ASM 磁盘	127
5.4.3	配置 ASM 实例	137
5.4.4	创建磁盘组	140
5.4.5	在数据库中使用 ASM 的磁盘组	140
5.4.6	如何从远程访问 ASM 实例	141
5.5	ASM 实例：将数据库迁移到 ASM	141

5.5.1	试验说明	141
5.5.2	操作步骤	141
5.5.3	最后验证	148
5.6	小结	148

第2部分 实战篇

第6章	RAC 维护工具集	150
6.1	Oracle Clusterware 工具集	150
6.2	节点层	150
6.3	网络层	151
6.3.1	公有网络	151
6.3.2	私有网络	151
6.3.3	私有网络的参数调整	154
6.3.4	oficfg 命令	155
6.3.5	工程中解决不能识别心跳网络的故障	157
6.4	集群层	159
6.4.1	crsctl	159
6.4.2	OCR 命令系列	164
6.4.3	一个 Bug 引起的系统重启	171
6.5	应用层	185
6.5.1	crs_stat	185
6.5.2	onsctl	188
6.5.3	srvctl	193
6.5.4	恢复	199
6.6	小结	202
第7章	HA 和 LB	203
7.1	什么是高可用性	203
7.2	Failover	204
7.2.1	Client-Side Connect Time Failover	204
7.2.2	TAF (Transparent Application Failover)	204
7.2.3	Client-Side Failover 和 TAF 的对照试验	205
7.2.4	Server-Side TAF	208
7.2.5	深入 TAF 细节	214
7.3	Oracle Clusterware HA 框架	221
7.3.1	术语介绍	221
7.3.2	配置命令	223
7.3.3	完整实例	228
7.4	LoadBalance	232
7.4.1	Connection Balancing	232
7.4.2	利用 Service 分散负载	235