

БОЛЬШИЕ БАЗЫ ДАННЫХ FIREBIRD

Алексей Ковязин,

ak@ib-aid.com

IBSurgeon Ltd

IBSurgeon – 8 лет!

- Клиенты

- Carl Zeiss Meditec, USA
- Внешторгбанк, Россия
- Wells Fargo Bank, USA
- Watermark Software, UK
- Bas-X, Australia
- Victoria University, New Zealand
- Kingsway Management, UK

- Команда

- Дмитрий Кузьменко
- Алексей Ковязин
- Сергей Никитин
- Олег Матвеев и другие

- Консультанты

- Влад Хорсун, Дмитрий Еманов, Алекс Пешков

- Партнеры

- IBPhoenix

О чем пойдет речь?

- Что такое БОЛЬШАЯ база данных Firebird?
 - База в 1 Terabyte (1024Gb) и немного истории
 - Реальные примеры и наша статистика
- Особенности больших баз данных
 - «Бэкап, рестор и...»
 - Ком проблем
- Решения IBSurgeon
 - Отлов плохих запросов и индексов, раскрытие «чужого» SQL-кода, тюнинг
 - Логгирование и аудит
 - Автоматическое обслуживание

Что же это такое – большая БД?

Немного истории

- 2005
 - В IBSurgeon обнаружили лимит в 37Гб на таблицу для Firebird 1.0-1.5 и всех InterBase (вплоть до 2007 версии)
- 2006
 - Firebird 2.0 исправляет ошибку с 37Гб/таблицу
- 2008
 - InterBase 2009 исправляет ошибку с 37Гб (но не с 1.6 миллиарда записей на таблицу – исправлена в XE)
- 2009 – IBSurgeon создает и тестирует базу данных Firebird 2.1 размером в 1Терабайт

<http://www.ib-aid.com/articles/item104>

Table name	Record counts	Size, gb	Index creation time	Tmp file size, Gb	Index size, Gb
ORDER_LINE	3720051796	359	12h 6m 18s	182.0	29.3

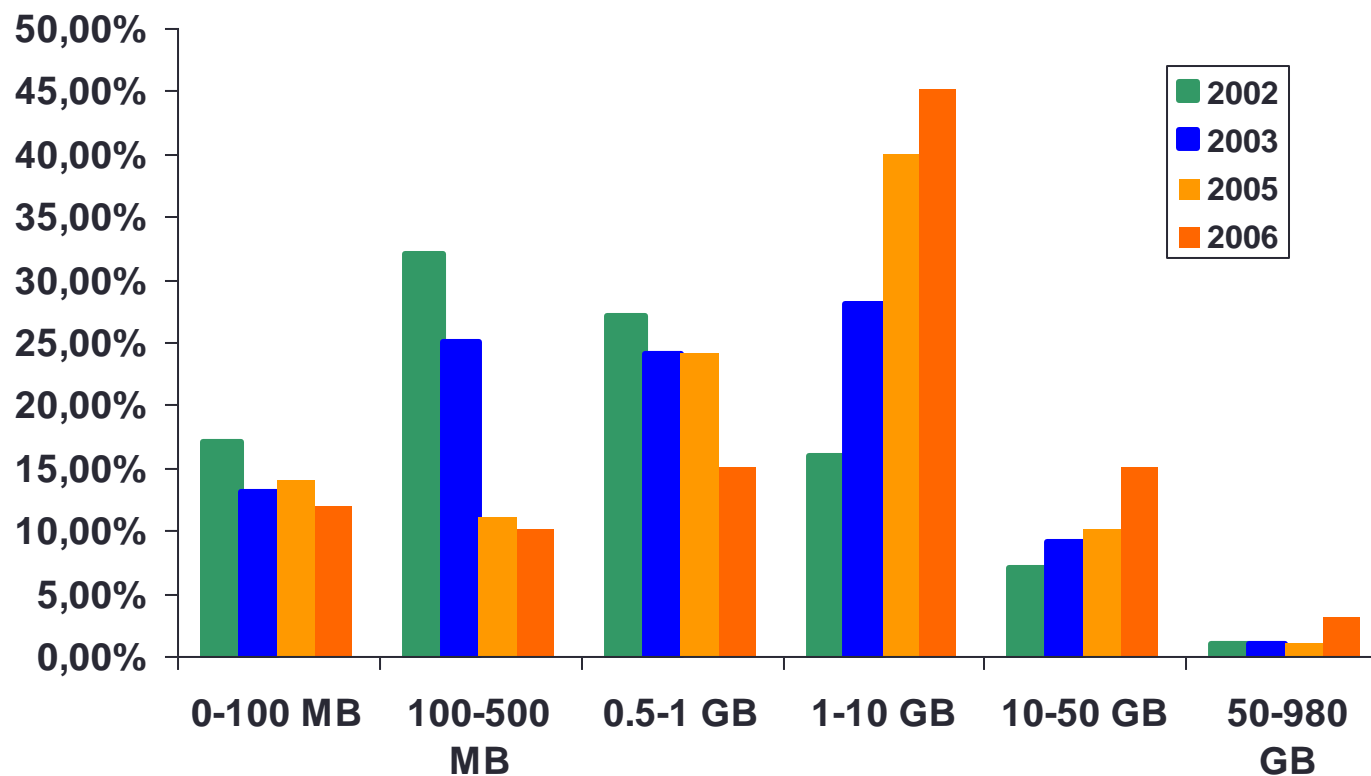
Что такое большая БД Firebird?

Несколько примеров

- **BAS-X**, Австралия, ERP и складской учет
 - 450 GB, без BLOBов, 300 пользователей
- **Watermark**, Великобритания, сканированные документы и полнотекстовый поиск
 - 300+ GBs, BLOBы
- **Profitmed**, Россия, фармдистрибутор
 - 65GB, без BLOBов, 250 пользователей
- **Avarda ERP**, компания Ansoft, Россия, ERP
 - Бай-Центр, ECCO, FinnFlare, Лидер, etc
 - До 100Gb, без BLOBs

Рост баз данных

(данные по клиентам IBSurgeon's)



КРИТЕРИИ ДЛЯ БОЛЬШИХ

Рассмотрим базу данных Firebird

Обычно БД изображают так:

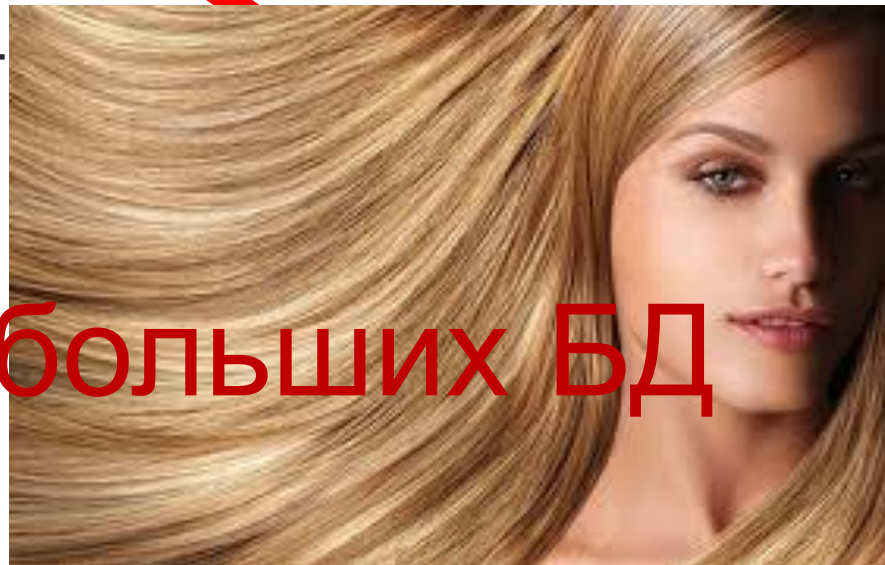


как будто это что-то совсем простое.

Нулевое обслуживание?

- Настройки по умолчанию в 90% случаев
 - firebird.conf достаточно хорош для большинства случаев
 - Параметры: sweep, buffers, etc также хороши почти для всего
 - Размер страницы в 4096 байт - Ok!
 - Настройки ОС (семафоры, кеш) - по умолчанию
- Обслуживание - простое
 - Просто делайте backup/restore..

Неверно для больших БД



Добро пожаловать в мир больших баз данных

- Цикл Backup/restore может занимать часы
 - Зависит от HDD и плотности данных
 - Плохой пример - 1 час для рестоора 3.6Гб БД
 - Хороший пример – 12 часов для backup 450Гб БД
- Sweeper может занять от 30 минут до нескольких часов
- gfix 37Гб базы - 1.5 -2 часа

Большие базы – большая ответственность

- Профитмед
 - ~250 одновременных пользователей
 - 65 Гб БД
 - Режим работы - 12 часов x 7 дней
- Высокие требования к надежности
 - «Предупредите нас за 5 минут до начала проблем, пожалуйста»
 - Потери от 1 дня простоя - \$XXXXXX
- Firebird 1.5 Classic 64 bit
- 64-bit CentOS Linux, 64Gb RAM, SCSI
- Backup в эксклюзивном режиме – 30 мин
- Multi-user backup – 2 часа
- Restore - 4 часа
- Sweep
 - Только ручной
 - От 30 минут до 1 часа
 - 2 раза в день

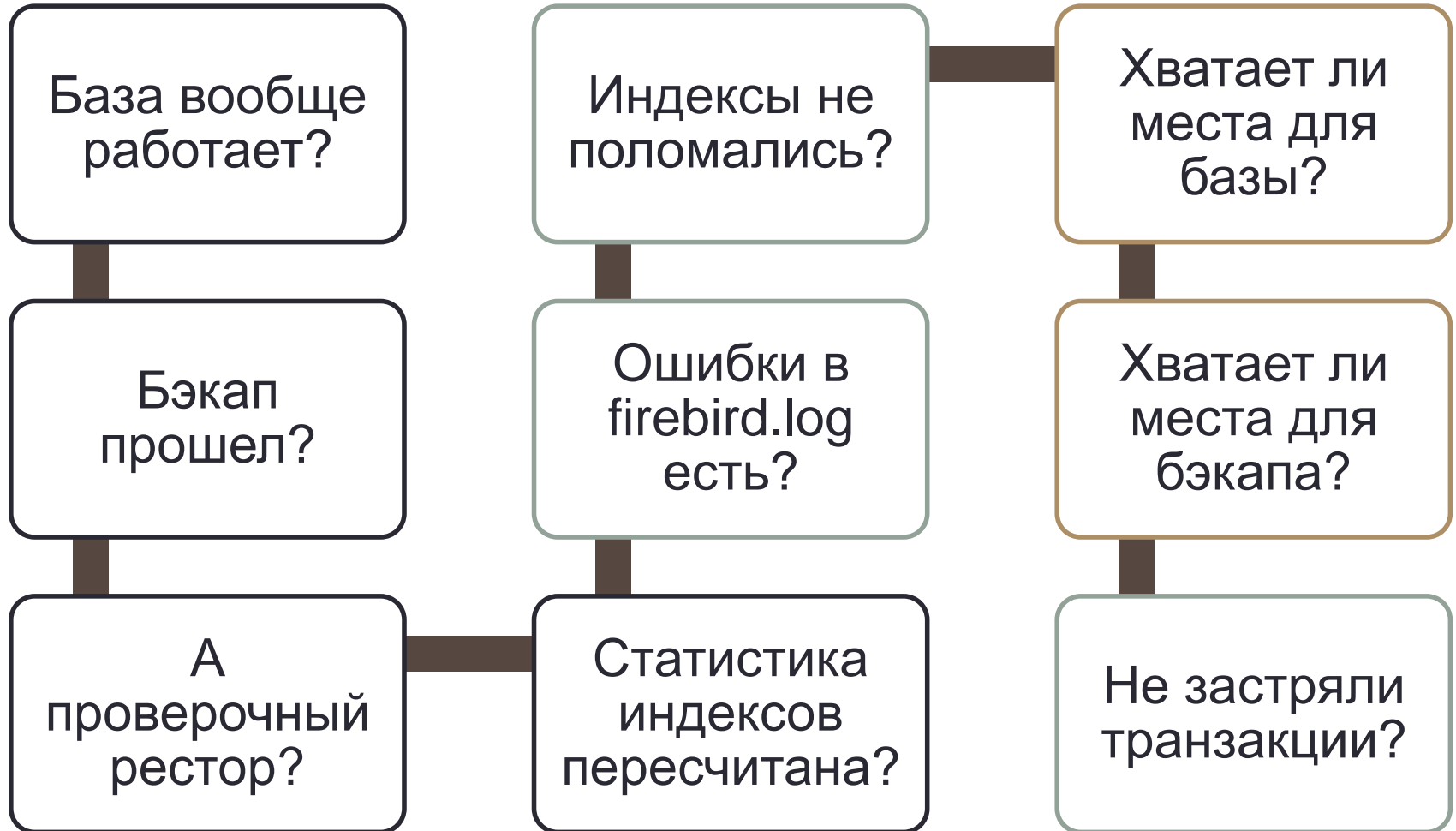
Ключевое отличие №1 между маленькой и большой БД

- **Маленькая БД – делайте backup/restore и не напрягайтесь**
 - Просто делайте регулярные циклы backup/restore – БД каждый раз пересоздается «с нуля»
 - Проблемы с транзакциями, накоплением мусора просто не успевают проявиться
- **Большую базу данных нужно проверять и обслуживать гораздо тщательнее...**

Большая БД нуждается в плане обслуживания


- План зависит от размера БД и режима работы
- Схема бэкапов выбирается индивидуально
- Обязательные тестовые ресторы в отдельный файл (а может, и сервер)
- Что нужно постоянно проверять
 - Ошибки – в firebird.log и ошибки чтения метаданных и данных в живой БД
 - Метаданные – проверка
 - Лимиты метаданных
 - Данные и BLOBы – проверка чтения БД, проверки сегментации таблиц и BLOBs
 - Indices – проверка «здоровья» и пересчет статистики
 - Транзакции – «дырки» и лимиты

Минимальный план обслуживания



**Без сомнения, Ваш администратор
именно так и делает...**



A photograph of a server room. In the foreground, a server rack is filled with various components: a monitor, a keyboard, a clock, and several server units. To the right, another rack contains more server units, two of which have telephones on top. The background shows more racks and a dense network of cables. A person's arm is visible on the left side of the frame.

**...особенно если у
Вас много
серверов, ...**

**...расположенных в
удаленных офисах.**



Ключевое отличие №2 – нарастание снежного кома проблем

- **«Я понятия не имею, что случилось»** (с)
Неизвестный DBA, поднятый среди ночи
- **«У нас на тестовом сервере все работает»**
(с) Из письма разработчика перед увольнением

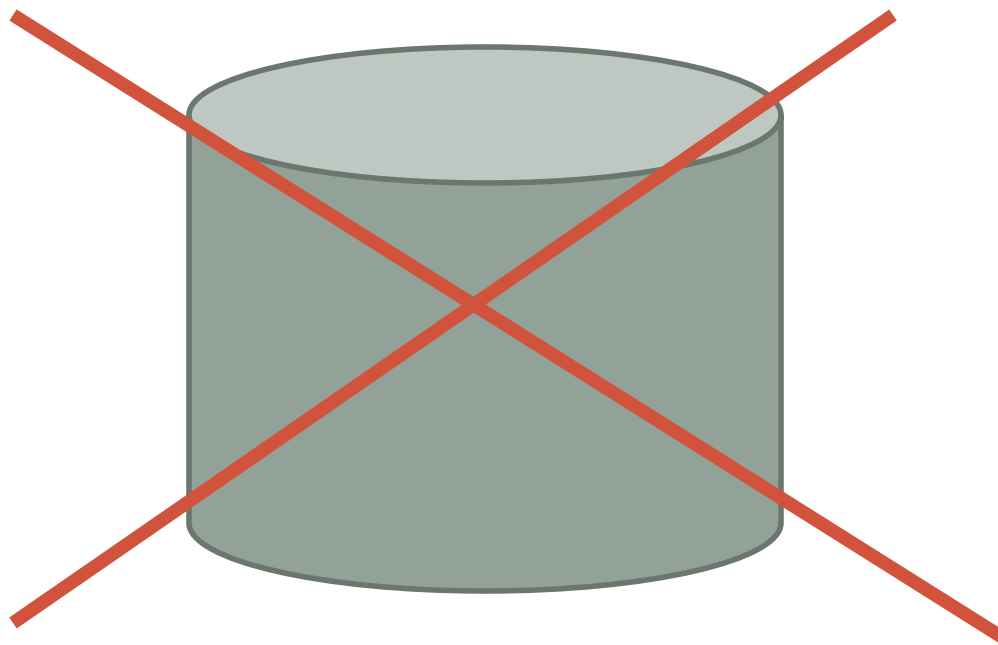
Очевидно, что в БД что-то происходит, но как узнать что?

Неприятные вопросы

- Какой-то запрос тормозит. *Откуда я знаю какой?*
- Новое приложение (новая версия) где-то тормозит. *Кто виноват?*
- **Нужны новые индексы.** *Или старые как-то взбодрить?*
- **Аудит.** *Что там внутри происходит?*

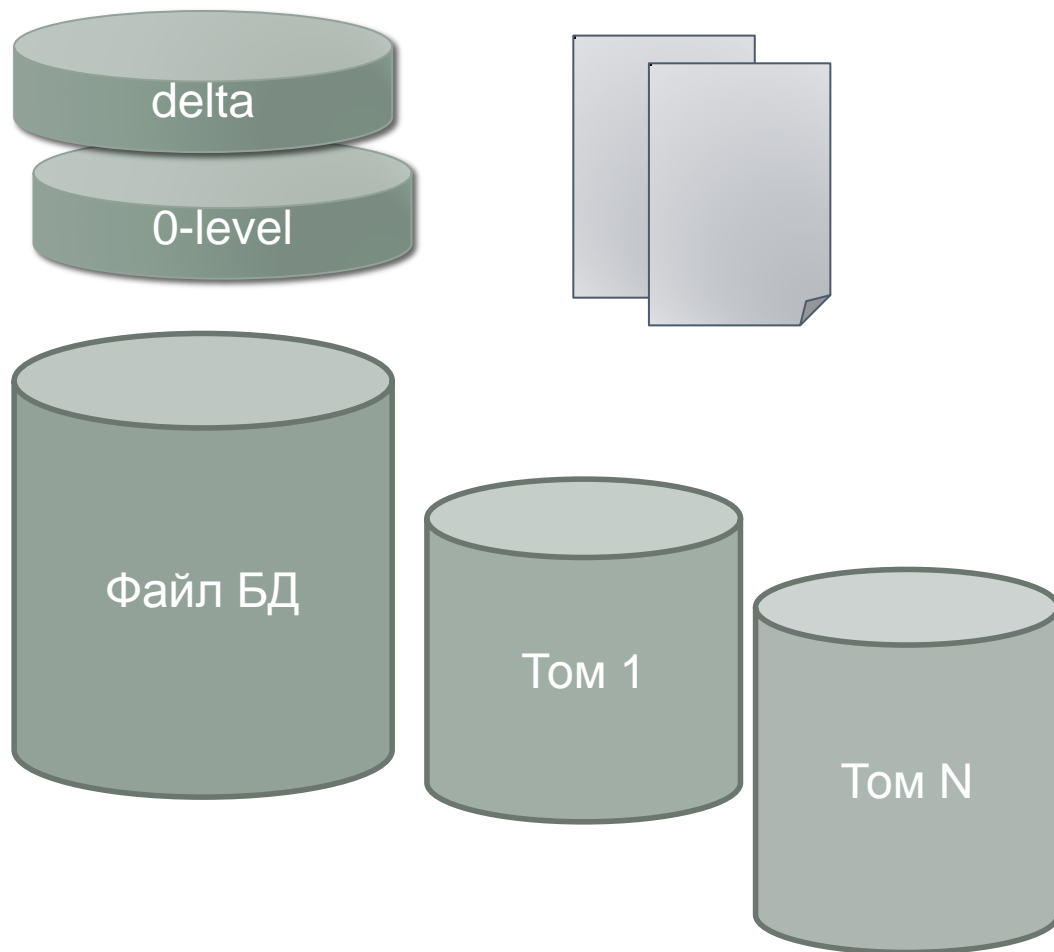
РЕШЕНИЯ IBSURGEON

База – это не просто!



Firebird в деталях

Файловые структуры



- Основной файл БД
- Файлы томов
- delta-файлы (nbackup) инкрементальные бэкапы
- Внешние таблицы

Firebird в деталях

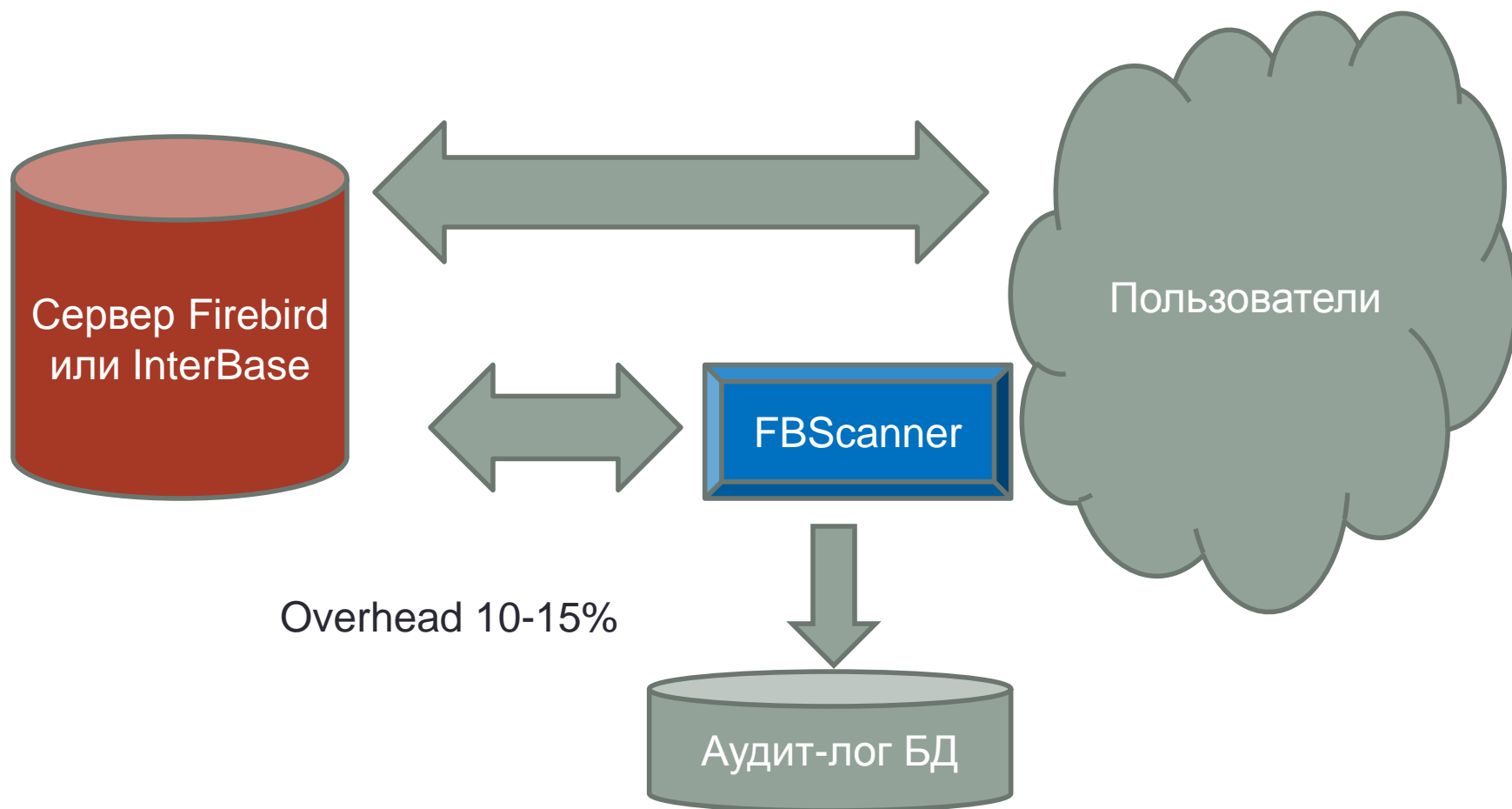
Внутренняя структура



Продукты

- **Продукты для починки баз данных на низком уровне**
 - IBFirstAID/FBFirstAID
 - IBBackupSurgeon, IBUndelete
- **Продукты для решения проблем больших баз данных**
 - FBScanner
 - IBAnalyst
- **Автоматическое обслуживание критически важных БД**
 - FBDataGuard

FBScanner



Функции FBScanner

- Поддержка InterBase 4.0-2009, Firebird 1.0-2.5
- Логгирование и аудит с сохранением лога во внешней БД
 - Аудит пользователей, все SQL запросы и транзакции (не как в MON\$)
- Теги – возможность помечать коннекты, пользователей и SQL запросы метками
- Извлечение планов запросов на лету
- Определение:
 - Слишком долгих запросов
 - Длительных транзакций
 - Длительных коннектов
- **Живая демонстрация – на стенде IBSurgeon**

FBDataGuard – защита и оборона

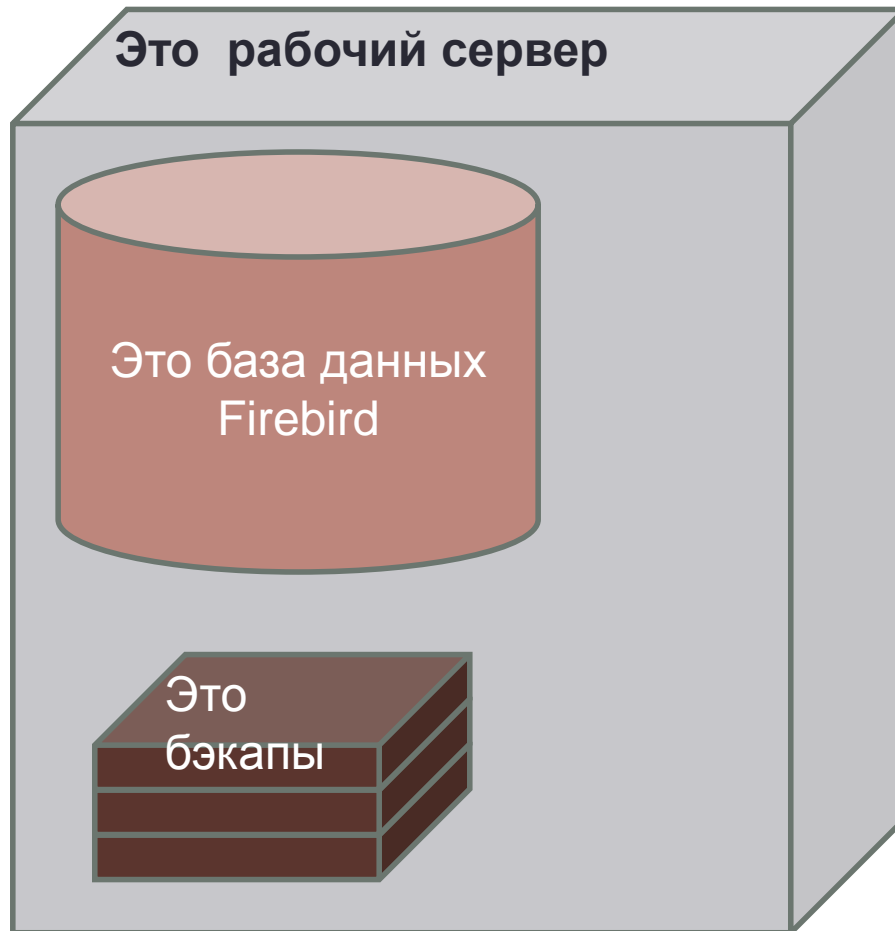
- Наблюдение за базой
- Предупреждения и советы
- Автоматизация обслуживания баз данных
- Гарантия восстановления в сложных случаях

**РАНЬШЕ БОРОЛИСЬ С ПОСЛЕДСТВИЯМИ,
ПОРА ВЗЯТЬСЯ ЗА ПРИЧИНЫ.**

Что делает FBDataGuard

- Следит за сервером
- Следит за базой данных
- Следит за окружением

Типичная рабочая среда Firebird

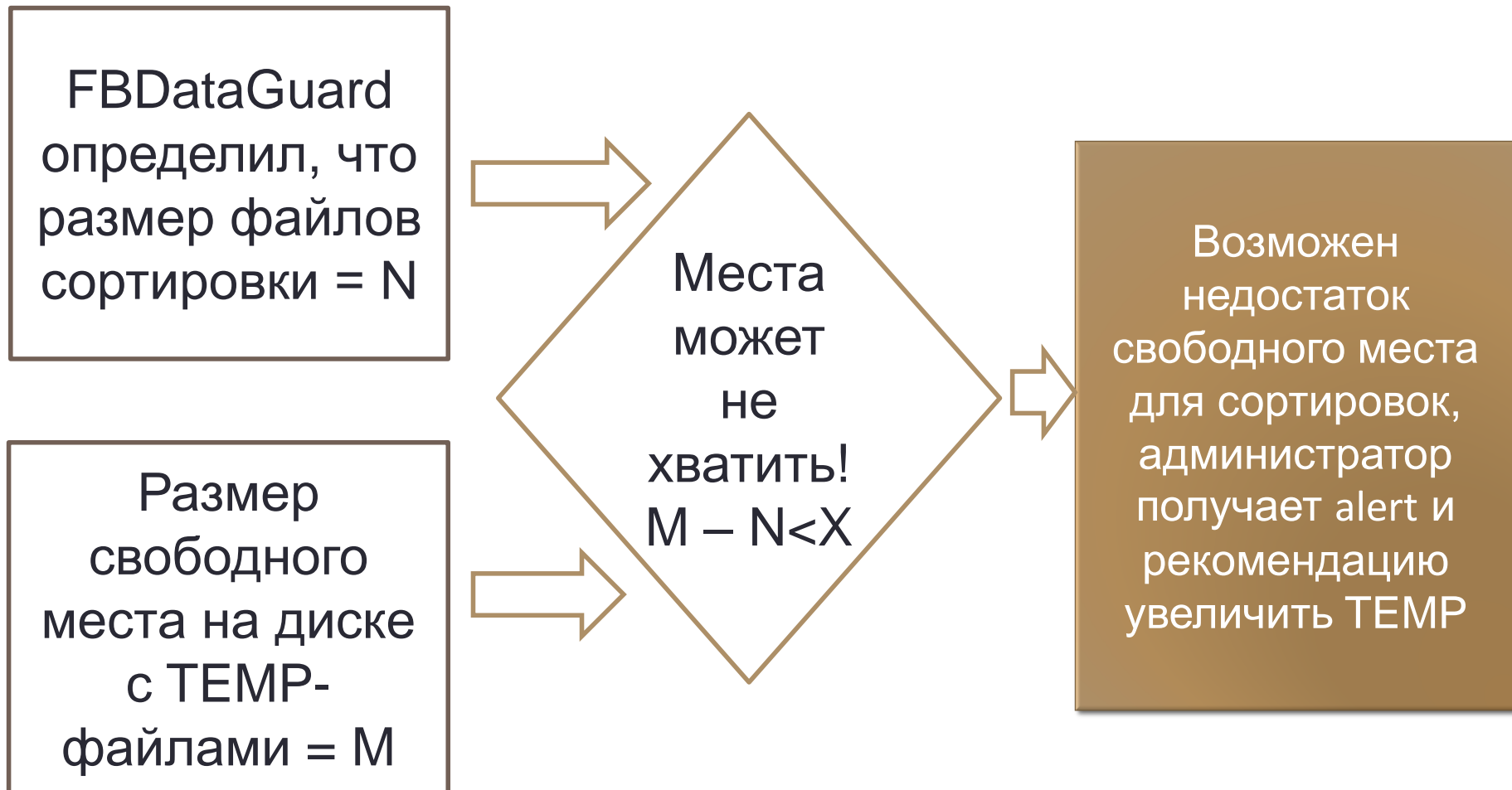


Сервер Firebird

7 параметров,
которые могут
сообщать о
проблемах с
базой данных и
сервером

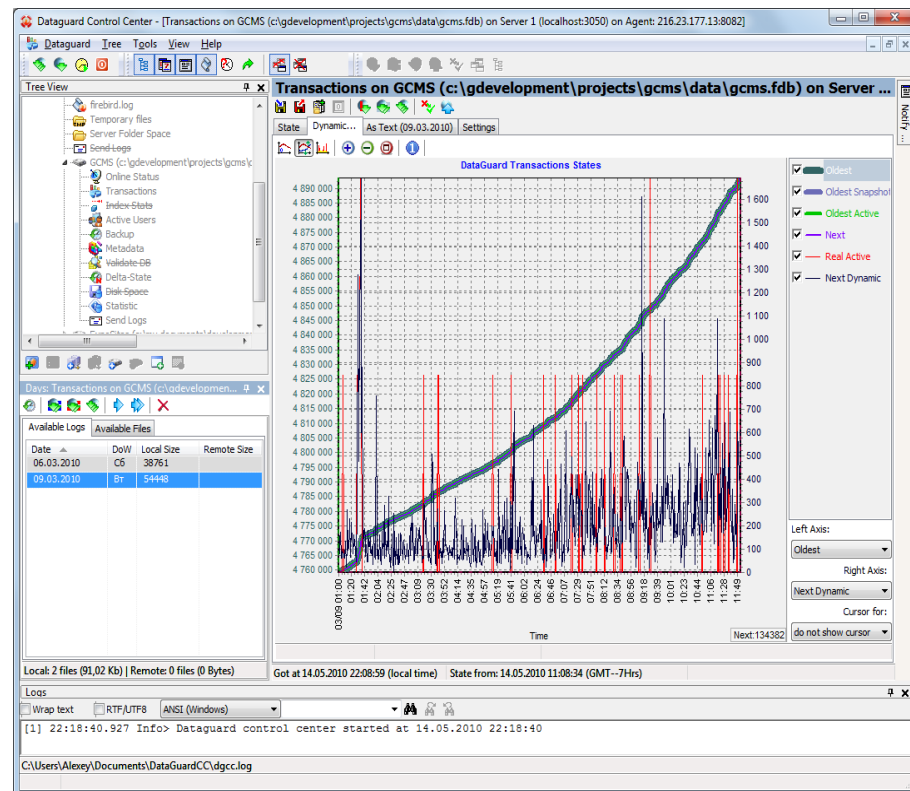
1. Доступность сервера
2. Размер RAM сервера
3. Количество временных файлов
4. Размер временных файлов
5. Записи в логе
6. Размер логов
7. Версия сервера

Пример разрешения проблемы с сервером



Ретроспективный анализ

- Все логи хранятся на сервере и позволяют анализировать события, происшедшие в прошлом
- Инструментарий для удобного просмотра логов



Пример разрешения проблемы с базой данных Firebird

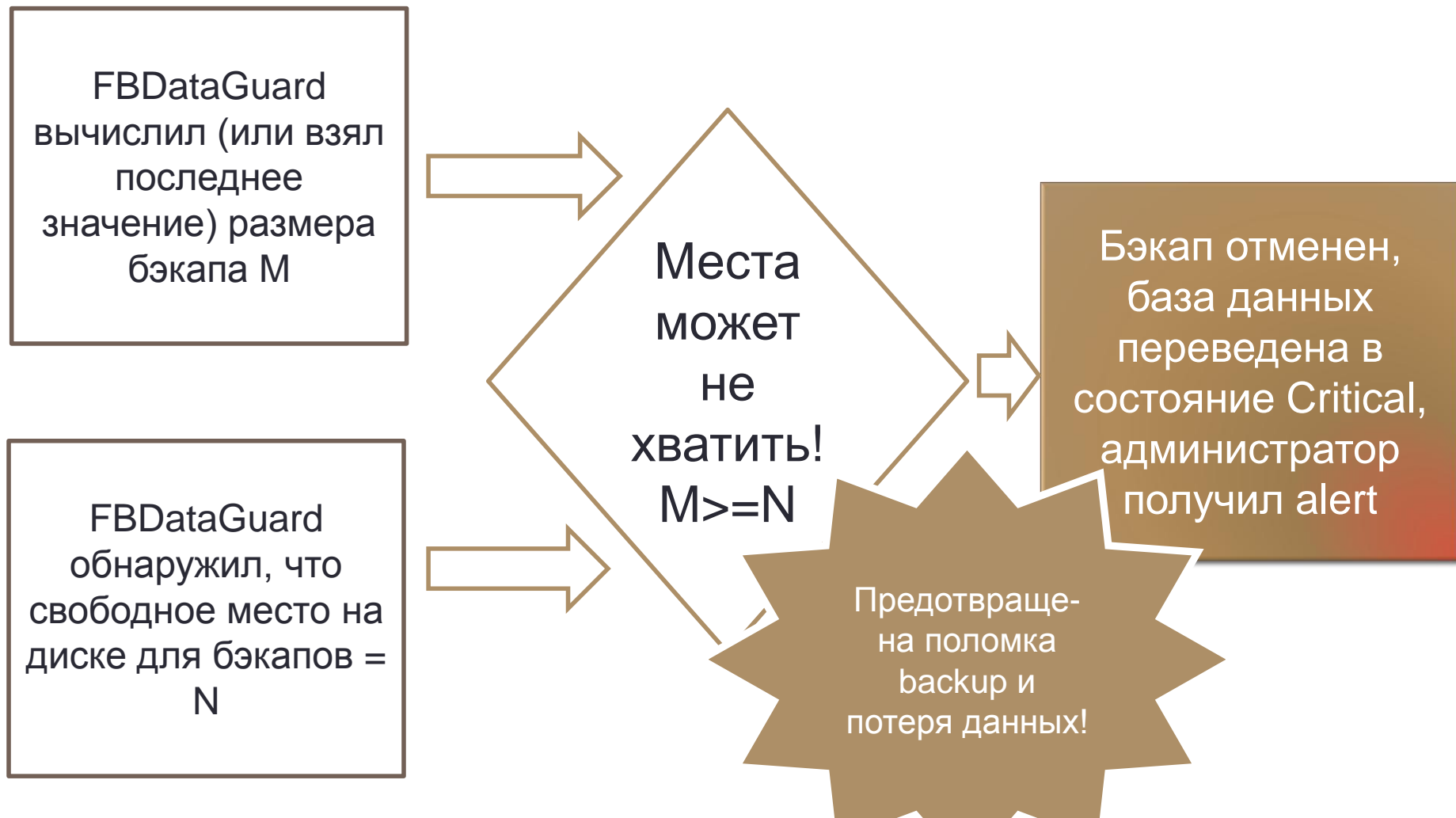
FBDataGuard
определил, что
после restore
индекс не
активирован

non-activated
индексы могут
указывать на
повреждения БД,
SQL запросы
могут
«тормозить»

Администратор
получает alert и
рекомендацию
проверить
индексы

Предотвращена потеря
производительности!

Пример разрешения аварийной ситуации с бэкапами



Катастрофические поломки



Серверы (как любые сложные устройства)
—
ЛОМАЮТСЯ.

Что может сломаться в железе?

Наиболее опасны для базы данных следующие поломки:

- Жесткий диск (HDD)
- Flash-накопители
- Память (RAM)
- Контроллеры SCSI/SATA и другие подобные устройства

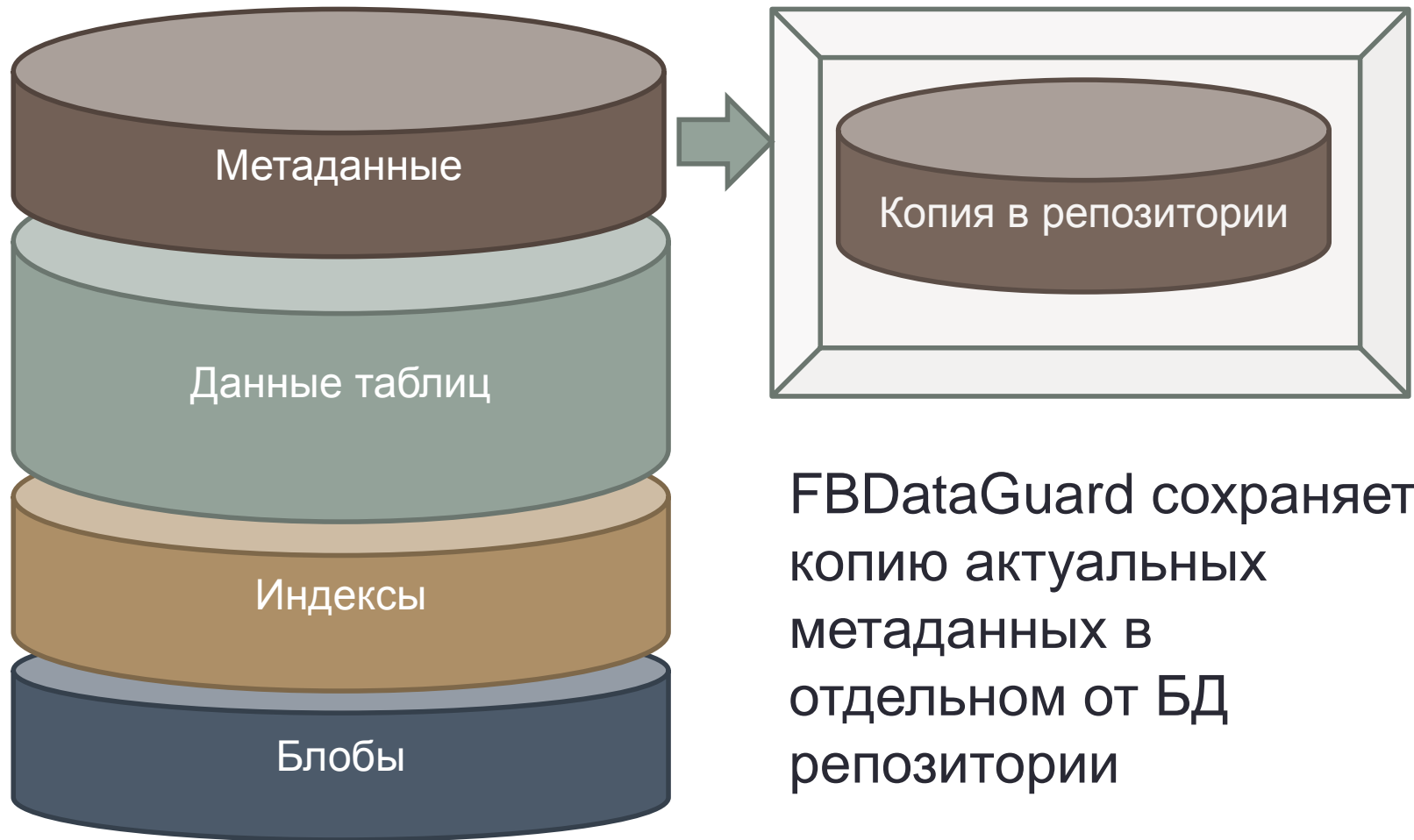
Типичные проявления поломок «железа»:

- Жесткий диск:
 - Потерянные и смешанные страницы (wrong page type)
 - Ошибки в цепочках записей (Cannot find record fragment)
- Память:
 - Ошибки на уровне записей (Wrong record length)
- Flash-накопители и Контроллеры
 - Сдвиги страниц (база не открывается в isql)
 - Ошибки страниц и ошибки в записей

Как FBDataGuard защищает от поломок железа?

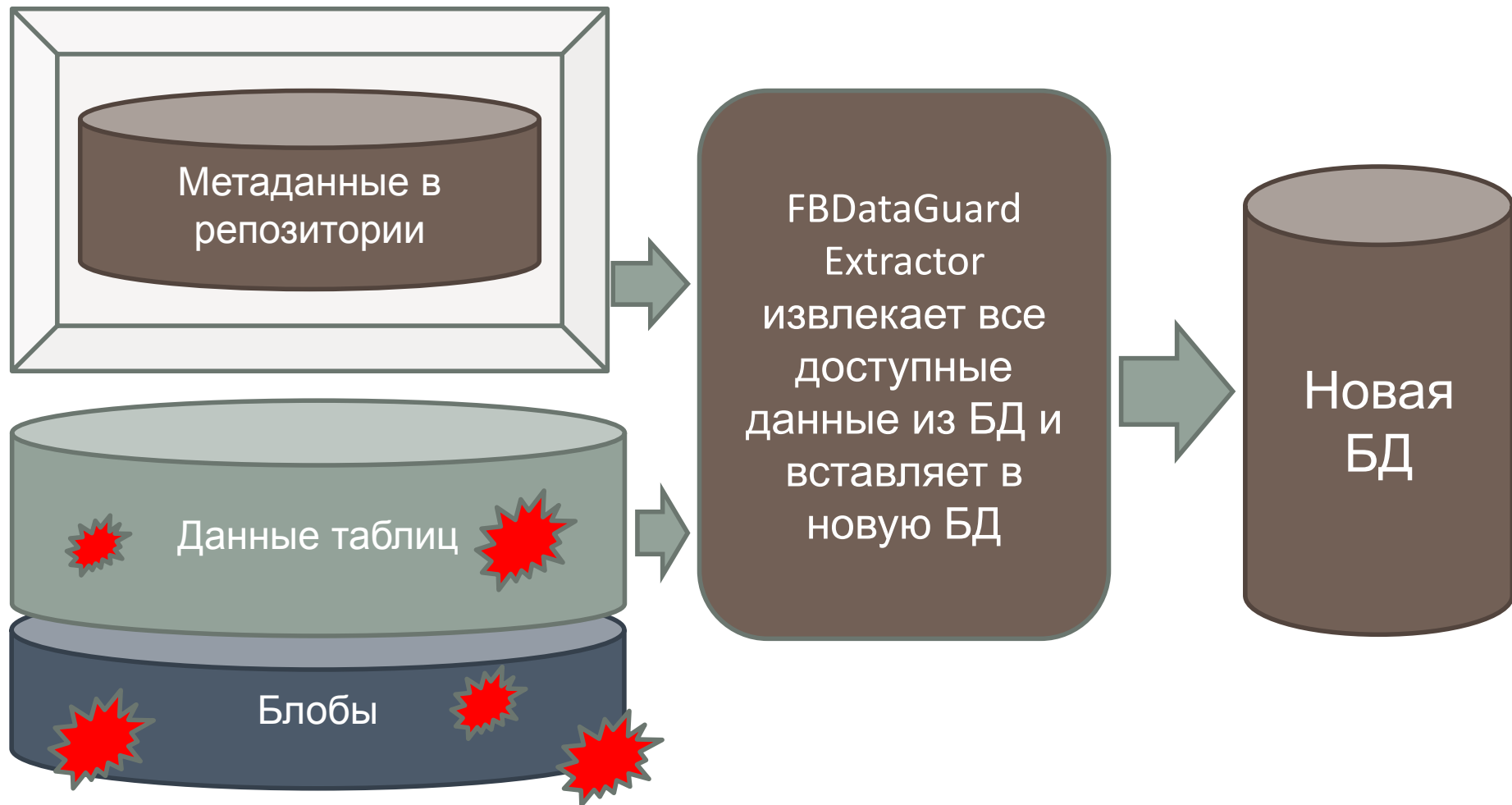
- Во-первых, верификация данных и индексов (выборка данных, пересчет статистики индексов)
 - Позволяет предупредить о появлении ошибки
- Во-вторых – ЗАЩИТНЫЙ РЕПОЗИТОРИЙ МЕТАДААННЫХ
 - Позволяет спасти данные даже в случае очень тяжелых повреждений

Защитный репозиторий метаданных



FBDataGuard сохраняет копию актуальных метаданных в отдельном от БД репозитории

В случае поломки железа:



Последний рубеж защиты

- FBDataGuard спасет оставшиеся данные
 - в случае потери метаданных
 - Данные из поврежденного delta-файла
 - В случае поломки жесткого диска, контроллера или flash-накопителя
 - Вытащит данные даже из «обрывка» БД

Но лучше не доводить ситуацию до крайности, не так ли?

Firebird DataGuard

- Наблюдение за 26 важными параметрами базы данных и сервера
- Предупреждения о потенциальных и реальных проблемах по email
- Правильная автоматизация обслуживания баз данных
- Возможность встраиваться в существующие приложения
- Windows, Linux, MacOS, Firebird 1.5-2.1
- FBDataGuard включает сервисы ремонта и оптимизации базы данных (в зависимости от лицензии)

Мониторинг-1

- Ошибки сервера и базы данных
 - Доступность сервера и базы данных
 - Анализ firebird.log (изменения)
 - Периодический опрос метаданных
- Транзакции
 - Отслеживание разницы между 4 маркерами
 - Лимит транзакций (2млрд. транзакций до бэкап-рестора)
- Пользователи
 - Минимальное/максимальное количество пользователей

Мониторинг-2

- Файлы базы данных
 - Однофайловые и многофайловые
 - Расположение – предупреждения о путях и пересечении с местом хранения бэкапов
 - Размеры и отслеживаемые пределы роста (задаются пользователем)
- Файлы delta (nbackup)
 - Время жизни и размер delta-файлов
- Файлы бэкапов
 - Наличие, размер и прогноз роста

Мониторинг-3

- Временные файлы Firebird
 - Общий размер и количество
- Количество форматов на таблицу
 - Не более 255 (лимит реализации)
 - В production-базе данных нежелательно
- Наличие неактивных и неактивированных индексов в базе данных
 - Неактивные – явно отключены
 - Неактивированные – не включились в результате сбоя рестора

Мониторинг-4

- Автоматизированный сбор статистики gstat
 - Задается cron-выражением
- Версия сервера Firebird, размер папки сервера
- Размер логов сервера
- Размер логов FBDataGuard
 - Логи изменений всех параметров хранятся для анализа
 - Автоудаление логов – по умолчанию хранятся 60 дней
- Автообновление
 - – предупреждение о выходе новых версий

Обслуживание-1

- Бэкапы
 - Револьверные (сдвигающиеся по времени) бэкапы
 - Настраиваемая глубина хранения бэкапов
 - Тестовый рестор с анализом результатов
 - Прогноз роста (при недостатке места бэкап не происходит, статус БД ->критический)
 - Контроль времени операции бэкапа (по умолчанию Max time =120 min)
 - Опциональное сжатие небольших бэкапов (до 4Гб)

Обслуживание-2

- Индексы
 - Пересчет статистики индексов
 - Всех (не рекомендуется для баз более 1Гб)
 - Выбранных (список индексов)
 - Всех, кроме исключенных
 - Проверка статусов индексов
 - Проверка «здоровья»
 - - алерт в случае обнаружения поврежденных индексов
 - По умолчанию джоб отключен

Обслуживание-3

- Валидация БД с помощью стандартных средств (Gfix)
 - Перевод базы в shutdown – работа в выделенном режиме
 - Анализ результатов (в т.ч. вывода в firebird.log)
- Валидация метаданных
 - Проверка ключевых системных таблиц
- Лог Firebird
 - Перенос частей лога в отдельные файлы по достижению установленного размера

Обслуживание-4

- Для серьёзных конфигураций баз данных необходимы:
 - Регламент обслуживания
 - План аварийного восстановления
- FBDataGuard помогает решить много технических проблем, значительно расширяя возможности администратора
 - Алерты и рекомендации
 - FBDataGuard сообщает о найденных или ожидающихся проблемах, выдает рекомендации
 - Устанавливает статус базе данных и серверу в целом

Защита

- Защитный репозиторий метаданных
 - Сохранение копии «сырых» метаданных, необходимых для расшифровки
 - Дополнительная проверка метаданных
- Расшифровка критически поврежденных файлов БД (в т.ч. при сбоях HDD и другого железа)
- Экспорта данных из поврежденных delta-файлов
- Восстановление удаленных данных (в том числе DROP TABLE) (только версия Enterprise)

Вопросы?

- dataguard@ib-aid.com
- Живая демонстрация и общение с разработчиками на стенде IBSurgeon (стол с большим монитором)