

# Распределенные системы хранения и обработки данных

Владислав Белогрудов, EMC

[vlad.belogrudov@gmail.com](mailto:vlad.belogrudov@gmail.com)

# Лекция 8

Непрерывность бизнеса, резервное копирование и восстановление

# Содержание лекции

- необходимость
- метрики непрерывности
- методы борьбы с SPOF
- топологии резервного копирования
- дедупликация
- резервное копирование в виртуальной среде

# Доступность информации

Информация может быть недоступна из-за

- наводнений, пожаров, землетрясений
- сетевых атак
- ошибок персонала
- сбоев оборудования
- обновления ПО и оборудования
- архивирования и восстановления



Организации теряют большие деньги

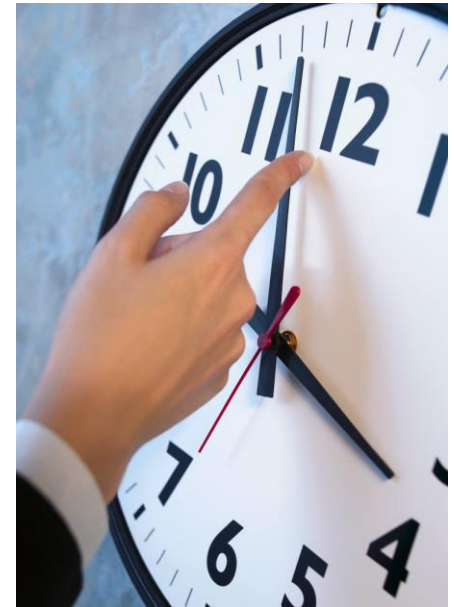
# Непрерывность бизнеса

- Business Continuity (BC)
  - процесс, охватывающий все предприятие
  - действия для уменьшения влияния простоев
  - обеспечение безопасности
  - профилактика
  - оценка рисков
  - защита данных

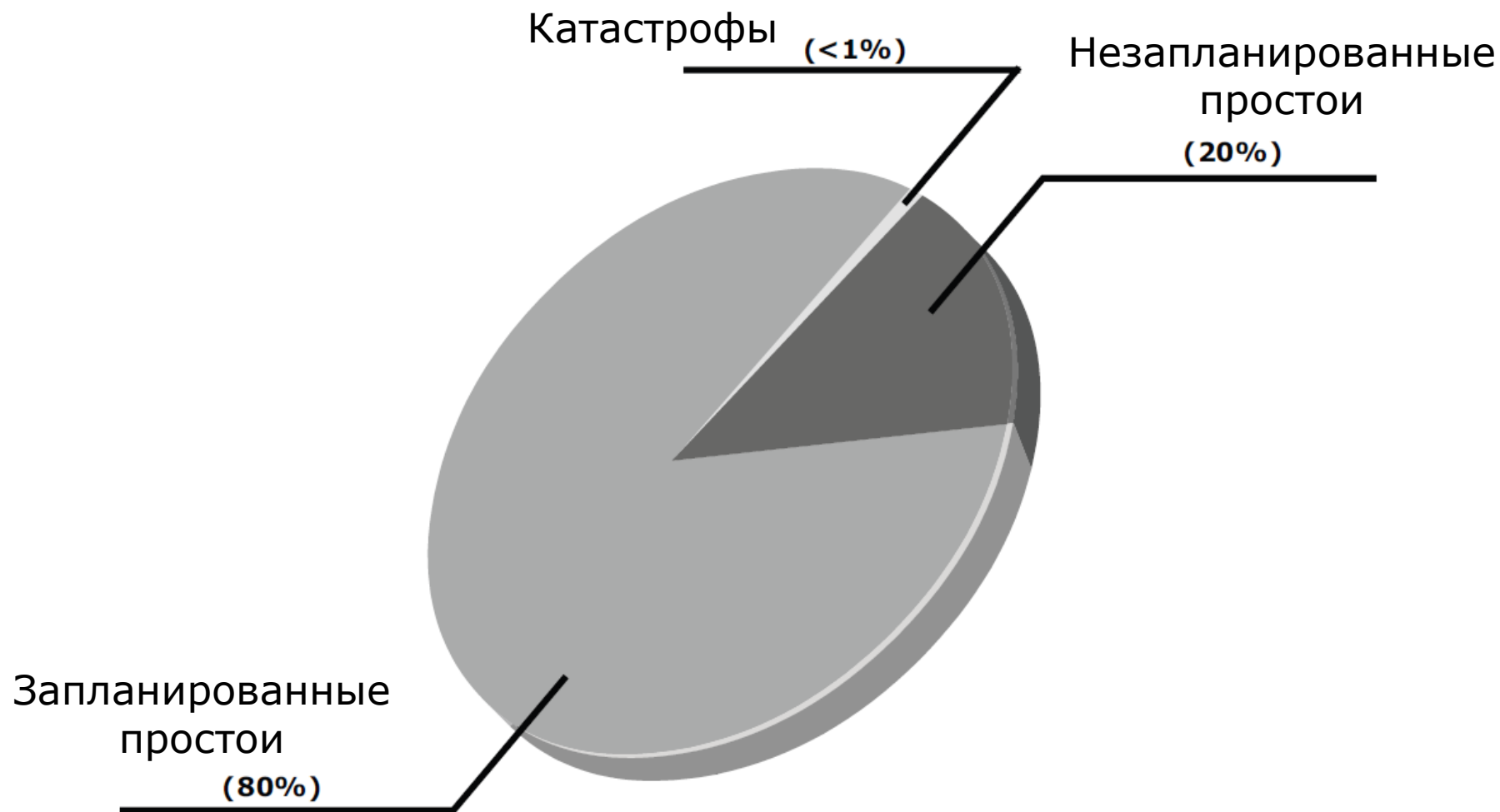


# Доступность информации

- Возможность инфраструктуры функционировать
  - в течение ожидаемого времени
  - в соответствии с требованиями
- Определяется
  - надежностью
  - возможностью доступа
  - своевременностью



# Причины недоступности



# Измерение доступности

- Среднее время между отказами (MTBF)
  - Компонент работает без сбоев

- Среднее время восстановления (MTTR)

MTTR = время на обнаружение

+

время на диагностику

+

время на восстановление

+

время на тестирование





# Расчет доступности

- Доступность =  
$$\text{uptime} / (\text{uptime} + \text{downtime})$$
  
или  
$$\text{MTBF} / (\text{MTBF} + \text{MTTR})$$
- Описывается девятками (99.9%, 99.999%)

# Процент доступности

Доступность %	Время простоя в год	Время простоя в месяц	Время простоя в неделю
90% ("одна девятка")	36.5 дней	72 часов	16.8 часов
95%	18.25 дней	36 часов	8.4 часов
98%	7.30 дней	14.4 часов	3.36 часов
99% ("две девятки")	3.65 дней	7.20 часов	1.68 часов
99.9% ("три девятки")	8.76 часов	43.2 минут	10.1 минут
99.99% ("четыре девятки")	52.56 минут	4.32 минут	1.01 минут
99.999% ("пять девяток")	5.26 минут	25.9 секунд	6.05 секунд
99.9999% ("шесть девяток")	31.5 секунд	2.59 секунд	0.605 секунд

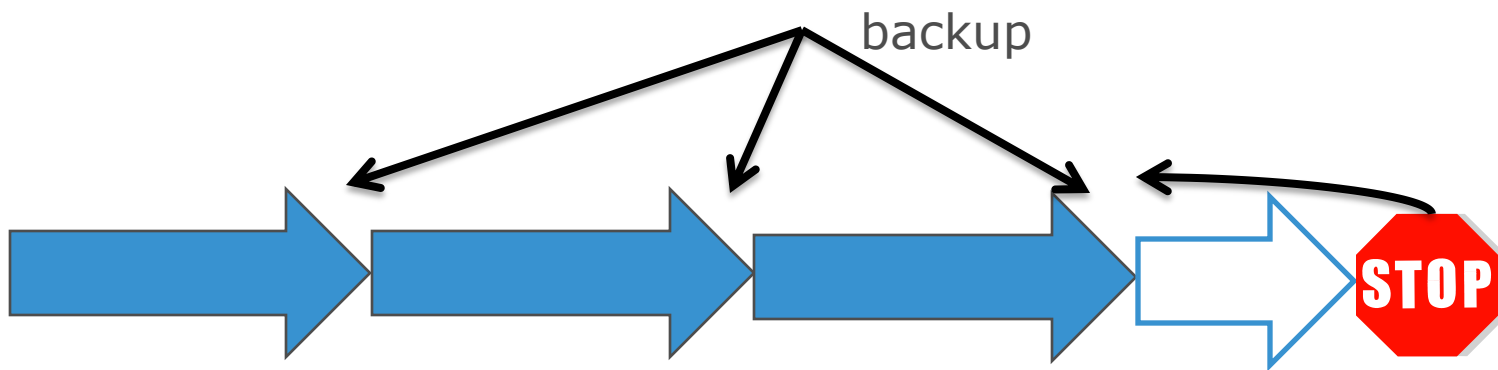
## Еще немного терминов

- Disaster Recovery (DR) – процесс аварийного восстановления систем и данных (замена компонентов, копирование из архивов и т.п.)
- Disaster Restart (DR) – процесс перезапуска бизнеса с зеркальных копий данных и приложений

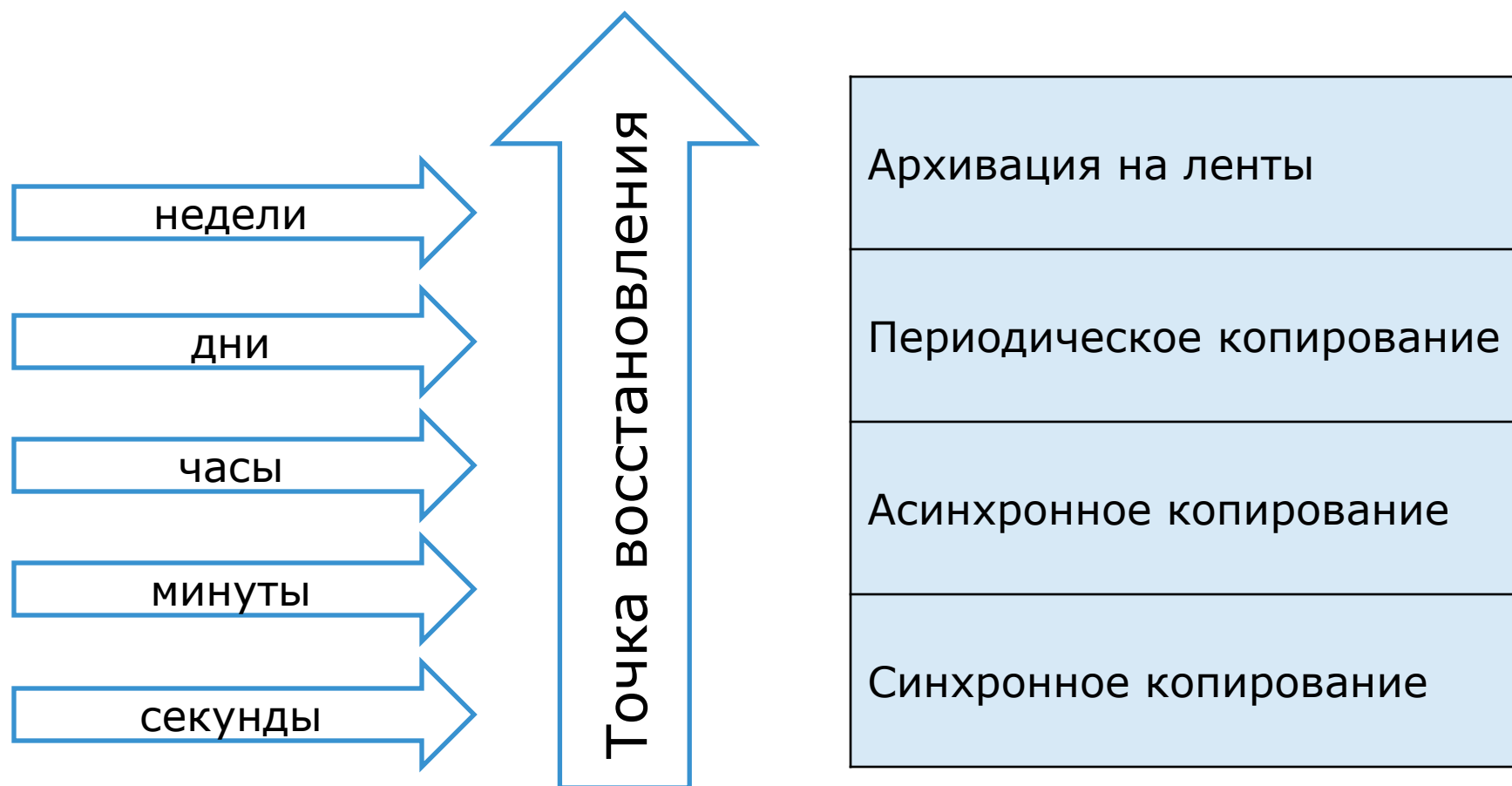


# Точки и время восстановления

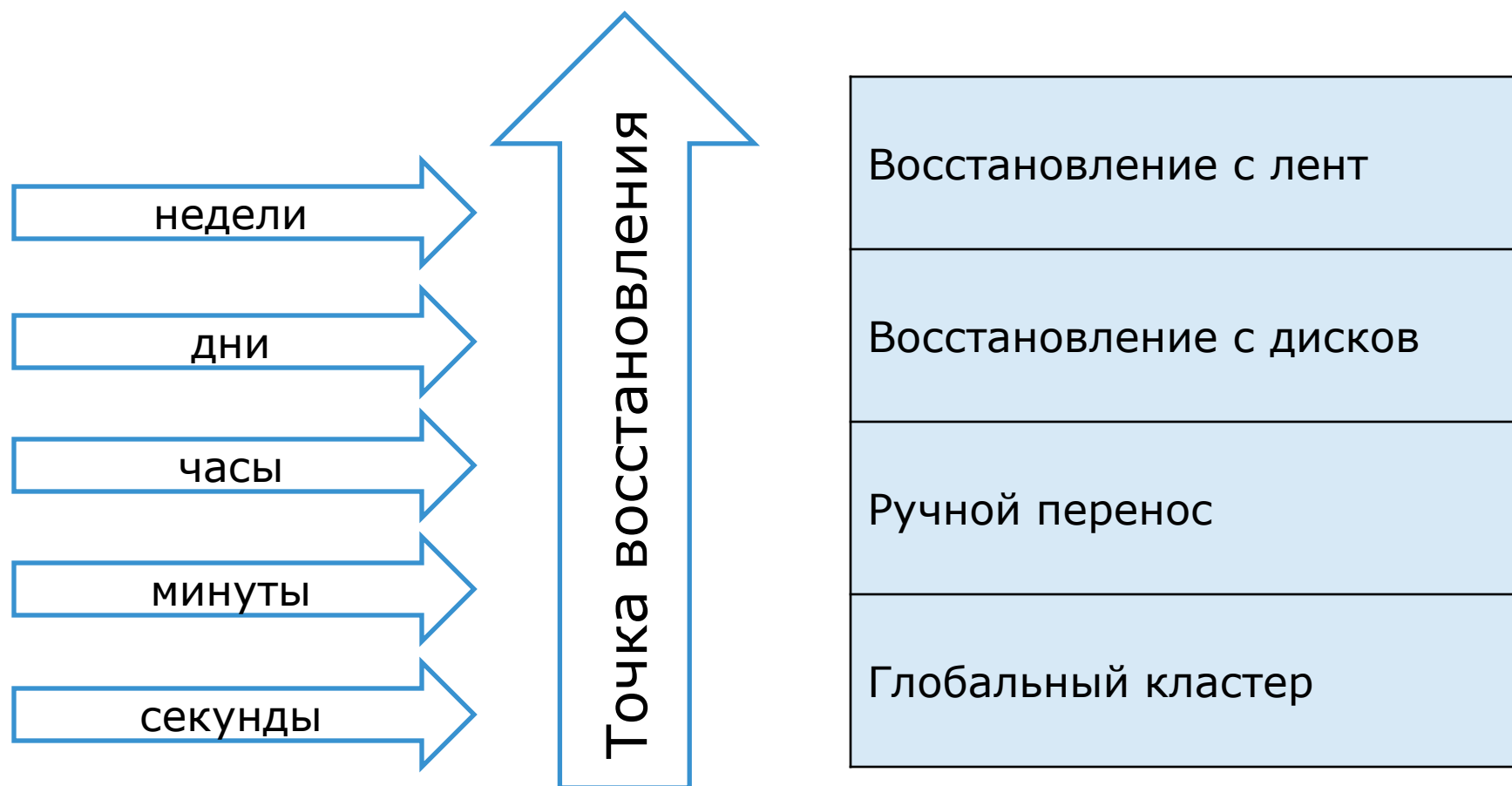
- Recovery Point Objective (RPO) – период времени, за который «не жалко» потерять изменение данных
- Recovery Time Objective (RTO) – период времени, за который система должна быть восстановлена



# Стратегии достижения RPO



# Стратегии достижения RTO

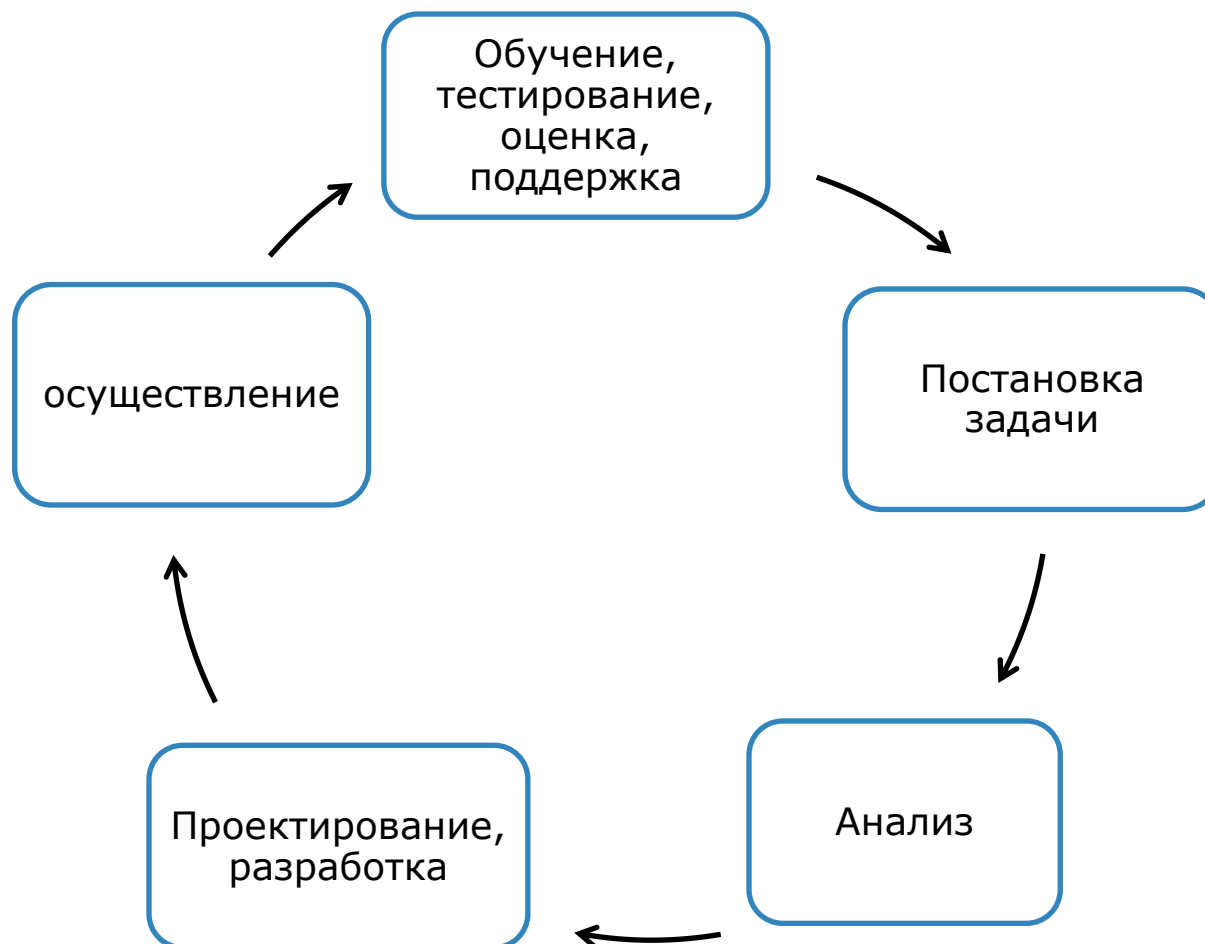


# Способы увеличения доступности

- Интеллектуальные СХД
- «Горячий» офис
- «Холодный» офис
- Кластер



# Планирование непрерывности





# ВС: Постановка Задачи

- Определение требований
- Оценка возможностей и бюджета
- Набрать команду экспертов
- Создать правила и политики

1

# ВС: Анализ

## 2

- Собрать информацию о
  - данных
  - процессах
  - зависимостях
  - инфраструктуре
- Установить приоритеты
- Проанализировать риски
- Создать стратегию смягчения рисков
- Проанализировать стоимость решения

# ВС: Проектирование и разработка

- Сформировать команды, распределить роли
- Разработать сценарии действий
- Проработать процедуры восстановления и перезапуска

3

# ВС: Внедрение

- Обеспечить архивирование, копирование данных, управление ресурсами
- Подготовить запасные места
- Проанализировать и исключить единые точки отказа



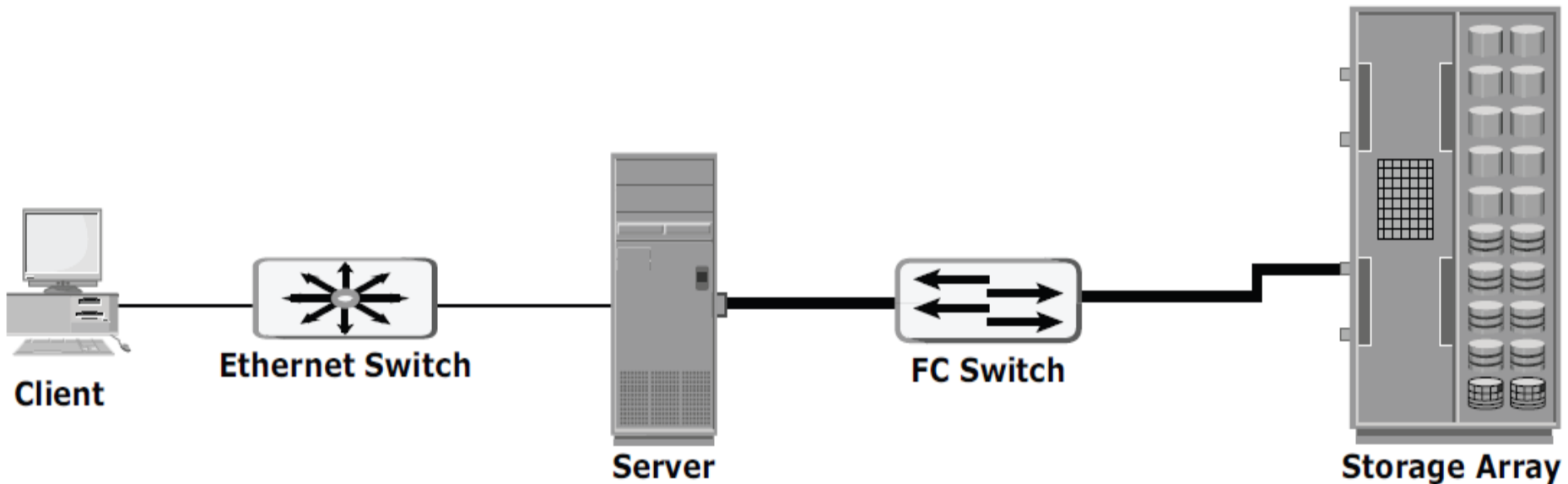
# ВС: Обучение, тест, поддержка

- Обучить персонал
- Обучить команду восстановления
- Протестировать план
- Провести оценку
- Вернуться на шаг 1 😊

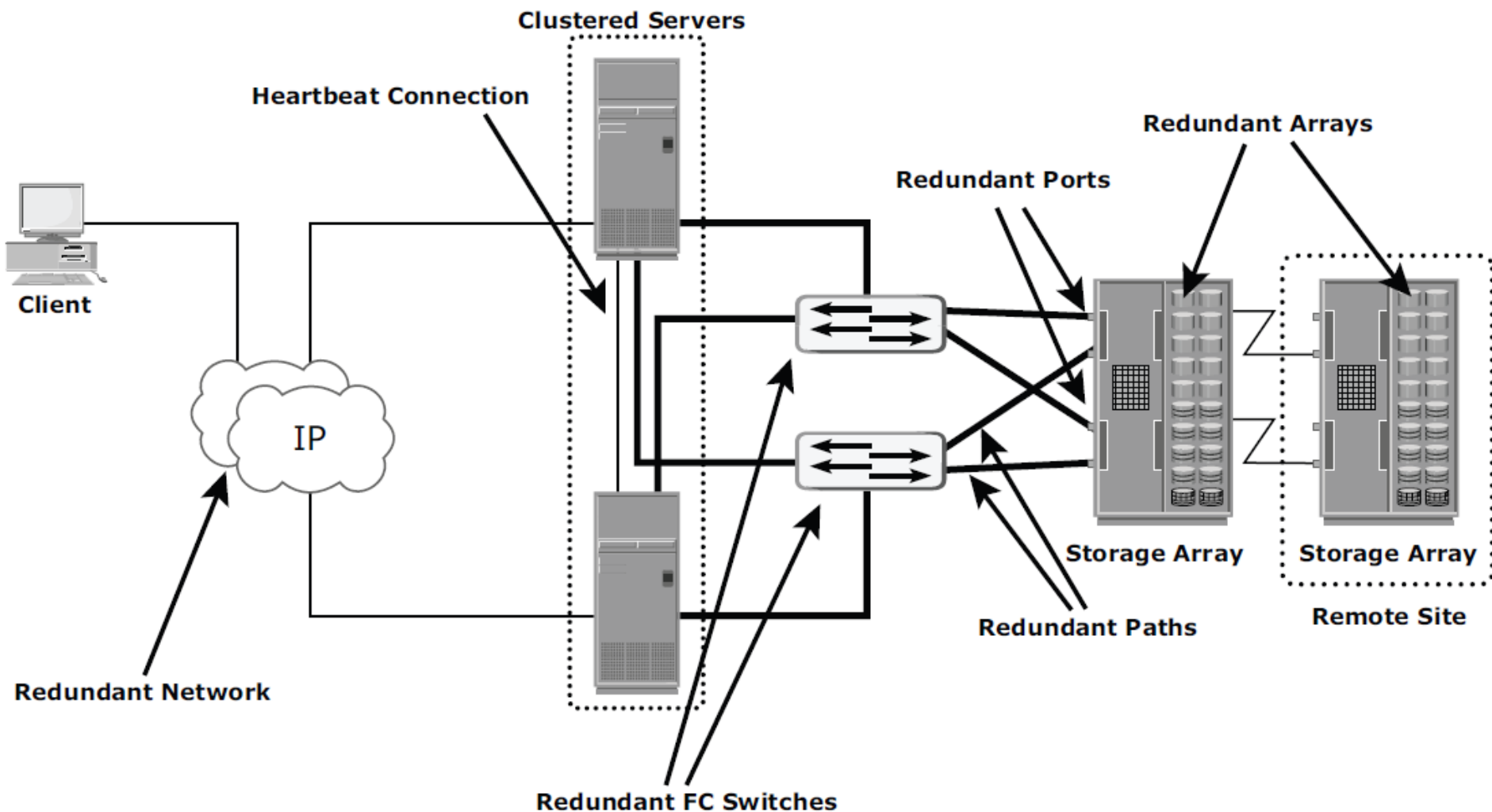
5

# Единая точка отказа

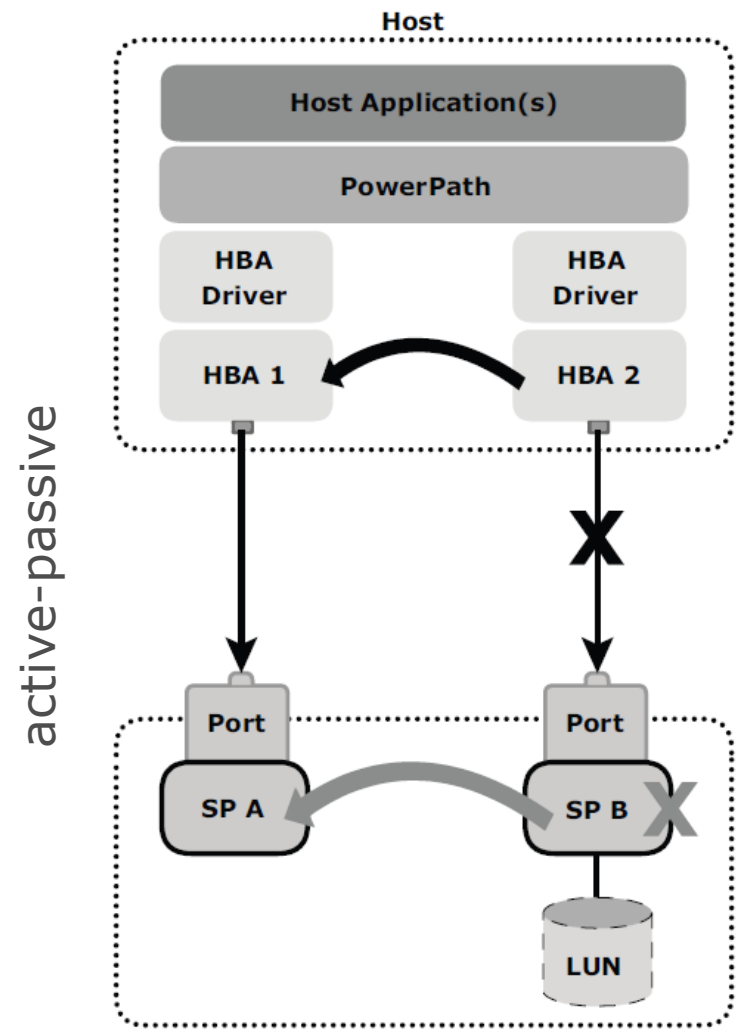
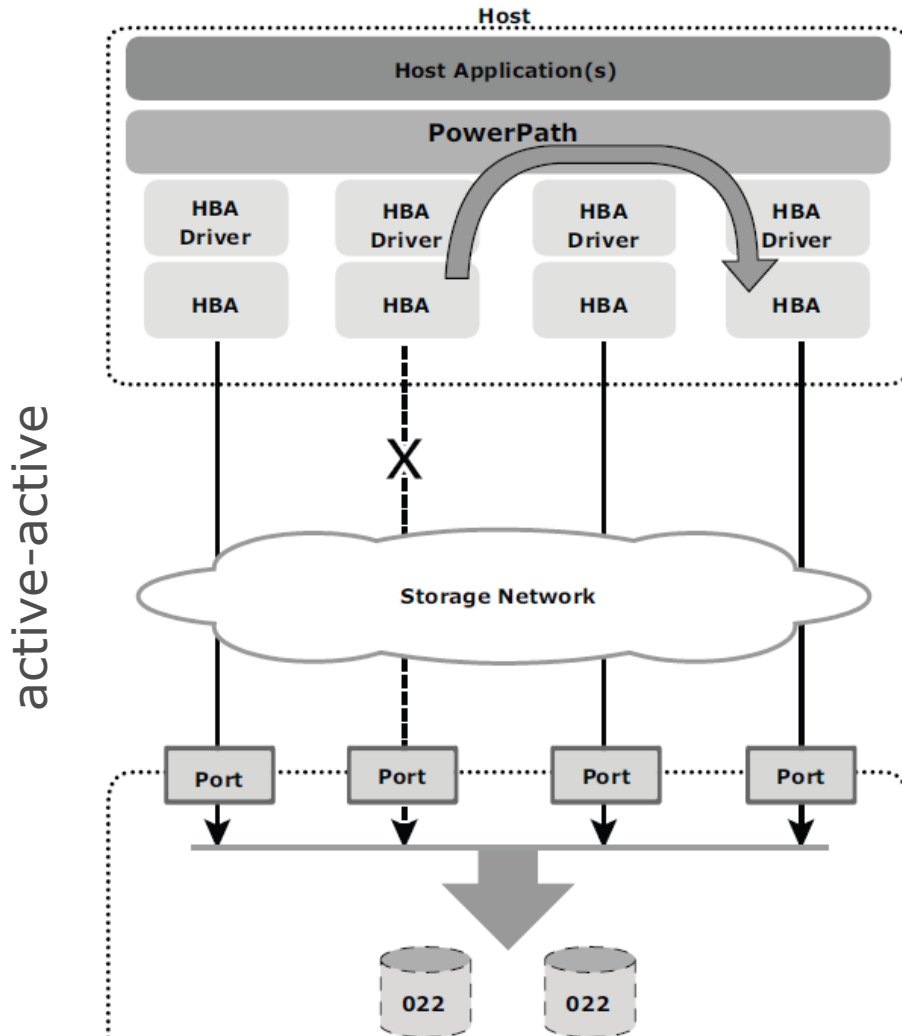
- Single Point of Failure (SPOF)
- Компонент, отказ которого приведет к отказу всей системы



# Отказоустойчивость



# EMC PowerPath - multipathing

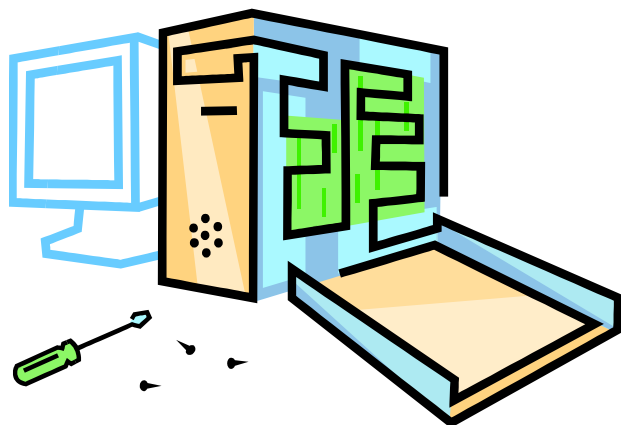




# Резервное копирование

- Цели

- восстановление после отказа
- операционное копирование
- архивирование



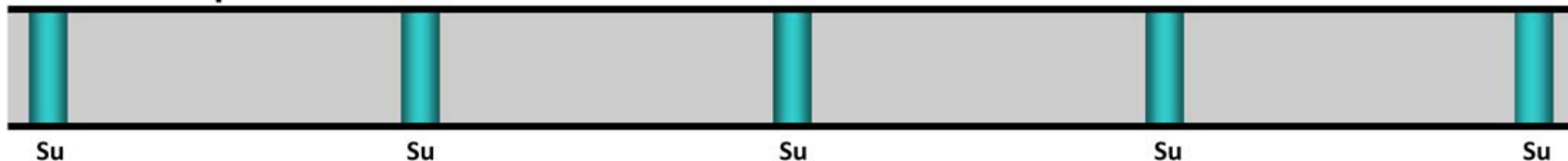
# Принципы резервного копирования

- RTO & RPO
- Период хранения копий
- Вид носителя
- Гранулярность
- Сжатие и дедупликация
- Местоположение, размер и количество файлов
- Время создания копий

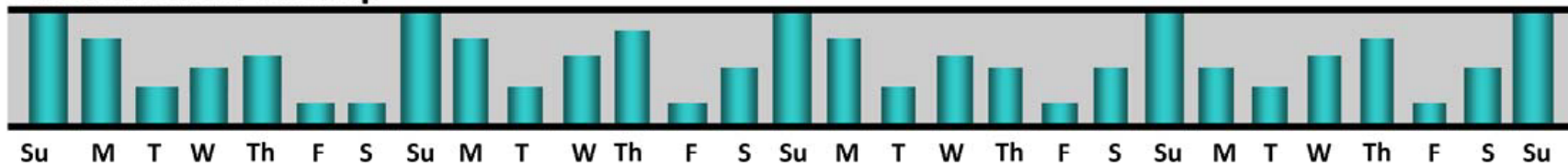


# Гранулярность копирования

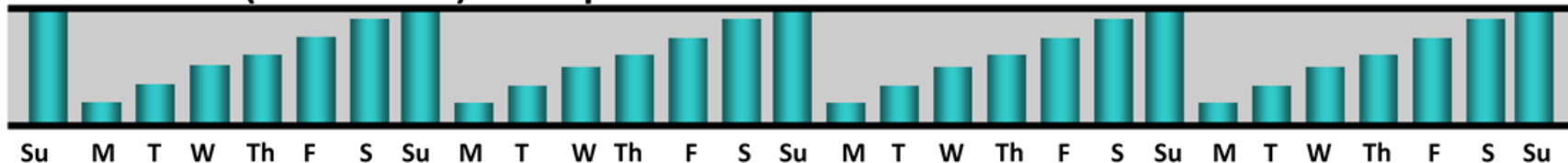
**Full Backup**



**Incremental Backup**

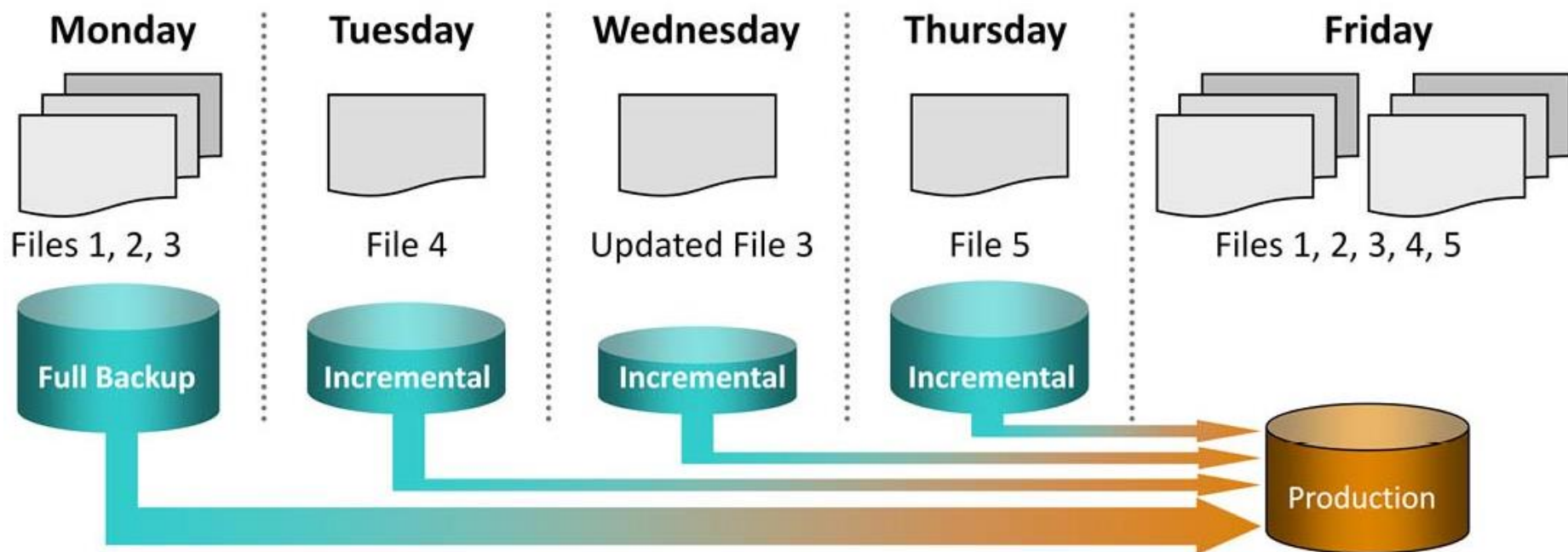


**Cumulative (Differential) Backup**

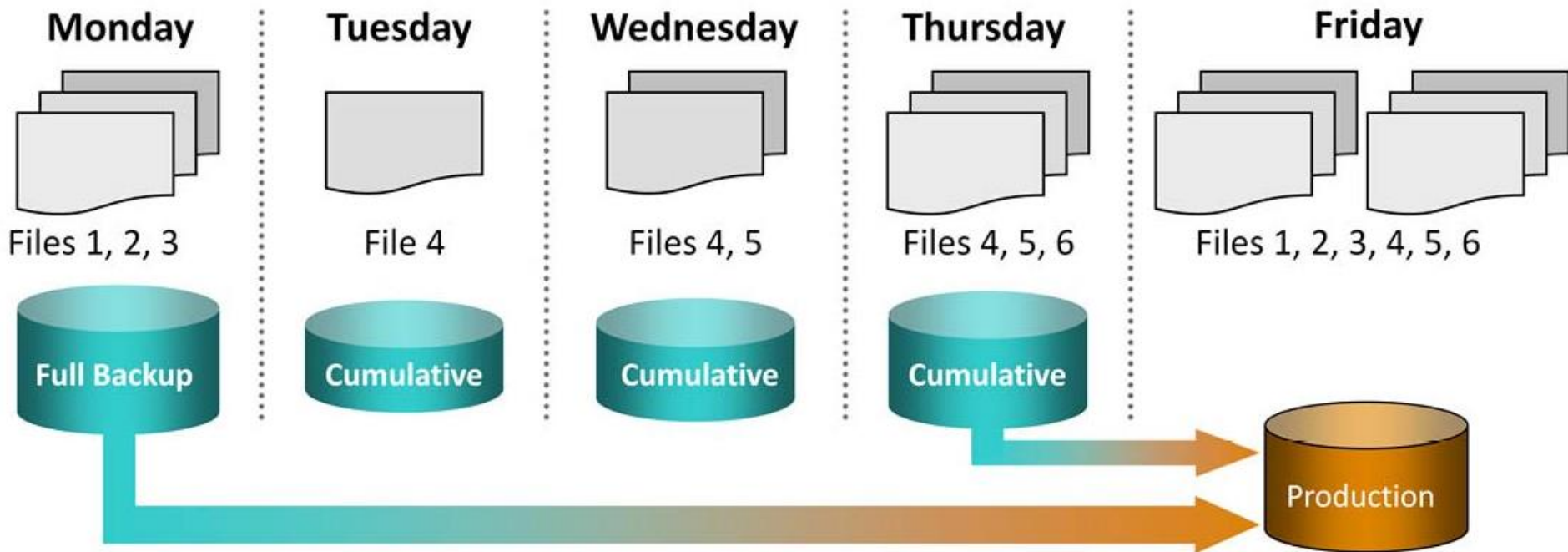


**Amount of Data Backup**

# Инкрементальные копии - восстановление

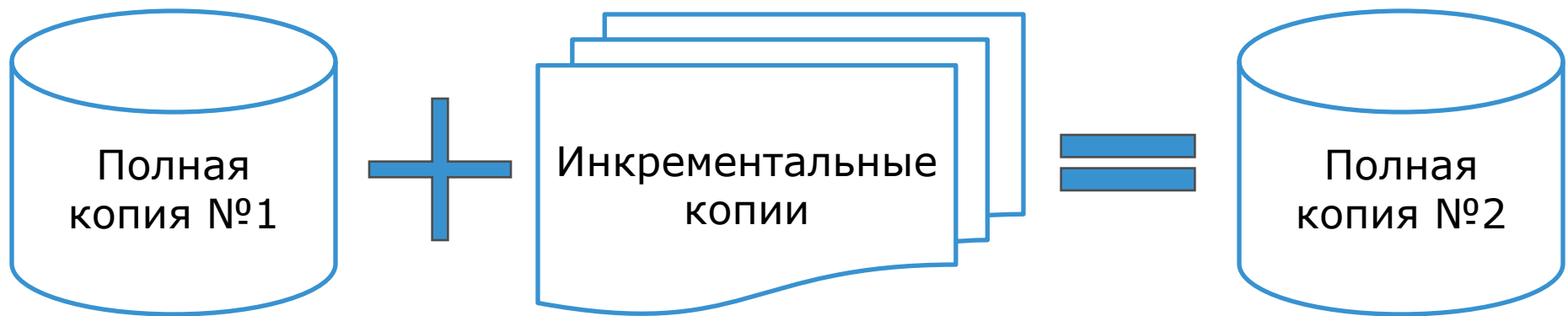


# Дифференциальные копии - восстановление



# Синтетические полные копии

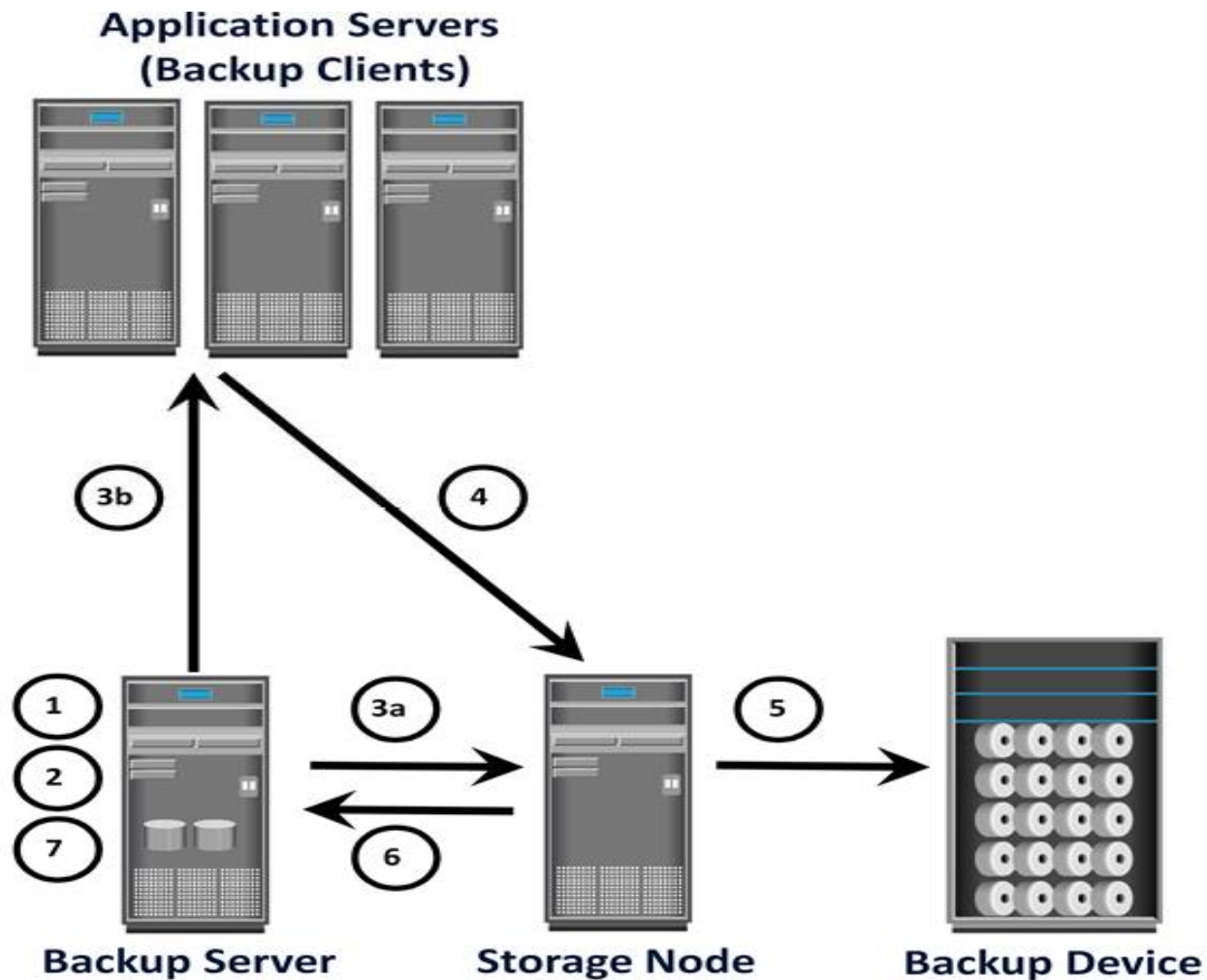
- Освобождение ресурсов
  - сеть
  - сервер



# Способы резервного копирования

- Горячие копии
  - необходимы агенты запущенных приложений
- Холодные копии
  - полный останов программ, проще, надежнее
- Резервирование файлов
  - система копирования ждет, пока файлы «освободятся». Можно не дожидаться 😊
- Point in Time (PIT) – мгновенные снимки системы
  - инкремент хранится в другом месте (файл, LUN)

# Процесс резервного копирования

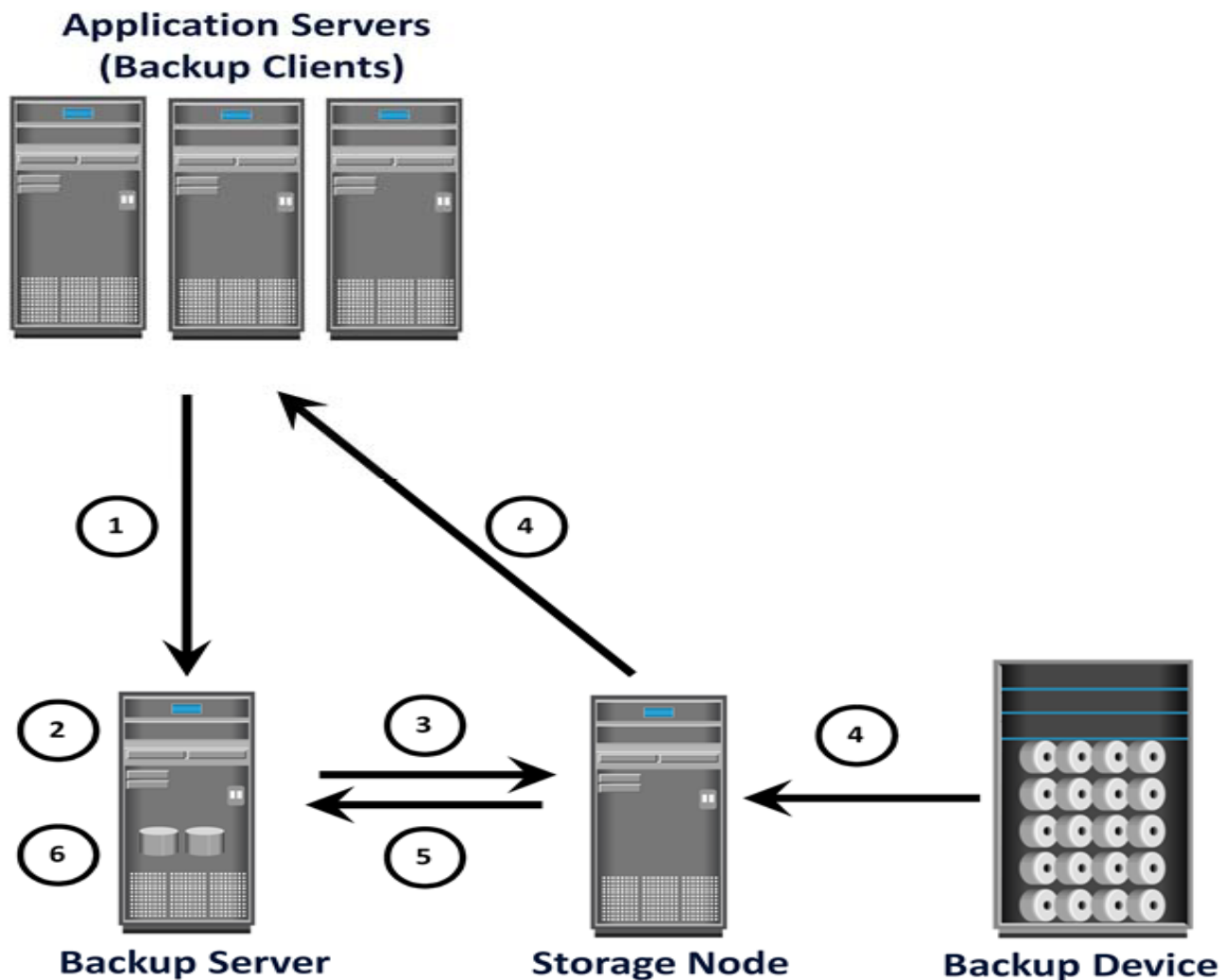




## Процесс резервного копирования (2)

1. Начало процесса по графику
2. Чтение каталога
3. Загрузка носителей, начало передачи данных
4. Передача данных
5. Запись данных
6. Передача метаданных
7. Обновление каталога

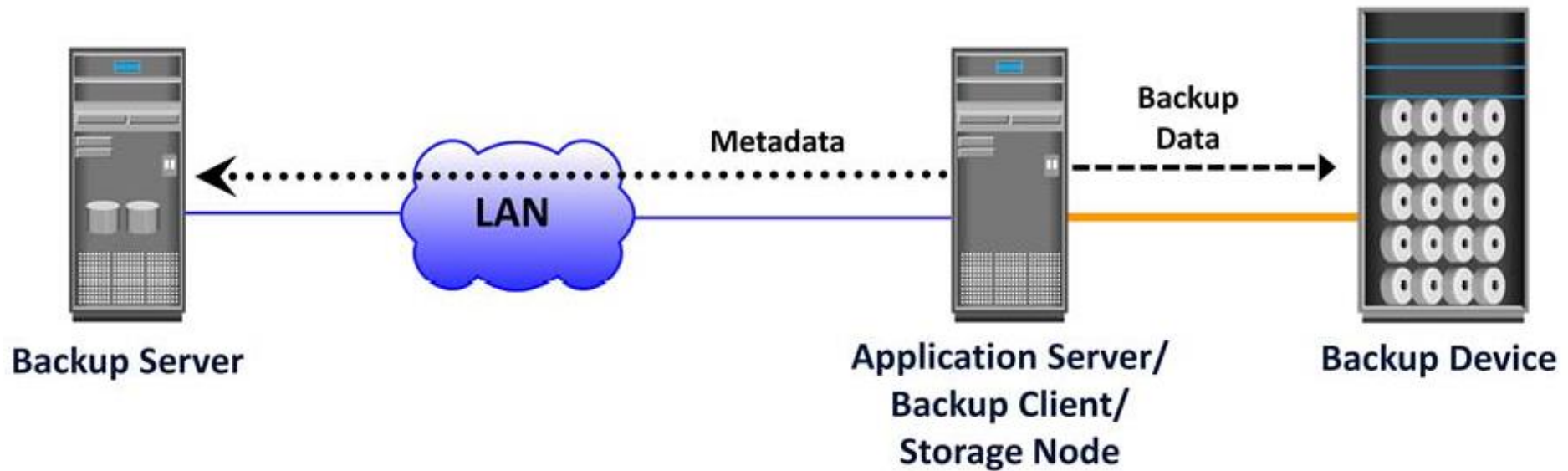
# Процесс восстановления



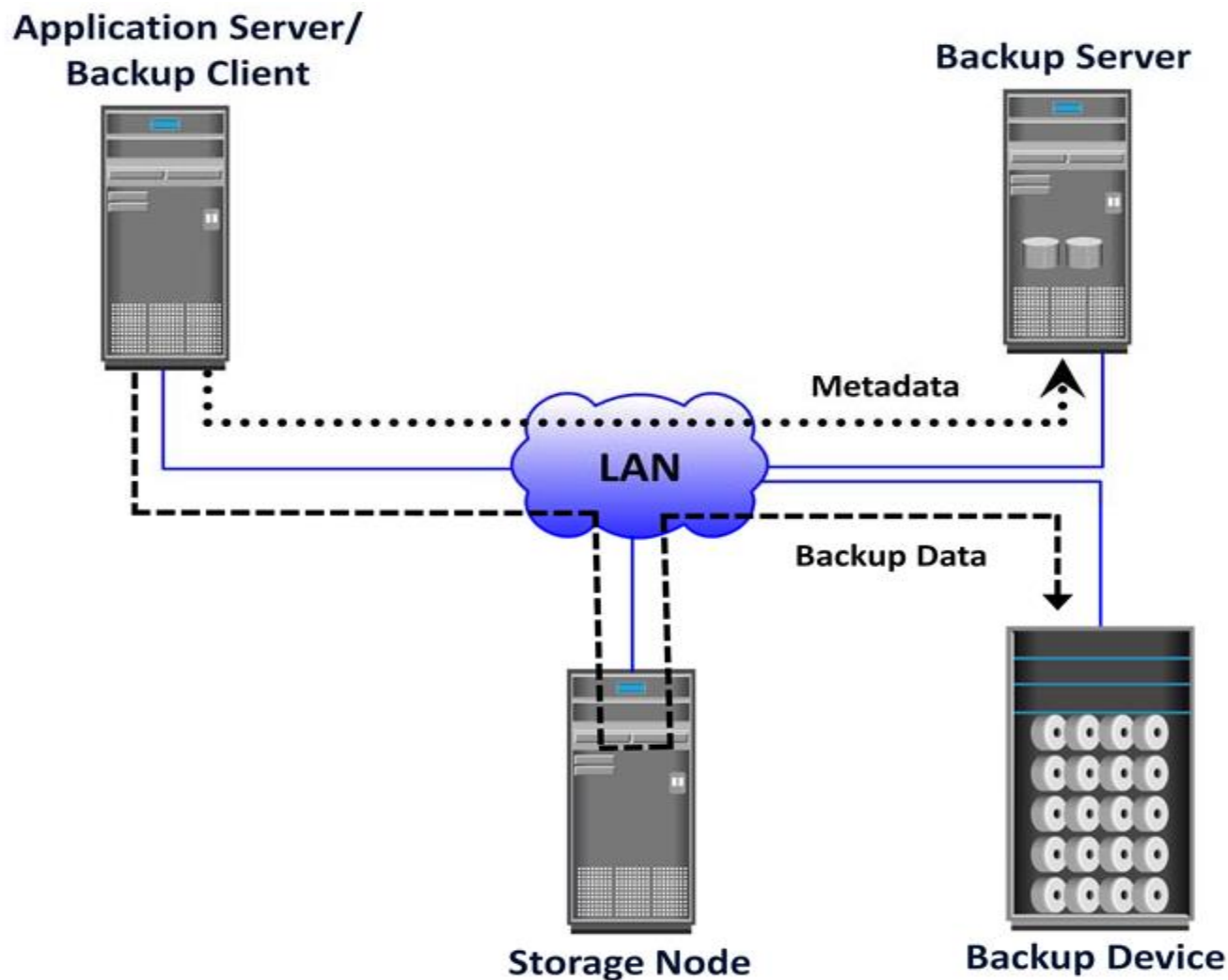
## Процесс восстановления (2)

1. Чтение метаданных файлов из каталога
2. Загрузка носителя данных
3. Передача файлов клиенту
4. Передача метаданных в каталог
5. Обновление каталога

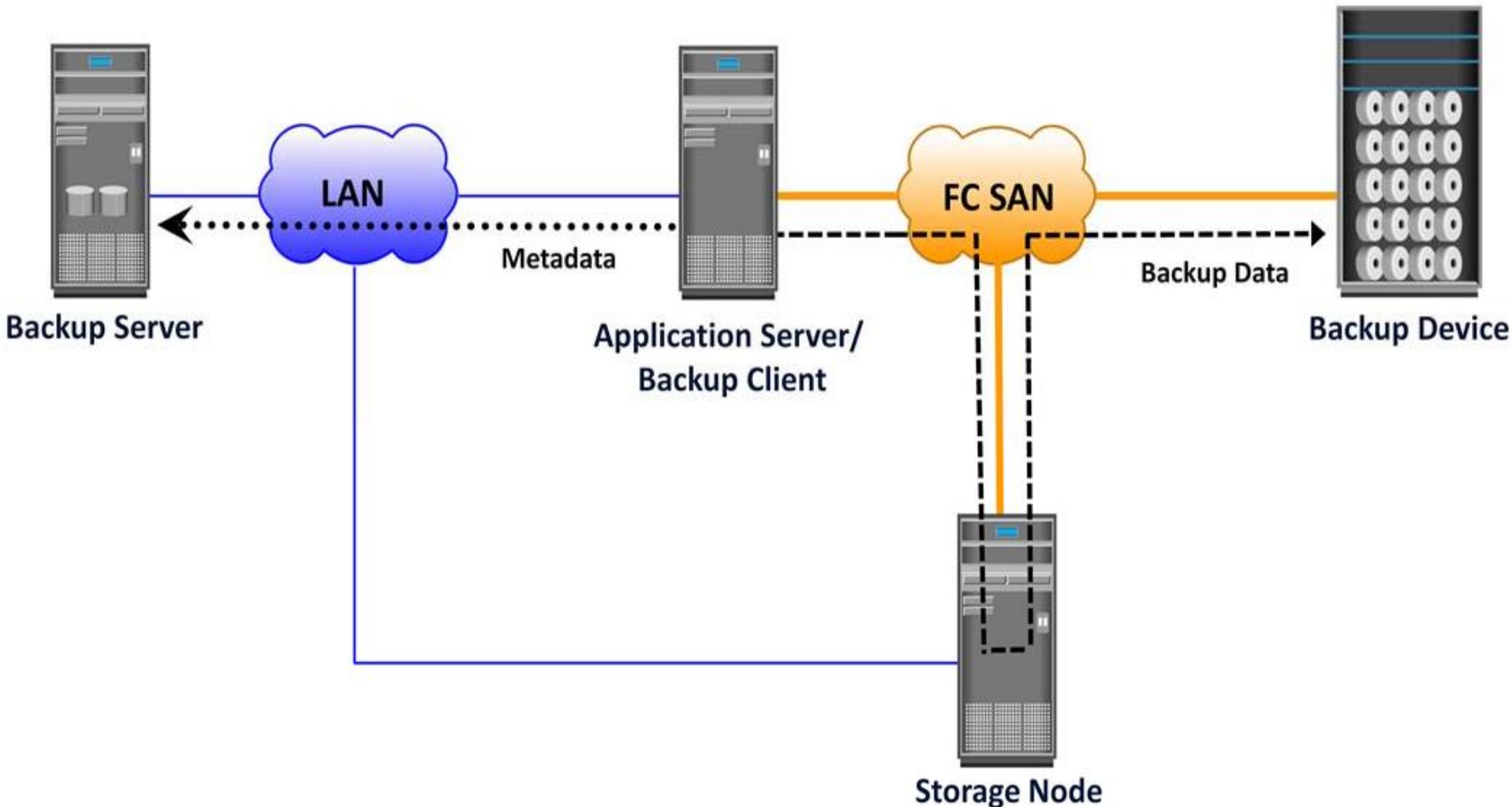
# Топологии - прямое подключение



# Топологии – локальная сеть



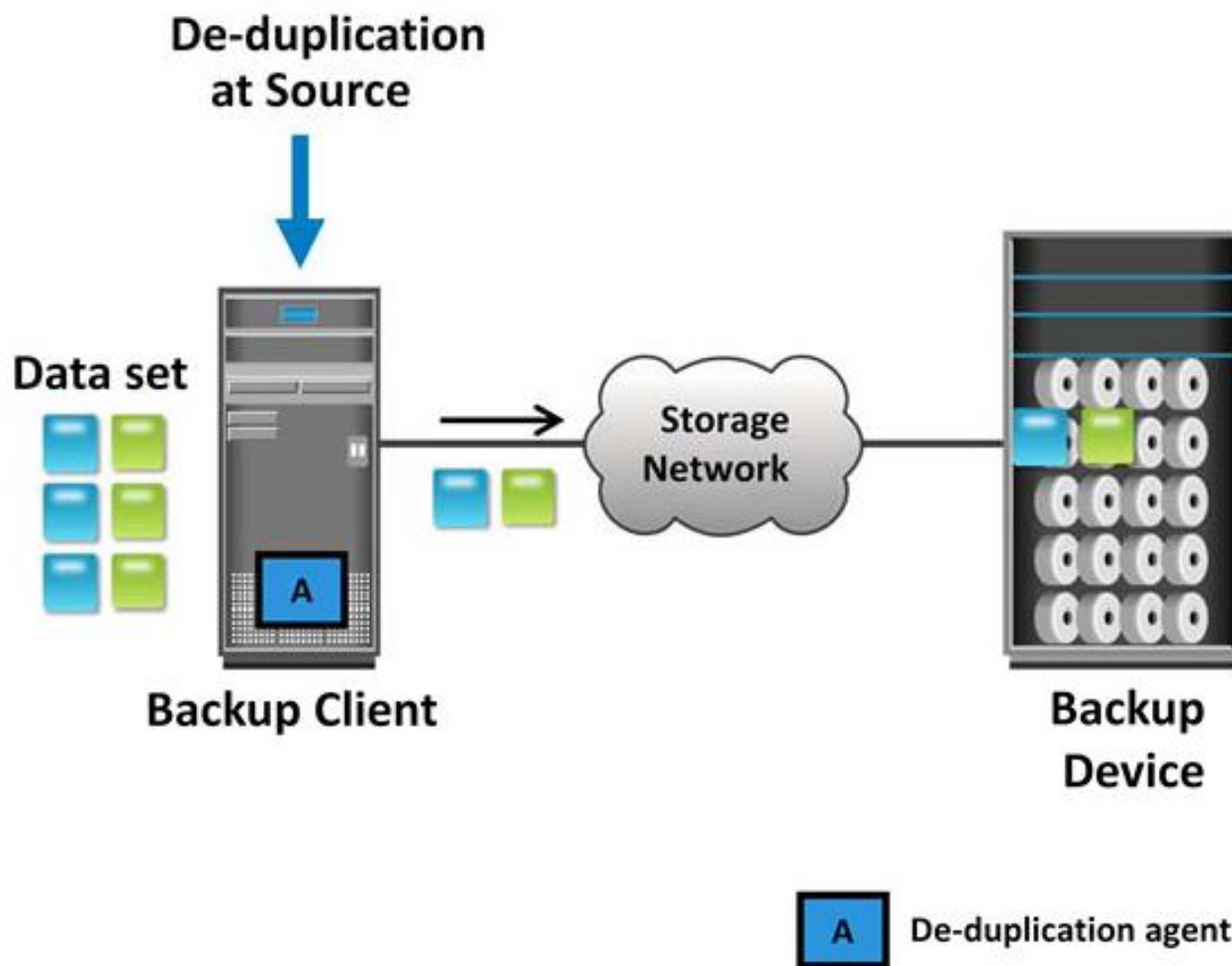
# Топологии – LAN free (SAN based)



# Дедупликация

- Хранение одинаковых объектов в одном экземпляре
  - файловая
  - блоковая
- Файловая дедупликация
  - полные файлы
  - фрагменты (фиксированный размер или плавающий)
- Хэш-функции для обнаружения

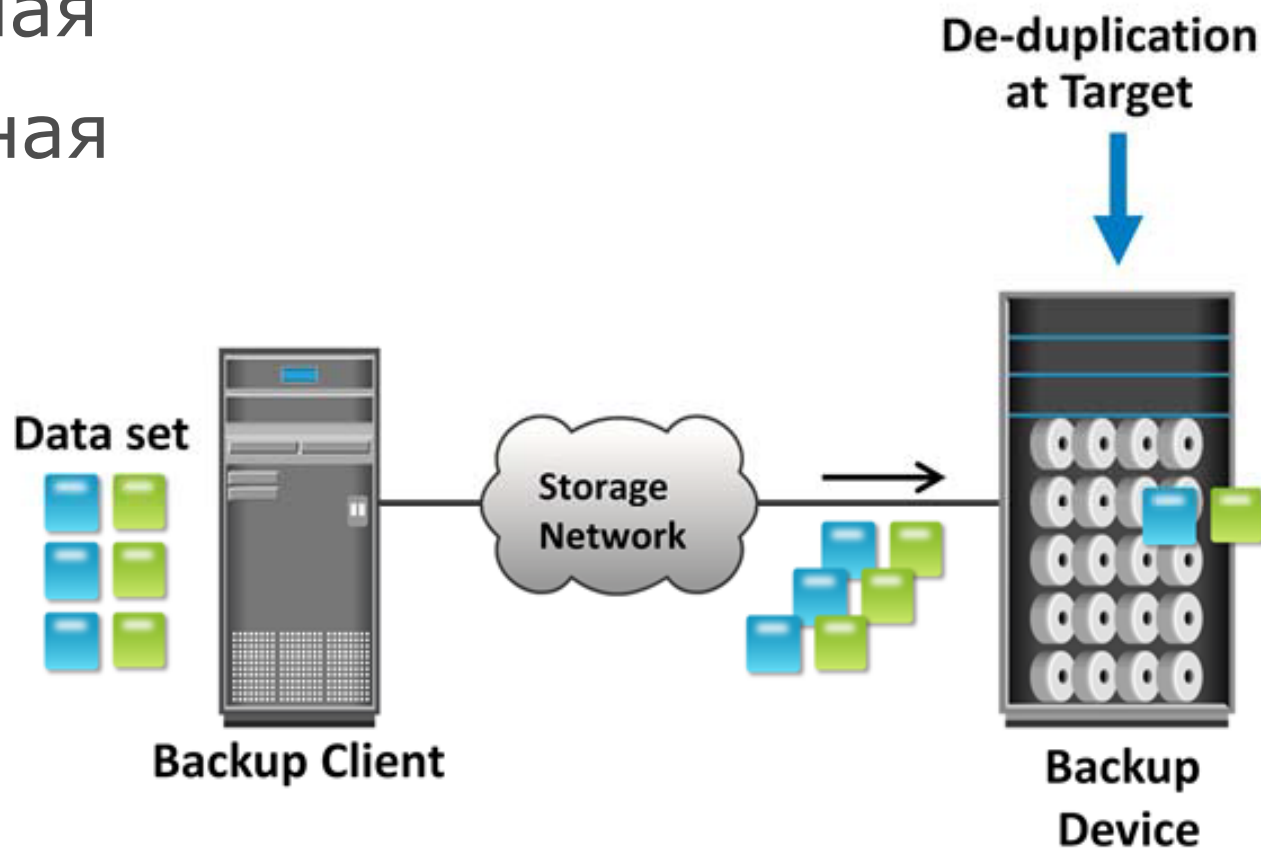
# Дедупликация на клиенте





# Дедупликация в СХД

- Мгновенная
- Отложенная

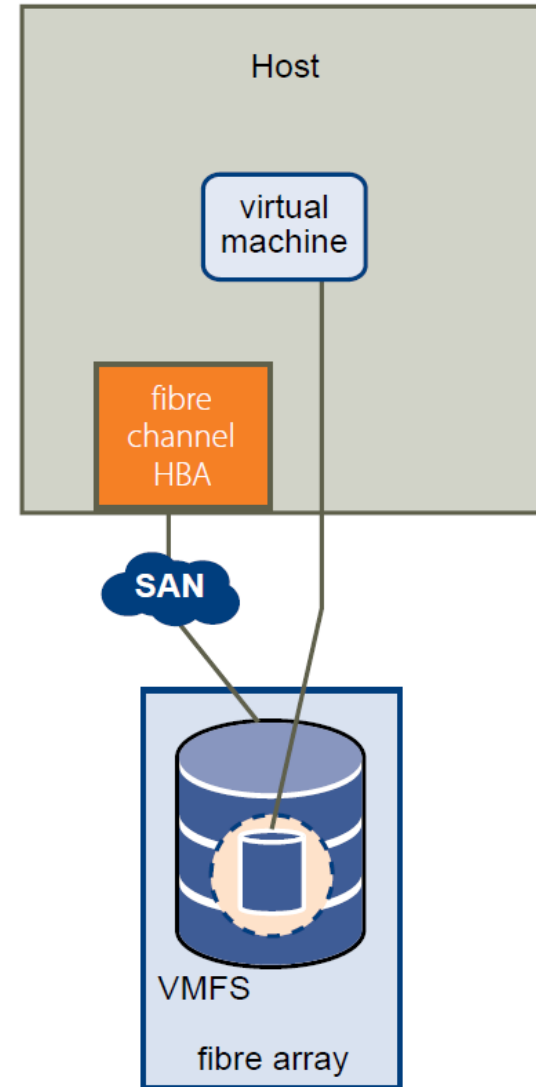


# Преимущества дедупликации в резервном копировании

- Уменьшение стоимости хранения
- Увеличение периода хранения
- Быстрее
- Меньше нагрузка на сеть (при дедупликации на клиенте)

# Резервное копирование виртуальных машин

- Агент в VM
- Агент в гипервизоре
- NAS based
- SAN based



# Спасибо!

EMC<sup>2</sup>®