

СУБД UMS-FAD™

Компания «Х-Технология»

Москва 2011



Системотехническое решение СУБД **UMS-FAD™**



Административный клиент GAI

Драйвер LCD

Фреймворк UMS:

веб-сервер UNET
виртуальная машина UVM
компилятор байт-кода UBC
сетевой шлюз UGW

Двигатель FAD™

Репозиторий `execute.erm`

Особенности СУБД UMS-FAD™



Расширенная реляционная модель данных ERM

Денормализованные отношения

Ассоциативные массивы данных

Служебные домены первичных ключей данных

Мультиверсионная архитектура

Отсутствие блокировок объектов базы данных

Отсутствие журналов транзакций и сегментов отката

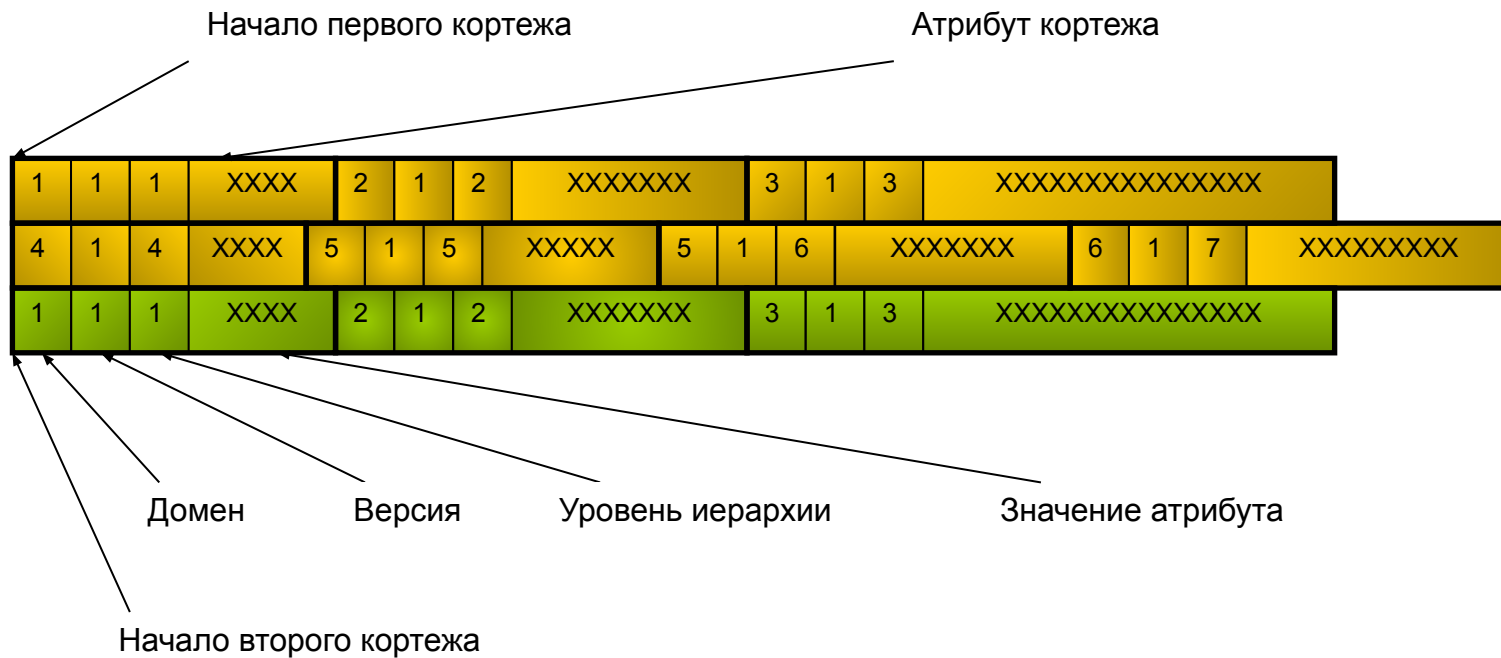
Сериализуемый уровень изоляции транзакций

Механизм поиска в виде матричных деревьев



Ассоциативный массив

Реляционное отношение



Расширенная реляционная модель данных ERM



- Денормализованные отношения – максимально полные информационные образы типов объектов предметной области
- В типах объектов поддерживаются логические связи между составляющими доменами (в виде иерархий)
- Первичные ключи – функционально независимые атрибуты служебного домена уникальных идентификационных номеров
- Вторичные ключи – соответствуют вторичным ключам базовой реляционной модели данных RM
- Манипуляционный и целостный аспекты модели ERM соответствуют подобным аспектам модели RM
- Доменный состав отношений модели ERM соответствует атрибувному составу классов объектно-ориентированной модели данных прикладных решений

Кортеж данного - мультимножество



УИН - 27498752063

Фамилия - Кузнецов

Имя - Иван

Телефон - Рабочий

Номер - 495 365 2987

Телефон - Мобильный

Номер - 916 120 6582

Телефон - Неизвестно

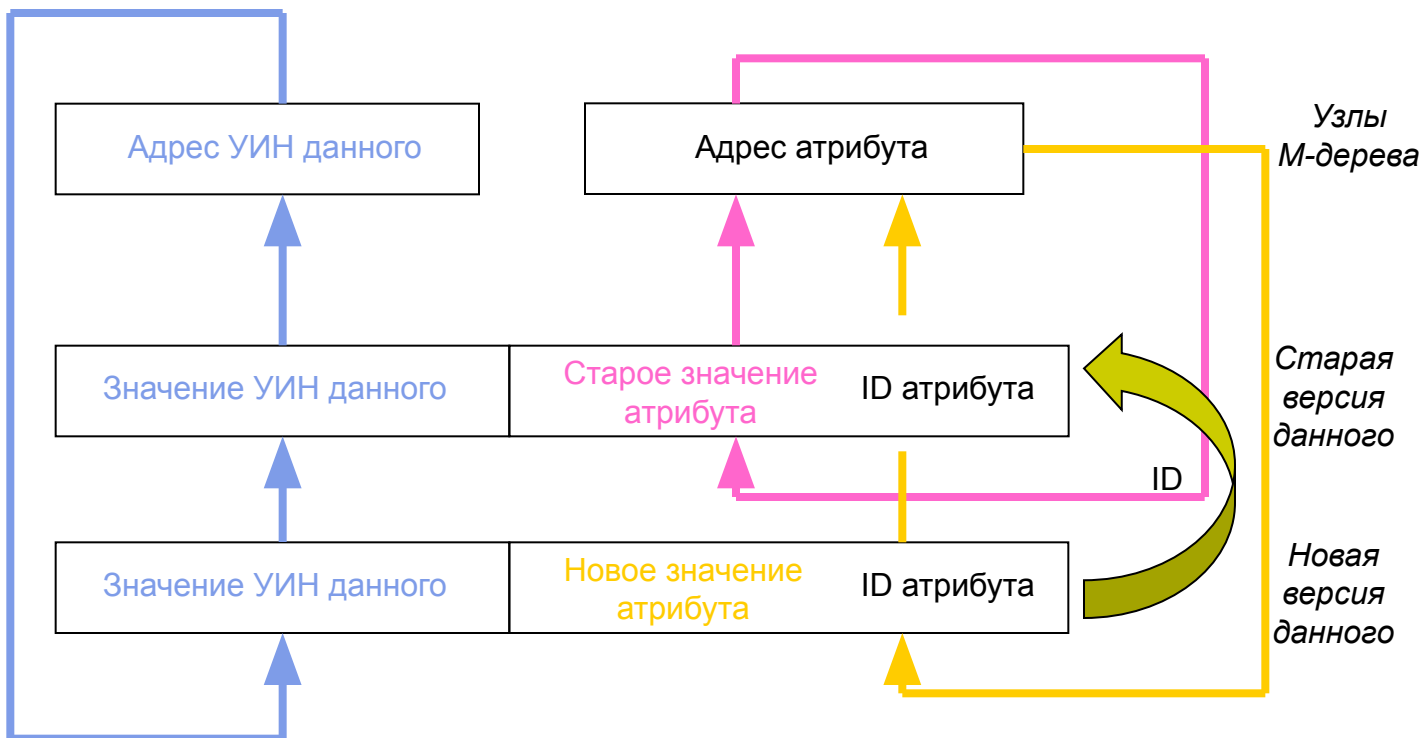
Номер - 499 268 5049

Структура кортежей данных



- Атрибуты каждого кортежа связаны в иерархическую структуру
- Арность кортежей в общем случае не равна арности отношений
- Кортежи могут содержать повторяющиеся атрибуты или группы атрибутов
- Каждое отношение содержит головной кортеж, содержащий стандартный набор атрибутов, имеющих неопределенные значения, и соответствующий одному из типов объектов предметной области
- Максимальный размер одного кортежа ограничен возможностями файловой системы
- Кортеж может содержать ссылки на фрагменты неструктурированной информации, хранимые в файле базы данных или в отдельных файлах

Мультиверсионная архитектура СУБД UMS-FAD™

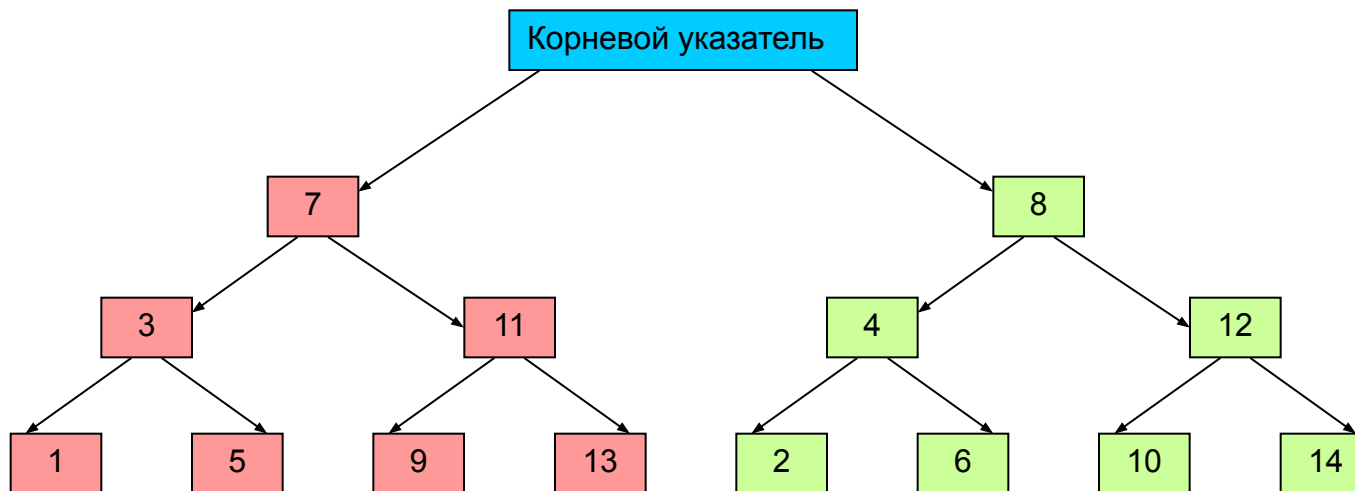




Структура версий данных

- Каждое данное состоит из первичной версии и набора дельта-версий, логически связанных совпадающим уникальным идентификационным номером (УИН)
- Версии содержат номера транзакций записи
- Модифицированные атрибуты в составе версий данных логически связаны служебным идентификатором экземпляра атрибута (ID)
- Запись новых версий данных осуществляется без чтения и блокировки старых версий, на основании УИН данных и ID атрибутов
- Все версии данных хранятся в одном отношении
- Очистка базы данных от устаревших версий данных производится в фоновом режиме путем консолидации версий

Двоичное сбалансированное матричное дерево поиска



Значение ключа четного поддерева = $2^L \times Y$, где L – уровень дерева, Y - сомножитель

Значение ключа нечетного поддерева = $(2^L \times Y) - 1$

Позиция узла дерева на уровне = $(Y + 1) / 2$

Значение ключа общей вершины = $(K_1 + K_2) / 2$, где $K_{1,2}$ – значение ключей нижних узлов

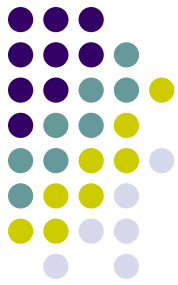
Особенности матричных деревьев поиска



- Матричные деревья поиска относятся к самобалансирующимся двоичным деревьям
- Включению элемента в матричное дерево предшествует присвоение номера элементу в порядке записи в базу данных
- Отдельное матричное дерево состоит из четного и нечетного поддеревьев (поиск во вдвое меньшем числе элементов)
- В процессе балансировки участвуют не более 4 элементов (в отличие от половины элементов любой другой разновидности самобалансирующихся деревьев)
- В случае монотонно возрастающего порядка записи элементов матричные деревья не требуют балансировки
- Механизм поиска состоит из нескольких ярусов матричных деревьев – типы данных, домены, дескрипторы атрибутов

Сравнение эффективности реализации и ERM моделей данных

RM



ER модель

A – индексируемые атрибуты, N – нормальная форма, F – сканирование файлов

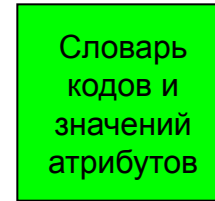
ERM модель

A = max, N = 1, F = 0



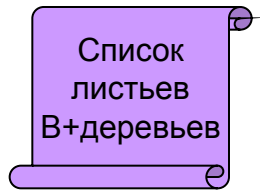
Обращения к диску

Запись ~ L/2
Чтение ~ 0



Обращения к диску

Запись ~ 0
Чтение ~ 0



Запись ~ 2
Чтение ~ 1



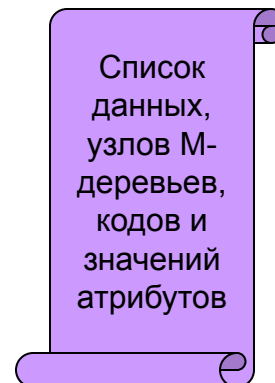
Запись ~ 0
Чтение ~ 0

Оперативная память

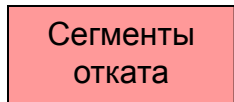
Энергонезависимая
память



Запись ~ N
Чтение ~ 1 (+ F)



Запись ~ 1
Чтение ~ 1



Запись ~ N
Чтение ~ 1 (+ F)

Итого:

Запись = $A \times (L/2 + 2) + 2N$
Чтение = $A \times 3 + 2F \times N$

Итого:

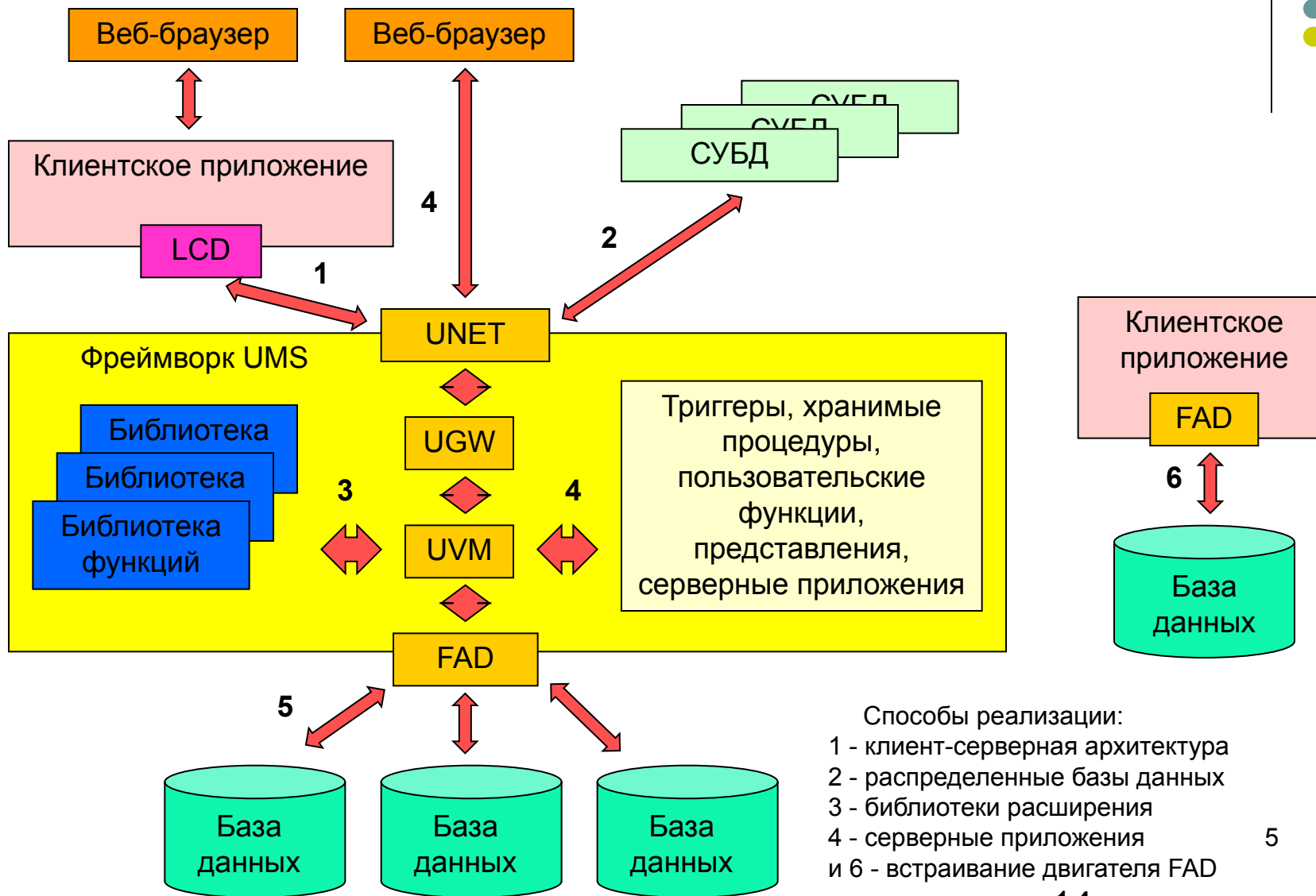
Запись = 1
Чтение = 1

Основные инновации СУБД UMS-FAD™



- Данные, метаданные и фрагменты неструктурированной информации входят в единую структуру данных с общей точкой входа
- Использование инкрементных версий данных
- Замена упреждающего резервного копирования метаданных на циклический процесс перезаписи их парных значений
- Модификация данных и метаданных на логическом уровне
- Консолидация версий данных («уборка мусора») в фоновом режиме

Компоновочные решения UMS-FAD™



СУБД **UMS-FAD™** реализована в клиент-серверной архитектуре



Серверная часть

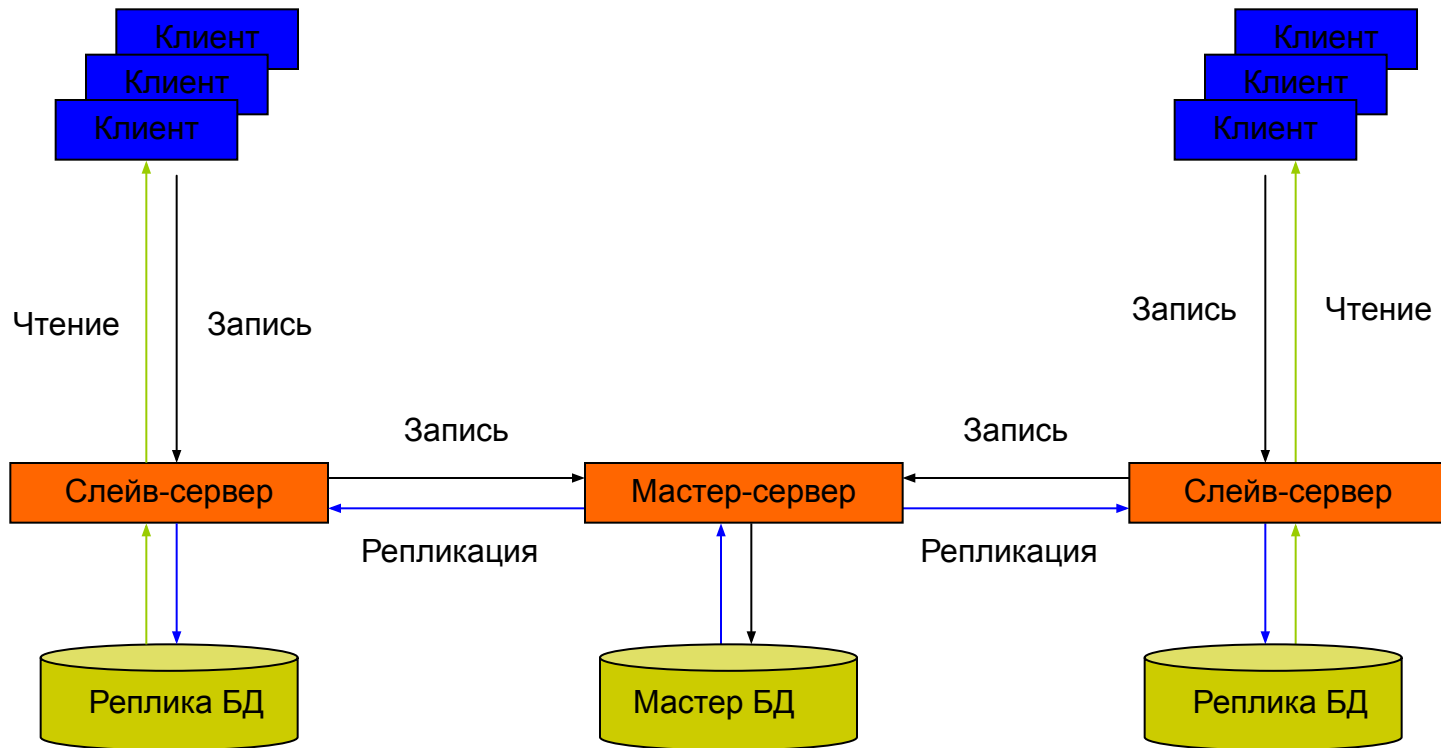
- работает на платформе 64-разрядных операционных систем UNIX/Linux
- обеспечивает обмен информацией в распределенной вычислительной среде в форматах UDTP и SOAP
- поддерживает взаимодействие с браузерами в формате HTTP

Клиентская часть

- работает на платформе 32/64-разрядных операционных систем Unix/Linux и Microsoft Windows
- обеспечивает обмен данными с клиентскими приложениями, аналитическими системами и табличными процессорами в формате XML



Масштабирование СУБД UMS-FAD™



Характеристики масштабирования и кластеризации СУБД **UMS-FAD™**



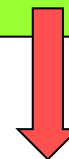
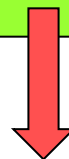
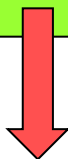
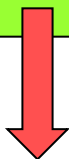
- Максимальное количество слейв-серверов, обслуживаемых одним мастер-сервером – до десяти тысячи единиц
- Максимальное количество серверов баз данных в кластере, обслуживаемых одним сервером приложения – до десяти тысяч единиц
- Скорость передачи данных:
 - между мастер-сервером и слейв-сервером 10 Гбит/сек;
 - между сервером приложения и сервером базы данных 16 Гбит/сек.



Кластеризация СУБД UMS-FAD™

Сервер приложения:

Данное # 1	Данное # 2	Данное # 3	Данное # 4	Данное # 10
Данное # 11	Данное # 12	Данное # 13	Данное # 14	Данное # 20
Данное # 21	Данное # 22	Данное # 23	Данное # 24	Данное # 30
Данное # 31	Данное # 32	Данное # 33	Данное # 34	Данное # 40



Сервер
базы
данных
1

Сервер
базы
данных
2

Сервер
базы
данных
3

Сервер
базы
данных
4

Сервер
базы
данных
.....

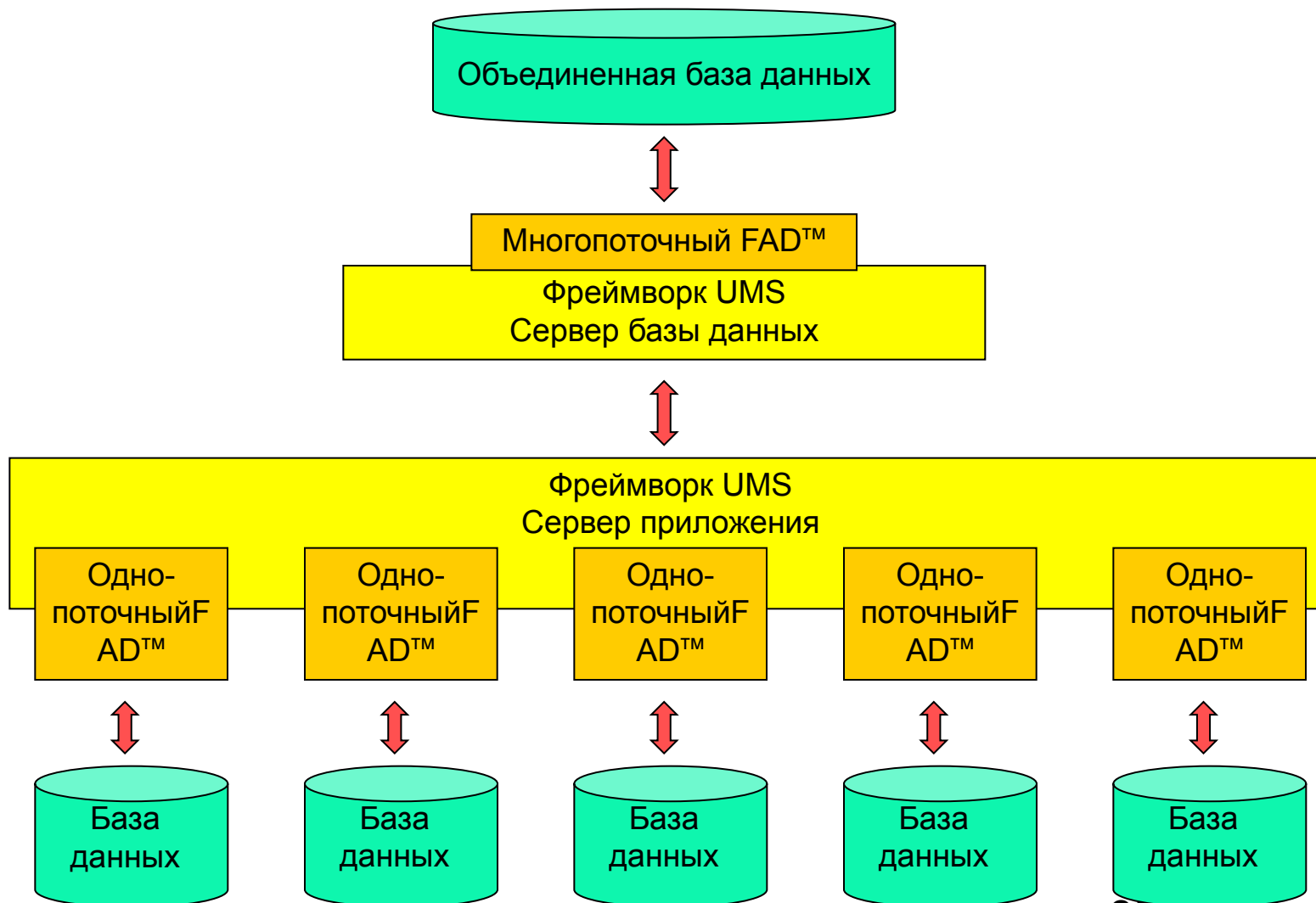
Сервер
базы
данных
10

Технология выполнения параллельных вычислений в среде UMS-FAD™



- Процесс верхнего уровня – генерация пользовательских сессий и консолидация промежуточных результатов параллельных вычислений
- Процесс нижнего уровня – выполнение параллельных вычислений и сохранение промежуточных результатов параллельных вычислений
- Использование на верхнем уровне одной многопоточной модификации двигателя баз данных FAD™
- Использование на нижнем уровне множества однопоточных модификаций двигателя баз данных FAD™
- Аппаратная платформа верхнего уровня – компьютер
- Аппаратная платформа нижнего уровня - суперкомпьютер

Параллельные вычисления на платформе СУБД UMS-FAD™

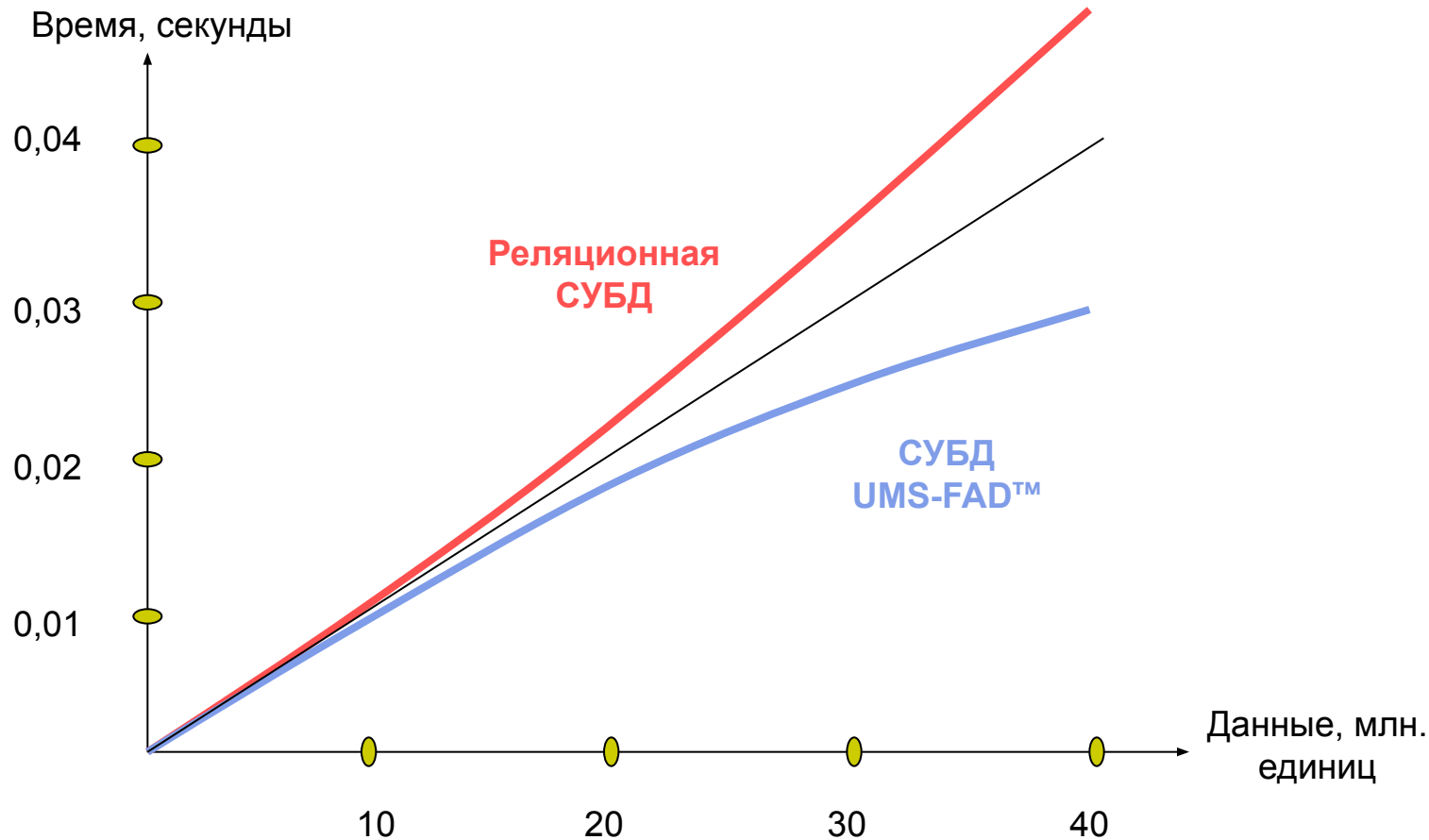


Индустриальный тест TPC-C по оперативной обработке транзакций OLTP



- Пиковая производительность СУБД **UMS-FAD™** составила 1 миллион транзакций в минуту в расчете на одно ядро процессора
- Цена одной транзакции в минуту СУБД **UMS-FAD™** составила единицы центов США

Прогрессивная и регрессивная зависимости времени чтения от объема баз данных, включающих 5, 10 15 и 20 млн. данных





Контакты

- ООО «Х-Технология»
- 127051, Москва, Малый Сухаревский пер., дом 9, строение 1, офис 36
- тел. +7 495 960 0050
- [http: //www.x-tex.ru/](http://www.x-tex.ru/)
- e-mail: info@x-tex.ru