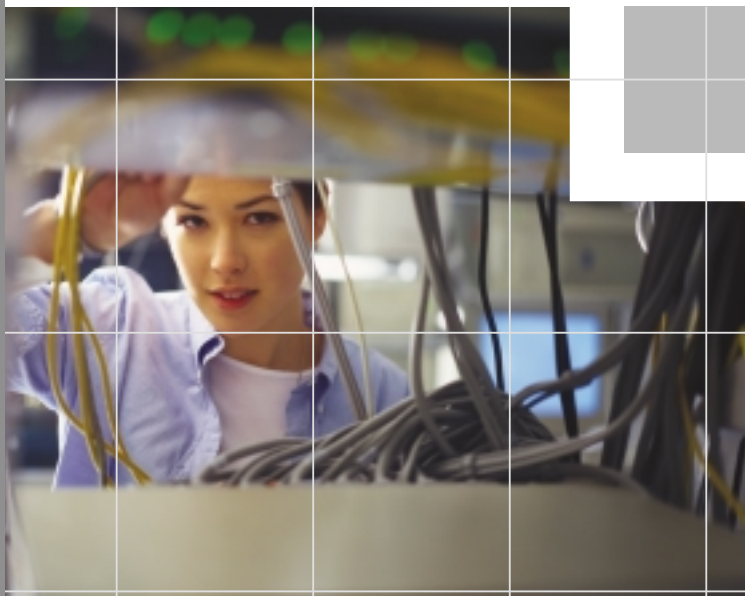


# КЛАСТЕРНЫЕ ТЕХНОЛОГИИ СУБД ORACLE

**ORACLE®**

**REAL APPLICATION  
CLUSTERS**



*Real Application Clusters  
Oracle Clusterware  
Automatic Storage Management*



**ORACLE IS THE INFORMATION COMPANY**

## Содержание

Введение .....	4
Real Application Clusters СУБД Oracle 10g .....	5
Архитектура Real Applications Clusters .....	5
Аппаратная архитектура .....	5
Oracle Clusterware .....	6
Файловые системы и управление дисковым пространством .....	7
Виртуальный IP адрес (VIP) .....	7
Сервисы (Services) .....	7
Быстрое оповещение приложений – Fast Application Notification (FAN) .....	8
Высокая доступность .....	8
Обеспечение высокой доступности для приложений .....	8
Бесперебойность при выполнении сервисного обслуживания .....	9
Масштабируемость .....	9
Балансировка рабочей нагрузки .....	10
Балансировка соединений .....	10
Рассылка рекомендаций по балансировке нагрузки .....	10
Динамически масштабируемое параллельное выполнение .....	11
Технология RAC в территориально-разнесенном кластере .....	12
Создание и управление кластерной средой Oracle .....	12
Утилита проверки кластера .....	13
Oracle Enterprise Manager 10g .....	13
Oracle Clusterware: Защита других приложений .....	14
Automatic Storage Management .....	15
Консолидация дисковых устройств хранения .....	16
Заключение .....	17
Приложение: Центры компетенции Oracle. ....	17

## Введение

Кластерные системы традиционно рассматриваются в качестве альтернативы большим компьютерам, созданным на основе архитектуры симметричной мультипроцессорной обработки (SMP). Преимущества кластеров над большими ЭВМ происходят из самой идеи объединения вычислительных мощностей множества независимых компьютеров в единую систему для решения одной задачи.

Даже самые большие вычислительные машины имеют ограничения на максимальное количество процессоров, объем памяти, карт ввода/вывода и т.д. Кластеры позволяют преодолеть физические ограничения одиночных компьютеров, используя эти компьютеры в качестве «кирпичиков» для построения системы требуемой вычислительной мощности.

Известно, что стоимость SMP-систем в перерасчете на один процессор стремительно растет при переходе от компьютеров одного класса к классу более мощных вычислительных машин. Объединение недорогих стандартных 2-х или 4-х процессорных компьютеров в кластерную систему может дать ощутимый экономический эффект, если сравнить ее стоимость с большой SMP-системой, которую можно применить для решения той же задачи. В этом случае аппаратная часть кластерной конфигурации может оказаться в несколько раз дешевле эквивалентной по мощности большой SMP-системы. Не стоит забывать и о немалых ежегодных отчислениях на контракты-поддержки, которые, как правило, требуются для минимизации простоев больших SMP-машин.

В информационной системе работают множество компонент, самыми важными из которых являются аппаратная часть, операционная система, системное и прикладное программное обеспечение. Сбой любой такой компоненты в отдельно взятой SMP-системе может привести к остановке обслуживания. Каждый узел кластера, как правило, может выполнять любые задачи выполняемые любым другим узлом, – такая взаимозаменяемость обеспечивает защиту от системных, аппаратных и программных сбоев.

Продукт **Oracle Real Application Clusters (RAC)** позволяет системе управления базами данных (СУБД) Oracle выполнять любые «коробочные», самостоятельно разработанные или заказанные приложения на кластере независимых серверов. Это обеспечивает для приложений высочайшие уровни доступности и масштабируемости. Выход из строя, какого либо из серверов не приводит к остановке СУБД Oracle – работа будет продолжена на оставшихся узлах. Достижение более высокой вычислительной мощности осуществляется простым добавлением в кластер требуемого количества серверов «на лету» – это не прерывает обслуживания пользователей.

Технология Oracle RAC позволяет системам, построенным с использованием недорогой аппаратной платформы, предоставлять высочайшее качество сервиса, сравнимое и даже превосходящее уровни доступности и масштабируемости самых дорогих SMP-систем и мэйнфреймов. Существенно сокращая расходы на обслуживание, и обеспечивая новые гибкие методы администрирования, программное обеспечение Oracle может быть использовано для создания среды Grid-вычислений предприятия.

## Real Application Clusters СУБД Oracle 10g

Oracle Real Application Clusters является опцией базы данных Oracle. Технология Oracle RAC была впервые представлена в Oracle 9i и на данный момент является уже проверенным решением, используемым тысячами заказчиков по всему миру во всех отраслях и для любых типов приложений.

Компонент Real Application Clusters расширяет СУБД Oracle, привнося в нее дополнительные возможности к уже встроенным механизмам управляемости, безопасности и надежности.

Oracle Real Application Clusters ключевой компонент Архитектуры Максимальной Доступности (Maximum Availability Architecture) компании Oracle, являющейся руководством по обеспечению высочайшей доступности приложений. При использовании Oracle RAC отдельный сервер больше не является компонентой, выход из строя которой, приводит к выходу из строя всей системы.

### Архитектура Real Applications Clusters

В «обычной» некластерной конфигурации к базе данных эксклюзивно имеет доступ один *экземпляр* программного обеспечения СУБД Oracle.

С кластерной базой данных, расположенной на общих дисковых устройствах, одновременно могут работать множество экземпляров СУБД Oracle, запущенных на различных узлах кластера. При увеличении вычислительных потребностей в кластер без остановки его работы можно добавить дополнительные узлы и экземпляры. Новые ресурсы могут быть задействованы в работу сразу после подключения.

Для объединения аппаратных компонентов в единую вычислительную систему необходимо кластерное программное обеспечение, управляющее членством узлов в кластере, осуществляющее мониторинг состояния и управление различных составляющих и служб кластера, предоставляющее механизм для взаимодействия между приложениями работающими на разных узлах и другие важные базовые функции.

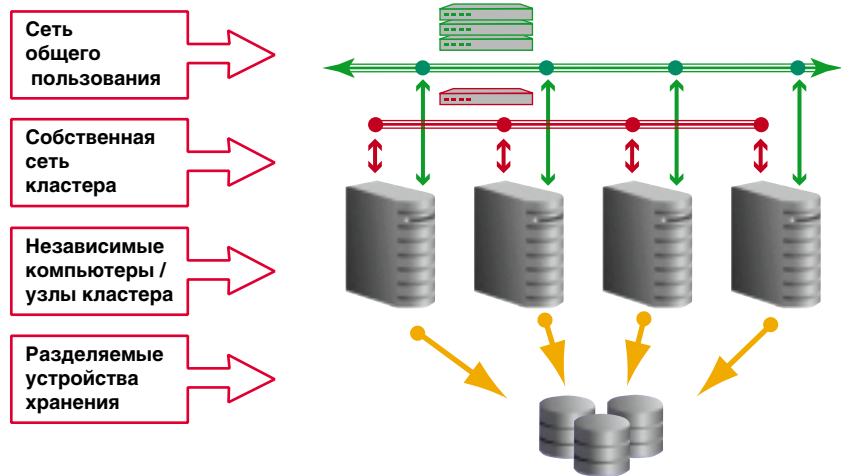
#### Аппаратная архитектура

Каждый узел кластера под управлением Oracle Clusterware должен иметь подключение к локальной сети, интерфейс для кластерного *межсоединения* и доступ к общему дисковому устройству хранения. Oracle Real Application Clusters строится по архитектуре с разделяемыми дисками. Все сервера кластера должны иметь полный и равноправный доступ ко всем устройствам хранения, на которых размещается кластерная база данных Oracle. В качестве таких дисковых устройств могут быть использованы дисковые устройства, подключаемые через сеть (NAS), специализированные сети устройств хранения (SAN) или SCSI-диски. На выбор типа устройства хранения часто влияет размер кластера, используемая платформа сервера и поддержка производителя аппаратного обеспечения. Устройство хранения должно обеспечить достаточную производительность операций ввода-вывода приложений и быть хорошо масштабируемым при добавлении в кластер дополнительных узлов.

Для работы кластера необходимо две изолированных друг от друга сети.

*Публичная* сеть для связи между клиентами и серверами кластера. Именно с использованием этой сети производится подключение клиентских сессий к базе данных, их балансировка между узлами и аварийное переключение в случае сбоя.

*Приватная* или *внутренняя* сеть, обычно называемая *межсоединением*, для передачи сообщений между узлами. Так, например межсоединение в RAC используется для реализации технологии «слияния» кэш различных узлов кластера – *Cache Fusion*. В большинстве случаев для обеспечения межсоединений в кластере вполне достаточно использование протокола UDP через Gigabit Ethernet. Дублирование сетевых интерфейсов увеличит надежность кластерной конфигурации.

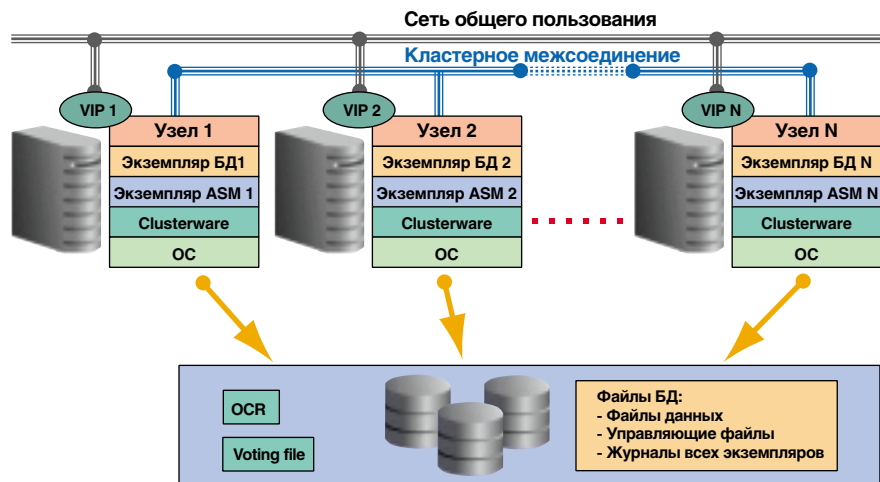


СУБД Oracle 10g Release 2, Oracle Clusterware и Real Application Clusters поддерживают до 100 узлов в кластере. Узлы кластера могут быть неидентичными, но должны иметь одинаковую аппаратную архитектуру, операционную систему и версию Oracle Clusterware.

Информация по сертификации использования Oracle Real Application Clusters может быть получена на Oracle Metalink (<http://metalink.oracle.com>) или Oracle Technology Network (<http://www.oracle.com/technology/support/metalink/index.html>).

### Oracle Clusterware

Начиная с СУБД Oracle версии 10g компания Oracle включило в комплект дистрибутива, специально разработанное и интегрированное с СУБД Oracle, программное обеспечение для кластеризации серверов – Oracle Clusterware. Теперь для того, чтобы иметь базу данных под управлением RAC, нет больше необходимости приобретать кластерное программное обеспечение от других производителей. Программное обеспечение Oracle Clusterware интегрировано, с хорошо знакомой каждому администратору базы данных, универсальной программой инсталляции Oracle (Oracle Universal Installer) и просто в установке. Тот факт, что у кластерного программного обеспечения и СУБД единый разработчик, упрощает настройку и техническую поддержку.



ПО Oracle Clusterware **необходимо** для работы Oracle RAC. Кластерное программное обеспечение является полнофункциональным и полностью самостоятельным продуктом – для своей работы Oracle Clusterware *не требует кластерного программного обеспечения других производителей*, но при необходимости может быть в него интегрировано для совместной работы.

Oracle Clusterware производит мониторинг и управление кластерными базами данных и другими программными компонентами, обеспечивающими их функциональность. При старте узла кластера Oracle Clusterware автоматически производит на нем старт и мониторинг всех экземпляров СУБД Oracle, прослушивающих процессов (listeners) и служб. В случае сбоя любой компоненты, Oracle Clusterware автоматически производит ее перезапуск, так, что обслуживание может быть восстановлено раньше, чем администратор заметит, что был сбой.

ПО Oracle Clusterware использует совместные дисковые устройства для хранения файла Oracle Cluster Registry (OCR), в котором описана конфигурация всех компонент кластера, и Voting File – файла используемого для голосования на случай физического разделения кластера на части в результате аппаратного сбоя. Кроме этого Voting File, так же как и кластерное межсоединение используется для прослушивания узлами «пульсов» других узлов входящих в кластер.

Oracle Clusterware 10g Release 2 предлагает программный интерфейс (API) для обеспечения высокой готовности процессов не только Oracle, но и других приложений. Когда администратор регистрирует такой процесс в Oracle Clusterware, он определяет правила старта, остановки и мониторинга процесса. Также должны быть определены правила перемещения процесса на другой узел кластера в случае сбоя того узла, на котором он выполнялся.

### Файловые системы и управление дисковым пространством

Oracle RAC строится на основе архитектуры с разделяемыми дисками, поэтому механизмы управления дисковым пространством и файловой системой в ОС на всех узлах должны поддерживать работу в кластере. Рекомендуется использовать встроенную в СУБД 10g Oracle систему автоматического управления дисковыми ресурсами для базы данных Oracle – Automatic Storage Management (ASM). ASM обеспечивает высокопроизводительные операции дискового ввода-вывода и простоту в управлении файловой системой и дисками. ASM автоматически производит оптимальное распределение данных между всеми дисковыми ресурсами для достижения наилучшей производительности, что исключает необходимость ручной настройки дискового ввода/вывода.

В качестве альтернативы Oracle поддерживает использование неформатированных дисковых разделов (сырых устройств) и некоторых кластерных файловых систем, например, таких как Oracle Cluster File System, доступная в операционных системах Windows и Linux.

### Виртуальный IP адрес (VIP)

В Oracle RAC 10g для каждого сервера в кластере требуется виртуальный IP-адрес. Виртуальный IP-адрес – это IP-адрес, который не привязан «жестко» ни к одному интерфейсу публичной сети кластера, но входит в адресное пространство той же подсети, что и статичные адреса этих интерфейсов. Этот адрес используется приложениями для подключения к кластерной базе данных. Если в узле происходит сбой, то его виртуальный IP-адрес перебрасывается на другой работающий в кластере сервер. Это предотвращает длительные ожидания окончания IP-таймаутов и позволяет быстро идентифицировать недоступность узла. В конечном итоге это увеличивает скорость переключения приложения на другой узел, повышая его доступность.

### Сервисы (Services)

Понятие *Сервис (Service)* в RAC Oracle Database 10g является фундаментальной основой обеспечения высокой доступности приложений и балансировки их нагрузки. *Сервис* позволяет классифицировать и объединять рабочие нагрузки приложений группированием на уровне сессий. К примеру, под *Сервисом* может подразумеваться группа сессий, выполняющих одно приложение или лишь какую либо его часть.

С помощью *Сервисов* группам сессий выделяются вычислительные ресурсы кластера. Администраторы баз данных определяют правила обеспечения вычислительными ресурсами каждого *Сервиса* – как для нормальной работы, так и для случая сбоя. Рабочая нагрузка создаваемая сессиями одного *Сервиса*,

распределяется и балансируется между узлами, которые назначены для обслуживания этого *Сервиса*. Интеграция с диспетчером ресурсов Database Resource Manager позволяет производить распределение ресурсов между сессиями различных *Сервисов* внутри каждого узла.

Производительность и качество выполнения каждого «*Сервиса*» отслеживается в автоматизированном репозитории данных о рабочих нагрузках Automatic Workload Repository, входящем в состав Oracle Database 10g. Контроль за качеством выполнения «*Сервиса*» может осуществляться установкой пороговых показателей производительности, в случае нарушений которых производится автоматическое оповещение или выполнение заданных процедур. По каждому *Сервису* собирается детальная статистика его выполнения. *Сервисы* интегрируются с механизмом Oracle Streams Advanced Queuing и системой выполнения фоновых пакетных задач.

Каждый *Сервис* может выполняться на одном или нескольких экземплярах СУБД Oracle, а каждый экземпляр может поддерживать несколько *Сервисов*. Количество и состав экземпляров, которые обслуживает тот или иной *Сервис* определяются администратором базы данных независимо от самого приложения. При возникновении сбоя выполнение *Сервисов* автоматически восстанавливается на работающих экземплярах.

### Быстрое оповещение приложений – Fast Application Notification (FAN)

Использование приложением новой возможности Oracle RAC 10g механизма быстрого оповещения приложения Fast Application Notification (FAN) позволяет ускорить реакцию приложения на сбой внутри кластера и улучшить качество распределения рабочей нагрузки между доступными вычислительными ресурсами. FAN обеспечивает интеграцию между базой данных RAC и приложением. Благодаря этому механизму приложение обладает информацией о конфигурации кластера в любой момент времени, поэтому приложения подключаются только к тем экземплярам, которые в текущий момент отвечают на запросы приложений. Инфраструктура высокой доступности Oracle RAC 10g немедленно передает событие FAN при изменении состояния кластера.

Клиентское приложение получает возможность быстрой реакции на события, происходящие в кластере. В качестве реакции на стороне сервера настраиваются вызовы внешних программ, которые можно, например, использовать для протоколирования проблем или уведомления администраторов о сбое.

Клиенты Oracle JDBC, ODP.NET и OCI интегрированы с FAN. Другие приложения могут использовать возможности FAN при помощи программируемого интерфейса приложений и прямой подписки на события FAN.

## Высокая доступность

СУБД Oracle известна своей надежностью. Real Application Clusters обеспечивает для СУБД Oracle еще большую доступность. RAC исключает ситуацию, когда выход из строя сервера или его программного обеспечения приводит к выходу из строя всей системы в целом. При сбое в узле или экземпляре СУБД база данных остается открытой, и приложения продолжают иметь доступ к данным через другие работающие узлы и экземпляры.

Программное обеспечение Oracle Clusterware производит мониторинг кластерных баз данных RAC и других компонентов кластера, быстро выявляет проблемы и само производит их разрешение. Кластерная СУБД Oracle автоматически определяет сбой любого экземпляра, после чего сразу же начинает процедуру его восстановления. Так, что сбой может быть устранен, до того как кто-либо заметит, что он имел место. Механизм быстрого оповещения приложений Fast Application Notification позволяет приложениям получать немедленное уведомление о неисправностях компонентов кластера и скрывать сбои от пользователя передачей операций другому работающему узлу кластера.



## Обеспечение высокой доступности для приложений

В базе данных Oracle содержится множество механизмов, облегчающих быстрое восстановление работы после всех видов сбоев. Такие механизмы, как Fast Application Notification, Fast Connection Failover и Transparent Application Failover, позволяют приложениям маскировать сбои компонентов кластера от пользователей.

Прозрачное переключение приложения Transparent Application Failover – механизм, который в случае сбоя экземпляра, обеспечивает автоматическое пересоздание соединения с другим экземпляром СУБД Oracle в кластере. Само переключение производится на уровне клиентского сетевого стека программного обеспечения Oracle прозрачно для приложения. При этом в случае, если в момент сбоя на экземпляре выполнялась транзакция, эта транзакция должна быть повторно выполнена инициировавшим ее приложением после переключения к другому экземпляру.

Для более тесной интеграции приложения с кластерной СУБД, приложение может подписаться на события FAN, что позволит мгновенно реагировать на любые изменения (в том числе планируемые) в конфигурации кластера, а не только на сбой, который уже произошел.

Автоматическая поддержка обработки событий FAN уже реализована для пулов соединений JDBC, ODP.NET, OCI – этот механизм называется Fast Connection Failover (FCF). Приложения, использующие такие пулы соединений, получают возможность реакции на события FAN без дополнительных усилий со стороны разработчиков.

## Беспереывность при выполнении сервисного обслуживания

Real Application Clusters обеспечивает непрерывное обслуживание, как при сбоях, так и при выполнении запланированных сервисных работ. Большинство сервисных операций с базами данных может быть выполнено без остановки в обслуживании и прозрачно для пользователя. Большинство других сервисных работ могут выполняться на узлах кластера поочередно, что приводит к значительной минимизации простоев.

В случае если обновление программного обеспечения Oracle сопровождается внесением изменений в общие структуры всех экземпляров кластера, например словаря базы данных, Oracle предлагает использовать логическую резервную базу данных Oracle Data Guard. Такой подход позволяет минимизировать остановку в обслуживании только на время переключения ролей между основной и резервной базами данных.

Процедура обновления включает в себя обновление логической резервной базы данных до следующего выпуска, запуск ее в смешанном режиме для проверки работоспособности обновления, изменение роли посредством переключения на обновленную базу данных, а затем обновление старой основной базы данных. При проведении тестирования в смешанном режиме возможна отмена установки обновления и восстановление старого программного обеспечения без потери данных. Для обеспечения защиты данных при этих процедурах можно использовать дополнительную резервную базу данных.

Благодаря возможности поочередных обновлений с минимальными задержками Data Guard уменьшает время обслуживания для большинства административных задач, тем самым обеспечивая непрерывную работу системы 24x7 (7 дней в неделю по 24 часа).

## Масштабируемость

Oracle Real Application Clusters обладает уникальной технологией масштабирования приложений. Типичный подход без использования кластеров заключается в замене сервера, которому уже не хватает мощности, новым сервером большего размера. Использование технологии RAC является альтернативой наращиванию вычислительной мощности одиночных серверов. Для того, что бы сохранить инвестиции, вложенные в уже имеющееся аппаратное обеспечение, можно просто добавить новый сервер к кластеру (или создать кластер). Приложения, обычно работающие на больших SMP-системах, могут быть переведены на кластеры, состоящие из небольших серверов.

Добавление новых серверов к кластеру, работающему под управлением Oracle Clusterware и RAC, не вызывает прерывания в обслуживании, а новые мощности можно использовать сразу после их подключения. Все серверы кластера должны работать в одинаковых операционных системах и иметь одинаковые версии Oracle, но могут различаться по вычислительной мощности. В настоящее время в мире эксплуатируются множество кластеров различной конфигурации, – начиная от кластеров, состоящих из недорогих серверов с 1-2 процессорами, и, заканчивая кластерами, количество процессоров, в узлах которых исчисляется многими десятками.

### Балансировка рабочей нагрузки

Для эффективного использования вычислительных ресурсов и обеспечения требуемого качества выполнения задач приложений, необходимо оптимальное распределение нагрузки между узлами кластера. Oracle Real Application Clusters 10g обладает инновационной технологией управления нагрузками, которая обеспечивает наилучшую производительность приложений и их высокую доступность на заданной конфигурации.

Пользователи приложений или серверы приложений промежуточного уровня соединяются с базой данных при помощи имени *Сервиса*. Oracle автоматически распределяет нагрузку пользователей среди множества узлов кластера обслуживающих запрошенный *Сервис*.

Выделяя вычислительные мощности для определенного *Сервиса*, администраторы баз данных имеют возможность определять, на каких узлах работают клиенты выполняющие тот или иной круг задач. Администраторы могут безболезненно добавлять вычислительные мощности при росте требований приложений.

### Балансировка соединений

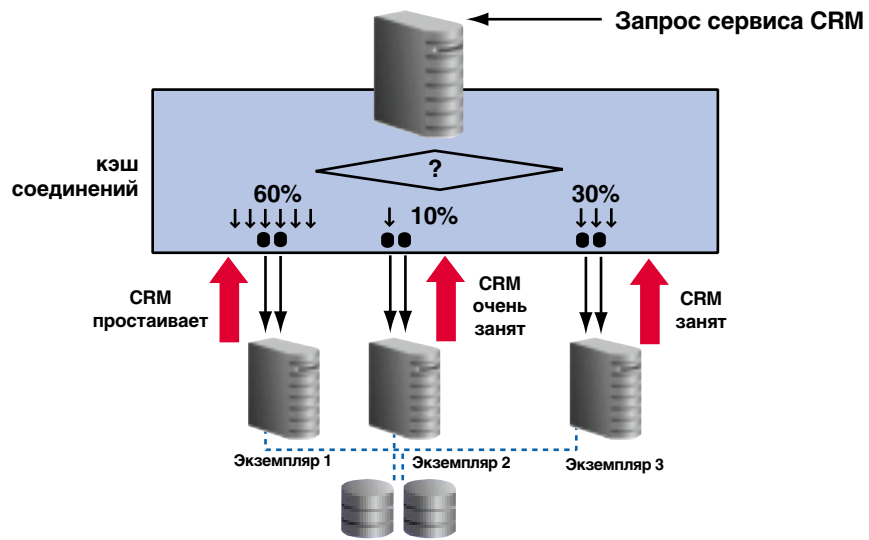
Сетевой стек программного обеспечения Oracle SQL\*NET, который обеспечивает взаимодействие между клиентом и сервером баз данных, обладает функцией балансировки нагрузки на уровне соединений. Балансировка нагрузки на уровне соединений осуществляется как на стороне клиента, так и на стороне сервера.

Для балансировки нагрузки на стороне клиента необходимо в описании подключения к сервису перечислить все узлы кластера. SQL\*NET случайно выбирает один из серверов. Если выбранный сервер недоступен, то производится попытка подключения к следующему серверу.

Балансировка нагрузки на стороне сервера производится прослушивающим процессом (LISTENER). Каждый прослушивающий процесс получает информацию от всех экземпляров кластера об обслуживаемых *Сервисах* и о качестве их выполнения. В зависимости от целей, определенной для конкретного *Сервиса*, прослушивающий процесс выбирает экземпляр, наиболее точно подходящий для данной задачи, и производит с ним соединение.

### Рассылка рекомендаций по балансировке нагрузки

Нагрузка сервера базы данных может изменяться с течением временем, также как может изменяться и конфигурация кластера, поэтому важно создавать соединения с базами данных и производить распределение работы внутри пулов соединений на основе самой оперативной информации. Oracle Real Application Clusters 10g Release 2 поддерживает механизм оповещений-рекомендаций по балансировке нагрузки Load Balancing Advisory. RAC постоянно производит мониторинг качества выполнения *Сервиса* для каждого экземпляра. Эта информация протоколируется в автоматическом архиве рабочей нагрузки (Automatic Workload Repository). На основе заархивированной информации производится анализ, результаты которого затем передаются приложениям при помощи событий FAN. В событиях FAN содержится текущее качество выполнения *Сервисов* и рекомендательная информация о том, какой процент заданий должен направляться на каждый экземпляр.



Интегрированные клиенты Oracle используют эти события для распределения нагрузки запросов от приложений. Без использования FAN большинство пулов соединений производят выбор соединения для выполнения запроса последовательно или случайным образом. При помощи событий FAN и рекомендациям по балансировке нагрузки пул соединений может выбирать соединение, обеспечивающее на данный момент наилучшее обслуживание. Пулы соединений Oracle JDBC и ODP.NET производят балансировку нагрузки с использованием FAN-рекомендаций.

### Динамически масштабируемое параллельное выполнение

Другим способом распределения нагрузки в базе данных Oracle является использование функции параллельного выполнения запросов. Параллельное выполнение (параллельный запрос или параллельный DML) распределяет работу по выполнению команды SQL между несколькими процессами. В среде Oracle Real Application Clusters эти процессы могут быть распределены между несколькими экземплярами. Оптимизатор по стоимости компании Oracle поддерживает функцию параллельного выполнения как фундаментальный компонент при оптимизации планов выполнения SQL-команд. В среде Real Application Clusters принятие решений о распараллеливании задачи зависит от возможностей внутриузлового и межузлового параллелизма. Например, если для конкретного запроса необходимо шесть процессов для его и шесть процессоров локального узла (узла, к которому подключен пользователь) свободны, то запрос будет обрабатываться только при помощи локальных ресурсов. Это демонстрирует эффективность внутриузлового параллелизма, при котором исключается распределение запроса между множеством узлов. Однако если в локальном узле доступны только два процессора, то для обработки запроса будут использоваться два процессора локального и четыре процессора другого узла. Таким образом, для ускорения операции с запросами используются как межузловой так и внутриузловой параллелизмы.

Оптимизатор Oracle способен изменять уровень запрашиваемого параллелизма в зависимости от загруженности системы. При недостаточном количестве свободных ресурсов, оптимизатор может понизить уровень параллелизма или даже вообще отменить распараллеливание выполнения. Это свойство является важным для сохранения стабильности работы системы, любая перегрузка системы может привести к резкой деградации производительности всех выполняемых задач.

## Технология RAC в территориально-разнесенном кластере

Территориально-разнесенный кластер – это архитектура, в которой узлы кластера находятся в разных Центрах Обработки Данных (ЦОД). Благодаря технологии Oracle RAC все узлы кластера, независимо в каком ЦОД они располагаются, включены в активную работу, как и в случае локального кластера. С этой точки зрения технология Oracle RAC является уникальной. Плюс к этому Oracle RAC в территориально-разнесенном кластере обеспечивает необычайно быстрое восстановление, в случае сбоя на какой-либо площадке. В связи с тем, что эта архитектура вызывает очень большой интерес и имеет успешные реализации, важно понять, где ее применение приносит наибольшую пользу с учётом расстояний, задержки в передаче сигнала и обеспечиваемую ей степень защиты.

Ограничением в применении этой технологии могут быть значительные задержки при передаче сигнала на больших расстояниях. Наилучшим образом эта архитектура может быть реализована, когда два центра данных находятся относительно близко друг другу (не более 100 км), а между центрами уже существуют достаточно дорогие прямые каналы связи без использования на их протяжении дополнительных ретранслирующих устройств.

Технология RAC для Extended Distance Clusters обеспечивает более высокую доступность, чем локальный RAC, но она может удовлетворять не всем требованиям мер по восстановлению в чрезвычайных ситуациях Вашего предприятия. Разнесение частей кластера на удаленные площадки является хорошей защитой от многих катастрофоподобных обстоятельств (локальное отключение электропитания, пожар, затопление помещения сервера, теракт), но не от всех. Такие стихийные бедствия, как землетрясения, ураганы или потопы, могут охватывать очень большую территорию. Компания должна произвести анализ и определить, могут ли оба центра обработки данных пострадать от одного и того же стихийного бедствия. Для всеобъемлющей защиты от различных угроз, включая защиту от повреждений данных и от региональных стихийных бедствий, специалисты компании Oracle рекомендуют использовать Oracle Data Guard совместно с технологией RAC, как описано в руководствах Oracle по архитектуре высокой доступности. Кроме этого Data Guard дает дополнительные преимущества, например поддержку поочередных обновлений версий Oracle.

Настроить территориально-разнесенный кластер сложнее, чем локальный. Особое внимание следует уделить месторасположению узлов, «кворуму» дисков и размещению устройств хранения с данными. При правильном применении эта архитектура может обеспечить более высокую доступность, чем база данных с локальным RAC. Для создания территориально-разнесенных кластеров можно использовать комбинацию программного обеспечения Oracle Clusterware, Oracle Real Application Clusters и Automatic Storage Management.

## Создание и управление кластерной средой Oracle

Программное обеспечение Oracle Real Application Clusters создает из кластерной конфигурации единый образ не только для клиентских приложений, но и для администратора, что упрощает настройку и управление. Кластерная база данных RAC может быть установлена и сконфигурирована с любого узла входящего в кластер. Работу с кластерной конфигурацией поддерживают все средства и программы для управления базой данных, включая универсальную программу инсталляции ПО Oracle (OUI), Oracle Enterprise Manager, программу-ассистента конфигурирования базы данных (DBCA), программу-ассистента по модернизации версии базы данных (DBUA), программу-ассистента по конфигурированию сети (NETCA) и утилиты командной строки, такие как srvctl.

## Утилита проверки кластера

В СУБД Oracle 10g Release 2 появилось новое средство проверки конфигурации кластера. Средство проверки кластера помогает устранять ошибки проведением ревизии до установки программного обеспечения, после его установки и при любых изменениях конфигурации. Оно также может использоваться для текущей проверки кластера. Это утилита может быть вызвана из интерфейса командной строки или через API при помощи других программ, например универсальной программы установки Oracle Universal Installer (OUI).

## Oracle Enterprise Manager 10g

Программный пакет Enterprise Manager традиционно предлагаются Oracle как средство управления инфраструктурой программного обеспечения Oracle. В новой версии Enterprise Manager 10g значительно переработан, так, например доступ к графическому интерфейсу управления теперь осуществляется через веб-браузер. Enterprise Manager 10g для управления СУБД Oracle предлагается в двух редакциях:

- ♦ *Database Control* – средство управления одной базой данных Oracle, автоматически настраивается при помощи DBCA во время создания базы данных.
- ♦ *Grid Control* – средство управления информационной инфраструктурой предприятия, включая базы данных Oracle, устанавливается с отдельного компакт диска, включенного в поставку СУБД Oracle.

Оба средства управления поддерживают работу с кластерами и обладают централизованной консолью для управления кластерной СУБД.

Страница управления кластерной базой данных «Cluster Database» позволяет:

- ♦ Следить за общим состоянием системы, например, за количеством и текущим состоянием всех экземпляров кластерной СУБД.
- ♦ Видеть предупреждения, полученные со всех экземпляров, а также углубляться в систему для детального изучения источника каждого предупреждения.
- ♦ Устанавливать пороги отслеживаемых показателей для генерации предупреждений на уровне кластерной базы данных.
- ♦ Выполнять мониторинг показателей производительности со всех экземпляров: агрегированных или отображенных друг с другом так, чтобы их можно было легко сравнить между собой. В случае необходимости можно произвести детализацию по отдельным компонентам.
- ♦ Контролировать статистику когерентности кэша кластера (например, *global buffer gets* и др.)
- ♦ Производить операции уровня всей кластерной базы данных: создание резервных копий и восстановление базы данных, запуск и остановка экземпляров и т. д.
- ♦ Управлять *Сервисами*, выполняя такие операции, как создание, изменение, запуск/остановка, включение/выключение или перемещение *Сервисов*, и производить мониторинг их производительности.

На странице «Cluster» пакета Grid Control можно просматривать состояние аппаратного обеспечения кластера и операционной системы в целом – это особенно полезно, когда кластер обеспечивает работу множество баз данных. Страница даёт обзор общего состояния компонент кластера с возможностью перехода к индивидуальным базам данных.

Grid Control 10g Release 2 может произвести автоматическое преобразование отдельного экземпляра базы данных Oracle в кластерную базу данных под управлением RAC. Кроме этого при помощи Enterprise Manager можно осуществить начальное создание кластера, включая установку программного обеспечения и конфигурацию Oracle Clusterware. Содержимое домашнего каталога программного обеспечения Oracle (Oracle Home), используемого для установки, может

храниться либо в Enterprise Manager как «Золотой образ» или на любом известном сервере. «Золотой образ» может быть создан из образцовой копии установленного программного обеспечения Oracle Clusterware или Real Application Clusters 10g Release 2. При клонировании программного обеспечения выполняются все без исключения действия по установке и конфигурации RAC и Oracle Clusterware, в том числе запуск «root.sh» под правами привилегированных пользователей и шаги выполняемые до и после инсталляции. Эта справедливо и при добавлении нового узла к уже существующему кластеру.

Для ОС Linux программное обеспечение от Oracle может установить «образ системы» на «голое железо». В такой «образ» может входить операционная система, агент Oracle Enterprise Manager, Oracle Clusterware и СУБД Oracle с Real Application Clusters, и его можно связать с профилем аппаратного обеспечения. Все компоненты «образа» хранятся как «Золотые образы» в Enterprise Manager. «Мастер по установке» позволяет производить выбор аппаратного обеспечения и устанавливать полный стек программного обеспечения на новый сервер. Новый узел автоматически добавляется к кластеру.

## Oracle Clusterware: Защита других приложений

Одной из самых значительных новинок Oracle Clusterware 10g Release 2 является возможность обеспечения для приложения среды так называемого «Failover» кластера или кластера, в котором конкретная задача работает только на одном узле и в случае его сбоя мигрирует на другой «живой» узел кластера. Это позволяет обеспечить высокую готовность не только приложений СУБД Oracle под управлением RAC, но и приложений созданных другими разработчиками, включая самостоятельно разработанные заказчиком приложения.

Включение приложения в среду высокой доступности Oracle Clusterware состоит из трёх этапов:

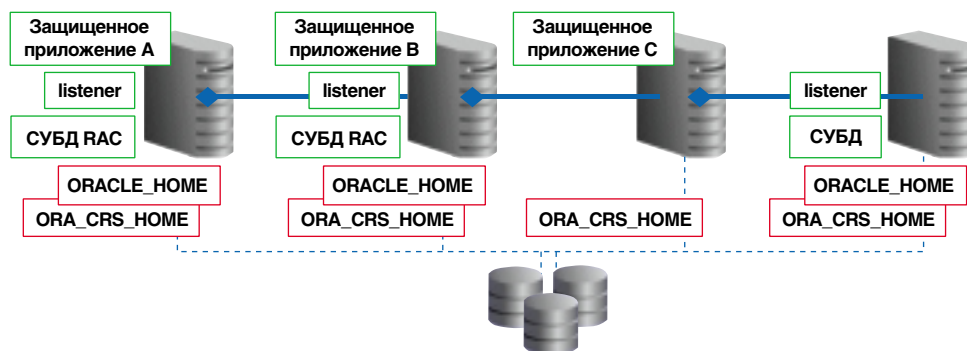
- ♦ Создание внешней программы для старта, мониторинга и остановки приложения, принимающей от Clusterware один из возможных аргументов – «start», «stop» или «check».
- ♦ Создание профиля приложения. В профиле кроме имени приложения и управляющей программы определяются узлы, на которых приложение может работать, и политика их выбора, зависимости приложения от других ресурсов кластера, в том числе VIP, временные параметры мониторинга, повторного старта и другие.
- ♦ Регистрация профиля приложения

Для большей интеграции приложений Oracle Clusterware предлагает программный интерфейс 'C'(API). С помощью этого интерфейса защищаемая программа во время работы может манипулировать содержимым Oracle Cluster Registry (OCR), и изменять поведение Oracle Clusterware по управлению приложением.

Даже без использования RAC кластерная конфигурация, построенная с помощью Oracle Clusterware, может обеспечить высокую готовность работающим на нем приложениям, в том числе некластерным экземплярам СУБД Oracle, позволить консолидацию дисковых систем и простой переход к RAC в дальнейшем, когда для баз данных Oracle потребуется еще большая доступность и гибкое масштабирование.

Oracle Clusterware является бесплатным программным обеспечением и лицензировано для использования на кластере при условии эксплуатации даже некластерной СУБД Oracle на любом из его узлов. Кластером считается группа компьютеров с выполняемым на них Oracle Clusterware и одним набором файлов OCR и Voting File.





Описание примера: В кластер с помощью Oracle Clusterware объединены четыре компьютера с одним набором файлов OCR и Voting File, расположенных на общих дисковых устройствах. На кластере под управлением Clusterware работают четыре приложения: приложение А на узле 1, приложение В на узле 2, приложение С на узле 3, некластерная СУБД Oracle работает на узле 4. Для всех четырех приложений любой другой узел может быть сконфигурирован в качестве резервного. Узлы 1 и 2 используются для работы кластерной СУБД Oracle.

## Automatic Storage Management

Механизм Automatic Storage Management (ASM) представленный в СУБД Oracle 10g объединяет в себе кластерную файловую систему и возможности менеджера томов. ASM входит в стандартный функционал СУБД Oracle и не требует дополнительного лицензирования, а значит является *бесплатным*. ASM сокращает стоимость владения системами хранения для файлов СУБД Oracle, автоматизируя множество дисковых операций. Механизм ASM производит балансировку распределения данных между дисковыми устройствами для оптимизации производительности и защищает данные при поддержке их избыточности.

Возможности ASM доступны как в одиночных экземплярах СУБД Oracle, так и в кластерных базах данных под управлением Oracle RAC. При этом ASM может использоваться по желанию, и его возможности могут быть применены в смешанных конфигурациях, когда одна часть файлов размещается на дисковых группах ASM, а другая на альтернативных файловых системах или на неформатированных разделах дисков.

ASM производит виртуализацию дисковых устройств – отдельные диски объединяются в дисковые группы, являющиеся единицами хранения файлов с точки зрения администратора базы данных и самой СУБД Oracle. Кроме сокращения количества единиц управления, СУБД Oracle может производить автоматическое именование файлов базы данных.

Механизм ASM производит оптимизацию распределения данных между дисковыми устройствами одной дисковой группы, используя технологию схожую с идеей чередования данных (striping), но по собственному алгоритму. Для этого ASM разбивает данные на экстенды размером в 1 мегабайт или 128 килобайт в зависимости от типа файла.

Механизм ASM позволяет изменять состав дисковых групп «на лету», без остановки доступа к данным на них расположенных. При добавлении или удалении дискового устройства из дисковой группы ASM производит автоматическую перебалансировку данных. Перебалансировка может осуществляться с разной степенью интенсивности, что позволяет избежать падения производительности производимых в этот момент операций ввода-вывода с базами данных.

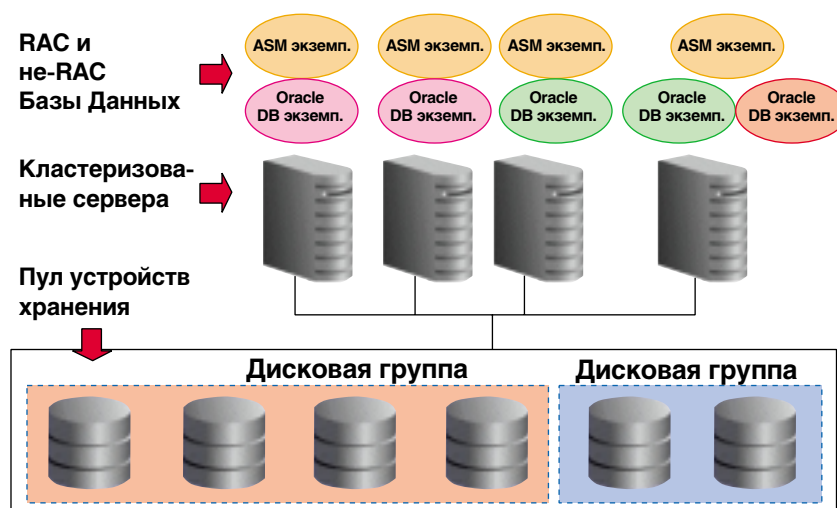
ASM поддерживает три режима избыточности данных:

- *External* – избыточность не поддерживается. Рекомендуется использовать при применении RAID массивов осуществляющих избыточность данных на аппаратном уровне;
- *Normal* – 2-х кратная избыточность. Поддерживаются две копии одного экстенда.
- *High* – 3-х кратная избыточность. Поддерживаются три копии одного экстенда.

Для защиты от сбоев аппаратных устройств обеспечивающих работу сразу множество дисков, в дисковых группах можно определить *failure* группы, при этом избыточность данных будет поддерживаться между дисками находящимися в различных *failure* группах. Это позволяет обеспечить «зеркалирование» данных между дисками находящимися под управлением разных контроллеров и даже между отдельными дисковыми массивами.

## Консолидация дисковых устройств хранения

Начиная с СУБД Oracle10g Release 2 программное обеспечение Oracle Clusterware не требует лицензии Oracle Real Application Clusters. Механизм ASM совместно с Oracle Clusterware и не кластерной СУБД Oracle может быть использован для консолидации всего дискового пространства предприятия, предоставляя единый кластерный пул устройств хранения для некластерных баз данных Oracle.



*Описание примера:* Для всех баз данных работающих на кластере, объединенном Oracle Clusterware, используется единый пул дисковых устройств. Файлы кластерной базы данных с экземплярами на узлах 1 и 2, кластерной базы данных с экземплярами на узлах 3 и 4 и не кластерной базы данных с экземпляром на узле 4 размещаются в общих дисковых группах под управлением кластерной конфигурации ASM.



## Заключение

Oracle Real Application Clusters предназначен для обеспечения высокой доступности и гибкой экономичной масштабируемости. Защищая от сбоев программной и аппаратной части, Oracle RAC обеспечивает непрерывный доступ к данным.

Технология RAC объединяет мощности нескольких компьютеров в единую вычислительную систему для решения задач по обработке и управлению информацией в базе данных. При этом Oracle RAC является *единственной жизнеспособной* альтернативой большим SMP-системам для приложений всех типов, включая хранилища данных и системы с массивной оперативной обработкой транзакций.

RAC и Grid-технологии дают возможность радикально снизить эксплуатационные затраты и обеспечить новый уровень гибкости, делая корпоративные системы более адаптивными и динамичными. Динамическое обеспечение узлами, устройствами хранения, центральными процессорами и оперативной памятью позволяет быстро и эффективно гарантировать необходимые уровни обслуживания при одновременном снижении затрат за счет лучшего использования ресурсов. Среда RAC полностью прозрачна для приложения, работающего с кластерной базой данных – его не требуется модифицировать.

Даже без использования RAC построение кластерной конфигурации на основе Oracle Clusterware может принести выгоды. Oracle Clusterware позволяет обеспечить высокую доступность для приложений разработанных другими производителями, а совместное его использование с Automatic Storage Management даёт возможность построить единый кластеризованный пул дисковых устройств предприятия.

**ORACLE®**

**PARTNERNETWORK**

ЦЕНТР КОМПЕТЕНЦИИ

## Приложение: Центры компетенции Oracle.

Успех сложных информационных проектов зависит от экспертизы и опыта специалистов компании – партнера Oracle. Для того чтобы помочь заказчику выбрать партнера, обладающего наибольшей экспертизой по тому или иному технологическому направлению Oracle создана программа Центров Компетенции. Наличие этого статуса подтверждает высокую квалификацию, экспертизу по направлению Центра Компетенции, надежной репутации и профессионализма партнера, проявленных в результате реализации проектов по внедрению решений Oracle. Он также подтверждает наличие высококвалифицированных специалистов, проводимую маркетинговую политику и активное продвижение новейших технологий Oracle.

## КОРПОРАЦИЯ ORACLE

Oracle Россия

Россия 119435, Москва

Саввинская набережная, 15

Тел.: +7 (495) 641 1400

Факс: +7 (495) 641 1414

E-mail: [oracle\\_ru@oracle.com](mailto:oracle_ru@oracle.com)

Internet: [www.oracle.com/ru/](http://www.oracle.com/ru/)

Oracle Украина

04070, Украина,

Киев, ул. Фроловская, 9 - 11

офисный центр «Swiss House»

Тел.: + 380 (44) 490 9050

+ 380 (44) 490 9051

Факс: + 380 (44) 490 9052

Oracle Казахстан

480099, Казахстан, Алматы,

микрорайон Самал2,

Самал Тауэрс, оф. 97, блок А2, бй этаж

Тел.: +7 (3272) 58 4748

+7 (3272) 58 4740

Факс: +7 (3272) 58 4744

Copyright © 2006 Oracle Corporation. Все права защищены.

Данный документ предоставлен исключительно в информационных целях и его содержание может быть изменено без уведомления. Этот документ не гарантирует отсутствие ошибок и не подразумевает никаких гарантий или условий, выраженных явно или подразумеваемых законом, включая косвенные гарантии и условия окупаемости или пригодности для решения конкретной задачи. Мы отказываемся от любой ответственности, связанной с этим документом, и никакие договорные обязательства не могут быть оформлены, прямо или косвенно, на основании данного документа. Этот документ не может быть воспроизведен или передан в любой форме и любыми средствами, электронными или механическими, для любых целей, без нашего письменного разрешения. Oracle, JD Edwards, PeopleSoft и Retek являются зарегистрированными товарными знаками корпорации Oracle и/или входящих в нее компаний. Другие наименования могут быть товарными знаками соответствующих владельцев.

Октябрь 2006

**ORACLE®**