

# Отказоустойчивые решения для Firebird 2.5 и 3.0

---

Алексей Ковязин, Дмитрий Кузьменко,  
[www.ib-aid.com](http://www.ib-aid.com), [www.ibase.ru](http://www.ibase.ru)



Advanced Firebird for Big Databases



# IBSurgeon

- Replication, Recovery and Optimization for Firebird and InterBase since 2002
- Платиновый спонсор Firebird Foundation
- Москва, Россия

[www.ib-aid.com](http://www.ib-aid.com)

# В чем проблема?

- Оборудование ломается
- Ошибки
- Человеческий фактор



## 4 ключевых момента

1. Время восстановления
2. % восстановленных данных
3. % потерянных данных с момента последнего бэкапа
4. Шанс неуспешного восстановления

# Варианты с бэкапами

- Вернуться к самому свежнему бэкапу
- Отремонтировать поврежденную БД
- 12 типичных ошибок при бэкапе баз данных
  - <https://habrahabr.ru/post/267881/>

# Итог по бэкапам

- Время восстановления – несколько часов
  - Зависит от размера БД и скорости диска
- % восстановленных данных с последнего бэкапа – от 0 до 90%
  - Зависит от момента повреждения. Чем позже повреждение – тем больше потеря данных
- Шансы невосстановления - низкие low
  - Check [12 Common Mistake in Databases Backups](#)

# Ремонт поврежденных баз

- Инструменты для ремонта:
  - Стандартные: gbak+ gfix
  - Расширенные: **IBSurgeon FirstAID**
- Время восстановления – на 1 гигабайт от 1 минуты до нескольких часов
  - Зависит от скорости диска (и процессора)
  - Часто требуется backup/restore или полный экспорт/импорт данных
- % восстановленных данных – 70-99%
- Иногда может быть комбинировано с данными из бэкапа
  - Увеличивает объем восстанавливаемых данных до 90-99%

# Итог по инструментам восстановления

- Время восстановления – несколько часов
  - Обычно дольше чем восстановление из бэкапа
  - Зависит от размера БД и скорости диска
- % восстанавливаемых данных от 70 до 90%
  - Зависит от сложности повреждения
- Шансы невосстановления - средние
  - Некоторые повреждения невосстановимы

# Нужно помнить

- Восстановление из бэкапа
  - Исходная база могла иметь дефекты (невосстановимый бэкап)
  - Восстановить можно IBackupSurgeon
- Ремонт БД
  - Чаще всего после ремонта требуется backup/restore
  - Количество итераций backup/restore может потребоваться >1
  - Иногда после ремонта требуется экспорт (FirstAid Extractor)
- Изложенное справедливо для gbak и nbackup

HIGH AVAILABILITY

---

# Что такое High Availability

- Standby кластер
- Cold standby, warm standby, hot standby
- Мастер/слэйв, или первичный-вторичный
- Репликация
  - Физическая (уровня страницы) и логическая (уровня записи)
- Другие средства
- Виртуальные машины
- Синхронность и асинхронность

# DRBD

- Решение
  - Distributed replicated block device
  - Удаленно-зеркалированная ФС
  - Работает прозрачно
  - Синхронный и асинхронный режимы
  - Open-source (GPL)
- Особенности
  - Невозможность онлайн-валидации
  - Неизвестно, что происходит в онлайн
  - Только для Linux
  - Запись дублируется
  - Чтение не масштабируется

# Shadow – теневая копия

- Решение

- Команда `CREATE SHADOW`
- Раздел с тенью подключается по NFS / SMB
- `RemoteFileOpenAbility` → `true`
- Работает синхронно

- Особенности

- Независимость от платформы
- Запись на диск дублируется
- Неизвестно, что происходит в онлайне
- Чтение не масштабируется

## Итог по DRBD/Shadow

- Время восстановления – от минут до часов
  - Требуется вмешательство администратора
- % восстановленных данных – от 0 до 100%
  - Зависит от типа повреждения
- Шансы невозможности восстановления - средние
  - Статус неизвестен до окончания restore

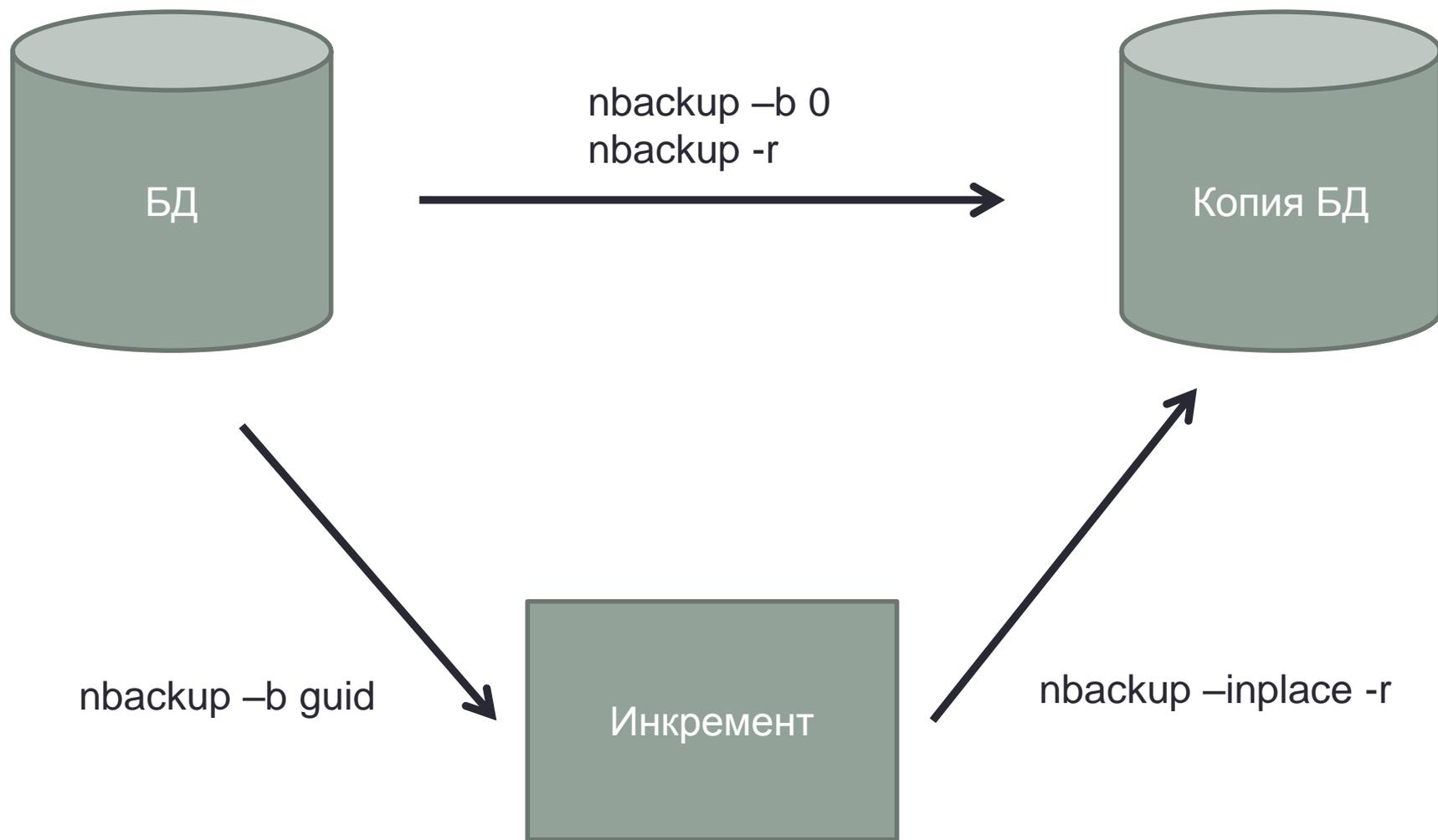
# HA и Бэкапы виртуальных машин

- Поставщики VM вносят путаницу между HA и Backup
- Бэкап VM это почти nbackup, только не БД, а файловой системы
- Бэкап без VSS похож на «снимок файловой системы в момент reset»
  - Без поддержки VSS база имеет шансы считаться «поврежденной» на этот момент
- VM бесполезна для защиты от повреждений БД
- Hqbird содержит VSS-провайдер

# Итог по бэкапам ВМ

- Время восстановления – от минут до часов
  - То же, что и nbackup
- % восстановленных данных – от 0 до 100%
  - Зависит от момента повреждения и скорости дисков
- Шанс невозможности восстановления - высокий

# Nbackup (Firebird 4.0)



# Репликация

- Физическая
  - Некоторые решения уже существуют
  - Но они страдают от одинаковых проблем
    - Ошибки физической структуры реплицируются
  - Требуется очень «широкого» канала
- Логическая
  - Выполняется на уровне записей
  - Меньше поток данных
  - Ошибки памяти и физической структуры не реплицируются
  - Возможен read-only доступ к слэйву - можно масштабировать чтение

# Репликация на триггерах

- Решение

- Журнал в виде таблицы, поддерживается триггерами
- По запросу (асинхронная)
- Гибкая конфигурация
- Возможен режим «мульти-мастер»

- Особенности

- Очистка журнала → сборка мусора
- Генераторы реплицируются вручную
- DDL не реплицируется в принципе
- Таблица логов – узкое место (производительность, версии, мусор)

# Встроенная репликация

	<b>HQbird</b>	<b>RedDatabase</b>	<b>Avalerion</b>	<b>Firebird 4</b>
Статус (Декабрь 2016)	Production	Production	Production	Alpha
Поддерживаемые версии Firebird	2.5, 3.0	2.5	3.0	4.0
ODS	11.2, 12.0	11.3	12.1	13.0
Средства миграции баз Firebird 2.5 и 3.0	Не нужны (та же ODS)	Backup - Restore	ALTER statements	Backup/ Restore
Синхронная	Да	Да	Нет	Да
Асинхронная	Да	Да	Да	Да

HQBIRD

---

# HQbird Enterprise

- Встроенная репликация
- Инструменты оптимизации
- Резервное копирование, восстановление, sweeper, интеграция с VM
- Мониторинг – транзакции, индексы, метаданные, свободное место, temp, дельта, ...

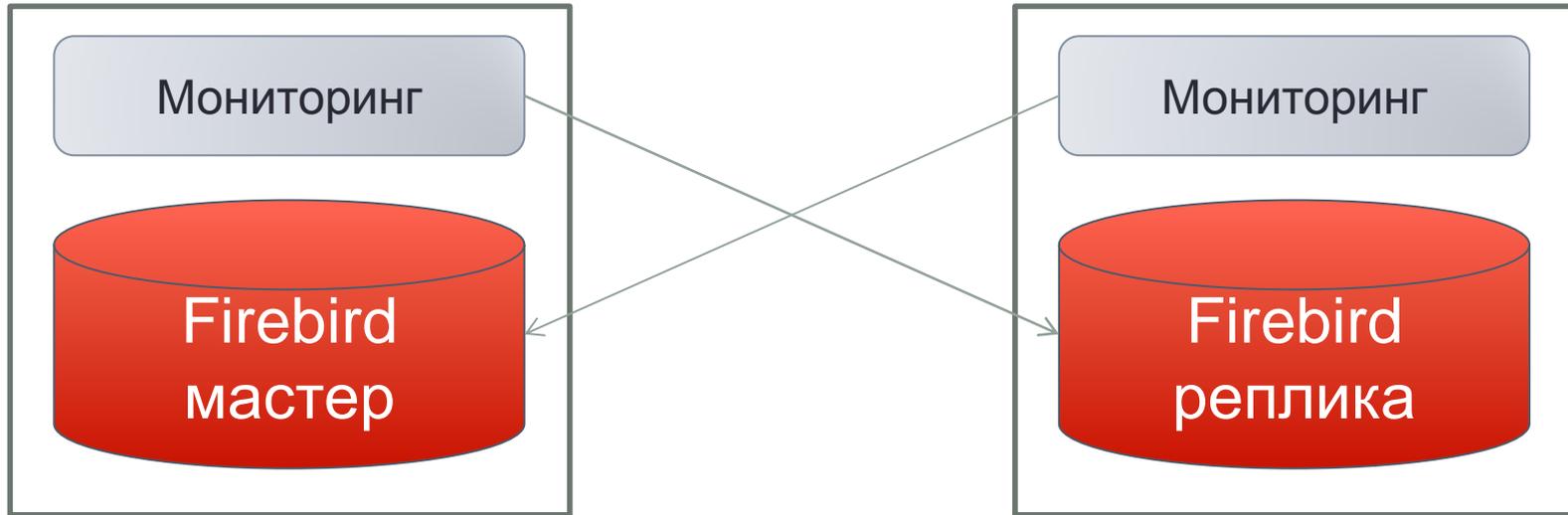
## Кластер БД не является:

- 1) Автоматическим решением
  - Требуется настройка (2-3 часа работы)
- 2) Действительно масштабируемым решением
- 3) Репликацией мульти-мастер
  - Реплика – только для чтения
- 4) Географически распределенным решением
  - Требуется быстрое соединение

# Кластер с точки зрения приложения

1. Приложения работают с БД Firebird, как обычно
2. Если база становится недоступной, производится попытка реконнекта к тому же серверу
3. Приложения должны использовать кэширование (например, `CachedUpdates`) для исключения потери несохраненных данных

# Как работает кластер: Сервер

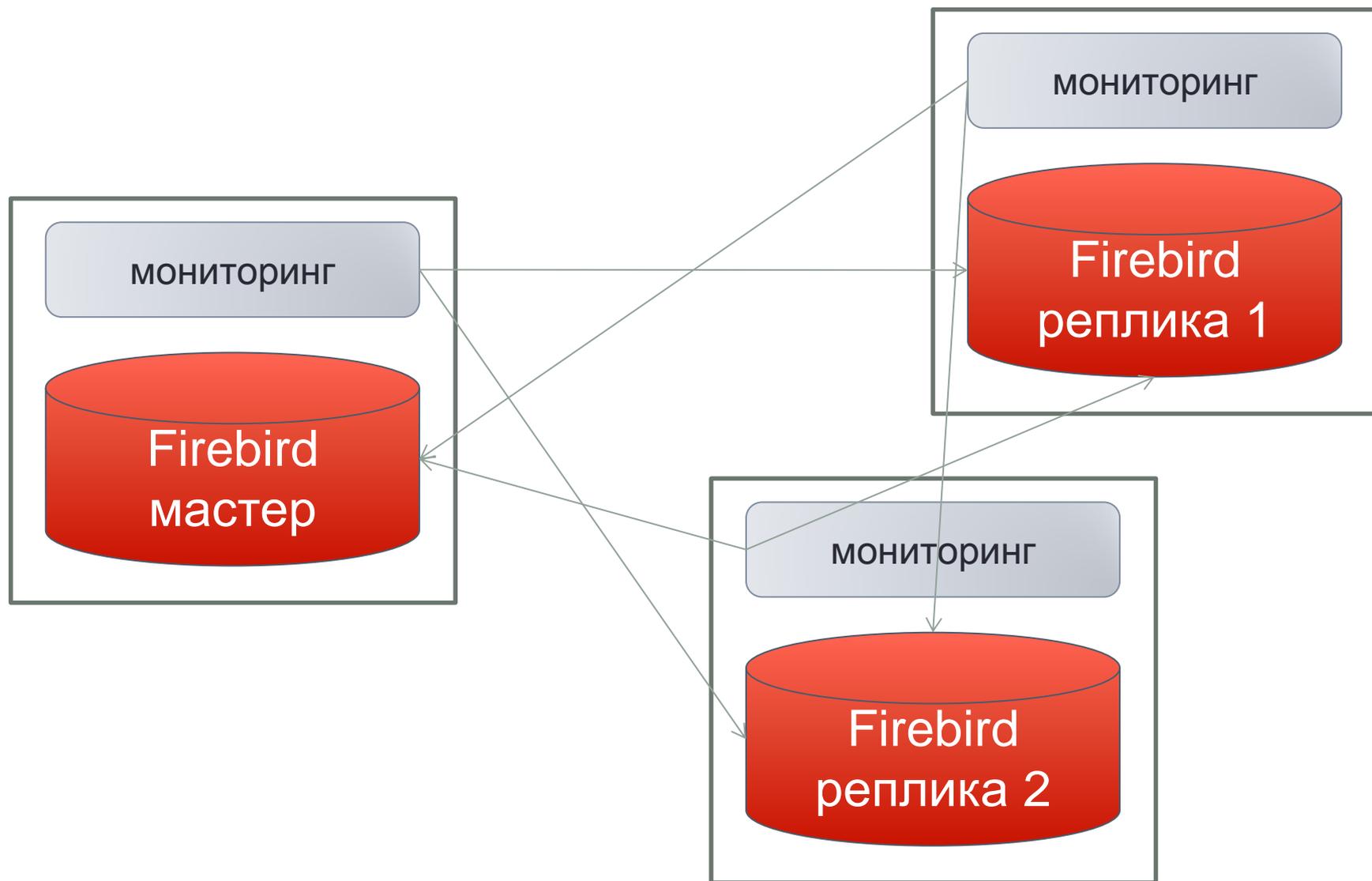


1. Узлы следят друг за другом
2. Если отказ на мастере, реплика
  - а) «пристреливает» мастер
  - б) подменяет собой мастер – скрипт + DNS

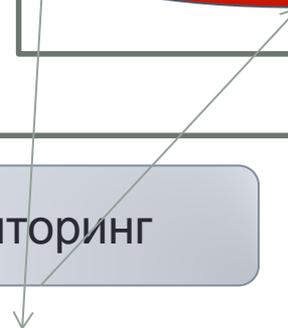
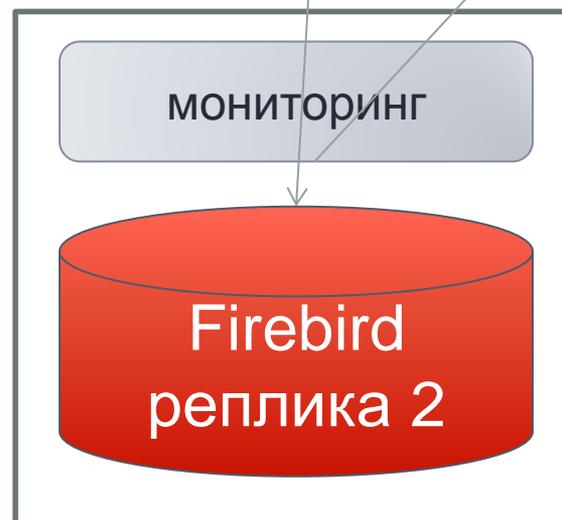
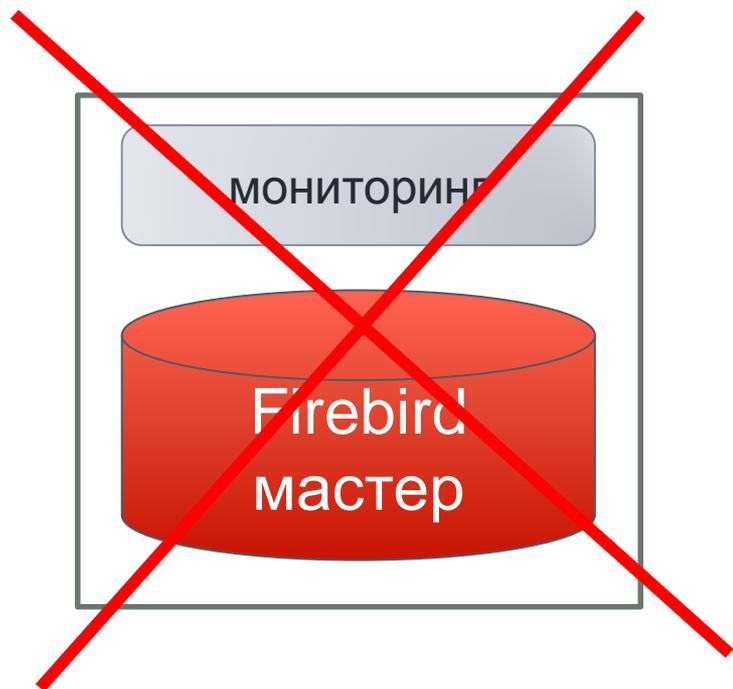
После сбоя кластера уже нет!  
Есть только одна БД!



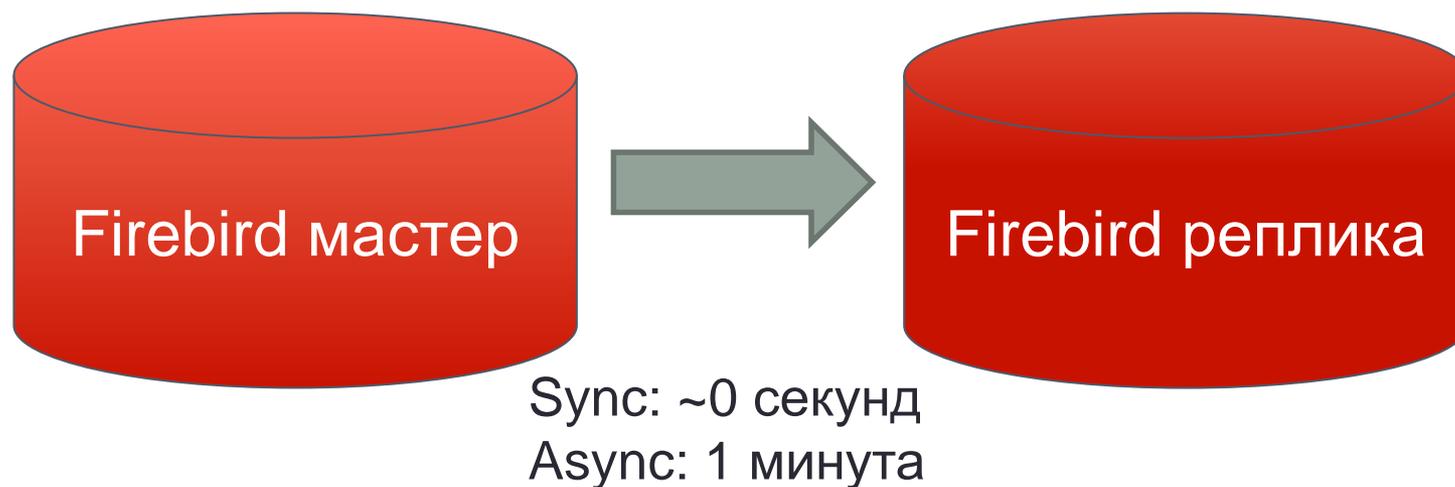
# Отказоустойчивый кластер



# Отказоустойчивый кластер



# Репликация в HQbird



1. Backup/restore не требуется
2. Опционально Sync или Async
3. Простая настройка

# Шаги по настройке

## Синхронная

1. Остановить Firebird
2. Скопировать БД на реплику
3. Сконфигурировать репликацию на мастере
4. Сконфигурировать репликацию на реплике
5. Запустить Firebird на реплике и мастере

## Асинхронная

1. Остановить Firebird
2. Скопировать БД
3. Сконфигурировать репликацию на мастере
4. Запустить Firebird
5. Скопировать БД на реплику, сконфигурировать

# Настройка репликации

Database replication configuration: test ✕

Replication role  Off  Master  Replica

Working mode  Synchronous  Asynchronous

Log archive directory

Optional parameters

Verbose

Cancel

Save

Cancel

Save

# Время простоя при настройке

- Синхронная
  - Время копирования БД с сервера на сервер + время настройки узлов
- Асинхронная
  - Время копирования БД
  - Подходит для географически распределенных систем

Асинхронная настраивается гораздо быстрее!

# Мониторинг и переключение с мастера на реплику

- HQbird (Windows, Linux)
- Расетmaker (Linux)
- Другие инструменты (разработанные самостоятельно)

## Fail-Safe Cluster – сложное решение

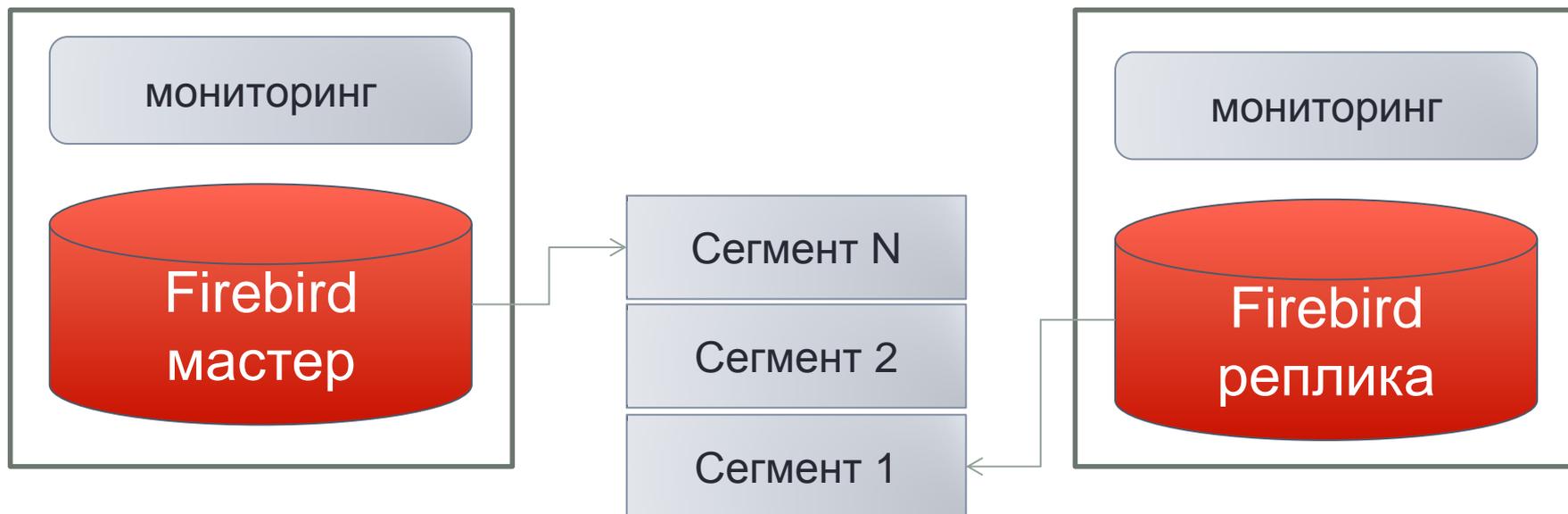
- Hqbird дает варианты, вам решать, как их использовать
- Требуется тщательное планирование
- Требуется инфраструктура
  - 2 или более реплики
  - Отключение сети на 10 секунд приведет к устареванию реплик
  - Быстрая сеть и быстрые диски

## Итог по отказоустойчивому кластеру

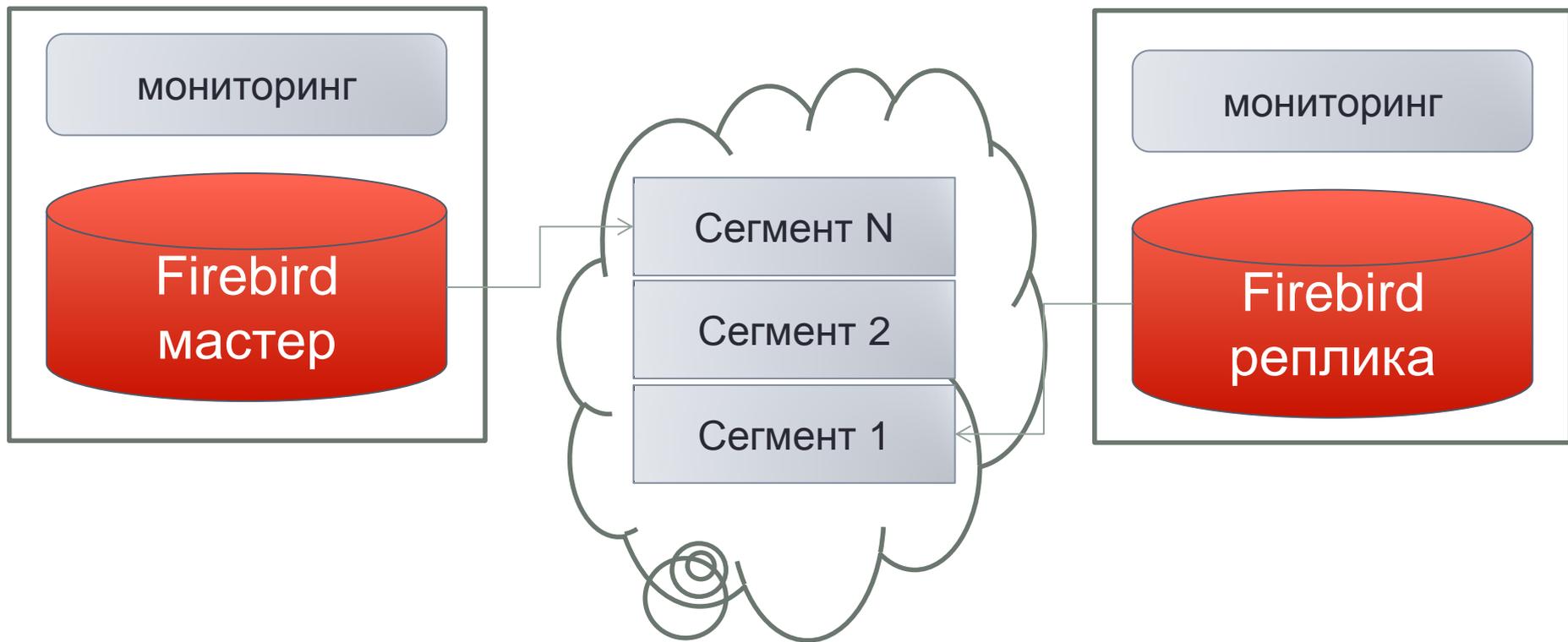
- Время восстановления – 30-120 секунд
  - Приложения должны переподключиться
- % восстановленных данных – 100%
  - Теряются только несохраненные изменения
- Шанс невосстановления - низкий

# Warm-standby (зеркало)

- Базируется на асинхронной репликации
- Простая и быстрая настройка



Может быть распределен  
географически (FTP, Amazon, cloud  
backup, VPS)



## Итог по Warm-standby

- Время восстановления – минуты
  - Приложения должны переподсоединиться
- % восстановленных данных – 100%
  - Теряются только несохраненные изменения
- Шанс невосстановления – низкий

# Варианты

	Время восстановл ения	% восстанавл иваемых данных	Шансы невосстановл ения
Возврат к бэкапу	Часы	0-90%	Низкие
Ремонт БД (в комбинации с бэкапом)	Часы	50-99%	Средние
HA-решения			
Virtual Machine High Availability	Минуты	0-99%	Высокие
DRBD	Минуты	0-99%	Средние
Failover-cluster	Секунды	100%	Низкие
Warm standby	Минуты	100%	Низкие

Спасибо за внимание!



Advanced Firebird for Big Databases

<http://www.ibase.ru/hqbird/>  
[support@ibase.ru](mailto:support@ibase.ru)