Тенденции развития коммерческих СУБД

Московская секция ACM SIGMOD, 31 января 2008 г

Марк Ривкин
Российское представительство Oracle
Mark.Rivkin@oracle.com

ORACLE

План презентации

- Почему я не согласен с утверждениями М Стоунбрейкера (в изложении С Кузнецова) о грядущей революции в области СУБД
- Почему можно говорить об основных тенденциях развития коммерческих СУБД на ближайшие годы, опираясь на Oracle 11g
- 13 наиболее важных тенденций
- Вопросы и ответы

Стоунбрейкер предсказывает:

- Универсальные коммерческие СУБД устарели, грядет эра специализированных СУБД (потоковых, с векторным хранением, супербыстрых, XML и т д)
- Грядет революция
- Старые СУБД очень сложны, никто не знает и не использует весь функционал
- Ядро переписать нельзя
- Надо отказаться от AdHock запросов, журналов, распределенных транзакций, работы с дисками, SQL, сериализации, многоуровневых архитектур и т д

Революции не будет
(в ближайшие годы)

- **Производительность важна, но это не главное, выбирают по совокупности характеристик**
- **Заказчик при выборе руководствуется:**
  - Performance
  - HA
  - Security
  - Manageability
  - Scalability
  - VLDB support
  - Богатство функций
  - Наличие специалистов (DBA, Developer)
  - Проверенность продукта (refferences)
- **Всем нужна совокупность свойств (на примере Spatial option)**
- **Никто не знает, что ему понадобится завтра**
- **Выбирают не специалисты в СУБД и берут по примеру коллег**
- **Нужна не СУБД, а платформа для разработки, на кот доступны разработчики**

Революции не будет

- Избыточные функции СУБД – не проблема, о них можно забыть и не использовать СЕЙЧАС
- Требования к HW не очень важны (память, диски, скорость CPU … растут очень быстро)
- Побеждает не лучшая технология, а деньги, маркетинг, стратегия и т д. Многие уже погибли
- С нуля нельзя создать хороший продукт быстро, а за это время идеи реализуют монстры
- Монстры не допустят революции – встроят новые функции, перепишут ядро, купят
- Революция в СУБД возможна лишь при революции в HW (отказ от дисков (flash), компьютер=сеть и т д)

Как предсказать тенденции

- Жесткая конкуренция в большой тройке, нельзя не реализовать осн. функции
- IBM и MS догоняют Oracle
  - Политка лабораторий в Торонто
  - Функции последних версий MS SQL
  - Возможно, представители MS и IBM добавят
- Поэтому берем за основу Oracle 11g + конкурентов => предсказание на основе анализа рынка СУБД (практической реализации), а не теоретических предложений (Лоуэллский отчет)

Оговорки

- Субъективно
- Не все, а наиболее важные тенденции
- Разные по значимости
- Список не полон (чертова дюжина)
- Терминология !!!!! – извините
- Прогноз только на ближайшие годы (2-3 версии)
- 2 группы:
  - тенденции развития СУБД
  - тенденции развития систем на основе СУБД, влияющие на функции СУБД

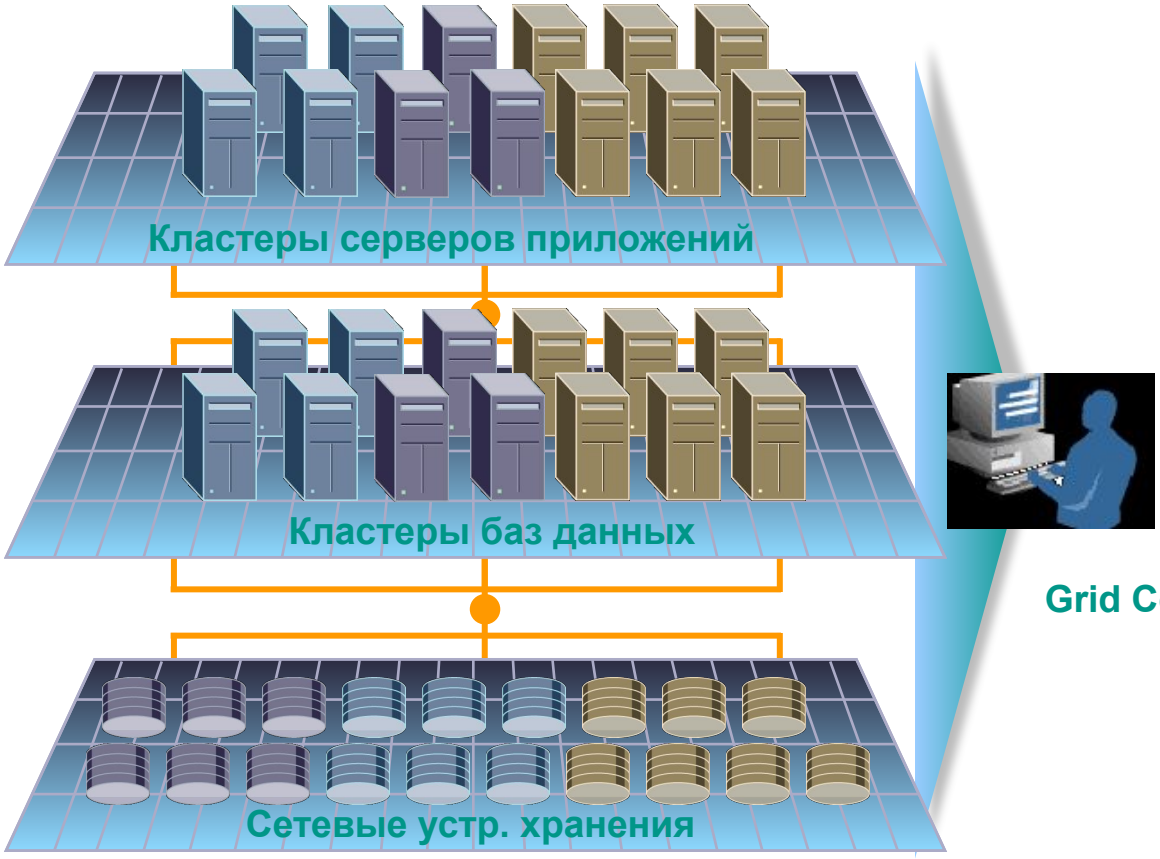
Чертова дюжина

1. Виртуализация ресурсов и GRID-технологии
2. ILM (Information LifeCycle Management) в СУБД
3. Самоуправление, самодиагностика, самолечение
4. Real Application Testing – механизмы промышленного тестирования версий и изменений
5. Совершенствование архитектур макс доступности (разные режимы Standby, Active standby, Snapshot standby, минимизация времени плановых простоев, модификация приложений и версий СУБД, online redefinition)
6. Включение измерения времени в СУБД
7. Новые типы данных (XML, RFID, Semantic Web, геном, медицина, быстрые LOB и т.д.)
8. Умные механизмы сжатия и устранения избыточности
9. Совершенствование защиты данных (Data Vault, Audit Vault, Access&Identity)

Тенденции развития систем на основе СУБД, влияющие на функции СУБД

1. Искажение данных
2. In-memory СУБД реального времени как кэш для коммерческих СУБД
3. Интеграция данных, универсальные и специализированные подходы (динамические DW, BDW – Banking Data Warehouse), MDM
4. Включение баз данных в пространство поиска поисковых систем

1. Виртуализация и Grid Computing



- Storage Grid
- Database Grid
- Application Grid
- Grid Control

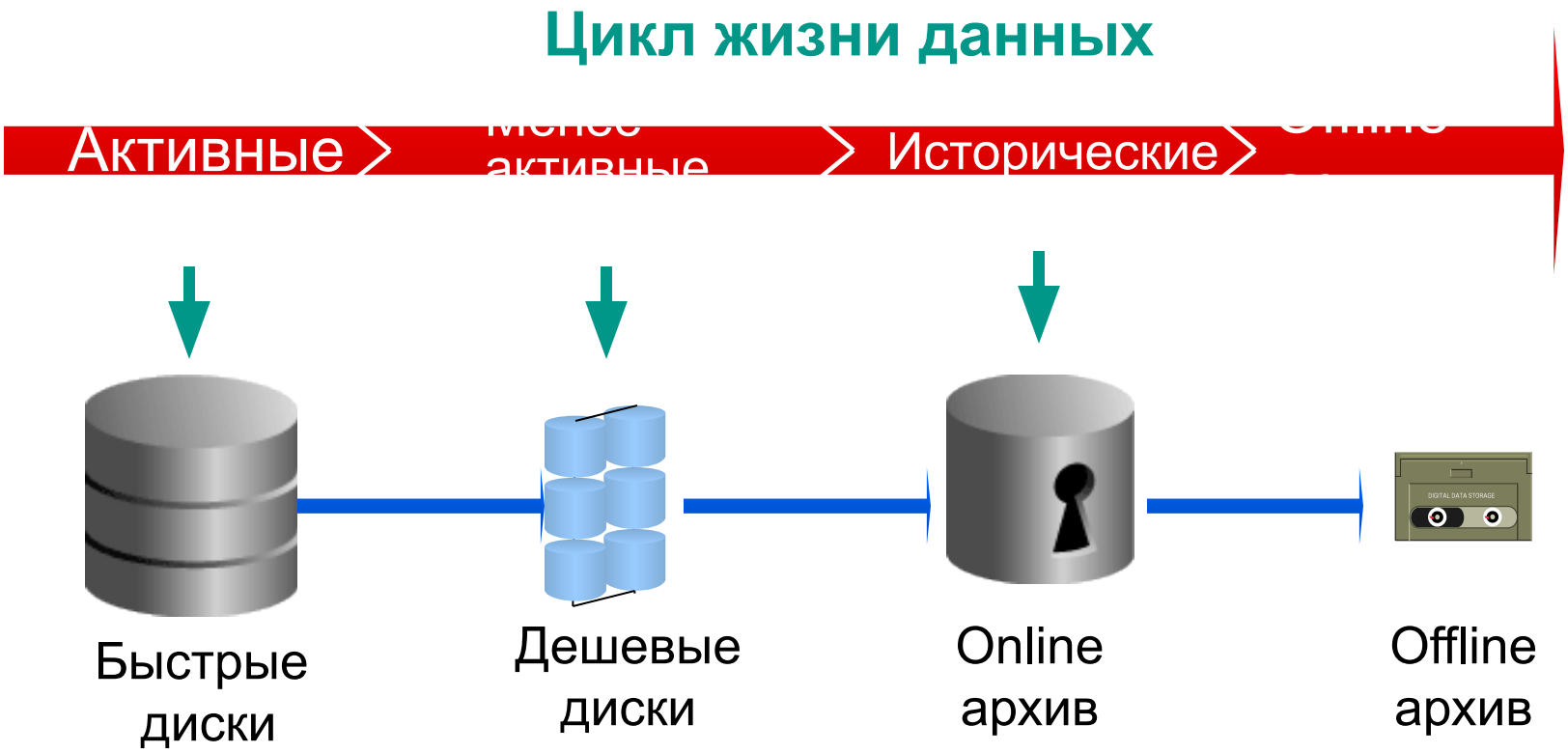
GRID

- Виртуализация ресурсов и простота управления
- Неограниченная мощность на основе дешевых элементов
- Добавление/удаление на лету, клонирование, патчирование
- Гибкость перераспределения ресурсов
- Динамическая адаптация к изменению условий эксплуатации на основе политик
- Независимость от места выполнения программ
- Надежность
- Более эффективное использование ресурсов
- Вычислительный ресурс как коммунальная услуга
- Без спец программирования
- Автоматические балансировка, зеркалирование, перераспределение ресурсов

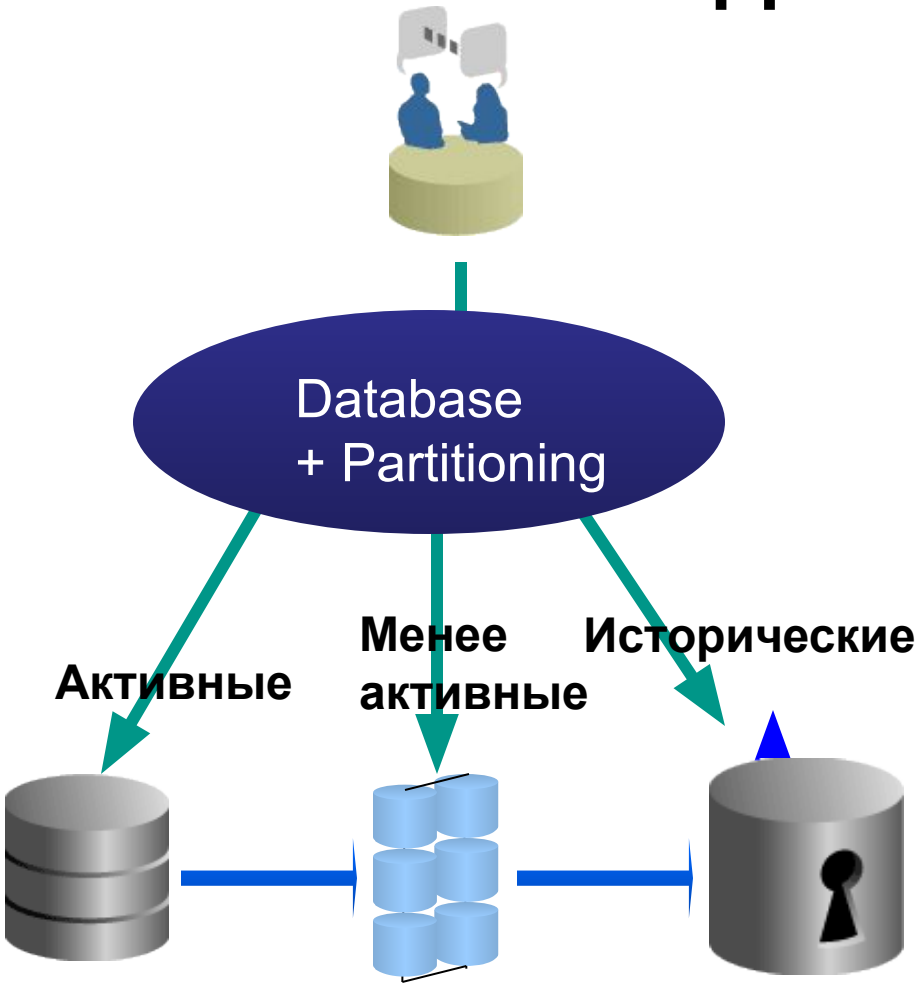
Подтверждение

- Бизнес – привлекательность
- Refferences (Amazon, e-bay Латинской Америки - Mercado Libre, EDS – ABNAMRO …)
- Oracle 11.2 (Динам настройка на среду, AS+HTTP+DB+ASM – делят ресурсы (VM), scaledown – VM, виртуализация БД и файлов)
- Sybase ASE Cluster Edition
- IBM Grid, Globus toolkit

2. Цикл жизни данных



Partitioning – идеальное средство для ILM



- Классы данных
- Области хранения
- Политики хранения/удаления /архивирования/сжатия

- Понимание бизнес данных
- Независимо от Hardware
- Прозрачно для приложений
- Дешево – Можно использо-вать ASM для управления множеством областей хранения (storage tiers)

ILM Assistant

Partitioning for Table - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

ORACLE ILM Assistant

Skip Navigation

Print Logout

User: ADMIN

Lifecycle Setup | Lifecycle Management | Reports | Compliance & Security | Help

Logical Storage Tiers | Lifecycle Definitions | Lifecycle Tables | Preferences

Partitioning for Simulated Table TL.SALES

Cancel | Preview Simulation | Migration Script | OK

Constructs a partition strategy based on a lifecycle definition for a table. If you have not purchased the Partitioning Option or wish to test a partitioning strategy, then the ILM Assistant will simulate partitioning and management of the table. If you have the Partitioning Option, then a SQL script can be generated and used to implement the partitioning strategy.

Lifecycle Definition: All Sales

Partitioning Column: TIME_ID

Partition Date Interval: 1 Month

Number of Rows: 98656400000 reset

Average Row Length: 41 reset

Estimated Compression Factor: 2.14 reset

Data Range Start Date: 01-Jan-1998 reset

Data Range End Date: 21-Oct-2006 reset

Without Lifecycle Management	Single-Tier	With Lifecycle Management	Multi-Tier	With Lifecycle Management	Multi-Tier (Compressed)
Storage Size (Gb)	3,767	Storage Size (Gb)	3,767	Storage Size (Gb)	2,234
		Storage Cost	$25,517	Storage Cost	$17,679
Cost Per Gb	$72.00	Average Cost Per Gb	$6.77	Average Cost Per Gb	$7.91
Storage Cost	$271,233	Storage Cost Savings	$245,716 (91%)	Storage Cost Savings	$253,553 (93%)

Partition Details | Storage Cost Details

Filter Options

Partition Name	Tablespace Name	Low Key Value	High Key Value	Lifecycle Stage Name	Stage Earliest Date	Stage Latest Date	Storage Tier	Storage Size (Gb)	Storage Cost
PFUTURE	Q4_ORDERS	01-Nov-2006	-	Unassigned	-	-	High Performance	0	$0
P20061001_20061031	Q4_ORDERS	01-Oct-2006	31-Oct-2006	Current Sales	01-Aug-2006	31-Oct-2006	High Performance	36	$2,559
P20060901_20060930	Q4_ORDERS	01-Sep-2006	30-Sep-2006	Current Sales	01-Aug-2006	31-Oct-2006	High Performance	36	$2,559
P20060801_20060831	Q4_ORDERS	01-Aug-2006	31-Aug-2006	Current Sales	01-Aug-2006	31-Oct-2006	High Performance	36	$2,559
P20060701_20060731	2005_ORDERS	01-Jul-2006	31-Jul-2006	Old Sales	01-Nov-2005	31-Jul-2006	Low Cost	36	$498

Done

3. Самоуправление, самодиагностика, самолечение

- Сбор, анализ статистики и проактивное реагирование
- Выявляет причины, а не следствия
- Память, ввод/вывод, структуры данных
- Проблемы с безопасностью, производительностью, пространством в БД, нарушение правил, отклонение от стандартов поведения
- Периодическая проверка структур блоков, словаря, контр сумм и т д
- Автоматическая реализация опыта и best practice
- Автонастройка на меняющиеся условия эксплуатации
- Отработка выхода из строя элементов инфраструктуры, советы по восстановлению БД (reparing adviser)
- Автоматические backup/restore

Самоуправление, самодиагностика, самолечение

- Автонастройка тяжелого SQL
- Совершенствование оптимизатора (учет взаимовлияния кардинальности, утверждение планов запроса, profiles, детальное тестирование решений и т д)
- Уменьшение числа ручек
- Задание бизнес параметров (время простоя, уровень сервиса, время отклика и т д)
- Принятие решения с учетом всех элементов инфраструктуры
- Контактирует с тех поддержкой при ошибках, критические патчи
- Остается и ручное управление для Guru
- DBA не умрет

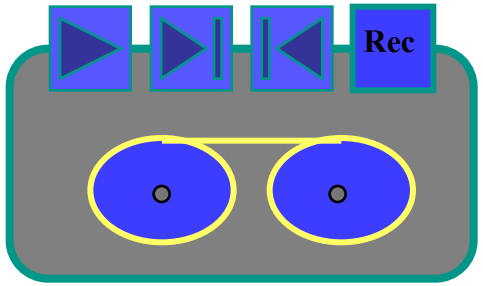
4. RAT – механизмы промышленного тестирования версий и изменений

- Изменения в ПО и оборудовании факт нашей жизни.
- Заказчики хотят оценить влияние изменений прежде, чем их осуществлять.
- Полноценное тестирование требует времени и денег.
- Результат может оказаться плохим несмотря на дорогое тестирование
  - Много проблем оказались не выявленными
  - Изменения негативно сказались на производительности и доступности системы
- Основная причина неудачного тестирования -
  - Неспособность воспроизвести реальную нагрузку
- RAT позволяет провести полноценное тестирование с РЕАЛЬНОЙ нагрузкой.

Захват и воспроизведение нагрузки СУБД - Database Replay

Beta

- Захват инфо о нагрузке
  - Записывает информацию о нагрузке СУБД, включая важную информацию об одновременности нагрузки
- Database Replay - проигрывание
  - Воспроизводит нагрузку в реальном времени (workload with actual timing)
- Анализ и отчетность
  - Отчеты об ошибках
  - Изменение данных
  - Изменение производительности
  - Использование ADDM для дальнейшего анализа производительности



Запиши и проиграй

Анализ и отчетность

Захват и воспроизведение нагрузки SQL – SQL Replay

- Фокус на проблемы с нагрузкой при выполнении SQL операторов
  - Детальный анализ производительности отдельного SQL оператора
- Захват информации про SQL Workload в эксплуатационной системе
  - Захватывает текст SQL, планы, bind variables, статистику выполнения
  - За указанный период времени
  - Можно захватить информацию для SQL в Oracle Database 10g Release 2
- Воспроизведение SQL Workload в тестовой среде
  - Тестовое выполнение SQL в тестовой среде
  - Выявляет SQL с изменившимися планами выполнения и SQL с ухудшившейся производительностью
  - Для ухудшившегося SQL можно провести настройку с помощью SQL Tuning Advisor (10g)
  - Производит анализ и отчеты

SQL Replay: анализ производительности

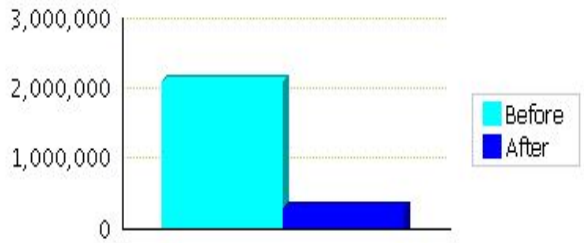
SQL Replay Analysis Result: STE_PAT

Task Name **STE_PAT**
Task Owner **SYSTEM**
Task Description

SQL Tuning Set Name STE_123
STS Owner **SYSTEM**

Global Statistics

Projected Workload Buffer Gets



Improvement Impact +86.692%
Regression Impact -0.080%
Overall Impact +86.612%

SQL Statement Count

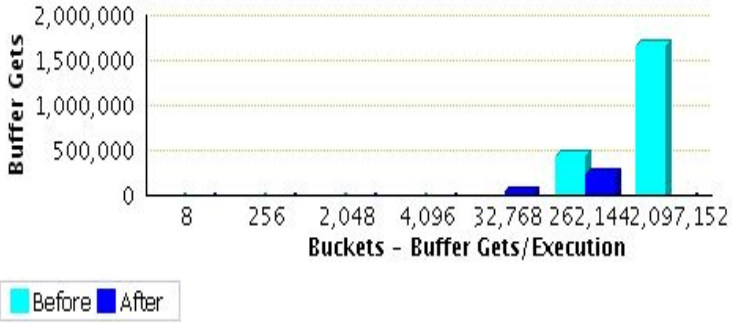
Plan Changed Plan Unchanged No Plan

Recommendations

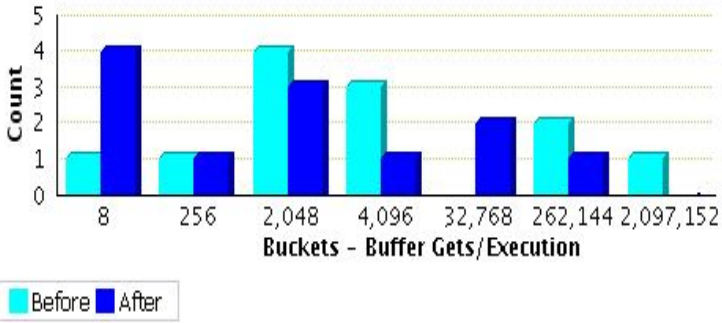
Run SQL Tuning Advisor to tune regressed SQL statements.

Run SQL Tuning Advisor

Projected Workload Buffer Gets Distribution

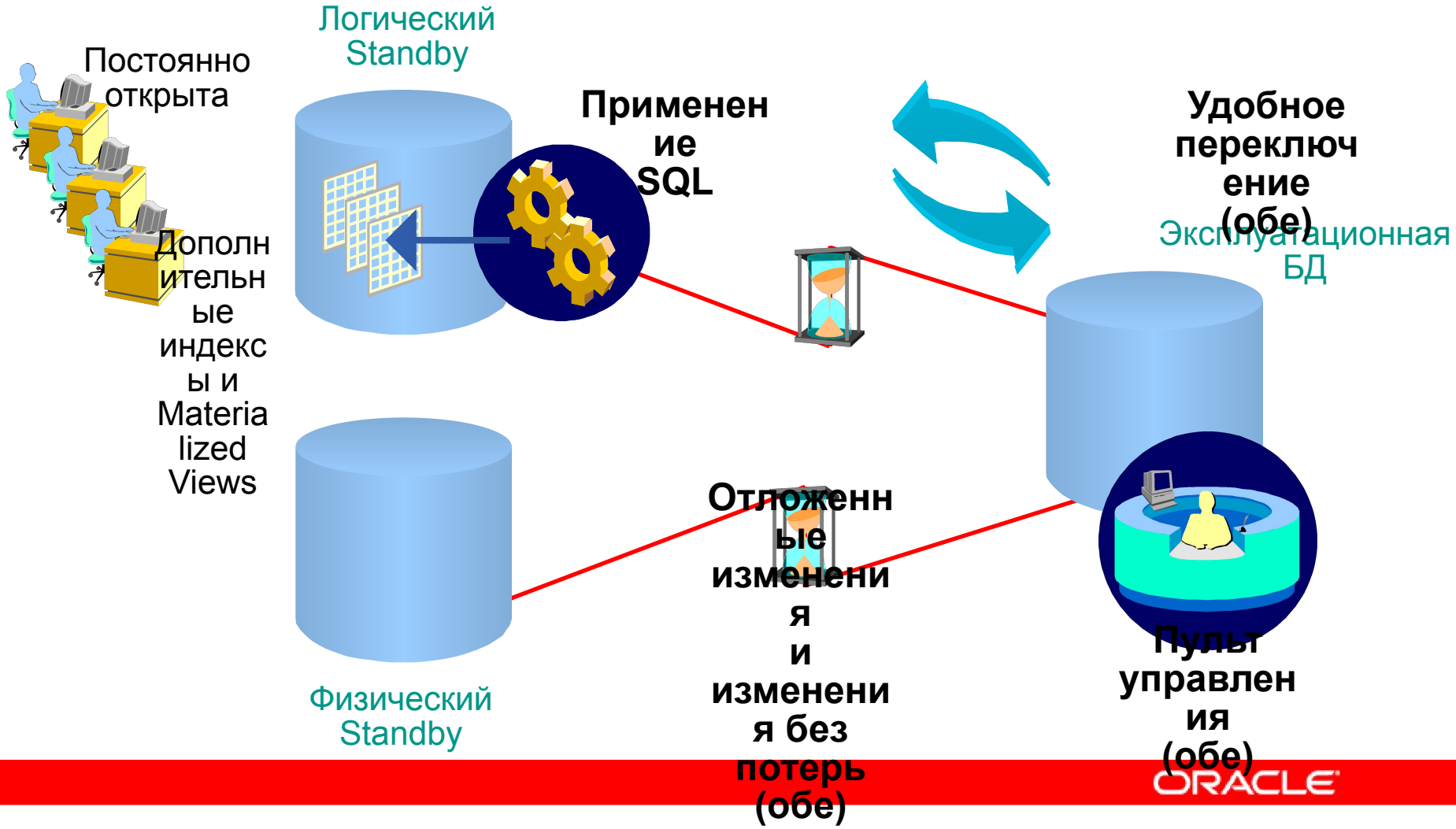


Single Execution SQL Statement Count Distribution

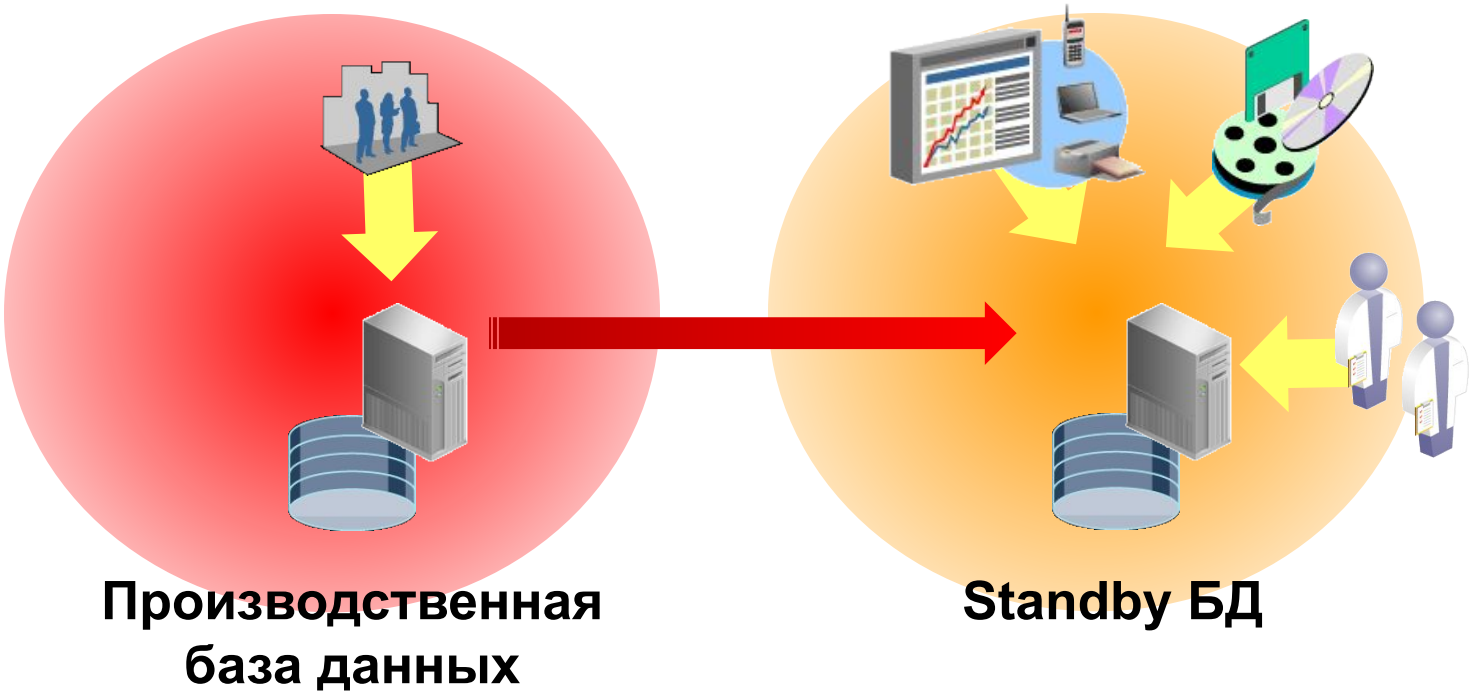


5. Совершенствование архитектур максимальной доступности

Защита от катастрофических сбоев (много разных режимов)



Активная резервная БД разгрузит производственную БД



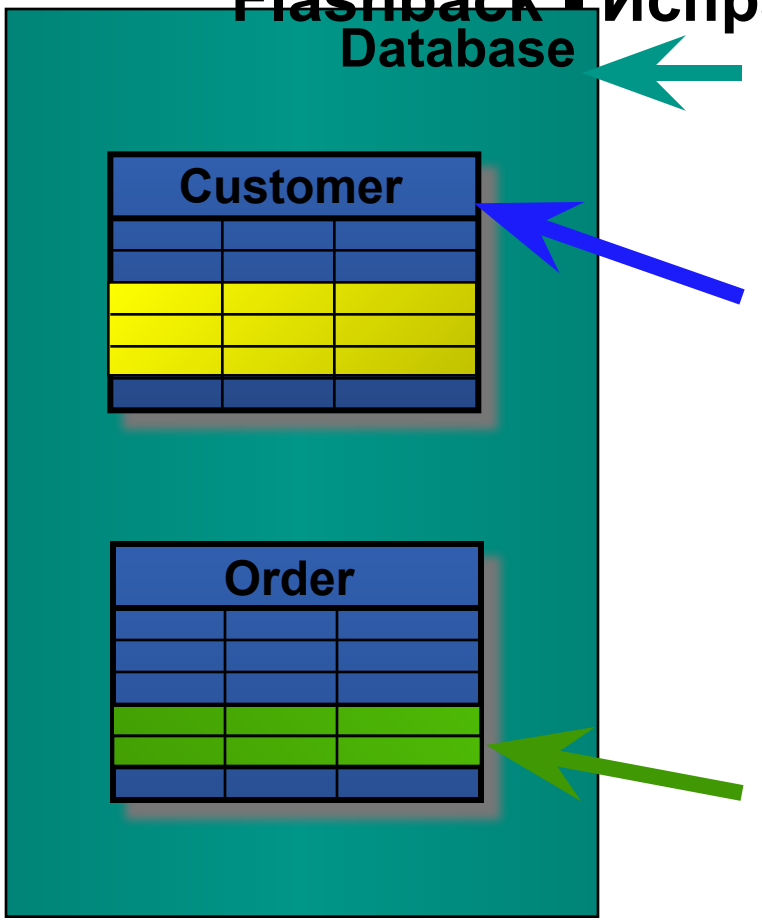
- Выполняйте тестирование изменений, печать отчетов, backup, read-only приложения на резервной базе данных

Совершенствование архитектур макс доступности

- минимизация времени плановых простоев
  - online redefinition структуры БД
  - Online патчи
  - Rolling Upgrade SW и HW
  - Изменение параметров
- модификация версий СУБД
- модификация приложений СУБД
  - Версионность таблиц
  - Версионность процедур, функций, views
  - Сосуществование старой и новой версии приложений, работающих с одними и теми же данными

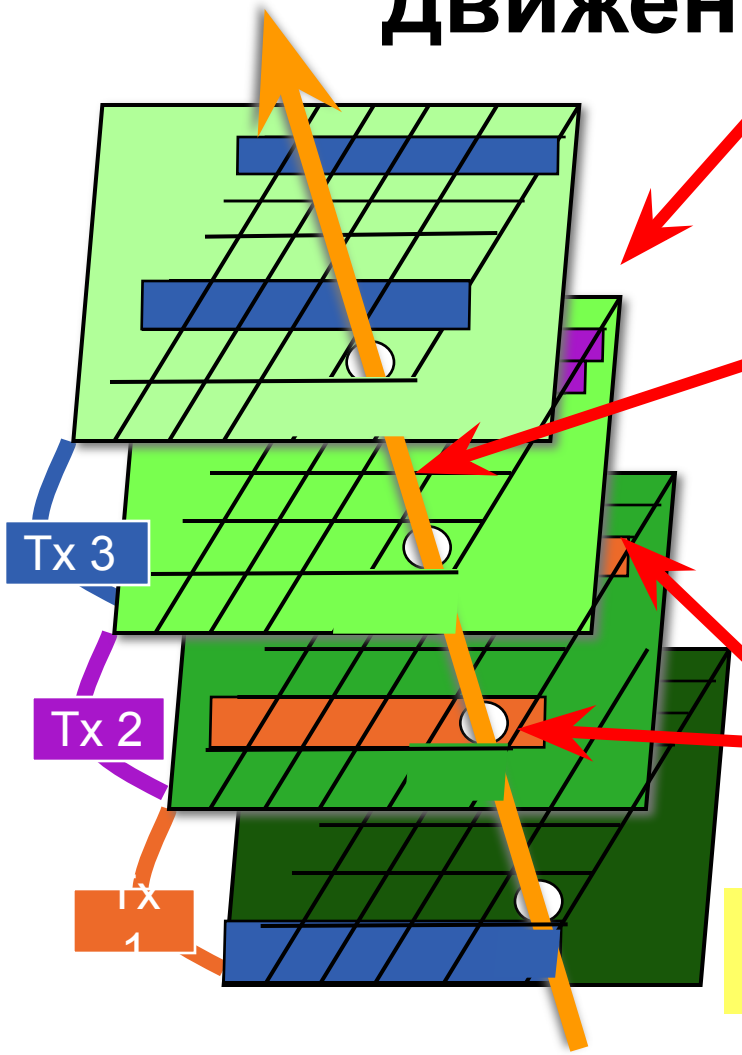
6. Включение измерения времени в СУБД

Flashback - Исправление Ошибок Человека



- Восстановление на всех уровнях к заданному моменту времени
- Уровень БД
  - Flashback Database восстанавливает БД целиком
    - Использует Flashback Logs
- Уровень таблиц
  - Flashback Table восстанавливает записи в наборе таблиц
    - Использует UNDO в БД
  - Flashback Drop восстанавливает удаленные таблицы и индексы
    - Использует Recycle bin
- Уровень записи
  - Восстановление конкретных записей
    - Использование Flashback Query
- New: Откат транзакций

Исследование проблем с помощью движения по времени



- Flashback Query
  - Запросить все данные на определенный момент времени

select * from Emp AS OF ‘2:00 P.M.’ where …

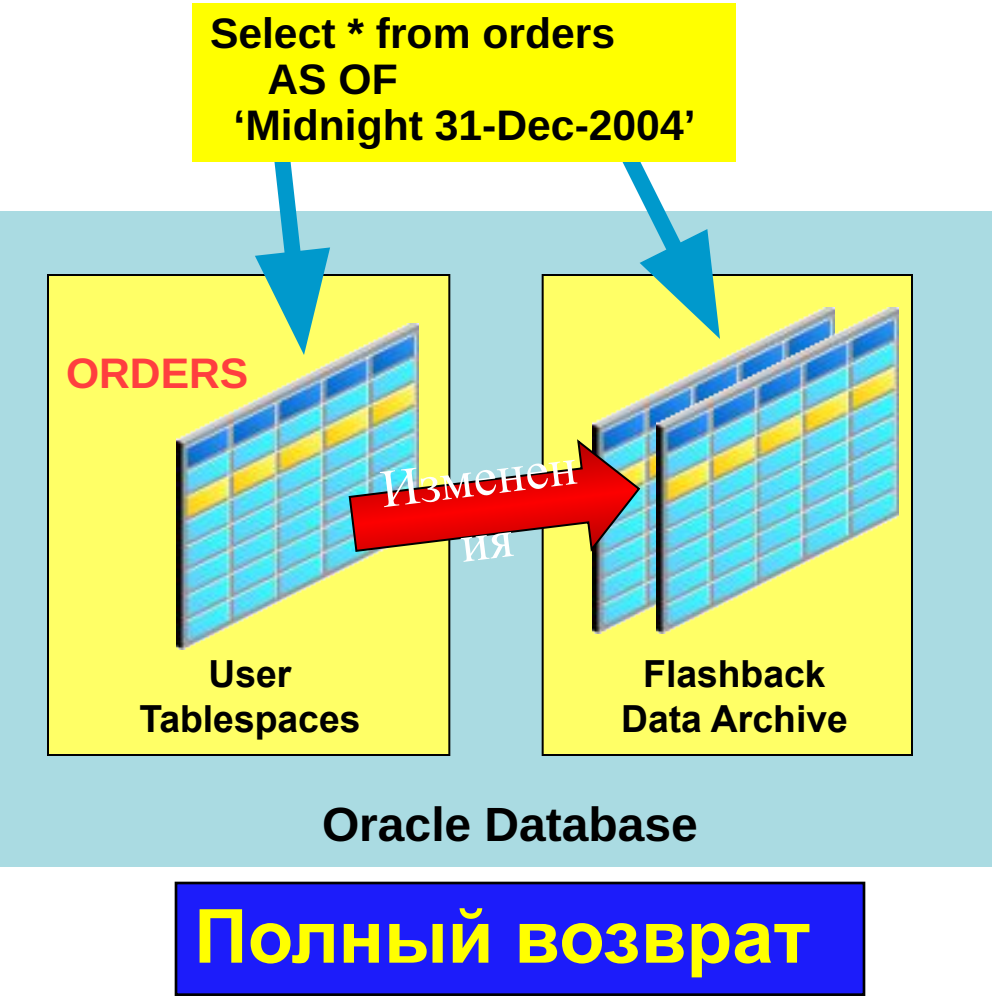
- Flashback Version Query
  - Посмотреть все версии строк в заданном интервале времени
  - Посмотреть транзакции, которые изменили строку

select * from Emp VERSIONS BETWEEN ‘2:00 PM’ and ‘3:00 PM’ where …

- Flashback Transaction Query
  - Посмотреть все изменения, сделанные транзакцией

select * from FLASHBACK_TRANSACTION_QUERY where xid = ‘000200030000002D’;

Flashback Data Archive – опция RECALL



Полный возврат

- **Измерение времени в БД**
- **Долговременное хранение - годы**
- **Автоматически сохраняет все изменения для выбранных таблиц в Flashback Data Archive**
  - Архив не может быть изменен
  - Старые данные удаляются в соответствии с политикой сохранения
- **Можно посмотреть содержимое таблицы на любой момент времени с помощью Flashback SQL**
- **Используется для**
  - Отслеживания изменений
  - ILM
  - Долговременной истории изменений
  - Аудита
  - Соответствия законадательству (Compliance)

7. Новые типы данных

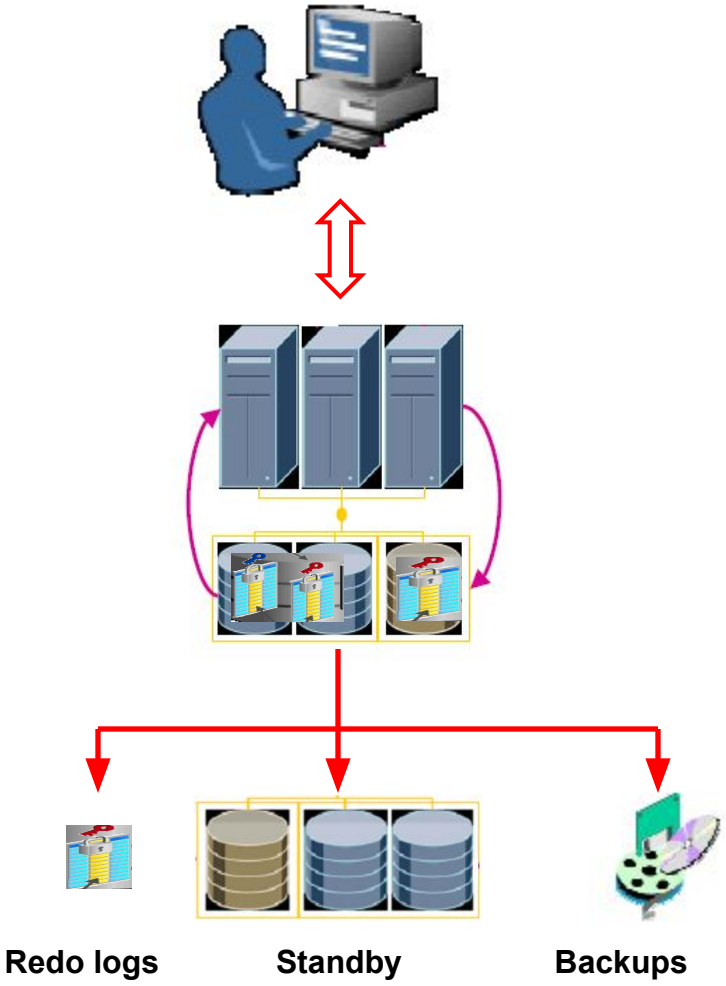
- RFID
- Semantic Web,
- геном
- Life Science (медицина, биология, химия, биохимия, иммунология, генетика, экология ....), алгоритм BLAST
- быстрые LOB
- XML
- и т д
- Специальные способы хранения, индексирования, оптимизации
- Набор стандартных операций

Семантические сети

- Oracle Database 11g – первая промышленная открытая, масштабируемая, безопасная и надежная СЕМАНТИЧЕСКАЯ СУБД со встроенной поддержкой стандартов RDF и OWL.
- Она может работать с наборами данных, превышающими по размеру в десятки раз объемы данных специализированных RDF и OWL СУБД.
- Oracle обеспечивает более продвинутые и точные средства запроса и позволяет легко расширить функционал существующих SQL приложений за счет подключения семантического поиска и семантических запросов.

8. Умные механизмы сжатия и устранения избыточности

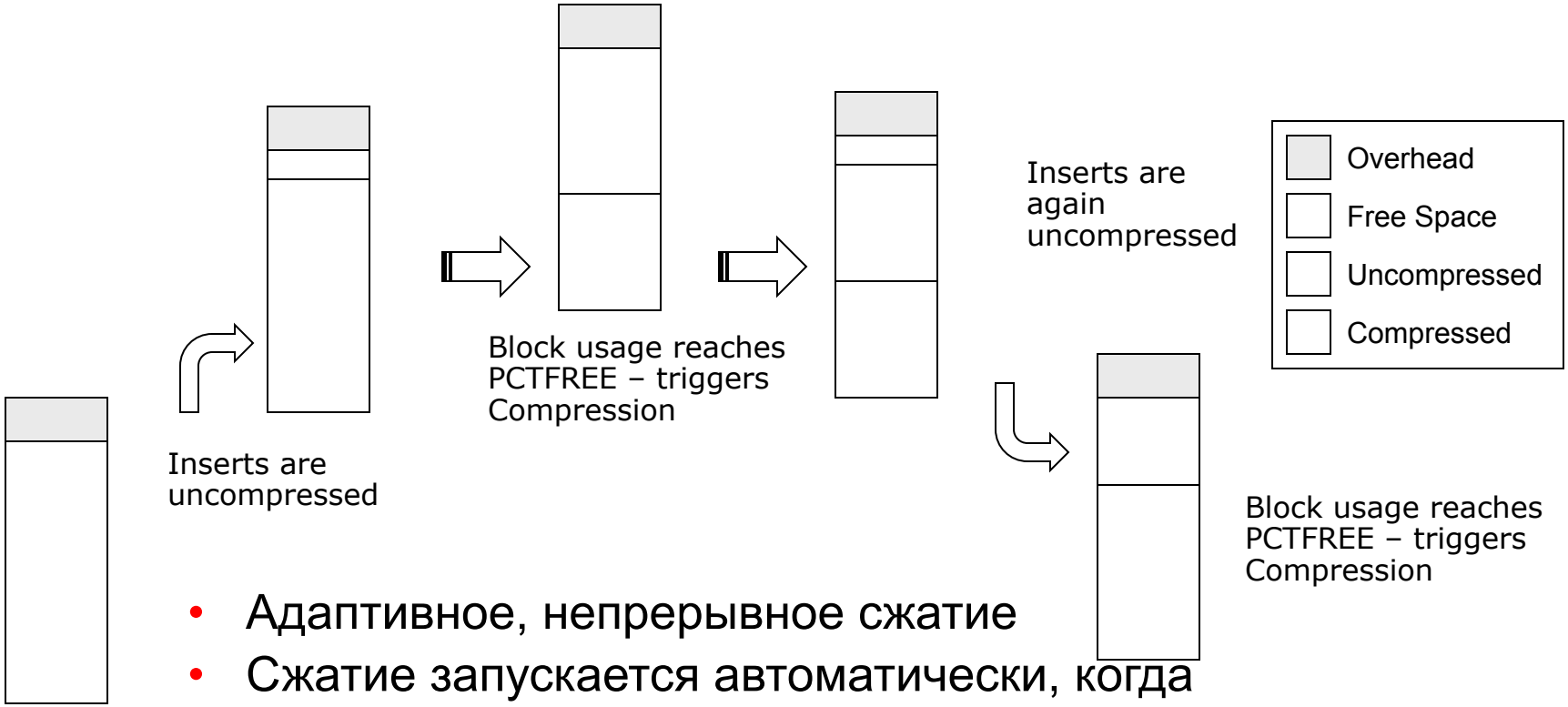
- Богатый набор механизмов сжатия
  - Сжатие структурированных/реляционных данных
  - Сжатие неструктурированных данных
  - Сжатие для backup
  - Сжатие сетевого трафика
  - Сжатие для Data Pump (на 75%)
  - Сжатие для Data Guard (gap resolution в 2 раза быстрее)
- Мин влияние на производительность
- OLTP и DSS
- Разные алгоритмы для разных данных



SecureFiles - Сжатие

- Автоматически определяет, что SecureFile можно сжать
  - Не выполняет сжатие для уже сжатых данных
  - Не выполняет сжатие, если экономия места будет минимальной или нулевой
- Два уровня сжатия
  - Уровни сжатия: MEDIUM (default), HIGH
  - Чем выше степень сжатия, тем больше задержка и нагрузка на CPU
- Сжатие SecureFiles независимо от сжатия таблиц и индексов
- Сжатие выполняется на сервере
  - При работе с частями файлов, разжимаются только необходимые блоки
  - Может быть задано на уровне partition

OLTP Table Compression



- Адаптивное, непрерывное сжатие
- Сжатие запускается автоматически, когда достигается % заполнения блока PCTFREE
- Сжатие исключает дырки, возникающие при удалении записей и максимизирует размер непрерывного свободного пространства в блоке

9. Совершенствование защиты данных

- **Audit Vault** – DW всей аудит информации предприятия
  - Много источников
  - Защищено, не удаляется
  - Стандартные отчеты, DataMining
- **Access&Identity Management** – вынесение механизмов аутентификации, авторизации, управления привилегиями и контроля доступа из СУБД
  - Более строгие механизмы
  - Workflow
  - Гибкие политики
  - Single Sign On
  - LDAP Directory
  - Централизация
- **Data Vault**

Oracle® Database Vault

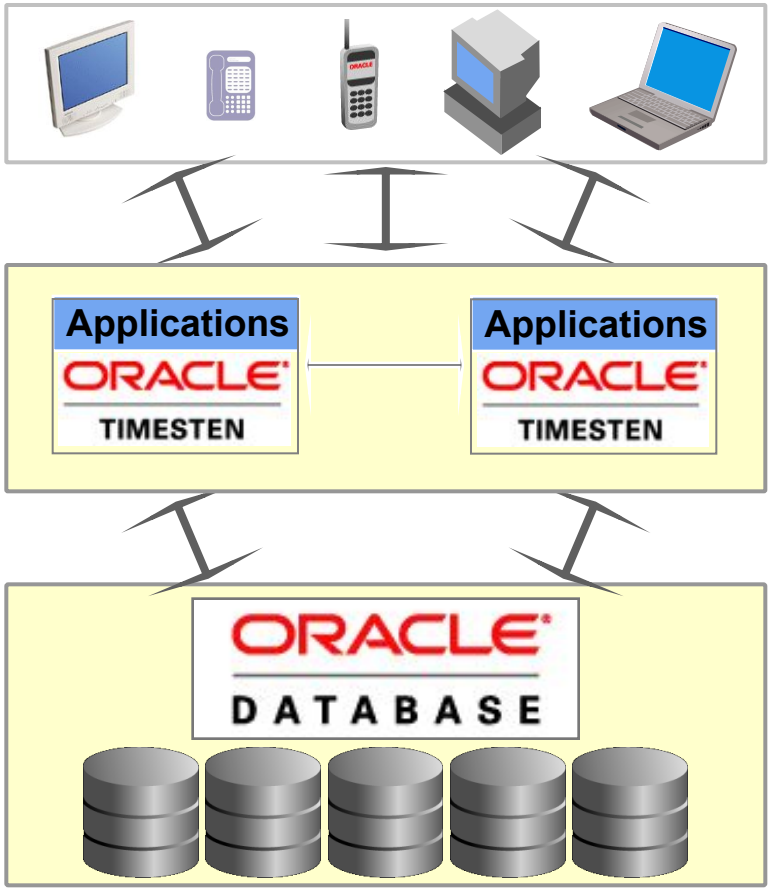
Функциональные элементы

10. Искажение данных

- При выдаче результата – подмена данных (военные, медицина ....)
- Передача БД для тестирования и разработки в другую компанию
  - Data Masking Pack
  - Замена на случайные значения, константы, перемешивание, значения из списка, пользовательских использование процедур
  - Маски, проверка и сохранение формата и ограничений
  - Ссылочная целостность базы и логики приложений
  - Библиотека стандартных форматов (имена, телефоны ....)
  - Уникальность
  - Быстро, в пакетном режиме

11. In-memory СУБД реального времени как кэш для коммерческих СУБД

- *Подтверждение – сегодня IBM купил Solid*
- *TimesTen – in-Memory СУБД*
- *Быстрый КЭШ над дисковой СУБД*
- *Обработка данных в режиме реального времени*
- *Без сброса на диск, надежность за счет репликации*
- *Можно отключить журналирование*

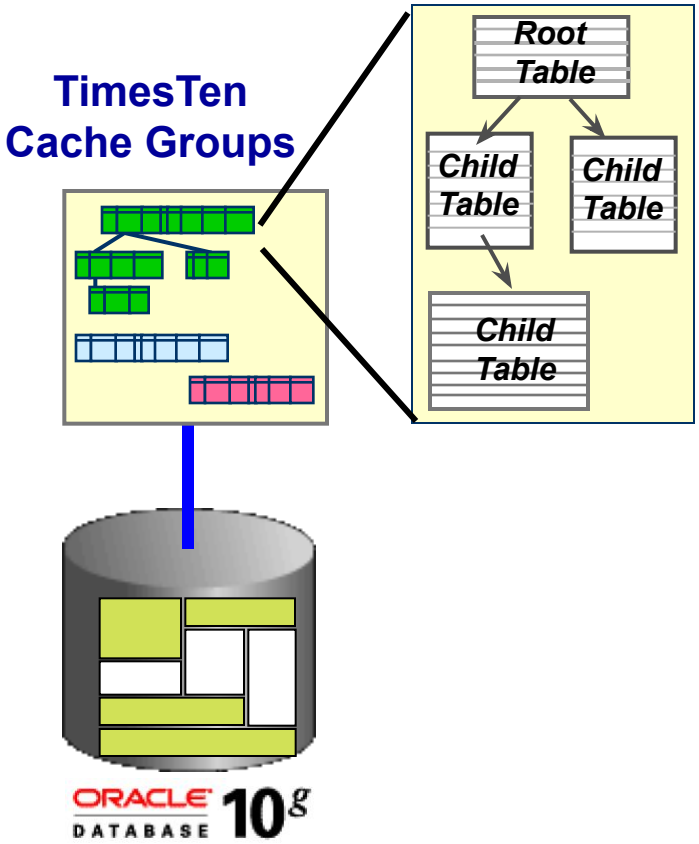


Основные достоинства TimesTen

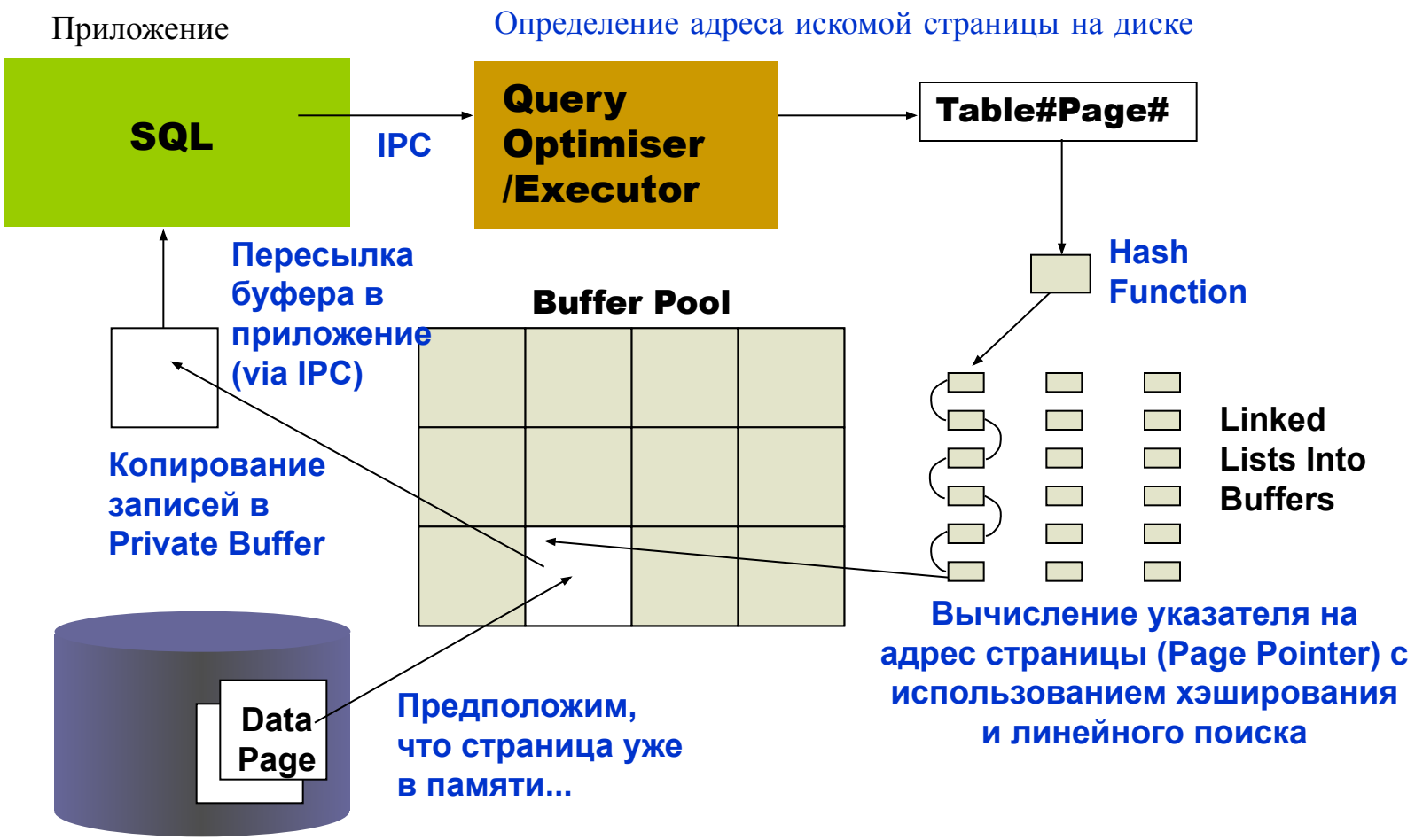
- In Memory Database (> 1 Tb)
- Время ответа – микросекунды (вместо миллисекунд)!
- Специальная архитектура и способы индексации
- Предсказуемое и постоянное время ответа
- Пропускная способность 100,000 TPS и выше
- Знакомая реляционная модель – разработчикам не нужно переучиваться (ODBC API)
- Вмешательство DBA - минимально

Кэширование данных Oracle Database

- Cache groups
  - Независимые или связанные таблицы
  - Вся таблица или подмножество строк и колонок
  - Поддержка отношения таблиц parent-child
- Read-only or updatable
- Двунаправленные обновления
  - Из TimesTen в Oracle
  - Из Oracle вTimesTen
- Несколько кэшей для одной Oracle DB

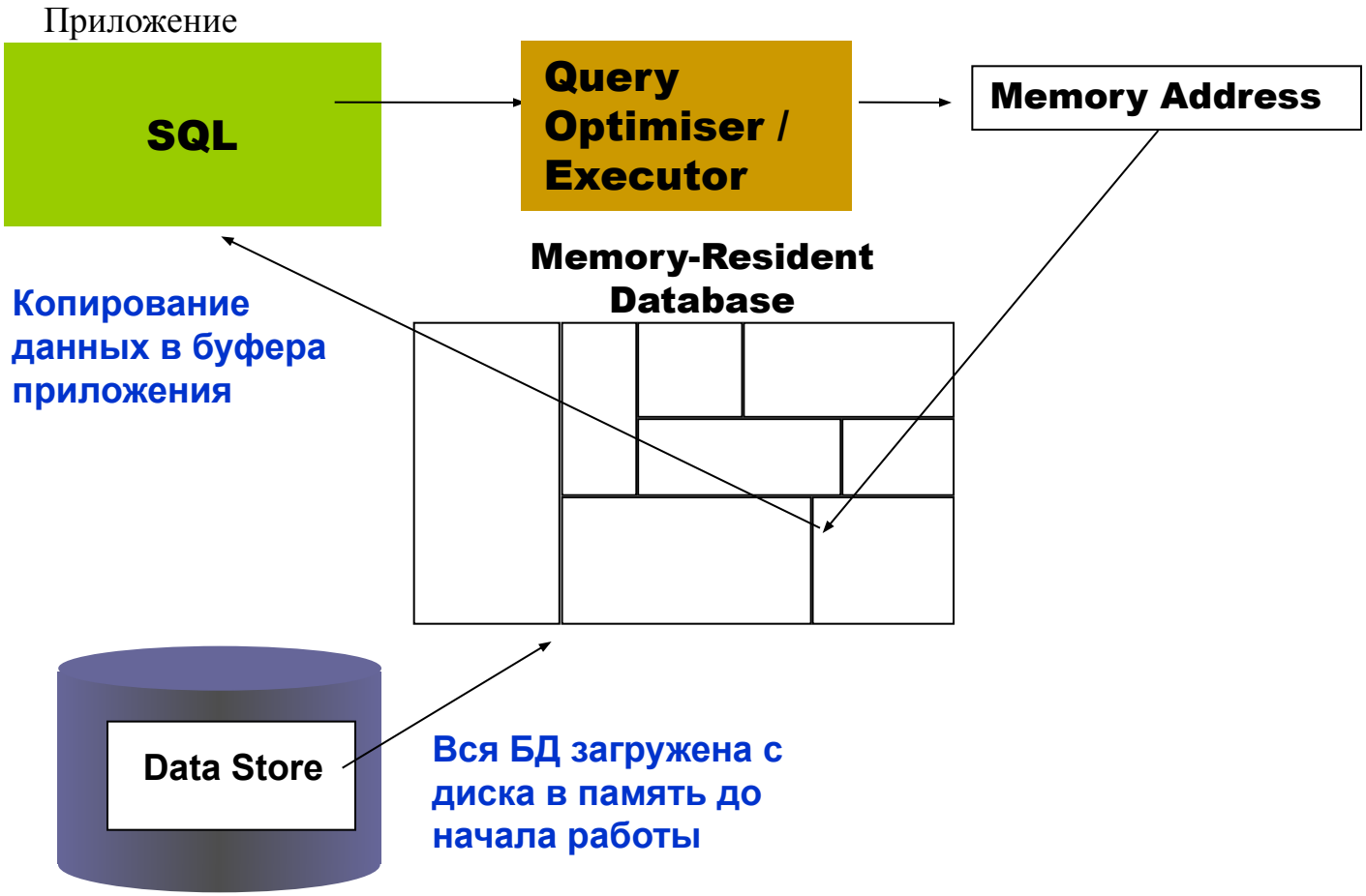


Почему TimesTen быстрее: Дисковая СУБД



Почему TimesTen быстрее: TimesTen

Вычисление прямого адреса в памяти для искомой записи



12. Интеграция данных, модели данных

- DW, Виртуальный DW, смешанная модель, RT DW
- Когда это лучше чем DW?
  - Нужна свежая информация
  - Небольшую часть данных лучше брать из online систем, нужны редко
  - Real Time Decision, IBM Dynamic DW, Sibel Server
  - Структурированная и неструктурированная информация в DW
  - Политики безопасности запрещают перемещать данные
- Единая интегрированная модель данных с точки зрения приложений
- Запросы преобразуются в запросы к источникам
- Это накладывает доп требования на СУБД, реализующую такой подход (IBM II)
  - Оптимизация с учетом источников, хар-к сети
  - Выполнение функций, не поддерживаемых источником, выталкивание обработки
  - Кэширование, федеративная СУБД
  - Единый язык запросов
  - Преобразование запросов и данных
  - Wrappers
  - Отображение нереляционных объектов в реляционную модель и т д

Модели данных для DW

- Универсальные индустриальные модели, как заготовка для модели данных DW (BDW, TDW, RDW, Sibel и т д
- Best practice, многое учтено, правильные структуры, постепенное наращивание, упрощен GAP анализ
- Нужны простые средства бизнес-редактирования и генерации оптимальных структур для конкретной СУБД
- Другое направление – MDM (Meta Data Modelling) – единые справочники (пользователи, продукты и т д)

13. Включение баз данных в пространство поиска поисковых систем

- Большинство поисковых систем ищет по сайтам и текстам
- Над искать и по БД
- С учетом прав доступа к информации
- Единый поисковый запрос к разнородной информации (архив док, СУБД, mail, Web …) mapping полей
- Свои правила определения релевантности
- Удобно реализовать в виде Web Service со стандартным интерфейсом

Вопросы

Как нас найти...

- Телефон в Москве
  +(7 495) 641-14-00
- www.oracle.com/ru
- www.oracle.ru
- Email
  Mark.Rivkin@oracle.com