

ORACLE®



**Сергей Данилов**  
**Oracle СНГ**

Ведущий консультант по  
Базовым технологиям  
Государственный сектор

**Анатолий Барышников**  
**Oracle СНГ**

Менеджер проектов  
Государственный сектор

# Введение в Oracle Enterprise Grid

*The best thing about the Grid is that it is unstoppable.*

The Economist, June 21, 2001

# 40 лет назад...

10<sup>g</sup>



**В 1964 году компания IBM создала первый mainframe под управлением OS/360**

**Gene Amdahl (Джин Эмдал) покинул IBM чтобы основать собственную компанию Admahl Computing**

**Начинается гонка по созданию больших серверов**

# Через 40 лет...

10<sup>g</sup>



В гонку включается даже компания Microsoft!

Продемонстрирована работа SQL Server на 64 процессорной машине под Windows

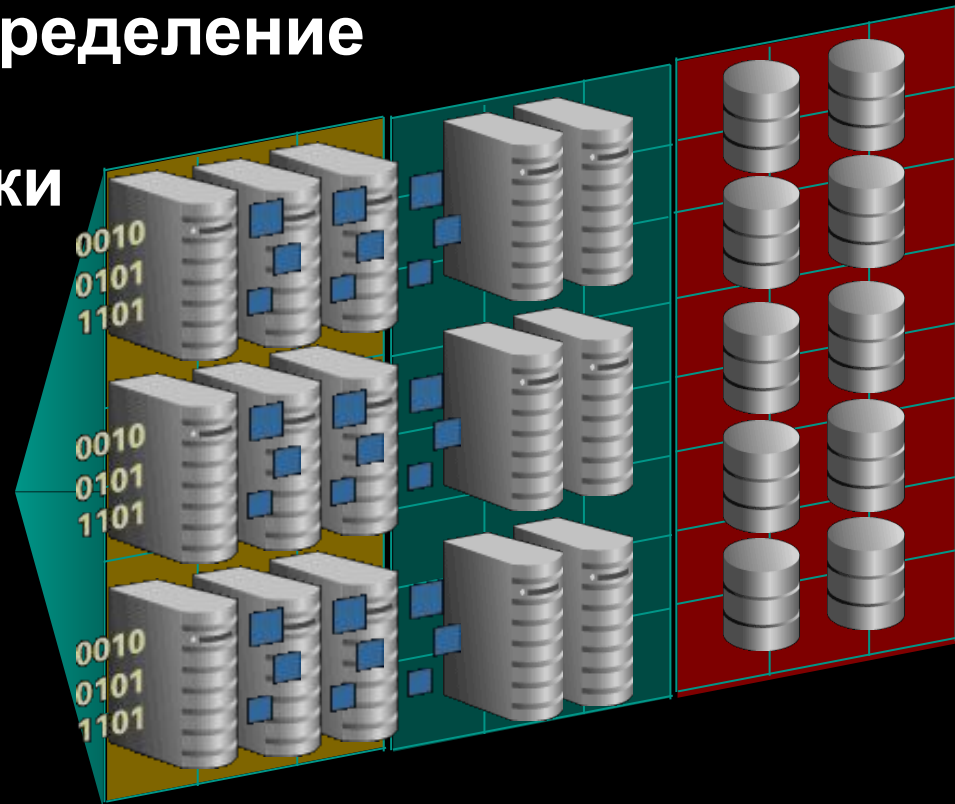
**Общая тенденция: Всем нужны все более и более крупные серверы**

# Недостатки такой архитектуры

1. Приложения перерастают даже самые большие компьютеры
2. Большие компьютеры дороги.  
Один процессор стоит 30-50 тыс. долларов.
3. Один большой компьютер –  
“недублированная точка отказа”  
(single point of failure) - Все равно не надежно

# Модель Grid вычислений Oracle (Enterprise Grid Computing)

- Объединение
- Виртуализация и распределение ресурсов
- Выравнивание нагрузки
- Качество сервиса
- Автоматизация





# Идея GRID Computing

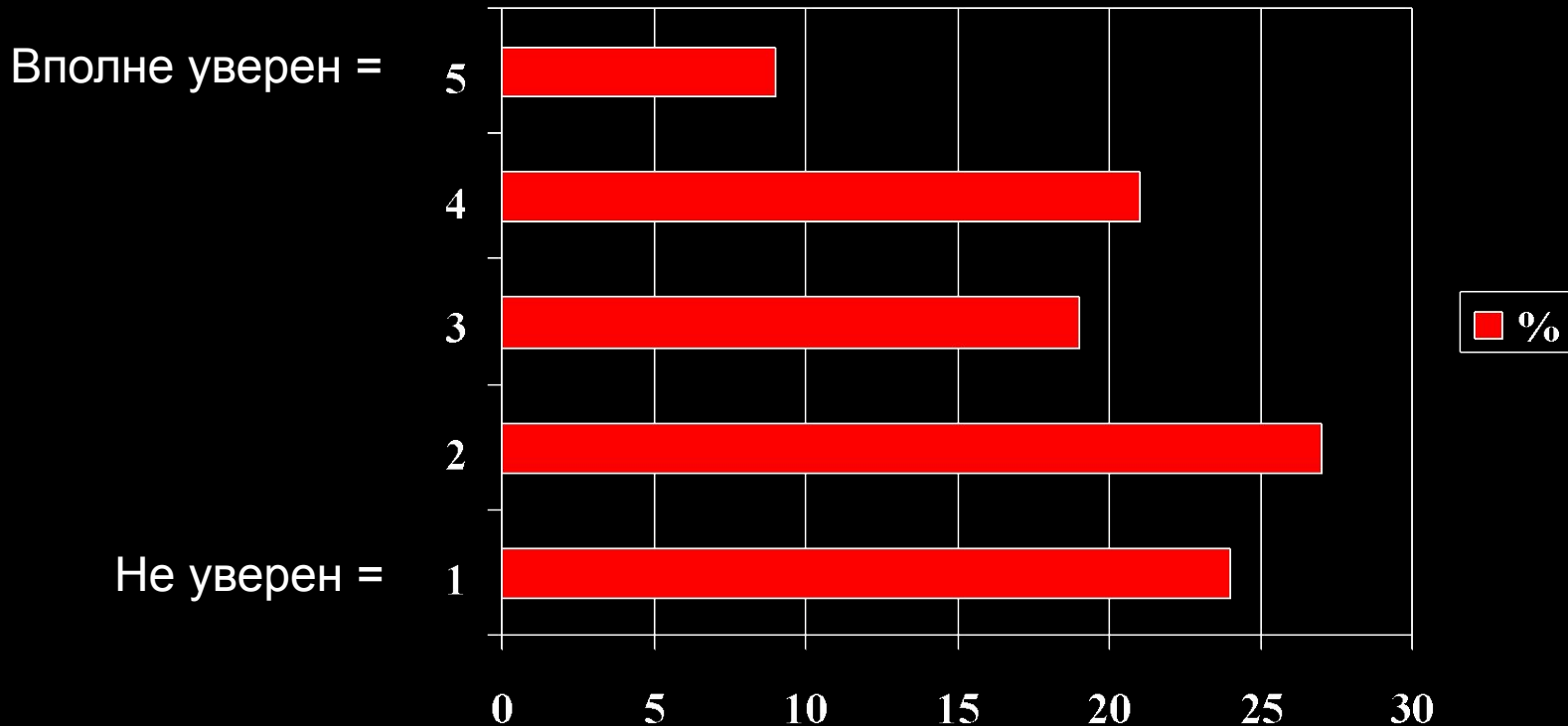
- Объединение множества компьютеров в один суперкомпьютер («Сеть – это Компьютер»)
- Вычисления как коммунальная услуга (Computing Utility)
  - Аналог: Электрические сети (power grid)
- Следующий шаг развития Internet (вторая, после Web, революция в Сети)
- С точки зрения потребителя вычислительных ресурсов:
  - Не важно, где в сети размещаются данные и какой компьютер обрабатывает запрос на их обработку
  - «Обработка данных по требованию»

“В мировом масштабе, компании тратят более чем \$49 миллиардов в год на серверы, а опросы IT-директоров показывают, что утилизация серверов составляет 60% или менее.”

— *Forrester Research, April 2002*

# Уровень информированности об утилизации инфраструктуры

10<sup>g</sup>

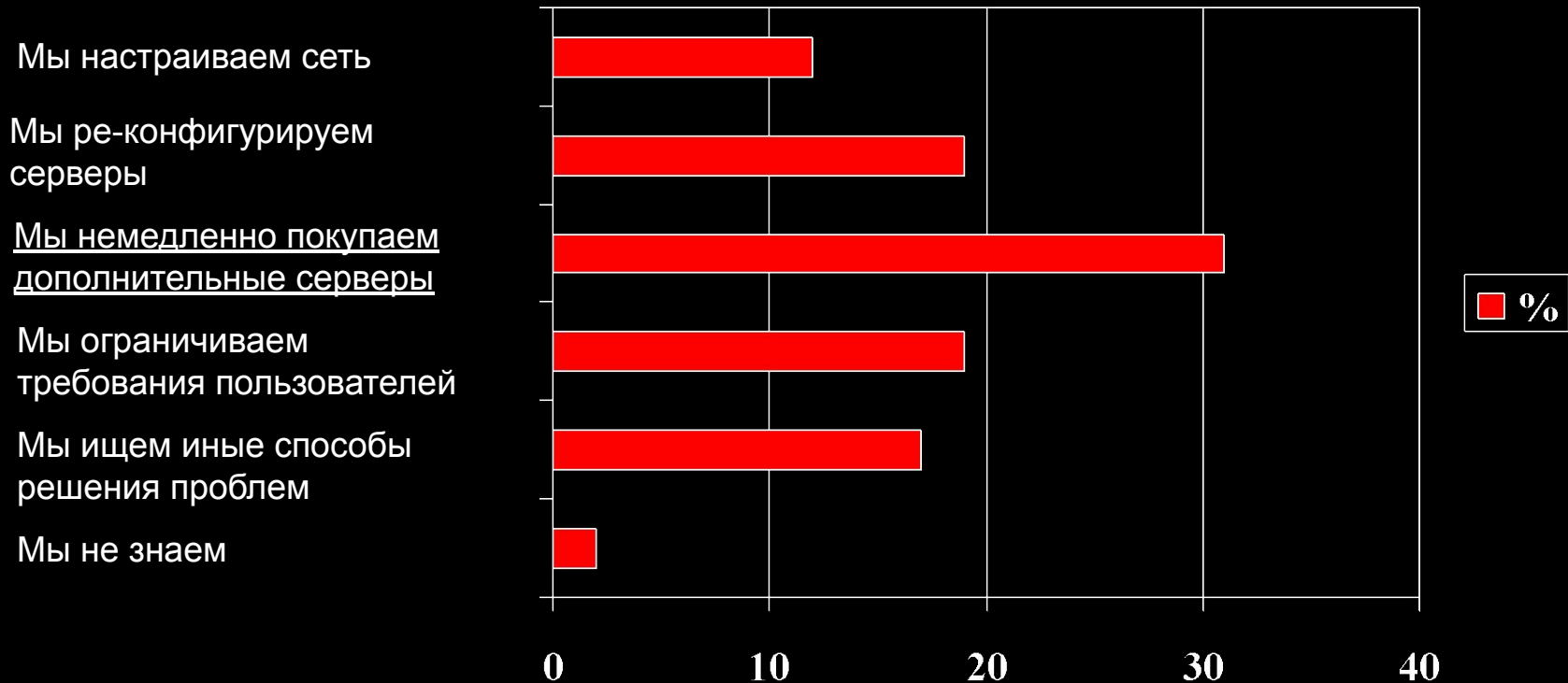


Насколько Вы уверены в знании уровня утилизации инфраструктуры?

ORACLE

# Стиль мышления, далекий от стратегического

10<sup>g</sup>



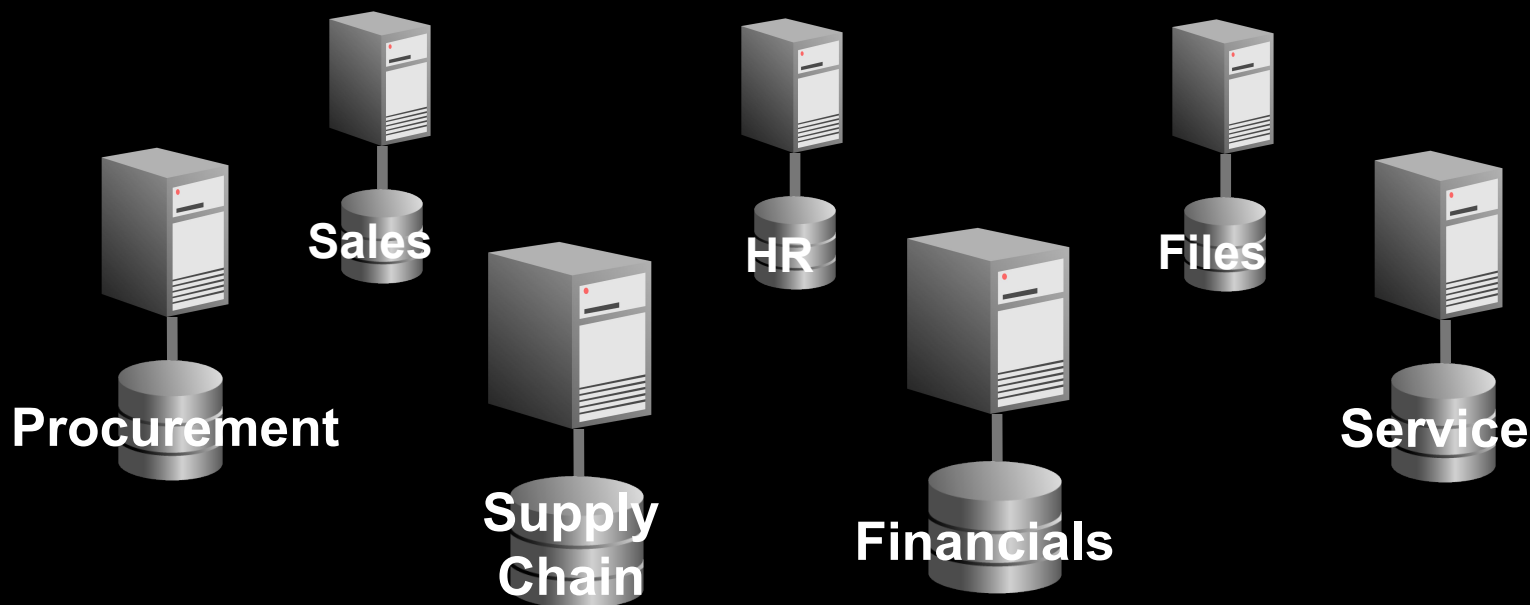
Какая из приведенных фраз в наилучшей степени отражает Ваш подход к решению проблемы перегрузки инфраструктуры?

ORACLE

# Проблема: «Островки» приложений

10<sup>8</sup>

- Готовность < 99.х%, недублированные точки возможного отказа, проблемы безопасности
- Ограниченная масштабируемость
- Специальное конфигурирование оборудования и ПО с учетом требований пиковой нагрузки



# Enterprise Grid решает проблему «островков» приложений

Coordinated use of many small servers acting  
as one large computer.

## Проблема

- Обособленные ресурсы
- Высокая стоимость
- Сложное реконфигурирование
- Возможны отказы
- Сложное администрирование



## Решение в GRID

- Разделяемые ресурсы
- Недорогие компоненты
- Динамическое выделение ресурсов
- Высокая надежность
- Унифицированное управление

# Определение Grid

“Grid представляет собой кластеры компьютеров или серверов, специальным образом связанных между собой, что позволяет использовать их как общий пул вычислительных ресурсов”

– “*Defining Grid Computing*”, Giga Research, August 2002

# Новые термины

- Grid – сеть распределенной обработки данных
- Grid Computing – технология распределенной обработки данных
- Enterprise Grid – корпоративная сеть распределенной обработки данных



# “Эволюция” GRID

10<sup>g</sup>



# Идеи GRID

- Утилизация неиспользуемых ресурсов
- Разделение ресурсов (resource pooling)
- Выделение ресурсов (resource provisioning)

# Решение: Computing On Demand

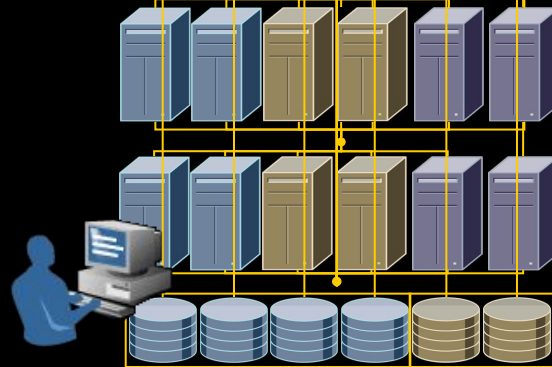
10<sup>g</sup>

## “Mainframe” Модель



- Партиционирование одного большого сервера
- Дорогие элементы
- Полное, интегрированное ПО
- Высококачественный сервис за большую цену

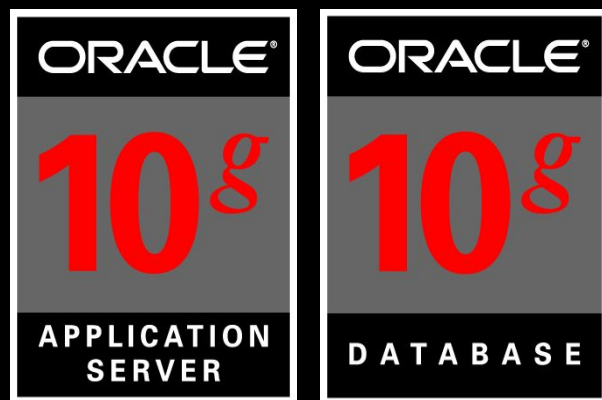
## Grid Computing Модель



- Скоординированное использование множества маленьких серверов
- Дешевые, стандартные, модульные элементы
- Открытое, полное, интегрированное ПО
- Высококачественный сервис за низкую цену

# Why Grid is the Next Big Thing?

- Новая экономика управления данными («affordability»)
- Развитие компьютерных платформ
  - Blade-servers (серверы-лезвия)
  - Blade-farms (фермы серверов-лезвий)
- Развитие операционных систем:
  - Высокая динамика развития Linux как серверной ОС
  - Linux – основная ОС для серверов-лезвий
- Внимание ведущих игроков рынка IT
  - HP (Computing Utilities), Sun (N1), IBM (On Demand)



Программная инфраструктура сети распределенных вычислений организации

# Технологический стек прикладных программных систем

10<sup>g</sup>



Концепция	Содержание
Самоконфигурация	Автоматическое конфигурирование компонентов и систем в соответствии с высокоуровневыми правилами
Самооптимизация	Компоненты и системы постоянно ищут возможность увеличить свою производительность и эффективность
Самовосстановление	Система автоматически выявляет, диагностирует и исправляет локализованные программные и аппаратные проблемы
Самозащита	Система автоматически защищается от вредоносных атак или ошибок, использует средства раннего предупреждения для прогнозирования сбоев

# Oracle Enterprise Grid

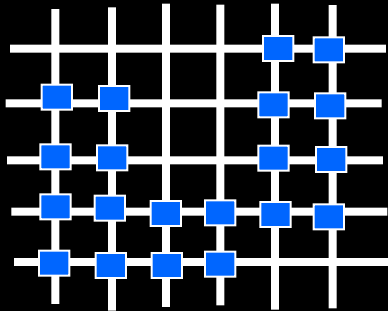
Корпоративная сеть  
распределенной обработки  
данных



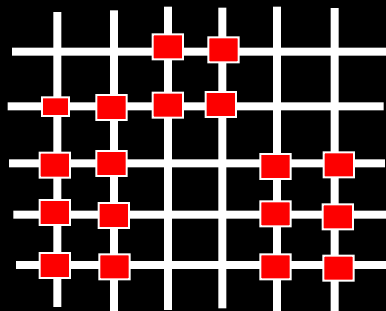
# Oracle Enterprise Grid

Согласованное использование  
множества серверов, работающих как  
один большой компьютер

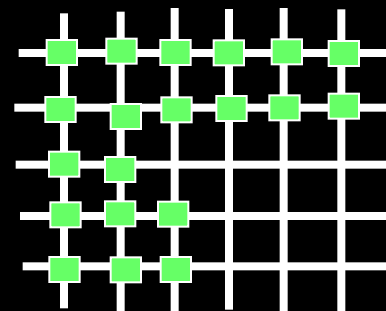
Storage Grid



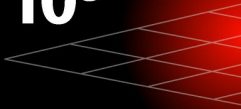
Database Grid



Application Server Grid



Grid  
Control

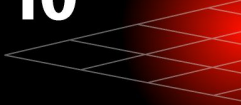


# Storage Grid

Корпоративная сеть устройств  
хранения данных

# Oracle Storage Grid

10<sup>g</sup>



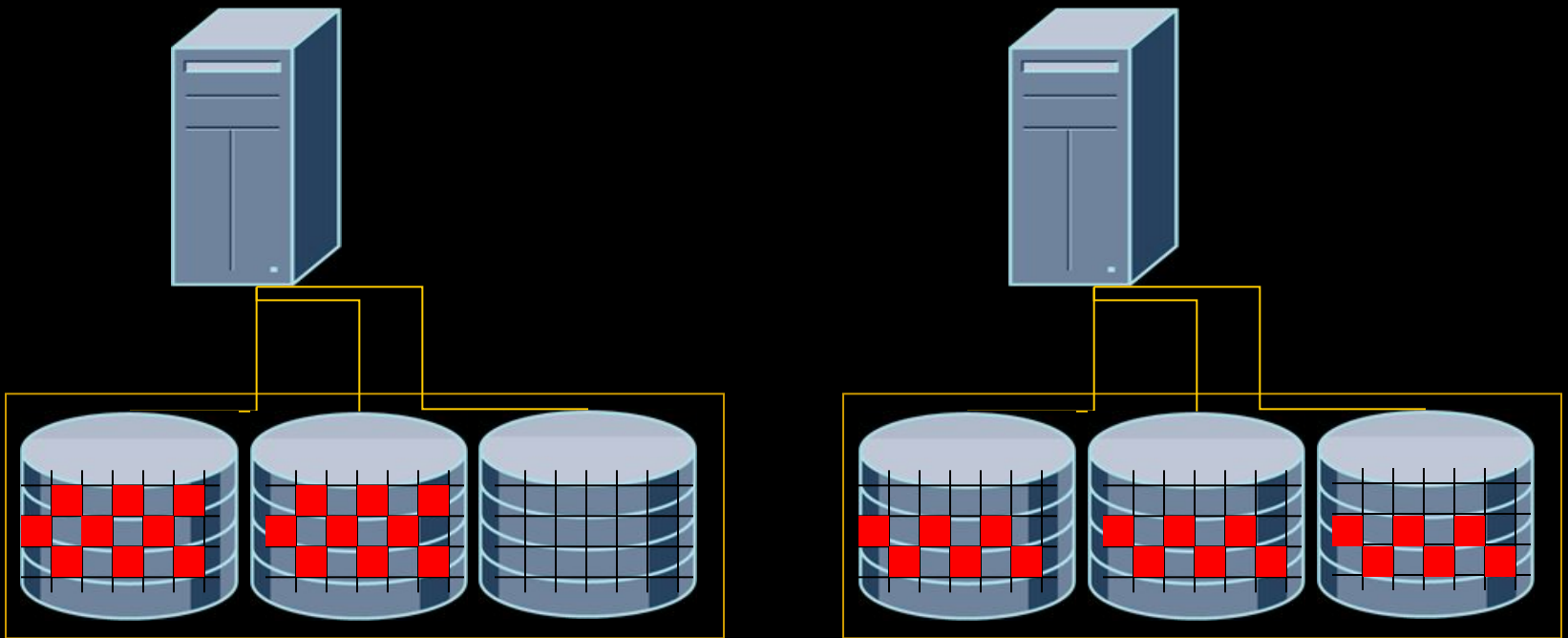
- Проблема: изолированные устройства хранения данных
  - “Утилизация емкости устройства хранения данных составляет 50%, но возрастает ежегодно на 30%”
- Решение: ASM

# Управление системой хранения данных\*

- Переносимая файловая система
- Нет томов: только виртуальный диск
- Высокая гранулярность
  - Диск разбит на фрагменты по 1 Мбайт каждый
- Исключается необходимость использования традиционных файловых систем и менеджеров томов (volume manager)
- Автоматическое зеркалирование (mirroring)
- Автоматическая настройка ввода/вывода
  - Распределяет данные по дискам, чтобы сбалансировать использование их емкостей

# ASM: Добавление нового устройства

10<sup>8</sup>



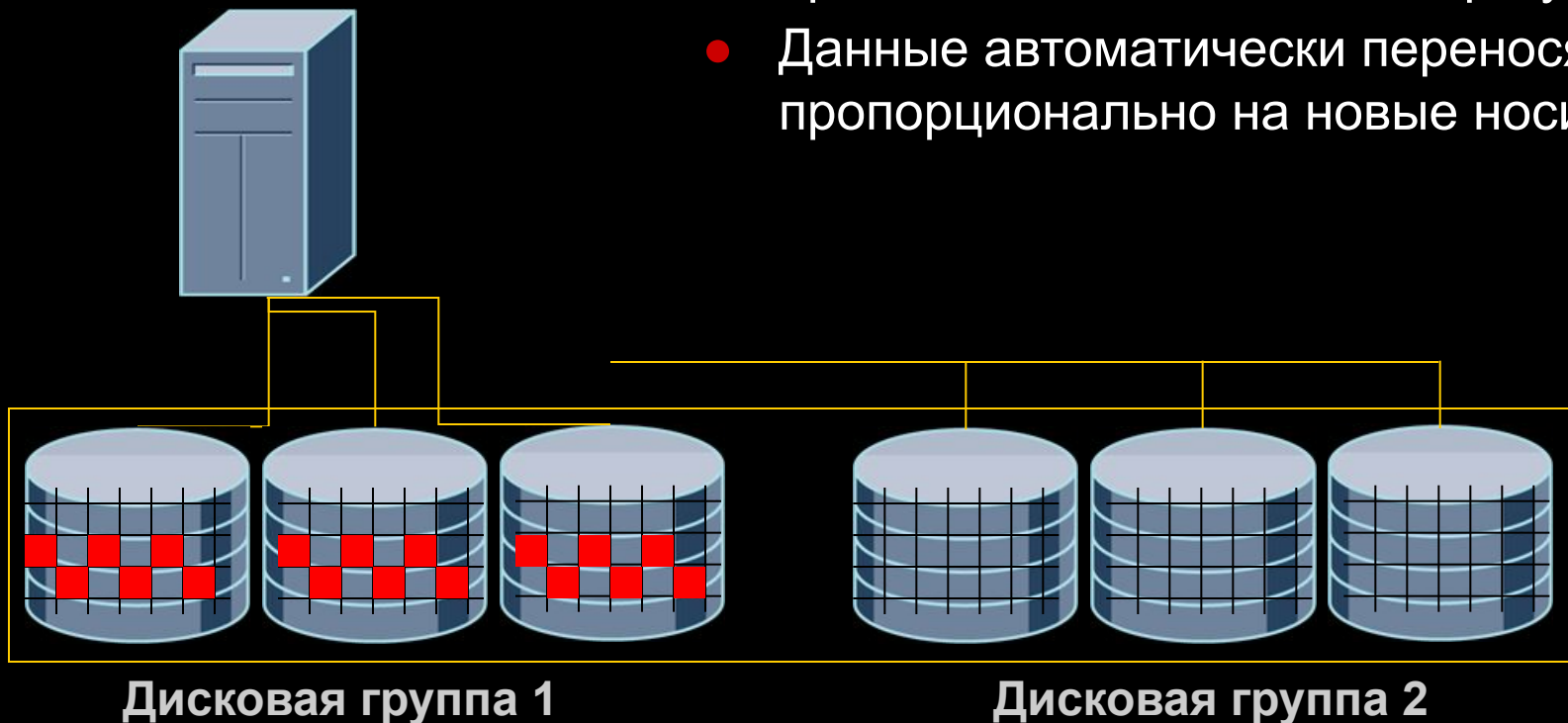
Сбалансированное использование  
суммарной емкости устройств  
хранения

# ASM - Добавление новой группы устройств хранения

10<sup>g</sup>

1

- Автоматическая балансировка «на лету» при любых изменениях конфигурации
- Данные автоматически переносятся пропорционально на новые носители

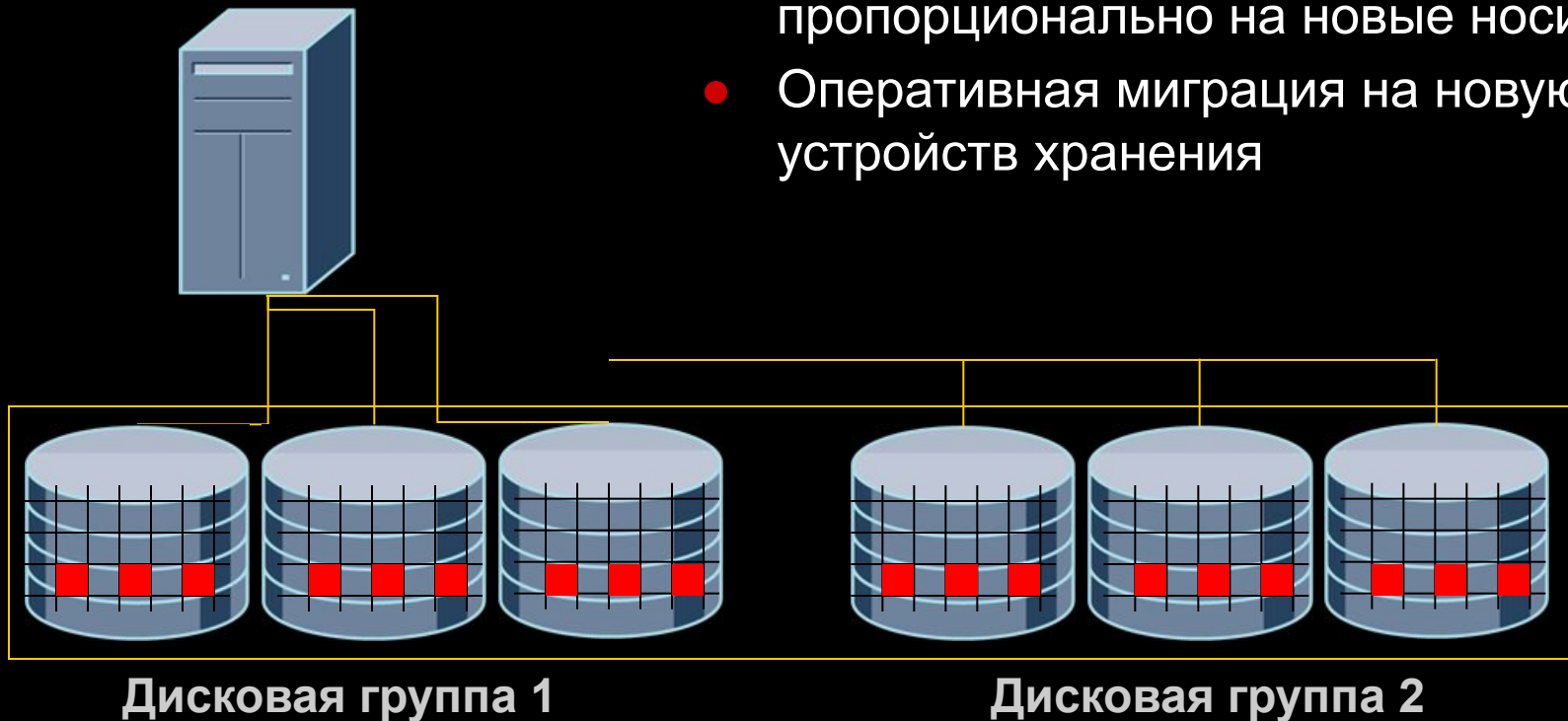


# ASM - Добавление новой группы устройств хранения

10<sup>g</sup>

2

- Автоматическая балансировка «на лету» при любых изменениях конфигурации
- Данные автоматически переносятся пропорционально на новые носители
- Оперативная миграция на новую группу устройств хранения

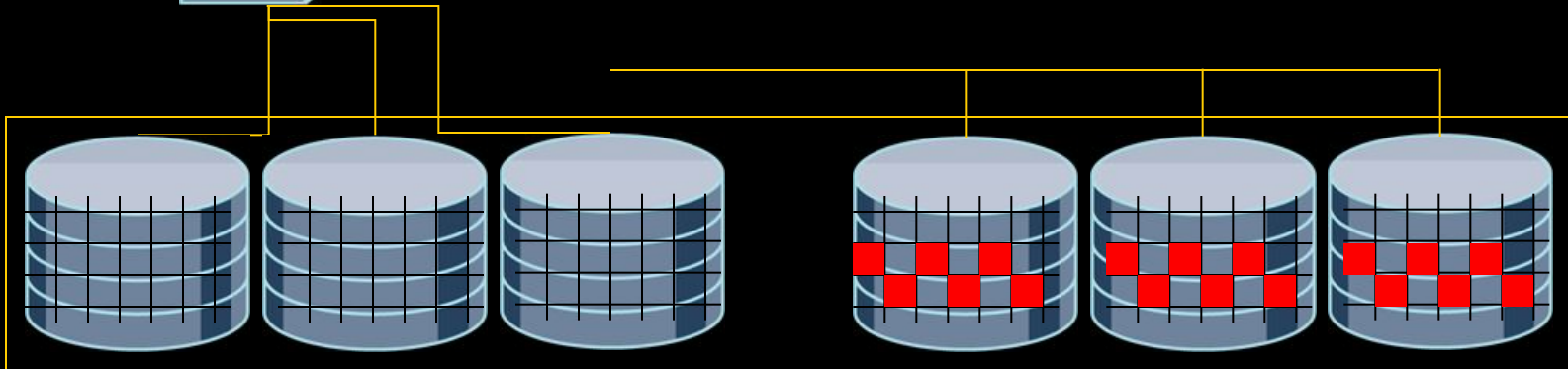


# ASM - Добавление новой группы устройств хранения

10<sup>g</sup>

3

- Автоматическая балансировка «на лету» при любых изменениях конфигурации
- Данные автоматически переносятся пропорционально на новые носители
- Оперативная миграция на новую группу устройств хранения



Дисковая группа 1

Дисковая группа 2

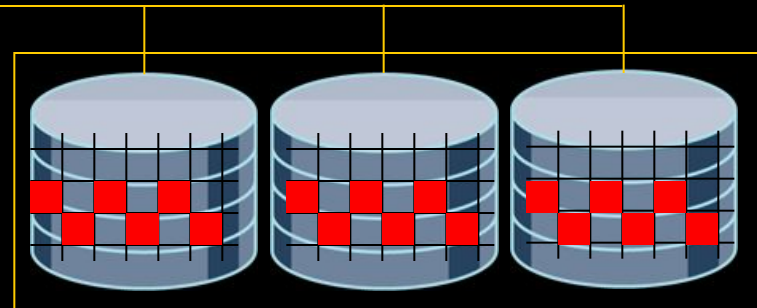


# ASM - Добавление новой группы устройств хранения

10<sup>g</sup>

4

- Автоматическая балансировка «на лету» при любых изменениях конфигурации
- Данные автоматически переносятся пропорционально на новые носители
- Оперативная миграция на новую группу устройств хранения

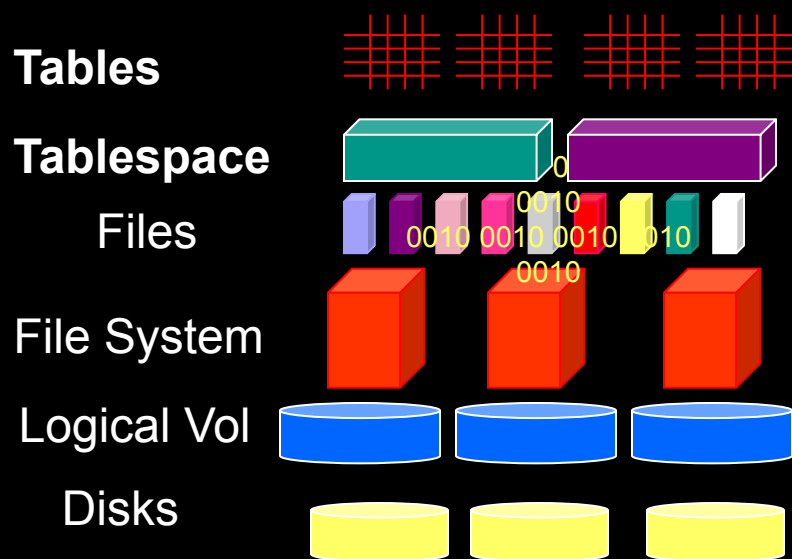


Дисковая группа 2

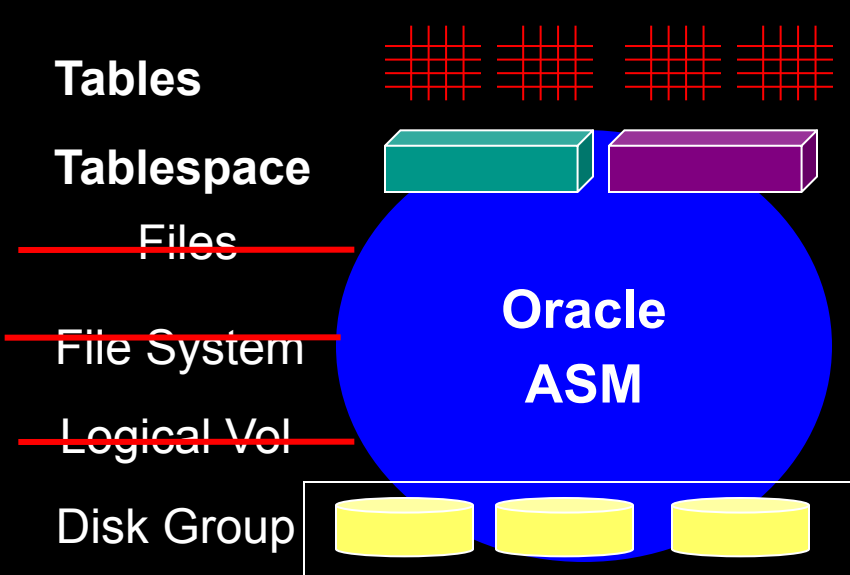
# Стек технологий

10<sup>g</sup>

## СЕГОДНЯ



## ASM



“Лучший способ снизить стоимость управления –  
уменьшить сложность”

# Сокращение затрат на управление системой хранения данных

10<sup>8</sup>



# Поддержка сверхбольших баз данных

10<sup>8</sup>



- Размер базы данных увеличен до миллионов терабайт (8 Exabytes)
- Неограниченный размер данных типа LOB (терабайты)
- Ultra Large Data File – терабайты в одном файле
- Хранение данных в файлах: тип данных BFILE
- Хранение таблиц в файлах: внешние таблицы (external tables)

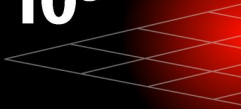
# Поддержка сверхбольших баз данных

10<sup>8</sup>

Организация	Платформа	Размер БД	#DBAs
France Telecom	HP Superdome	49 TB	2
Acxiom	HP Alpha GS320	12 TB	2
Acxiom	HP Alpha GS320	6 TB (9i RAC)	2
Amazon.com	HP Superdome	13 TB	2

ORACLE

[http://www.wintercorp.com/vldb/2003\\_TopTen\\_Survey/TopTenWinners.asp](http://www.wintercorp.com/vldb/2003_TopTen_Survey/TopTenWinners.asp)



# Database Grid

Корпоративная сеть серверов  
баз данных

# Oracle Database Grid

- Проблема: «островки» обработки данных
  - “Вычислительные ресурсы серверов баз данных используются только на 15%“
- Решение: консолидация процессов обработки данных, объединение серверов промышленных стандартов в вычислительные пулы
  - Стандартные серверы невысокой стоимости (Intel, Linux)
  - Вычислительные мощности по требованию
  - Кластеры баз данных (shared disk)

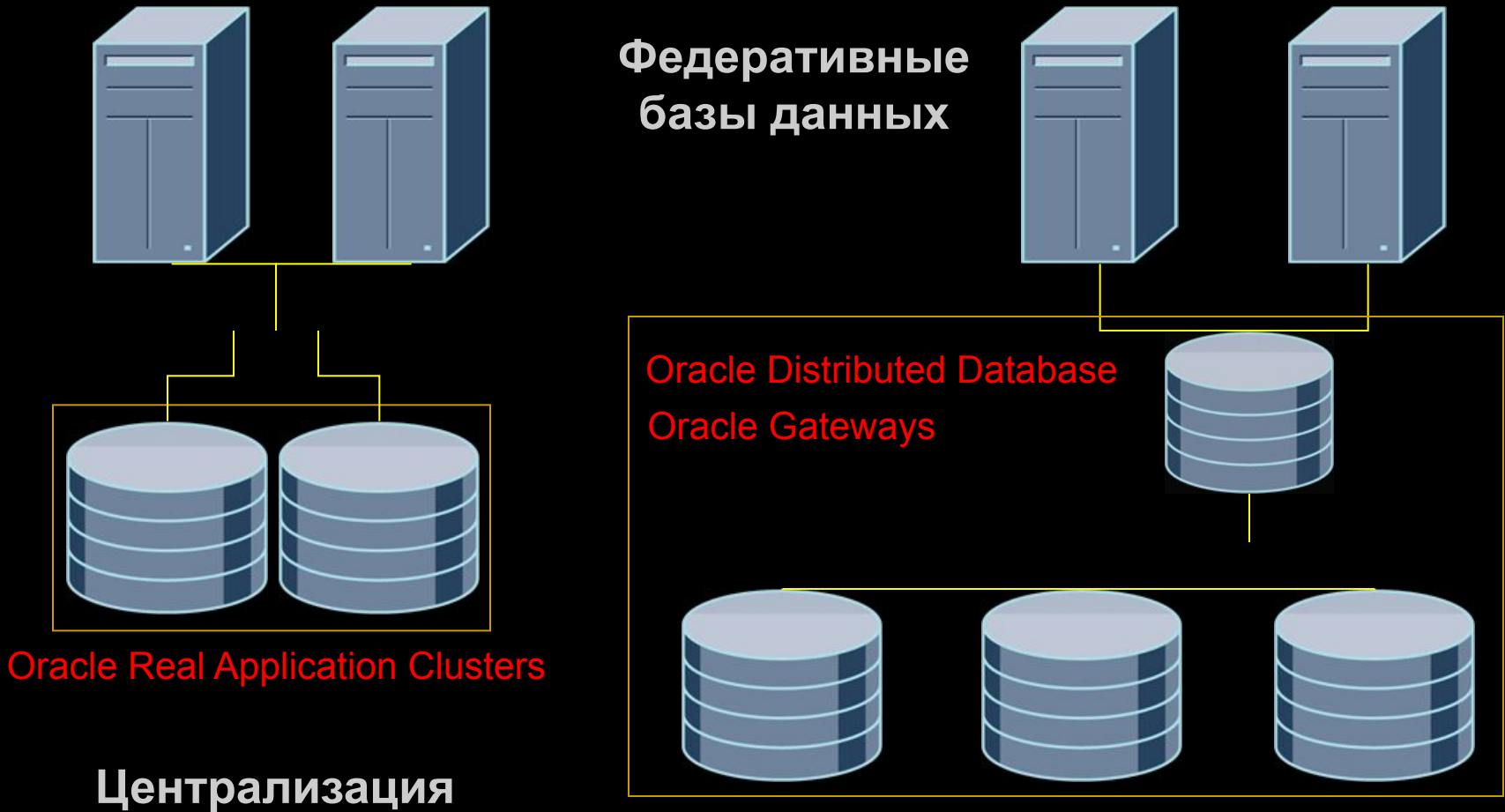
# Три способа организации данных в Enterprise Grid

10<sup>g</sup>

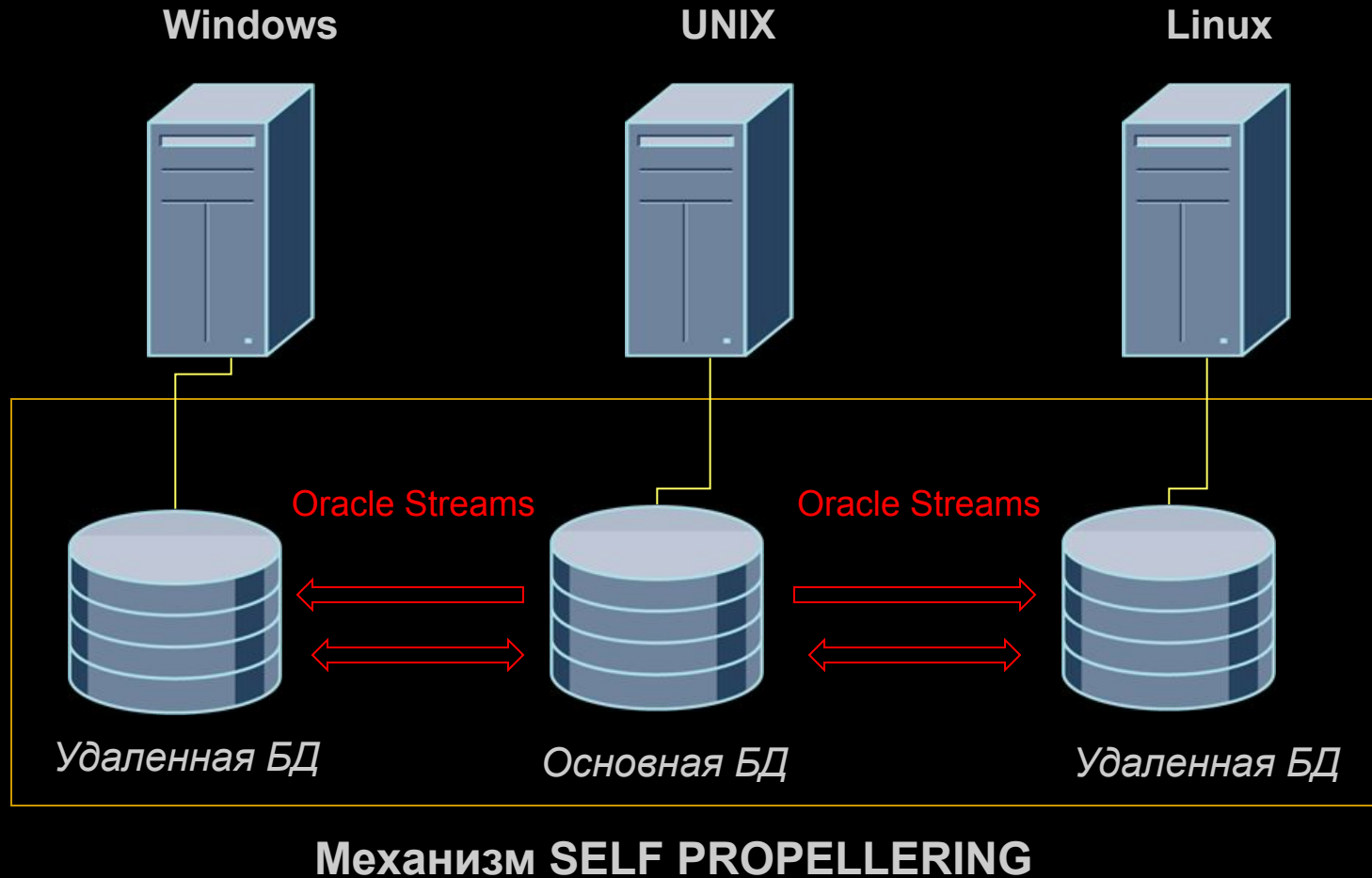
- Централизация данных в единой базе данных
  - Кластеризация на основе Oracle 10g Real Application Clusters
- Множество самостоятельных независимых баз данных
  - Федеративные (распределенные) базы данных
- Временный перенос данных из основной базы данных в удаленную
  - Перемещение блоков данных (табличные пространства) в удаленную базу данных
  - Последующая синхронизация блоков данных в основной и удаленной базах данных



# Топология Enterprise Grid

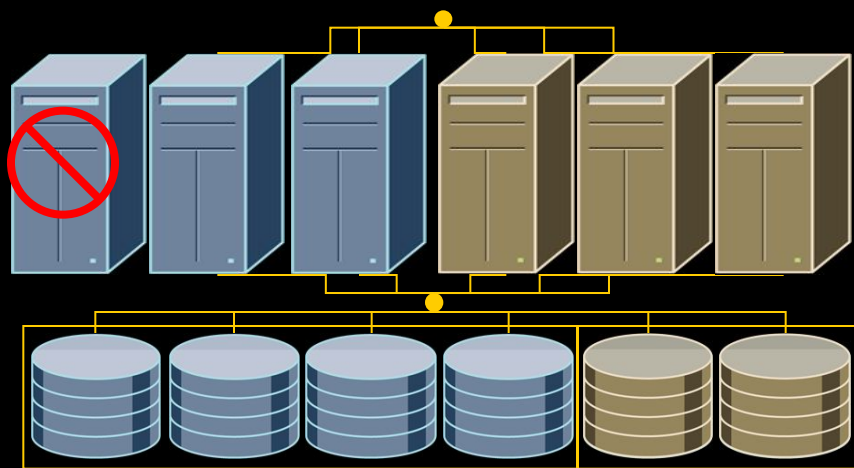


# Топология Enterprise Grid



# Enterprise Grid: СИСТЕМА ВЫСОКОЙ ГОТОВНОСТИ

10<sup>g</sup>



- Стандарт Enterprise Grid: недорогие узлы и модульные дисковые массивы
- Основа - Oracle Real Application Clusters
- Восстановление после сбоя может произойти на любом узле Grid
- Балансировка нагрузки в Grid

# Enterprise Grid: Управление нагрузкой

10<sup>g</sup>

1



# Enterprise Grid: Управление нагрузкой

10<sup>g</sup>

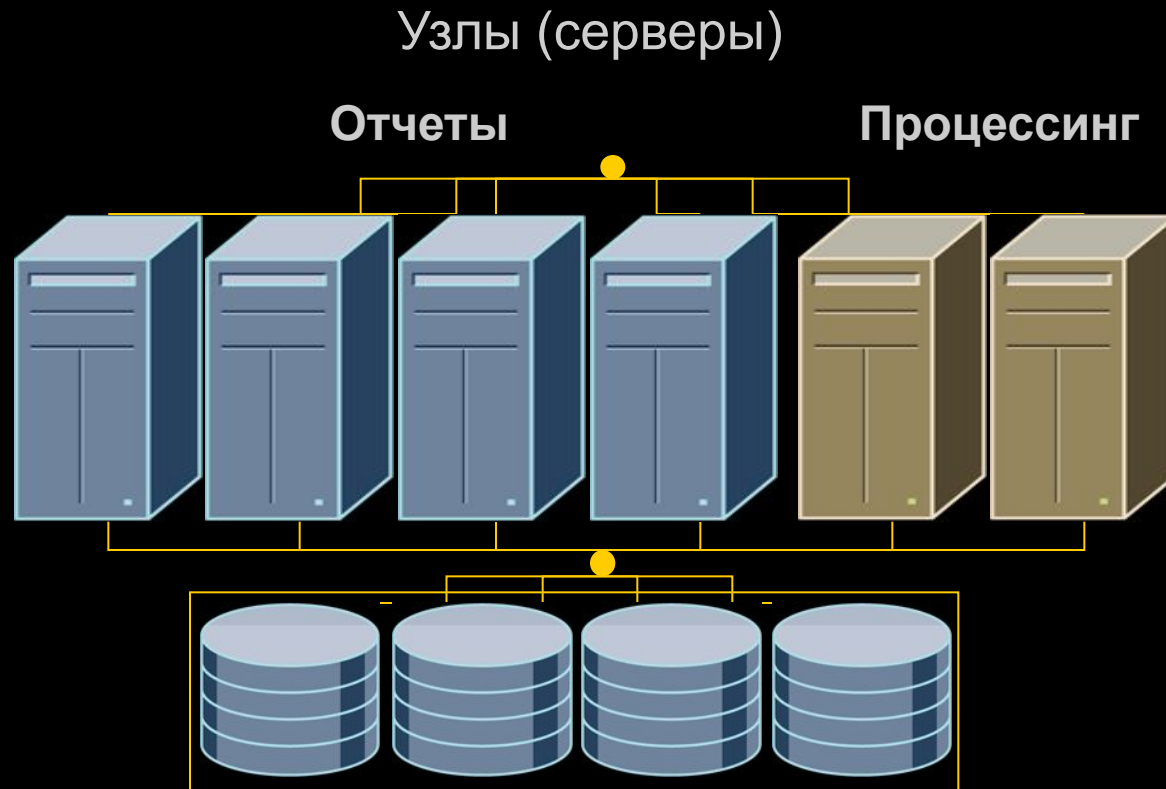
2



# Enterprise Grid: Управление нагрузкой

10<sup>g</sup>

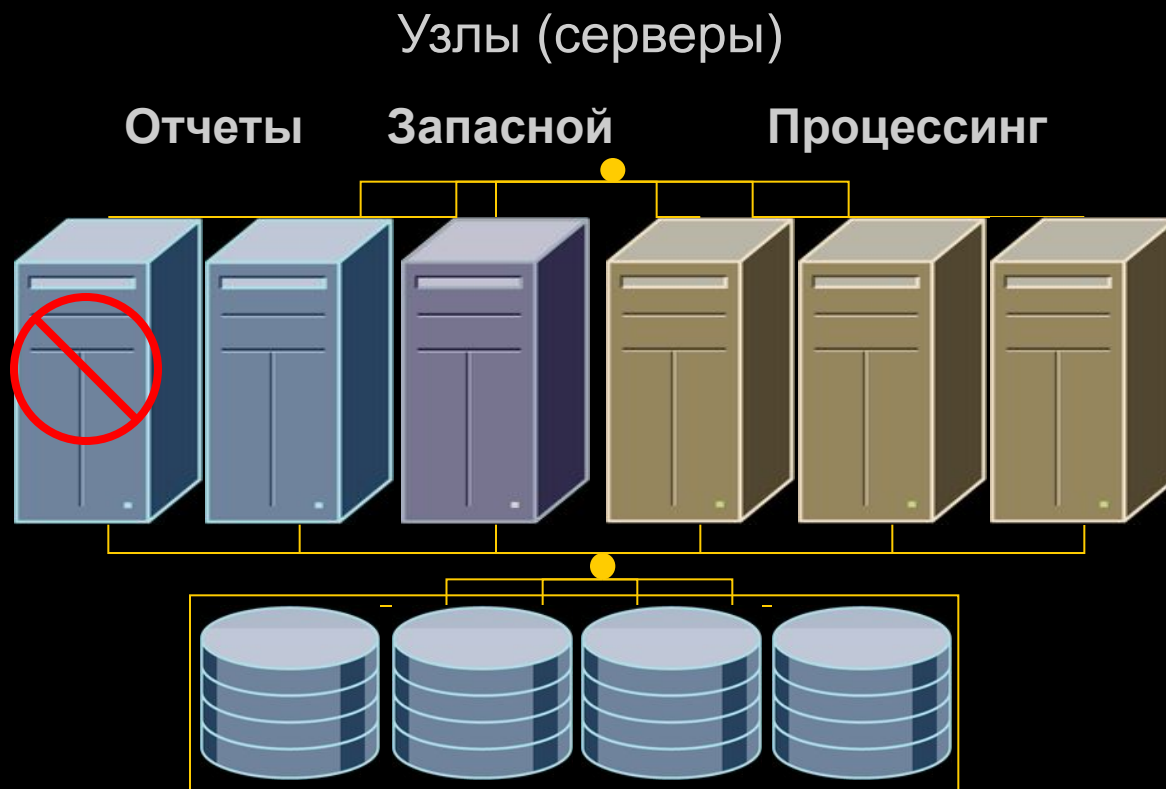
3



# Enterprise Grid: Восстановление после сбоя

10<sup>g</sup>

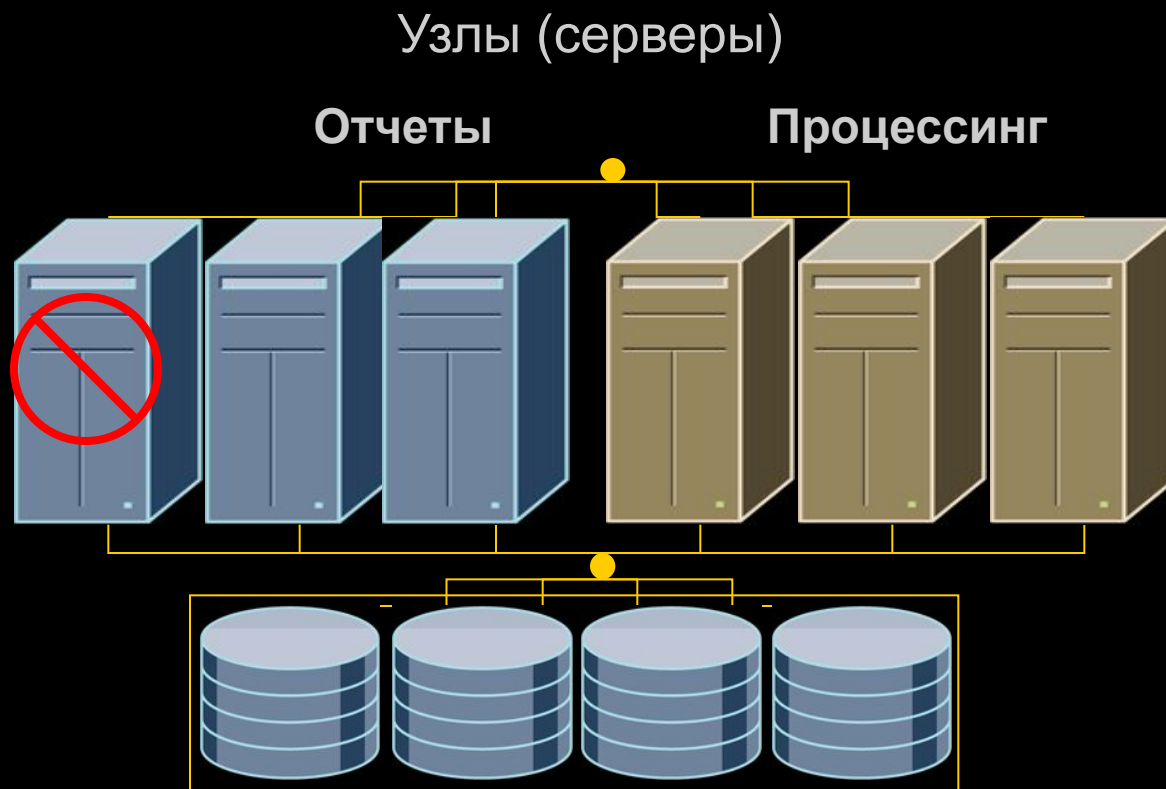
1



# Enterprise Grid: Восстановление после сбоя

10<sup>g</sup>

2



Миграция сервисов на резервный узел



# Enterprise Grid: Восстановление после сбоя

10<sup>g</sup>

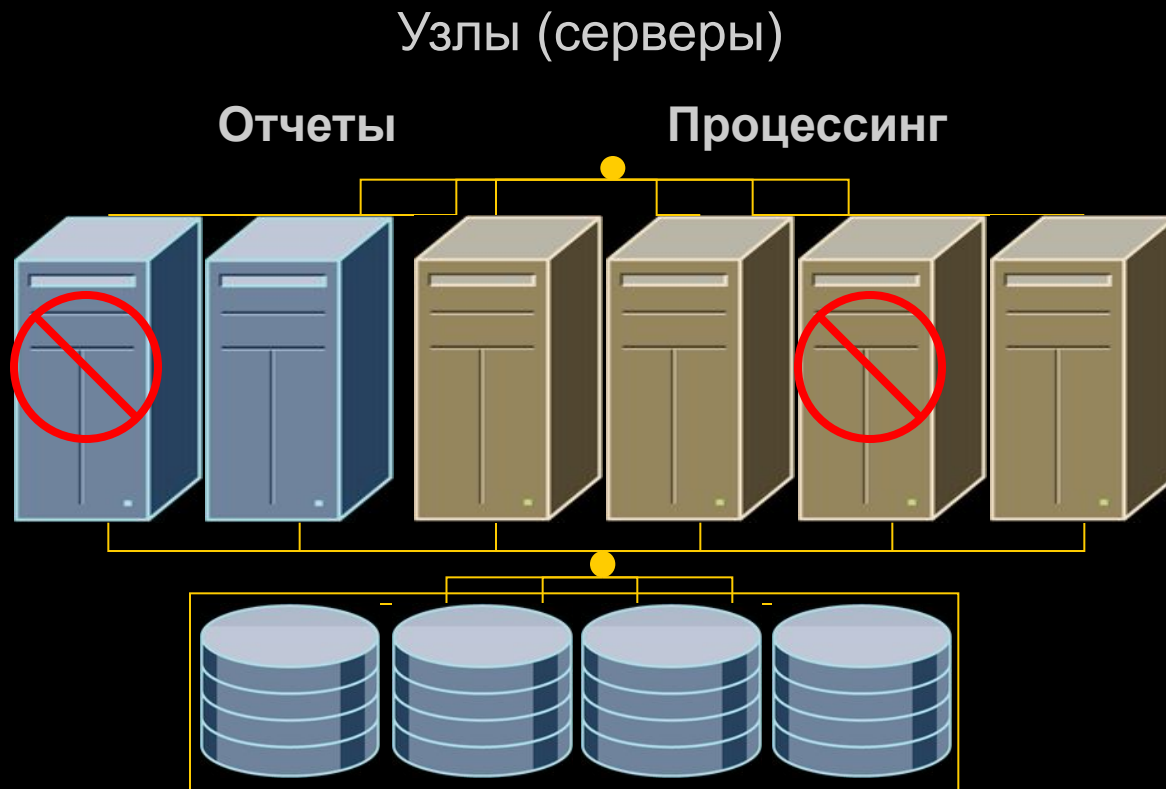
3



# Enterprise Grid: Восстановление после сбоя

10<sup>g</sup>

4



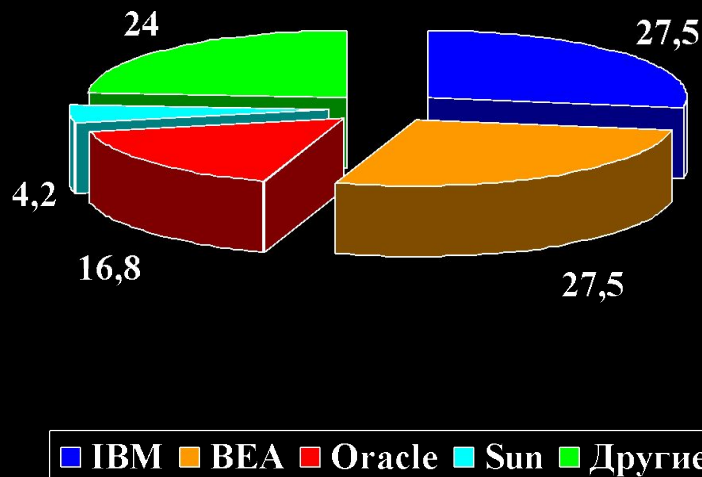
# Application Server Grid

Корпоративная сеть серверов  
приложений

# Oracle Application Server

10<sup>g</sup>

Worldwide Application Server Platform Software Market\*



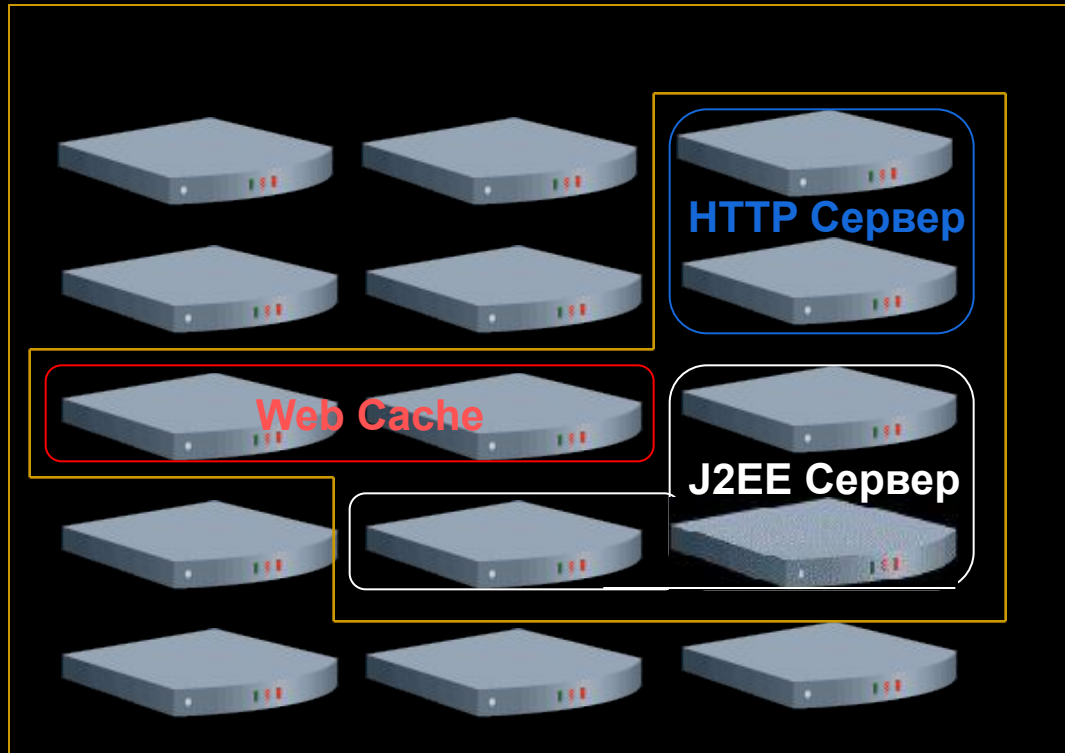
- Свыше 16 тыс. заказчиков
- Более 1000 компаний – разработчиков прикладного ПО
- Свыше 1400 системных интеграторов
- 99 технологических патентов

ORACLE

# Application Server Grid

- Кластеризация сервера приложений
  - Кластеризация всех типов серверов: HTTP, Cache, J2EE
  - Автоматическое удаление/добавление узлов в кластер
  - Автоматическое распределение нагрузки между узлами кластера
- Балансировка нагрузки на основе:
  - Метрик потребления ресурсов
  - Метрик, специфические для конкретных приложений
  - Планов и политик

# Виртуализация сервисов промежуточного слоя



Инфраструктура прикладной системы:  
объединение ресурсов и сервисов

- Серверы:
  - HTTP
  - Web Cache
  - J2EE
- Процессы
  - EJB
- Сервисы
  - Portal
  - Wireless Services
  - Web Services
  - Directory Services
  - Authentication
  - Authorization

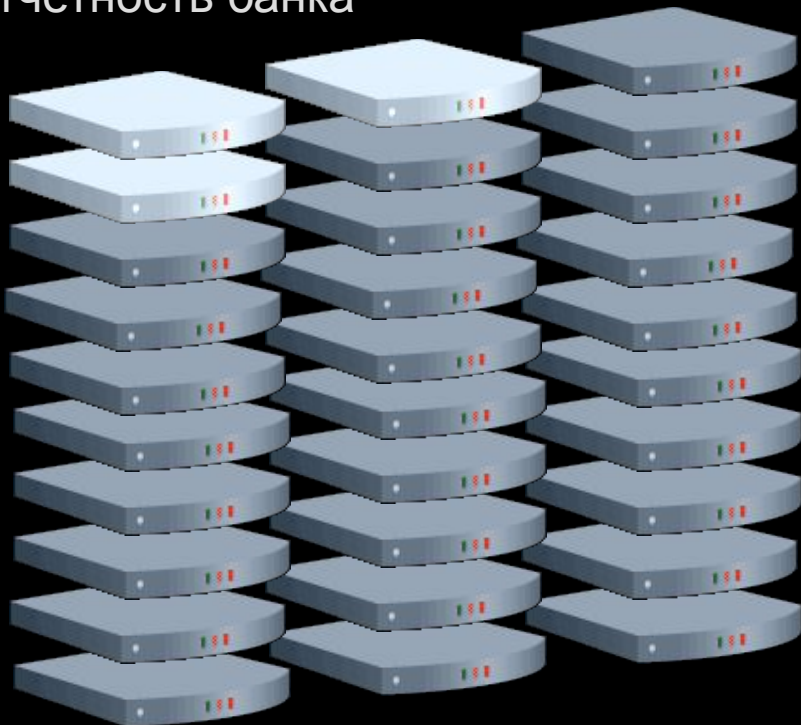
# Запланированное перераспределение ресурсов

10<sup>8</sup>

Начало месяца:



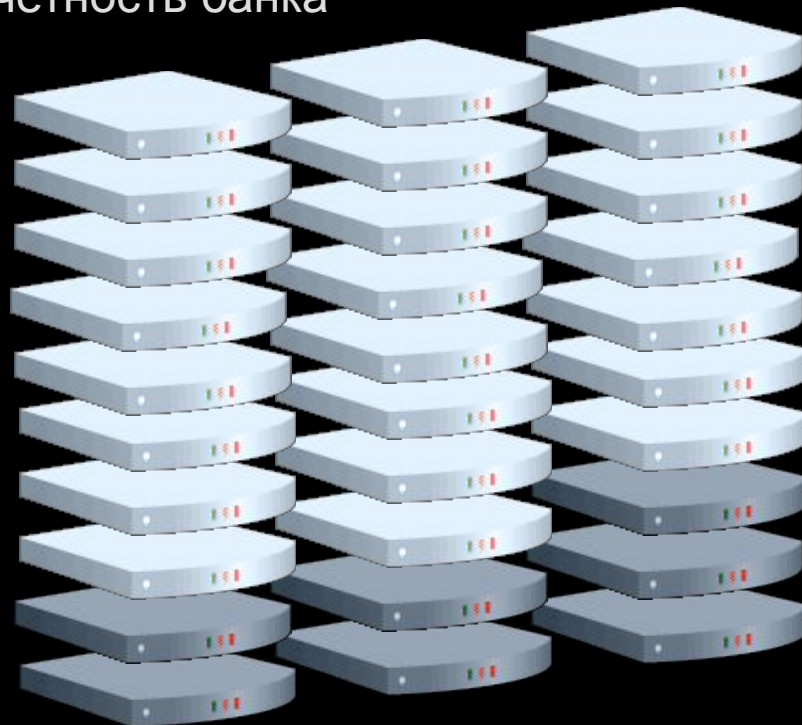
Отчетность банка



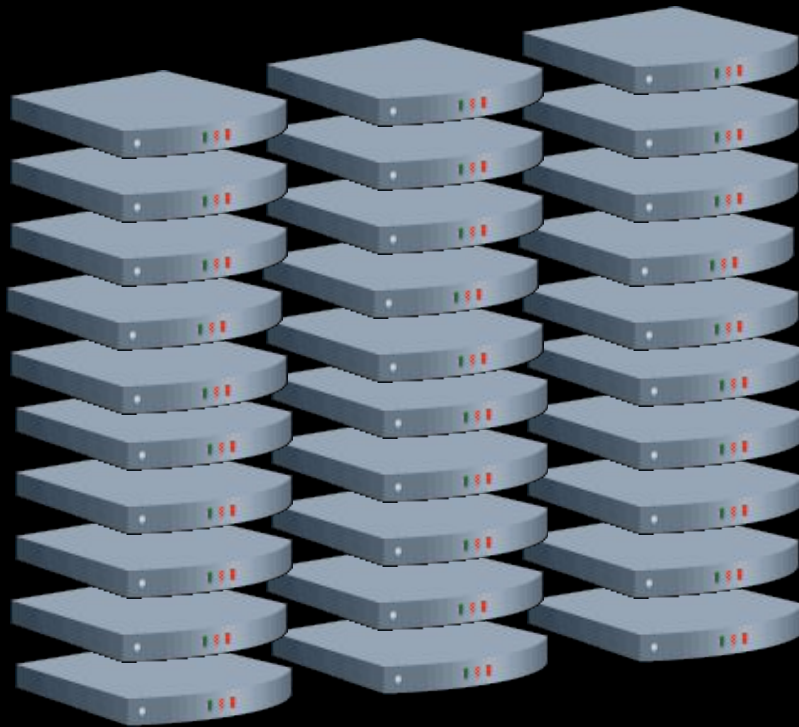
Конец месяца:



Отчетность банка



# Управление на основе политик



## Менеджер нагрузки

**Диспетчер и Планировщик**  
*Распределяют нагрузку на основе политик*

**Менеджер политик**  
*Управление репозиторием политик*

**Менеджер ресурсов**  
*Управление статусом и доступом к ресурсам*



# Информационная безопасность

- Централизованное администрирование и отзыв пользователей
  - Централизованное управление привилегиями пользователей для доступа к ресурсам и определение тождественности пользователя
  - Обеспечение возможности однократной регистрации для доступа к корпоративным приложениям
  - Разделение политик безопасности в рамках единой инфраструктуры

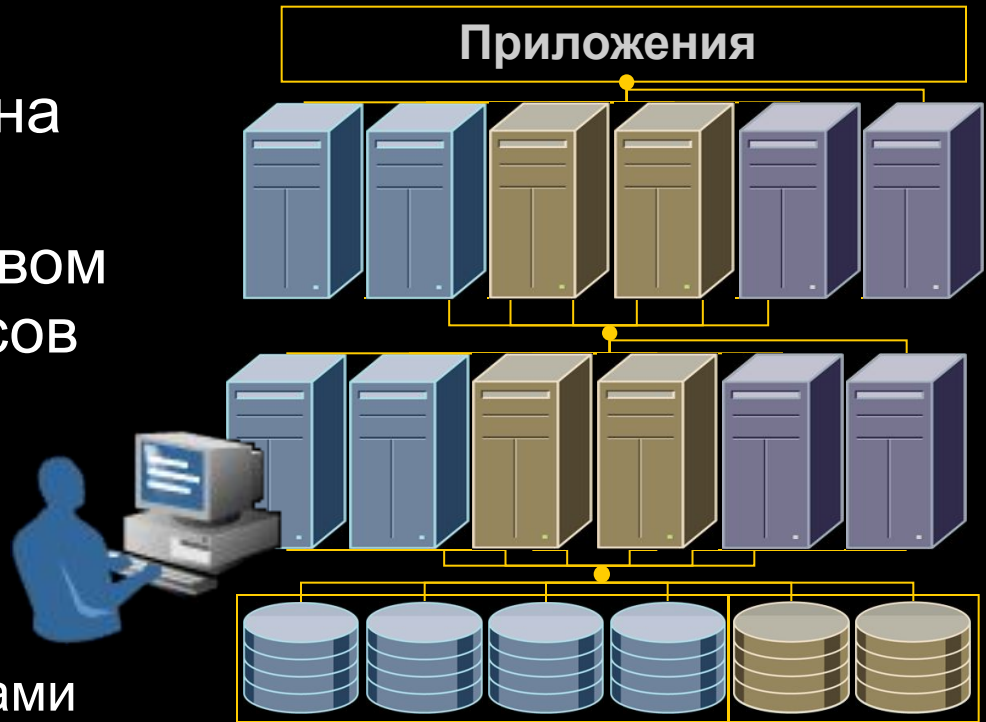
# Grid Control

Управление корпоративной  
сетью распределенной  
обработки данных

# Функции Grid Control

10<sup>g</sup>

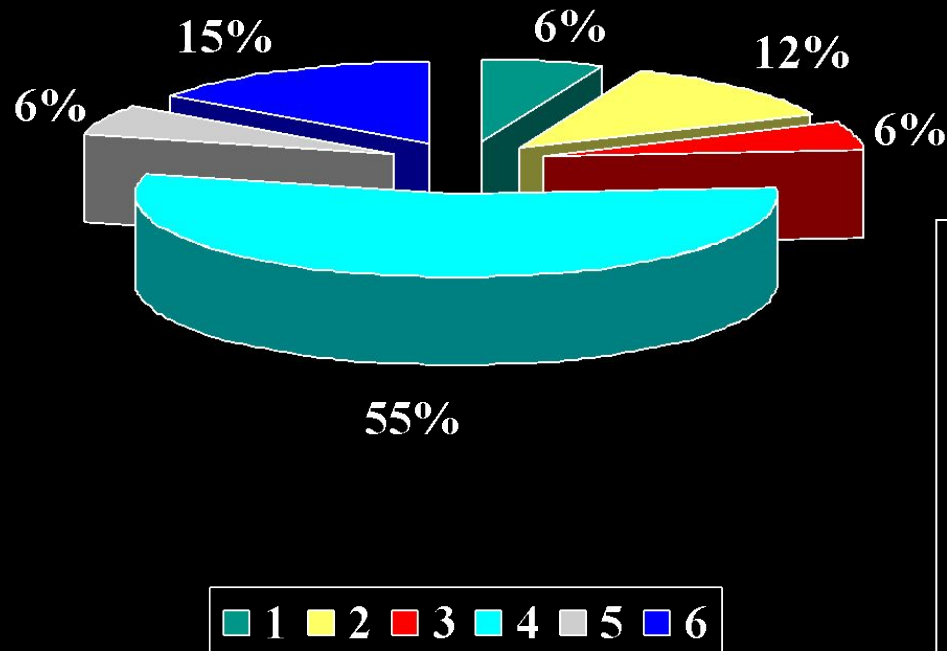
- Управление структурой Grid
- Распространяется на тысячи серверов
- Управление качеством прикладных сервисов
- Управление конфигурациями
  - Снабжение ПО
  - Клонирование
  - Управление патчами



# Управление конфигурациями

- Исследование и анализ конфигураций
  - Полная информация о всем ПО Oracle (версии, патчи)
  - Детали конфигураций продуктов Oracle
  - Поиск и сравнение конфигураций
  - Отслеживание изменений конфигураций
  - Эталонные конфигурации
- Автоматическое клонирование ПО
  - Массовое тиражирование эталонных конфигураций
- Автоматическое управление патчами
  - Оперативное обнаружение новых патчей
  - Автоматическая доставка и применение

# Затраты на сопровождение ПО



- 1 Установка
- 2 Создание и конфигурирование
- 3 Загрузка данных
- 4 Текущее управление системой и ее настройка
- 5 Поддержка программного обеспечения
- 6 Прочее

# Самоуправляемая СУБД

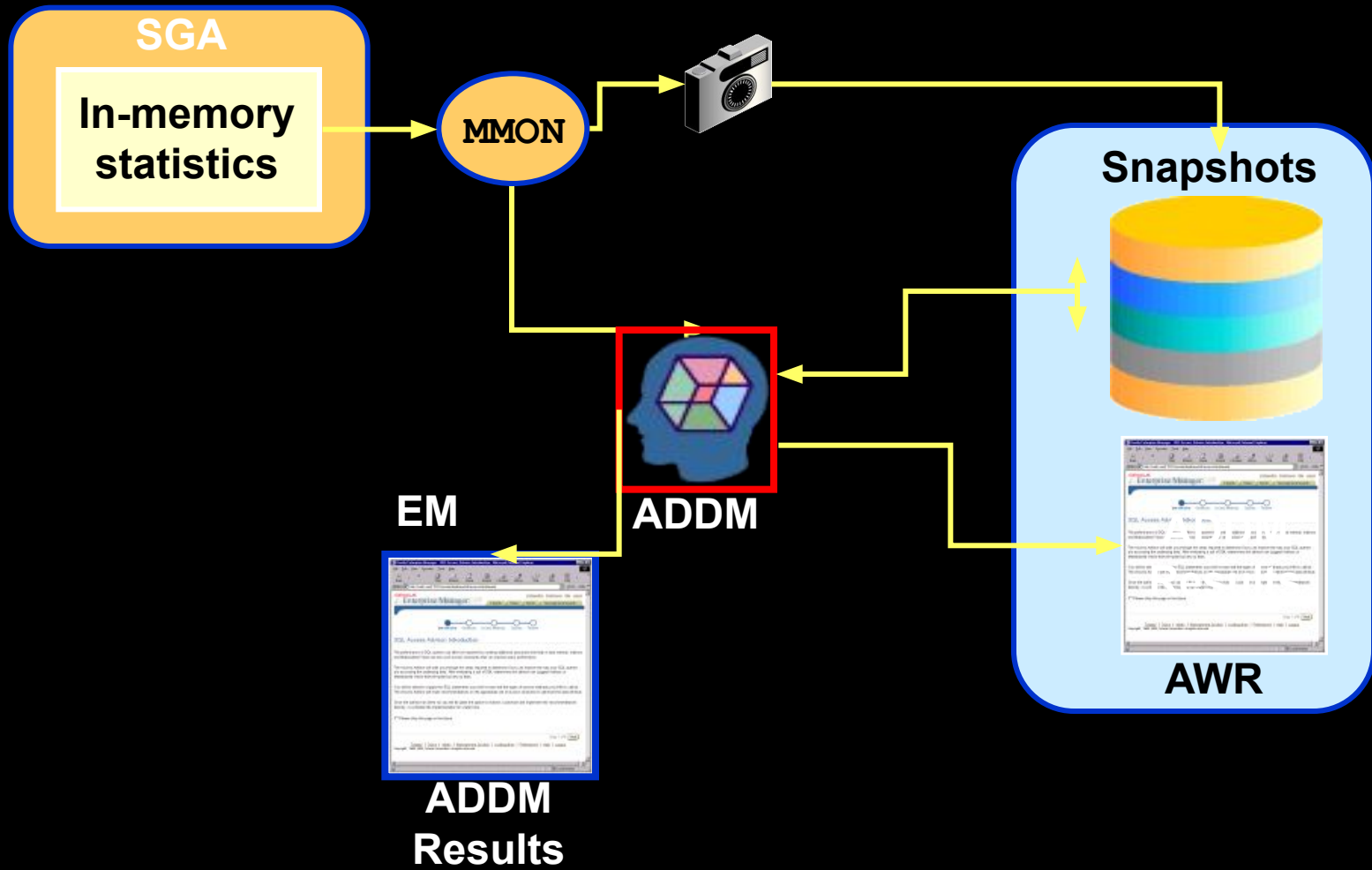


Инфраструктура самоуправления

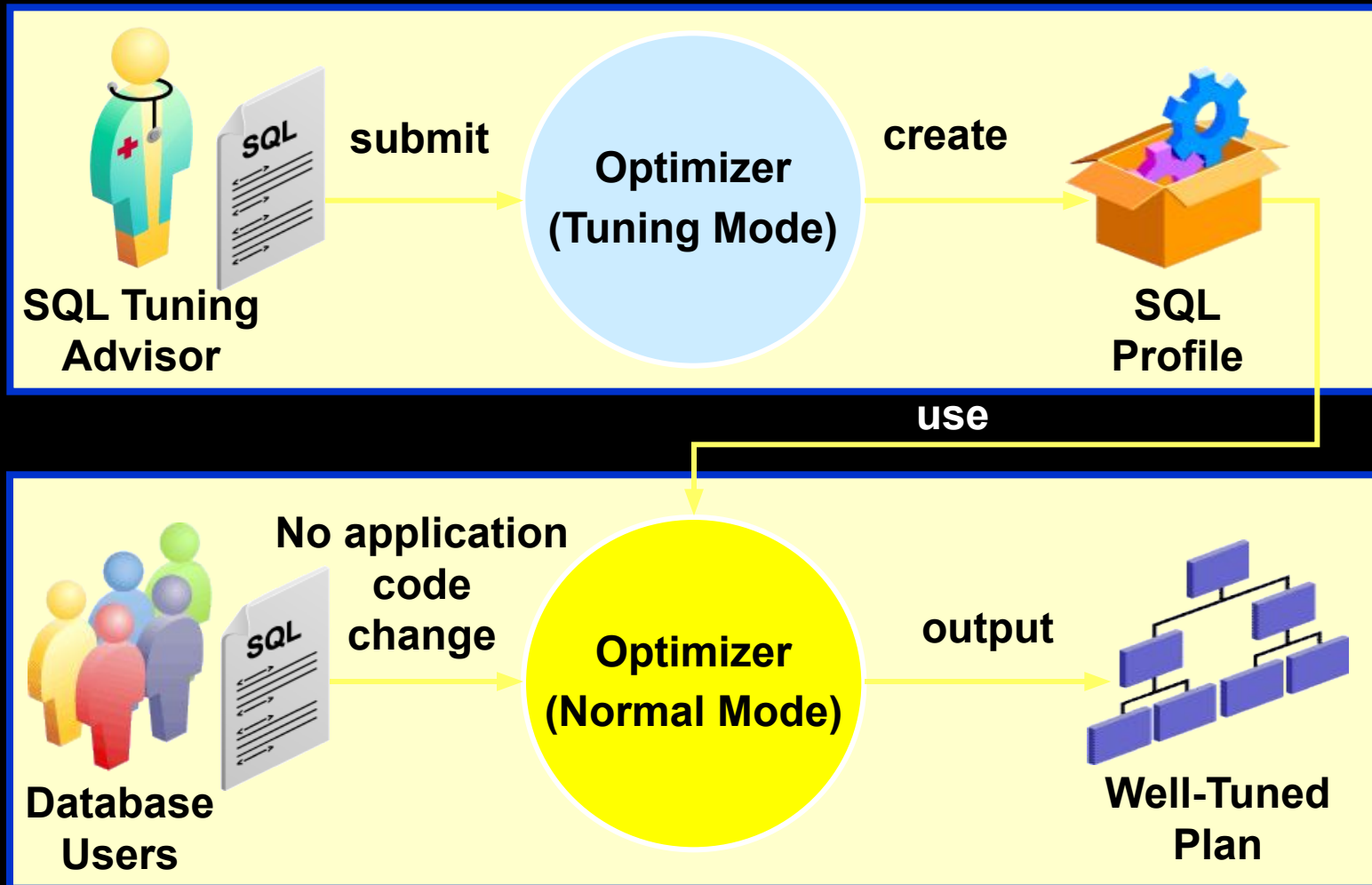
- Функции встроенной системы самоуправления
  - Настройка
  - Фиксация проблем
  - Извещения
  - Консультации

Ресурсы:  
Automatic Workload Repository  
Automatic Diagnostic Monitor  
Automatic Maintenance Task  
Advisory Infrastructure

# ADDM Performance Monitoring



# Plan Tuning Flow





# Database Performance

10<sup>g</sup>

## World Record TPC-C

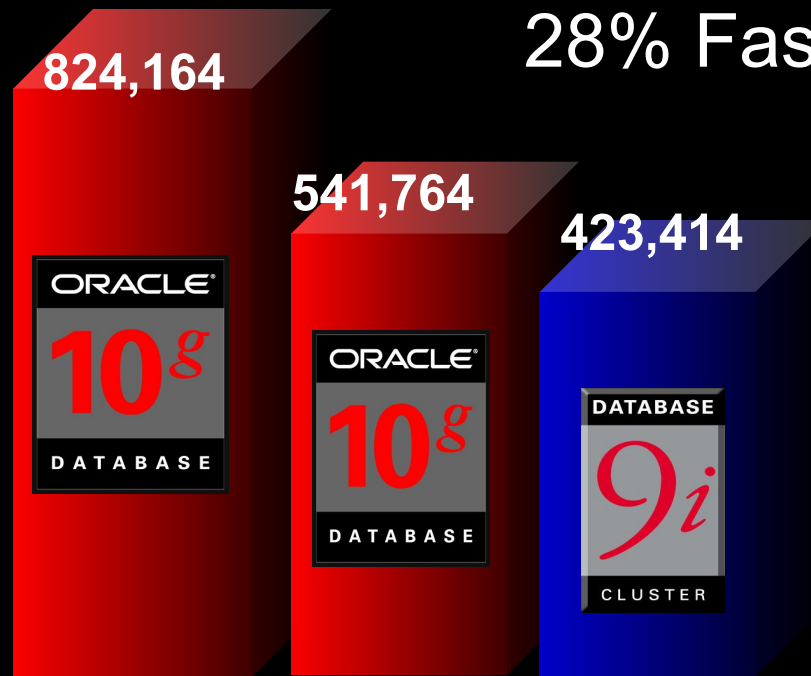


ORACLE

# Database Performance

10<sup>g</sup>

World Record TPC-C  
28% Faster than Oracle9i

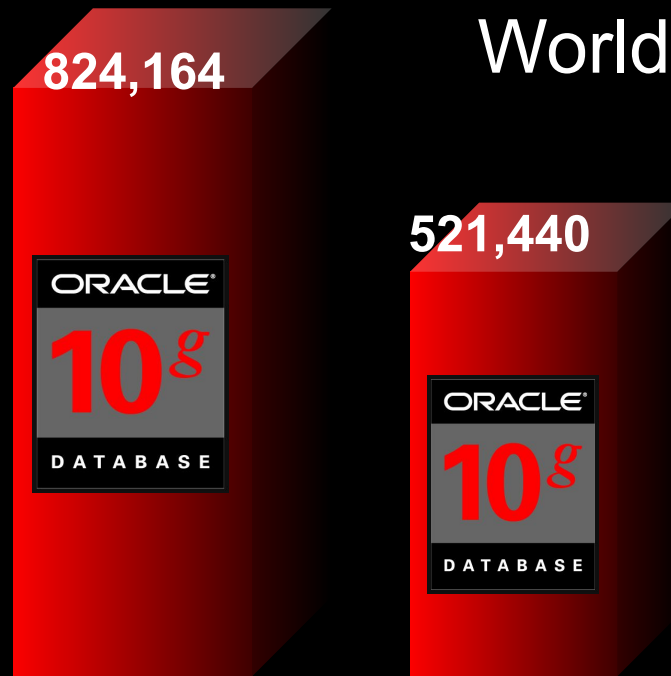


ORACLE

# Database Performance

10<sup>g</sup>

## World Record TPC-C World Record on Windows



ORACLE

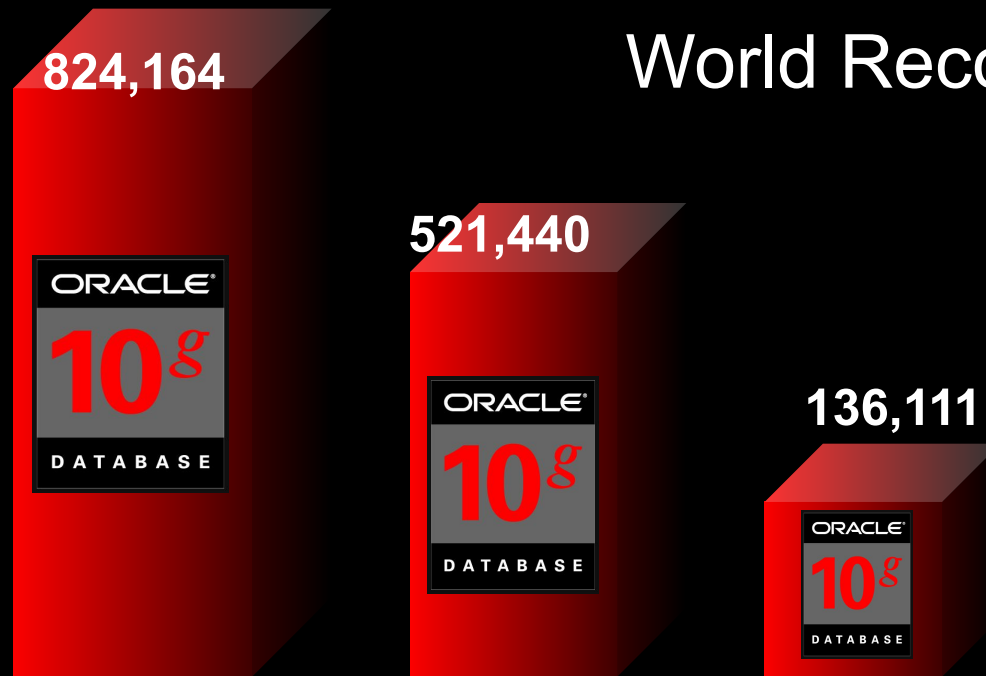
As of September 5, 2003: Oracle Database 10g Enterprise Edition, HP Integrity Superdome, 824,164.53 tpmC, \$8.28/tpmC, available 12/31/03. NEC Express5800/1320Xd, Oracle Database 10g Enterprise Edition, 521,440.53 tpmC, \$11.77/tpmC, available February 15, 2004.

Source: Transaction Processing Council (TPC) [www.tpc.org](http://www.tpc.org)

# Database Performance

10<sup>g</sup>

World Record TPC-C  
World Record on Linux



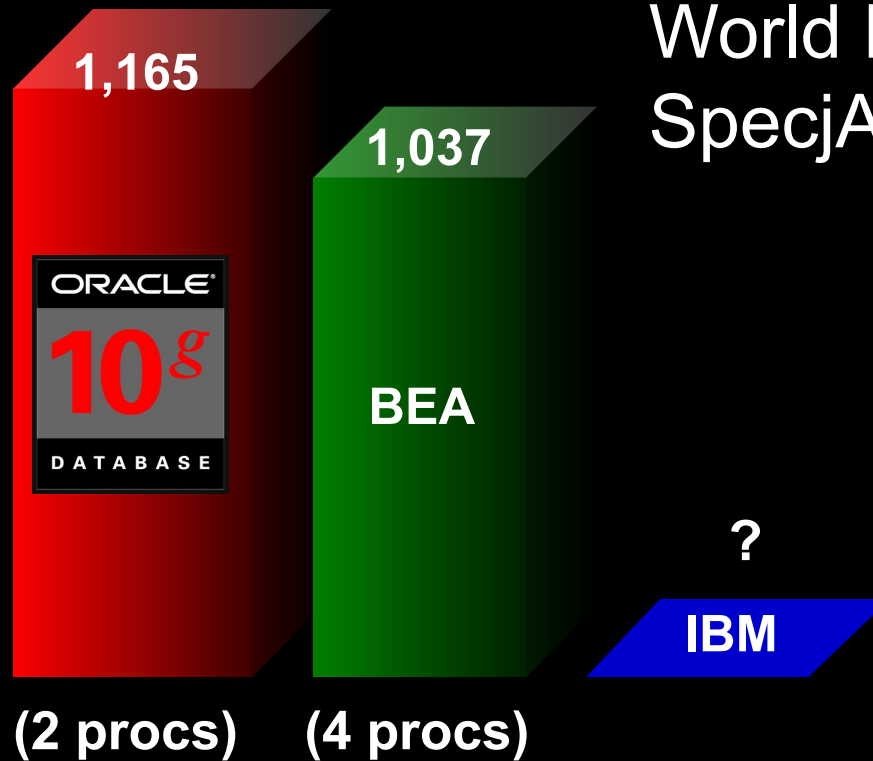
ORACLE

As of September 5, 2003: Oracle Database 10g Enterprise Edition, HP Integrity Superdome, 824,164.53 tpmC, \$8.28/tpmC, available 12/31/03. NEC Express5800/1320Xd, Oracle Database 10g Enterprise Edition, 521,440.53 tpmC, \$11.77/tpmC, available February 15, 2004. HP Integrity rx5670, (Oracle Database 10g Standard Edition), 136,110.98 tpmC, \$4.09/tpmC, available December 31, 2003.

Source: Transaction Processing Council (TPC) [www.tpc.org](http://www.tpc.org)

# Application Server Performance

10<sup>g</sup>



World Record Linux  
SpecjAS2002

ORACLE

# Oracle Grid

Примеры построения корпоративных  
систем распределенной обработки  
данных

# Electronic Arts

- Компания – производитель игровых систем
- Игра «The Sims Online» на Oracle
- 18 кластеров баз данных (Intel-Linux)
  - 100,000 - 150,000 пользователей
  - 30,000 SQL-операторов в секунду
- Существенная экономия (\$M) за счет использования серверов невысокой стоимости и подхода «pay-as-you-grow»
- Высокая надежность – неременный атрибут игровых систем

# CERN

- Европейская организация по ядерным исследованиям (Conseil Europeen de la Recherche Nucleaire - CERN)
  - На оборудовании CERN работает 7 тыс. ученых из 500 научных центров
- Большой Адронный Коллайдер (Large Hardon Collider – LHC)
  - Крупнейшая в мире установка для ускорения, накопления и столкновения пучков частиц сверхвысоких энергий
  - Длина вакуумного кольца – 27 км., глубина тоннеля – 100 м.
  - Запуск в 2007 году



# CERN Openlab for Grid Applications

10<sup>g</sup>



- Партнеры лаборатории:
  - Entrasys (сетевое оборудование), HP (серверы), IBM (устройства хранения), Intel (процессоры Itanium), Oracle (программное обеспечение Grid)
- Крупнейший в мире вычислительный комплекс\*
  - Около 3,300 серверов, от 50 до 10 тыс. процессоров
  - Проектный объем базы данных – минимум 15 Петабайт
- Используются программные продукты Oracle
  - Oracle9i Database
  - Oracle9i Real Application Clusters
  - Oracle9i Application Server

ORACLE

*\*По классификации GRID – разделение ресурсов*

# Oracle University Grid

10<sup>g</sup>

- 240 Linux-серверов для 6 тыс. студентов

## 1999 год

- 2 дня на установку недельного курса
- 108 часов в неделю на системное администрирование



## 2003 год



- 13 секунд на курс, одновременная установка всех курсов
- 2 часа в неделю

# Oracle Outsourcing Grid

- Центр аутсорсинга Oracle E-Business Suite (г.Остин, Техас)
  - Более 500 серверов поддерживают Outsourcing Grid
  - Свыше 1000 Oracle DB, объем баз данных - 80TB
- Каждому заказчику предоставлен набор серверов (slice of grid)
  - Дополнительные slice of grid выделяются из пула зарезервированных серверов
  - Использован Oracle9i Real Application Clusters
- По классификации Grid – выделение ресурсов

# Oracle Grid в России

- Значение Grid: Mainstream развития инфраструктуры
- На сегодняшний день:
  - Понимание целей и задач построения Grid
  - Стабильно высокий уровень интереса заказчиков
  - Первые пилотные проекты
  - Oracle Enterprise Grid – программная инфраструктура центров обработки данных
- Перспективы
  - В ближайшие 2-3 года: построение первых промышленных Grid
  - Эволюционный характер построения Grid: включение в процесс по мере готовности заказчиков

**“Из всех видов сильнейшим  
является тот, который лучше  
всех реагирует на изменения.”**

– Чарльз Дарвин

# Информация для контактов

**Сергей Данилов**  
Oracle СНГ  
Тел. 721-3261  
[sergey.danilov@oracle.com](mailto:sergey.danilov@oracle.com)

Ведущий консультант по  
Базовым технологиям  
Государственный сектор

**Анатолий Барышников**  
Oracle СНГ  
тел. 258-41-80  
[Anatoly.baryshnikov@oracle.com](mailto:Anatoly.baryshnikov@oracle.com)

Менеджер проектов  
Государственный сектор

ORACLE®