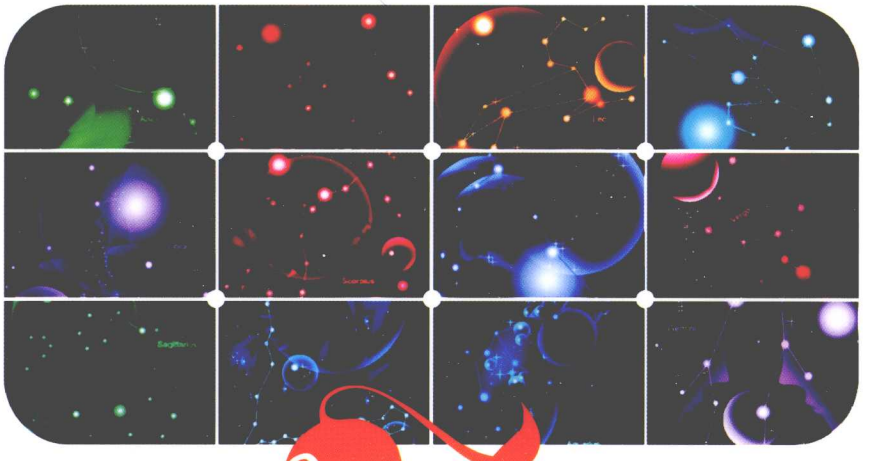
HZ BOOKS
华章科技

数据库技术丛书

多年管理和维护经验总结　案例实用方便读者快速上手



Oracle RAC 11g 实战指南

刘宪军◎编著

机械工业出版社
China Machine Press

数据库技术丛书

Oracle RAC 11g 实战指南

刘宪军◎编著

NLIC 2970671028

机械工业出版社
China Machine Press

Oracle RAC 11g是Oracle公司最新推出的集群数据库版本。本书从实用的角度出发，详细介绍了RAC的安装过程和管理方法。书中提到了许多在安装和管理RAC时需要注意的问题，这些问题都是作者在实际的工程实施过程中遇到过并亲手解决的。

本书主要介绍RAC 11.2在UNIX/Linux系统中的安装和管理方法。从内容组织形式上来看，本书共分为10章和2个附录：第1章介绍了系统高可用性的概念，第2章介绍了RAC的体系结构，第3章介绍在AIX系统中所需要做的安装前的准备工作，第4章～第6章介绍了RAC的整个安装过程，第7章介绍RAC集群环境的管理方法，第8章介绍ASM实例和磁盘组的维护方法，第9章介绍了对RAC数据库的常规管理方法，第10章介绍集群数据库的备份与恢复。附录A和附录B分别介绍了在Solaris和Linux两种操作系统中需要做的安装前的准备工作。

与以前的版本相比，RAC 11.2有许多重要的变化，如：彻底放弃了对裸设备的支持，ASM和Clusterware一起合成了Grid Infrastructure软件，引入了SCAN地址和ACFS文件系统，在ASM磁盘组中可以创建卷，OCR和Voting文件可以存储在ASM磁盘组中等。读者在学习RAC时，需要特别注意不同版本之间的差别。

本书既不是对Oracle官方文档的翻译，也不是对RAC安装过程的简单描述，而是作者实施和管理RAC的经验的结晶。本书不仅告诉读者怎么做，还告诉读者为什么这么做。在书中提到许多需要注意的问题，这些问题都是在RAC的实施和管理过程中经常遇到，并且在官方文档中可能找不到答案的问题。

本书不仅可以作为工程技术人员的参考手册，还可以作为培训中心的培训教材。

图书在版编目（CIP）数据

Oracle RAC 11g实战指南/刘宪军编著. —北京：机械工业出版社，2011.1
（数据库技术丛书）

ISBN 978-7-111-32877-3

Ⅰ. O…　Ⅱ. 刘…　Ⅲ. 关系数据库－数据管理系统，Oracle RAC 11g－指南
Ⅳ. TP311.138-62

中国版本图书馆CIP数据核字（2010）第256793号

机械工业出版社（北京市西城区百万庄大街22号　邮政编码　100037）
责任编辑：陈佳媛
北京市荣盛彩色印刷有限公司印刷
2011年1月第1版第1次印刷
186mm×240mm　·　13.25印张
标准书号：ISBN 978-7-111-32877-3
定价：39.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：（010）88378991；88361066
购书热线：（010）68326294；88379649；68995259
投稿热线：（010）88379604
读者信箱：hzjsj@hzbook.com

前　言

自从计算机诞生以来，在全世界就产生了一个全新的行业，无数工程师投身于这个新兴的行业。然而，在这个行业里挣扎多年以后，工程师们发现了一个他们不愿意承认的事实：计算机是不安全的。无论人们做什么样的努力，计算机还是不安全的。

计算机的安全实际上就是数据的安全。为了保证数据的安全，人们开发了许多附加于计算机之上的硬件和软件产品，如防火墙、杀毒软件、磁盘阵列、集群、异地容灾等。借助于这些产品，人们希望能够尽量保证数据不受诸如病毒感染、黑客攻击、硬件故障、极端分子破坏等各种事件的影响。

集群是一种软件和硬件相结合的应用环境，它的功能是向用户提供一个可以24小时不间断访问的高可用环境。目前各个大的计算机公司都有自己的集群软件，如IBM 公司的HACMP、HP公司的ServiceGuard、SUN公司的SUN Cluster，以及Oracle公司的RAC。其中前三种产品只能安装到各公司自己的操作系统中，而RAC可以安装到各个主流的操作系统中，可以应用于适合数据库运行的各种环境。

RAC是伴随Oracle 9i数据库产品开始出现的，当时的RAC功能比较简单，而且需要依赖于操作系统中的集群。从10g版本开始，RAC不再依赖于操作系统集群，可以独立地构成一个集群环境。经过几年的平稳发展之后，Oracle推出了功能更加强大的 RAC 11g。与以前的版本相比，RAC 11g增加了很多新的特性，即使在11.2和11.1两个子版本之间，也有很多不同的特点。

在RAC 11g中，Oracle对自己的存储管理技术——ASM的支持力度进一步加大。从11.2开始，ASM作为一个独立的产品，与Clusterware产品合在一起，构成了Grid Infrastructure软件。有了ASM技术，RAC中重要的共享文件（如OCR、Voting和数据库文件）就可以存储在ASM磁盘组中，而不再依赖于共享文件系统和裸设备。

尽管很多重要的数据库系统都处于或者将要处于RAC环境中，尽管RAC的功能非常强大，但是RAC的安装和管理却是非常困难的，而且这方面的资料非常少。虽然Oracle公司提供了详细的官方文档，但是读者面对一大堆的英文文档，可能理不出什么头绪来。而且在安装过程中可能会遇到很多疑难问题，在官方文档中根本就找不到这些问题的答案。在网上虽然也能找到一些关于RAC的文档，但是这些文档基本上都是RAC爱好者发布的一些并不完整的心得体会。按照这些文档，读者很难把RAC安装成功。

本书的编写目的，就是为读者提供一套“手把手”的技术指南，书中介绍了在AIX、Solaris和Linux三种主流操作系统平台下，RAC的详细安装过程和管理方法。需要注意的是，RAC毕竟是一个非常复杂的应用环境，读者在安装RAC之前，需要对Oracle数据库的常规管理以及Oracle所运行的UNIX操作系统的管理方法非常熟悉，而且在安装RAC的过程中需要反复

实践，不断地积累经验。一般来说，初学者在安装RAC时，失败的几率是100%。然而只要你能达到“为伊消得人憔悴”的境界，那么当你被失败折磨得痛苦不堪的时候，你可能距离成功只有一步之遥了。就在你“众里寻他千百度”之后，蓦然回首，发现“那人却在灯火阑珊处”，你将体会到“衣带渐宽终不悔”的舒畅。

如果说安装RAC是最复杂的工作的话，那么管理RAC则是最重要的工作。只要会安装RAC，那么基本上就知道怎样管理RAC了。然而RAC毕竟和单机的数据库有很大的区别，要管理好RAC并不是一件容易的事情。如果RAC出现了故障，工程师往往不知道如何下手去解决这样的问题。本书作者在解决这种问题时采取的方法是这样的：利用闲置的计算机，创建一个模拟的RAC环境，与生产系统的RAC结构完全相同，在对实际的RAC进行维护之前，先在模拟环境中试验一次，如果没有什么问题，再在实际的RAC中实施。

需要注意的是，不同版本的RAC之间都有较大差别，11.2版本与以前版本之间的差别更大。读者在安装RAC之前需要了解RAC的版本。只要按照本书介绍的内容，掌握RAC 11.2的安装，那么借助于其他文档，再安装其他版本的RAC，就是很轻松的事情了。

本书由刘宪军编写。本书的出版得到了机械工业出版社编辑的大力协助，没有他们，这本书不可能这么快和读者见面。在这里对他们所付出的劳动表示诚挚的感谢。

由于在RAC的安装和管理过程中涉及多个用户、多种工具，为避免混乱，在这里列出常见的命令提示符：

#	UNIX/Linux的SHELL提示符，表示root用户的登录
$	UNIX/Linux的SHELL提示符，表示oracle用户或grid用户的登录
ASMCMD>	ASMCMD工具的提示符
RMAN>	RMAN工具的提示符
SQL>	SQL*Plus工具的提示符

作　者

2010年11月

目　录

第 1 章　高可用性概述

计算机安全一直是困扰相关从业人员的一个重要问题。为了保证计算机的安全，尤其是保证计算机中所存储数据的安全，人们研究开发出无数的硬件和软件产品。例如：为了防止病毒感染，人们开发了杀毒软件；为了防止黑客的袭击，人们开发了防火墙；为了防止磁盘损坏而导致数据丢失，在重要的计算机中都使用磁盘阵列来存储数据；为了防止计算机出现故障，人们又研究开发了集群产品；为了防止计算机所运行的机房环境受到破坏，人们又开发了异地容灾产品。

总之，计算机的安全问题一直伴随着计算机的发展。计算机在给人们的生活带来翻天覆地的变化的同时，也给人们带来了无穷的烦恼。无数的聪明人在研究更先进、更智能、运算速度更快的计算机，又有无数的聪明人在研究如何保护这种脆弱的智能机器。

高可用性是计算机安全问题中的一个重要分支，它的目的是保证重要的计算机系统可以向用户提供不间断的访问。

1.1　什么是高可用性

许多重要的业务系统都需要提供7×24（即一天24小时，一周7天）不间断的服务，如银行、电信、保险、政府等部门的业务系统。对于一个普通用户来说，他最关心的问题是，他所需要的服务能否得到满足，比如在银行的ATM上能不能随时取出钱来，他所关心的股票能不能顺利地进行交易，在医院看病时能不能得到保险公司所支付的医药费等。而对于运营者来说，这些业务系统停止运行就意味着巨大的经济损失，更重要的是，这将失去用户的信任。

随着经济全球化的加速发展，人们对这种以计算机为核心的业务系统的依赖性越来越强。对于管理这些业务系统的部门来说，面临的重要任务是保证它们的高可用性，尽量减少停机时间，如果业务系统出现停机现象，应该在最短时间内对它们进行恢复。

业务系统的停机包括计划内停机和计划外停机两种。计划内停机是指管理员有意识安排的停机，比如在对硬件进行升级、对软件进行升级、更换损坏的硬件、对系统进行备份、系统的新功能测试时，可能需要停止业务系统的运行。计划外停机是指非人为的、因外界环境变化而引起的停机，比如当硬件出现重大故障、应用程序停止运行、计算机所运行的机房环境遭到灾难性的破坏时所引起的业务系统停止运行。

实践证明，在所有的引起业务系统停机的各种情况中，计划内停机约占85%，而由于硬件

故障所引起的停机现象大约只占1%。也就是说，绝大部分停机现象是人为安排的，这多少有点出乎人们的意料。实际上，随着工艺水平的不断提高，计算机硬件出现故障的几率越来越低，随着人们对业务系统的重视程度不断提高，计算机所运行的机房环境也越来越安全，所以，计划外停机对业务系统的影响是很小的。很多客户都有这样的经历：业务系统连续运行几年都没有出现任何故障，偶尔有一天对软件或硬件进行升级时，或者仅仅清理机箱后面的灰尘，却使业务系统无法重新正常运行。

高可用性的目的就是尽量减少停机时间，无论是计划内停机还是计划外停机。目前实现高可用性的方法基本上都是这样的：在一个业务系统中，对于重要的计算机组件，都有一个或多个对等的组件作为后备，一旦某个重要的组件由于人为原因或者因为故障而无法工作，后备的对等组件马上接替它的工作。在业务系统中，我们把这种由于重要的计算机组件没有后备的对等组件而可能引起的故障称为“单点失败”（即SPOF，Single Point Of Failure）。高可用性就是通过消除业务系统中的单点失败而减少停机时间。

1.2 如何获得高可用性

根据前面的描述，我们已经知道，通过消除单点失败，可以减少业务系统的停机时间，从而提高业务系统的高可用性。对于那些需要运行重要业务系统的计算机，厂商在硬件和软件方面都进行了一系列的强化，以保证计算机本身的高可用性。

在硬件方面，作为服务器的计算机大都具有以下特点：

1）多CPU

在计算机中至少有两颗CPU。通过多CPU，不仅可以保证多个进程能够真正实现并发执行，而且可以保证当一颗CPU出现故障时，整个计算机仍然能够运行。

2）冗余电源

在计算机中至少有两个电源，当其中一个电源出现故障时，其他电源仍然能够向计算机持续供电。

3）冗余网卡

在计算机中一般有两个以上网卡，用户可以通过任何一个网卡访问计算机中的服务。在网卡上可以指定可“漂移”的IP地址，当一个网卡出现故障时，这个网卡上的IP地址就漂移到另外一个网卡上。通过多个网卡还可以实现数据流量的均衡。

4）ECC内存

ECC（Error Checking and Correcting）内存就是具有错误检测和错误纠正的内存，这种内存可以检测到数据的错误，并对其进行纠正，从而使计算机系统更加稳定。

5）可热插拔的设备

许多外部设备都可以进行热插拔，如果这样的设备出现故障，工程师可以在不关闭计算机系统的情况下对其进行更换和修理。

6）磁盘阵列

对于一个业务系统来说，核心的部分是数据。为了保证数据的安全，在计算机中大多使用磁盘阵列作为数据的存储设备。磁盘阵列通过RAID（Redundant Array of Independent Disk）技术将多个独立的磁盘虚拟为一个大的存储空间，数据就存储在这个虚拟的空间中。如果一个磁盘出现故障，工程师可以在磁盘阵列中直接对其进行替换，这个磁盘中的数据也将被自动恢复。目前的磁盘阵列一般都使用光纤通道接口，为了保证磁盘阵列的访问路径是安全的，在一台计算机中一般有多个光纤卡，分别连接到不同的光纤交换机上，通过多条路径连接到磁盘阵列上，这样可以防止访问路径出现故障。在磁盘阵列中，对于同时出现故障的磁盘数量是有限制的。为了防止多个磁盘同时出现故障而导致数据无法恢复，一般都有一块或多块磁盘作为Hot Spare，这样的磁盘平常是闲置的，当其他存储数据的磁盘出现故障时，作为Hot Spare的磁盘将立即接替出现故障的磁盘。

在软件方面，操作系统一般都有以下特点：

1）稳定的文件系统

文件系统的稳定，对整个计算机系统的稳定起着至关重要的作用。目前各种操作系统一般都通过日志机制来保证文件系统的稳定性，如AIX的JFS/JFS2文件系统、HP-UX的VxFS文件系统、Solaris的ZFS文件系统、Linux的ext3文件系统等。

2）可动态修改的内核

许多操作系统的内核可以进行动态配置，配置结果可以立即生效，不用重新启动计算机系统。

3）应用程序监视

对于计算机中重要的应用程序，应该保证它时刻都在运行。在系统中应该有另外一个程序，对重要的应用程序的运行进行监视，如果发现它的运行意外中止，应该立即重新启动它。

4）数据的备份

备份无疑是保证数据安全的一种重要措施，当因为系统出现故障而导致数据丢失时，利用备份可以对数据进行恢复。

1.3　什么是集群

如果一台计算机出现故障，那么在这台计算机上运行的应用程序将停止运行。在一个业务

系统中，对于重要的计算机，也应该有一台或多台对等计算机作为后备，一旦运行重要应用程序的计算机出现故障，作为后备的计算机立即接替它的工作，应用程序将切换到后备计算机上继续运行，这样可以保证整个业务系统的高可用性。实现这种应用程序切换的解决方案就是集群。这里所说的应用程序，就是指运行在一台计算机上的、向用户提供某种特定服务的应用软件，如Web服务器、邮件服务器、数据库服务器、中间件服务器等。

集群是由一些特定的硬件和软件组成的应用系统，用来提供应用程序的高可用性。多台物理上相互独立的计算机通过网络连接在一起，每台计算机通过心跳信号探测其他计算机的状态，一旦一台计算机出现故障，在这台计算机上运行的应用程序将立即切换到另外一台计算机上。为了建立集群，需要配置以下硬件设备：

1）节点

节点就是物理上独立的计算机，在一个集群中至少需要两个节点，这些节点的硬件配置应该基本相当，在节点上需要运行相同的操作系统，安装相同的应用程序。

2）网络

节点之间至少需要通过两个网络相连，其中一个称为私有网络，另外一个称为公共网络。私有网络用于发送心跳信号，心跳信号是网络中一种特殊的数据包，每个节点通过心跳信号探测其他节点的状态。公共网络一方面用于发送心跳信号，另一方面用于提供用户的访问。

3）共享的存储设备

共享的存储设备用于存储应用程序的数据，所有节点必须能够同时访问这种设备。应用程序在其中一个节点上运行，一旦这个节点出现故障，应用程序将切换到另外一个节点上。无论应用程序在哪个节点上运行，它都能够访问共享存储设备中的数据。

集群的硬件连接情况如图1.1所示。

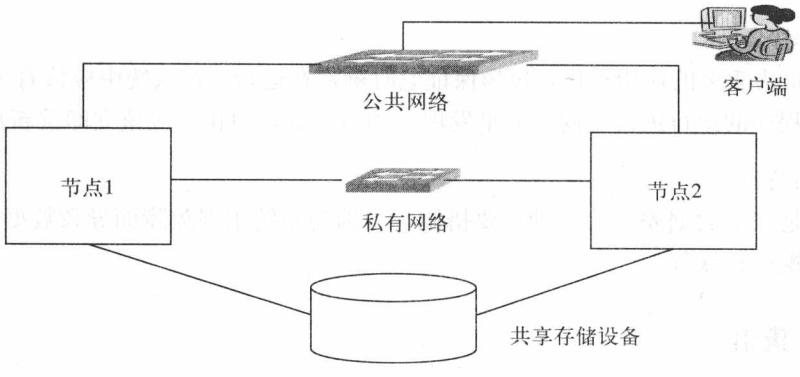


图1.1　集群的硬件连接

在每个节点上需要运行相同的操作系统，安装相同的集群管理软件和应用程序。当安装集群管理软件后，.在每个节点上将运行一个称为“集群管理器”的进程。集群管理器有两个基本功能：一是通过心跳信号监视每个节点的状态；二是进行资源组的切换，一旦发现某个节点出现故障，集群管理器将把这个节点上的资源组切换到另外一个节点上。资源组是资源的集合，每个资源组代表一个应用程序，在资源组中包含当前应用程序所需要的资源。资源组可以在集群范围内进行切换，这就意味着应用程序在不同节点之间的切换。

在集群中需要为每个应用程序定义一个资源组，资源组中需要包含以下资源：

1）IP地址

应用程序将监听这个IP地址，用户通过这个地址访问应用程序。这个IP地址是可以漂移的，如果当前节点出现故障，IP地址将随着资源组的切换而漂移到另外一个节点上。

2）应用程序的启动脚本和停止脚本

应用程序虽然安装在每个节点上，但是只在一个节点上运行，集群管理器通过启动脚本启动应用程序，通过停止脚本关闭应用程序。

3）存储设备

应用程序所需要的数据都存储在这个存储设备中，存储设备在应用程序所运行的节点上打开，如果这个节点出现故障，存储设备将随着资源组的切换而在另外一个节点上打开。在不同的操作系统中，对存储空间的称呼是不一样的，如在AIX和HP-UX系统中，这样的存储空间称为卷组，在Solaris中称为磁盘集。这里所说的存储设备，是指同时连接在所有节点上、所有节点都可访问的外部共享存储设备，如磁盘阵列。

为了保证应用程序的高可用性，在集群中需要为应用程序定义一个资源组，在资源组中至少需要包含三种资源：用于向用户提供访问的IP地址、应用程序的启动脚本和停止脚本、用于存储应用程序的数据的存储设备。当集群启动时，资源组在其中一个节点上打开，这就意味着应用程序在这个节点上运行。当这个节点由于故障而不可用时，资源组便切换到另外一个节点上，在另外一个节点上打开。资源组也可以在管理员的控制下手工切换到另外一个节点上。

资源组的切换情况如图1.2所示。假设某个应用程序先在节点1上运行，它所对应的资源组在节点1上打开。如果节点1出现故障，资源组便切换到节点2上，应用程序所用的IP地址也将漂移到节点2上，应用程序的数据所在的存储设备在节点2上打开，应用程序的启动脚本在节点上也将自动执行，于是这个应用程序便在节点上运行。当节点1的故障解决后，应用程序所对应的资源组可自动切换到节点1上，也可以继续在节点2上保持打开状态。如果资源组继续在节点2上处于打开状态，那么当这个节点出现故障时，资源组便切换到节点1上。

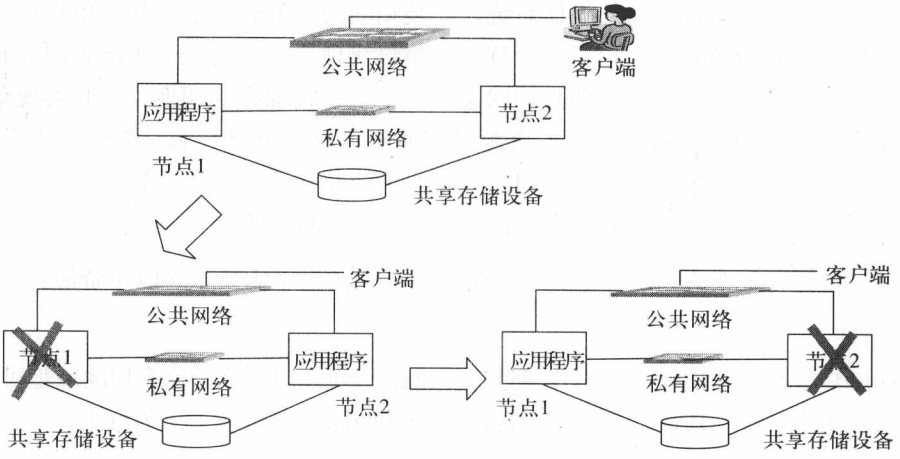


图1.2 集群的资源组切换

在图1.2中，资源组在其中一个节点上打开，其他节点都是空闲的。如果在每个节点上都运行一个应用程序，就可以充分利用这些节点的处理能力。在集群中运行多个应用程序的情况如图1.3所示：在集群中为每个应用程序分别定义一个资源组，每个资源组在不同的节点上打开。如果节点1出现故障，那么这个节点上的资源组将切换到节点2上，这时在节点2上将打开两个资源组。尽管节点2的负载大大加重，但是两个应用程序对用户都是可以访问的。当节点1的故障解决后，以前在这个节点上打开的资源组便可再次切换到这个节点上。

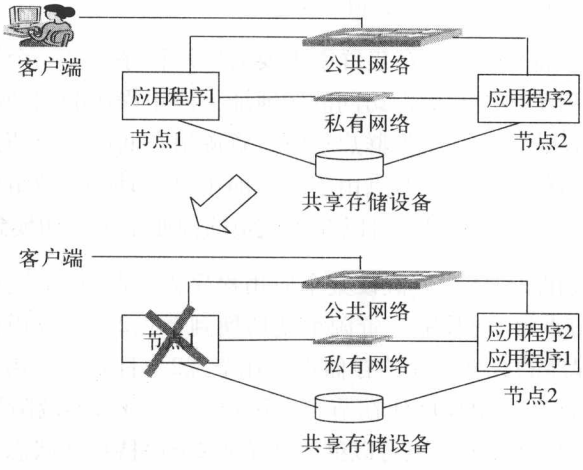


图1.3 两个资源组的切换情况

目前主流的集群管理软件主要有以下几种：

- IBM公司的HACMP
- HP公司的ServiceGuard
- SUN公司的Sun Cluster
- Oracle公司的RAC

其中前三种产品只能安装在各个公司自己的操作系统中。它们的工作原理基本上是相同的。它们的基本功能就是通过资源组切换的方式来保证应用程序的高可用性。RAC可以安装在适合Oracle数据库运行的任何操作系统中，它不仅可以实现资源的切换，而且可以向用户提供并发访问，用户可以通过任何一个节点访问数据库中的数据。

集群是一个完整的应用环境。在配置集群时，需要在每个节点中安装集群管理软件，还需要在集群中安装应用程序，如Weblogic、Websphere、Oracle等。上述前三种软件仅仅是集群管理软件，而RAC本身就可以提供一个完整的应用环境，其中集群管理器的功能由Clusterware来提供，而Oracle数据库就是运行在这个集群管理器之上的应用程序。

1.4　Oracle的高可用性产品

Oracle公司不仅向用户提供了一个优秀的数据库产品，而且提供了许多用于保证数据库高可用性的产品。具体来说，Oracle通过以下组件或产品来保证数据库的高可用性：

1）Oracle RAC

Oracle RAC是所有集群管理软件中功能最复杂、最难配置的一种。本书的主要内容就是介绍RAC的结构、功能和详细的配置过程。

2）Oracle Restart

Restart是从Oracle 11.2开始出现的一个组件，它的功能是在硬件或软件故障之后，或者在重新启动系统时，自动启动数据库、监听器以及其他相关组件，并且保证这些组件之间的正确启动顺序。Restart只能工作在单实例数据库中，在RAC环境中，Restart所具有的功能是由Clusterware来实现的。

3）Oracle Data Guard

Data Guard的功能是实现Oracle数据库的异地容灾，防止数据库所运行的站点发生灾难性的故障，如业务系统所在的机房出现火灾、水灾，或者遭到人为破坏。两个相互独立的数据库通过网络保持同步：其中一个数据库向用户提供正常的访问，这个数据库称为主数据库（Primary）；另一个数据库保持闲置，或者向用户提供少量的只读访问，如产生报表、对数据库进行备份等，这个数据库称为Standby。当用户在主数据库中执行DDL或DML语句修改数据

时，数据库服务器将产生重做日志，Data Guard通过网络将重做日志传输到Standby数据库，Standby数据库服务器通过重做日志产生同样的数据。当主数据库出现故障时，用户的访问将切换到Standby数据库。这两个数据库一般分开放在两个不同的城市，甚至两个不同的国家。

4）Advanced Replication

高级复制（Advanced Replication）主要用在分布式数据库中。多个相互独立的计算机可以同时向用户提供访问，用户一般就近访问其中一个数据库。如果用户修改了一个数据库中的数据，这些修改将被复制到其他数据库中。

5）Oracle Streams

Streams是一种特殊的高级复制技术，它提供了功能更强大、更加灵活的复制功能。

6）Oracle Flashback

利用Flashback技术，我们可以查询一个表在过去某个时刻的数据，或者在过去一段时间内在一个表上所发生的事务，可以把一个表或者整个数据库恢复到过去某个时刻，还可以还原对表的DROP操作。

7）Oracle ASM

ASM（Automatic Storage Management）是Oracle强力推荐的存储管理技术，它是Oracle公司提供的逻辑卷管理器，这种技术目前主要用在RAC环境中。

8）Recovery Manager

RMAN（Recovery Manager）用于对数据库进行备份和恢复，它是一种强有力的备份与恢复工具，利用这个工具，可以对整个数据库、一个表空间或者一个数据文件进行完全备份和增量备份。

9）LogMiner

LogMiner的功能是对重做日志文件进行分析，将其中存储的重做日志还原为文本格式。通过分析重做日志，可以确定一条DML或DDL语句的精确执行时间，可以跟踪用户在某个特定表上的DML或DDL操作，也可以获得DML语句所对应的反操作，通过这样的反操作，可以取消用户所执行的DML语句。

第 2 章　Oracle RAC 11g的体系结构

从第1章的描述我们已经知道，集群是一种应用环境。在集群中安装重要的应用程序，通过集群可以保证应用程序的高可用性。集群管理器的基本功能是：通过心跳型号监视每个节点的状态，如果发现节点出现故障，便进行资源组的切换，使应用程序在其他没有出现故障的节点上继续运行。

RAC也是一种集群环境，这个应用环境包括三部分：Clusterware、ASM和Oracle数据库。其中Clusterware的功能就是通过心跳信号监视节点的状态，并进行资源的切换。ASM的功能是对磁盘进行管理。Oracle数据库就是运行在这个集群环境中的应用程序，RAC的主要功能就是保证Oracle数据库的高可用性。

本章的主要内容就是介绍RAC这种特殊集群环境的结构。

2.1　Oracle RAC 11g的新特性

与以前的版本相比，Oracle RAC 11g有较大的变化，特别是在RAC 11.2中。其中最大的变化是ASM存储技术。在RAC 11.2中，ASM软件是和Clusterware软件一起被安装的，在安装这两种软件的同时，就可以创建ASM实例和ASM磁盘组，OCR和Voting文件可以存储在ASM磁盘组中。下面列出了在Oracle RAC 11g中出现的一些主要的新特性：

- Clusterware和Oracle数据库软件由两个用户分别安装，ASM实例和数据库实例也由两个用户分别进行管理。
- 增加了SYSASM权限。具有这个权限的用户就是ASM实例的管理员，这个用户就能够以“AS SYSASM”的方式登录ASM实例。
- 在RAC 11.1中，OCR和Voting文件可以存储在磁盘裸设备和磁盘块设备中。在RAC 11.2中RAC中，OCR和Voting文件可以存储在ASM磁盘组中。
- 从RAC 11.2开始，Clusterware软件和ASM软件合在一起组成了Grid Infrastructure软件。由于Grid Infrastructure软件是先于Oracle数据库软件被安装的，所以OCR和Voting文件可以存储在ASM磁盘组中。
- 从RAC 11.2开始，完全取消了对裸设备的支持。
- 从RAC 11.2开始，增加了集群时间同步服务（CTSS），利用这个服务，可以对多个节点之间的时间进行同步。

- 从RAC 11.2开始，客户端应用程序既可以通过VIP，也可以通过SCAN地址连接数据库实例。Oracle建议使用SCAN。
- 从RAC 11.2开始，在ASM磁盘组中可以创建卷和ACFS文件系统。在安装Oracle数据库软件时，可以将ACFS文件系统指定为软件的安装路径。

2.2 RAC集群的体系结构

RAC是一个完整的集群应用环境，它不仅实现了集群的功能，而且提供了运行在集群之上的应用程序，即Oracle数据库。无论与普通的集群相比，还是与普通的Oracle数据库相比，RAC都有一些独特之处。

RAC由至少两个节点组成，节点之间通过公共网络和私有网络连接，其中私有网络的功能是实现节点之间的通信，而公共网络的功能是提供用户的访问。在每个节点上分别运行一个Oracle数据库实例和一个监听器，分别监听一个IP地址上的用户请求，这个地址称为VIP（Virtual IP）。用户可以向任何一个VIP所在的数据库服务器发出请求，通过任何一个数据库实例访问数据库。Clusterware负责监视每个节点的状态，如果发现某个节点出现故障，便把这个节点上的数据库实例和它所对应的VIP以及其他资源切换到另外一个节点上，这样可以保证用户仍然可通过这个VIP访问数据库。

在普通的Oracle数据库中，一个数据库实例只能访问一个数据库，而一个数据库只能被一个数据库实例打开。在RAC环境中，多个数据库实例同时访问同一个数据库，每个数据库实例分别在不同的节点上运行，而数据库存放在共享的存储设备上。

通过RAC，不仅可以实现数据库的并发访问，而且可以实现用户访问的负载均衡。用户可以通过任何一个数据库实例访问数据库，实例之间通过内部通信来保证事务的一致性。例如，当用户在一个实例修改数据时，需要对数据加锁。当另一个用户在其他实例中修改同样的数据时，便需要等待锁的释放。当前一个用户提交事务时，后一个用户立即可以得到修改之后的数据。

RAC集群环境的基本结构如图2.1所示。

在创建RAC集群时，一般来说，Clusterware软件和Oracle数据库软件安装在每个节点的本地文件系统中，而那些要被所有节点访问的文件则存放在共享的存储设备中。在安装Clusterware软件时，需要在共享存储设备中创建OCR和Voting文件。其中，在OCR文件中记录RAC集群的配置信息，在Voting文件记录每个节点的成员资格信息。每个节点中的RAC集群在启动时，都需要读这两个文件，以确定当前节点的成员资格，并获得整个集群的配置信息。在创建数据库时，数据库文件、重做日志文件、控制文件、参数文件也存放在共享的存储设备中。