

Распределенные системы хранения и обработки данных

Владислав Белогрудов, EMC

vlad.belogrudov@gmail.com

Лекция 9

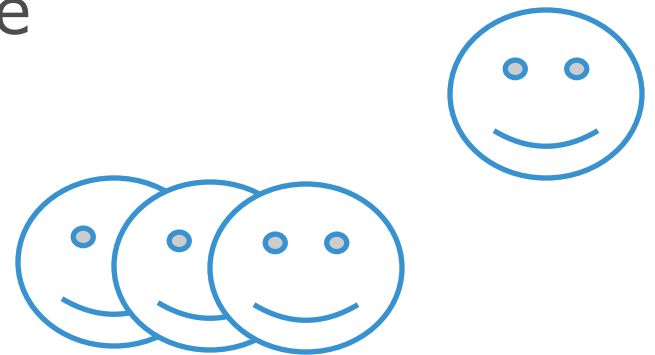
Локальная и удаленная репликация

Содержание лекции

- принципы
- методы
- топологии

Репликация

- Процесс создания точной копии данных
- Обеспечивает непрерывность бизнеса
- Позволяет минимизировать RPO и RTO
- Две категории
 - Локальная, в пределах СХД или ЦОД
 - Удаленная, в другом регионе



Источники и приемники

- Реплика – точная копия
- Копия может быть синхронной или асинхронной (нужна ресинхронизация)

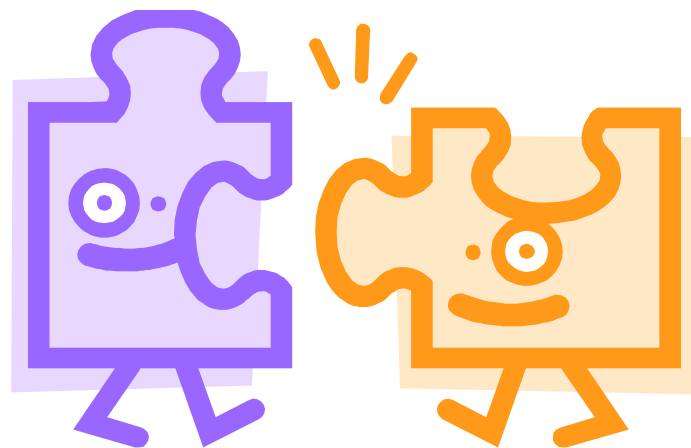


Использование локальных реплик

- Резервное копирование
 - Снижение нагрузки на рабочий диск/LUN
 - Сведение к нулю окна резервного копирования
- Быстрое восстановление (RTO -> min)
- Принятие решений, анализ данных
 - Снижение нагрузки на рабочий диск
- Платформа для тестирования
- Миграция данных (LUN 1 -> LUN 2)

Согласованность данных

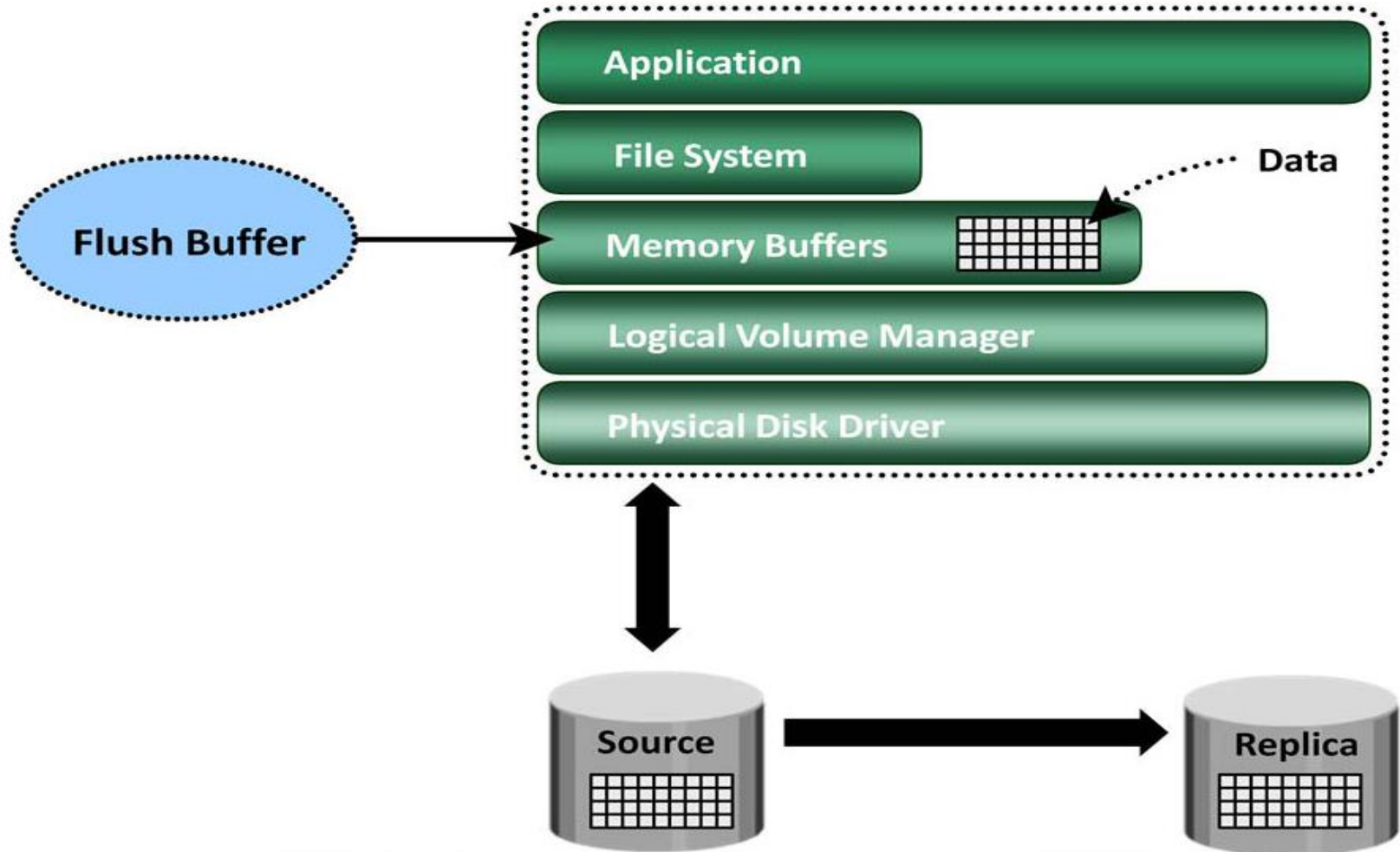
- Реплики должны быть согласованными (полными)
- Согласованность на уровне
 - Файловых систем
 - Баз данных



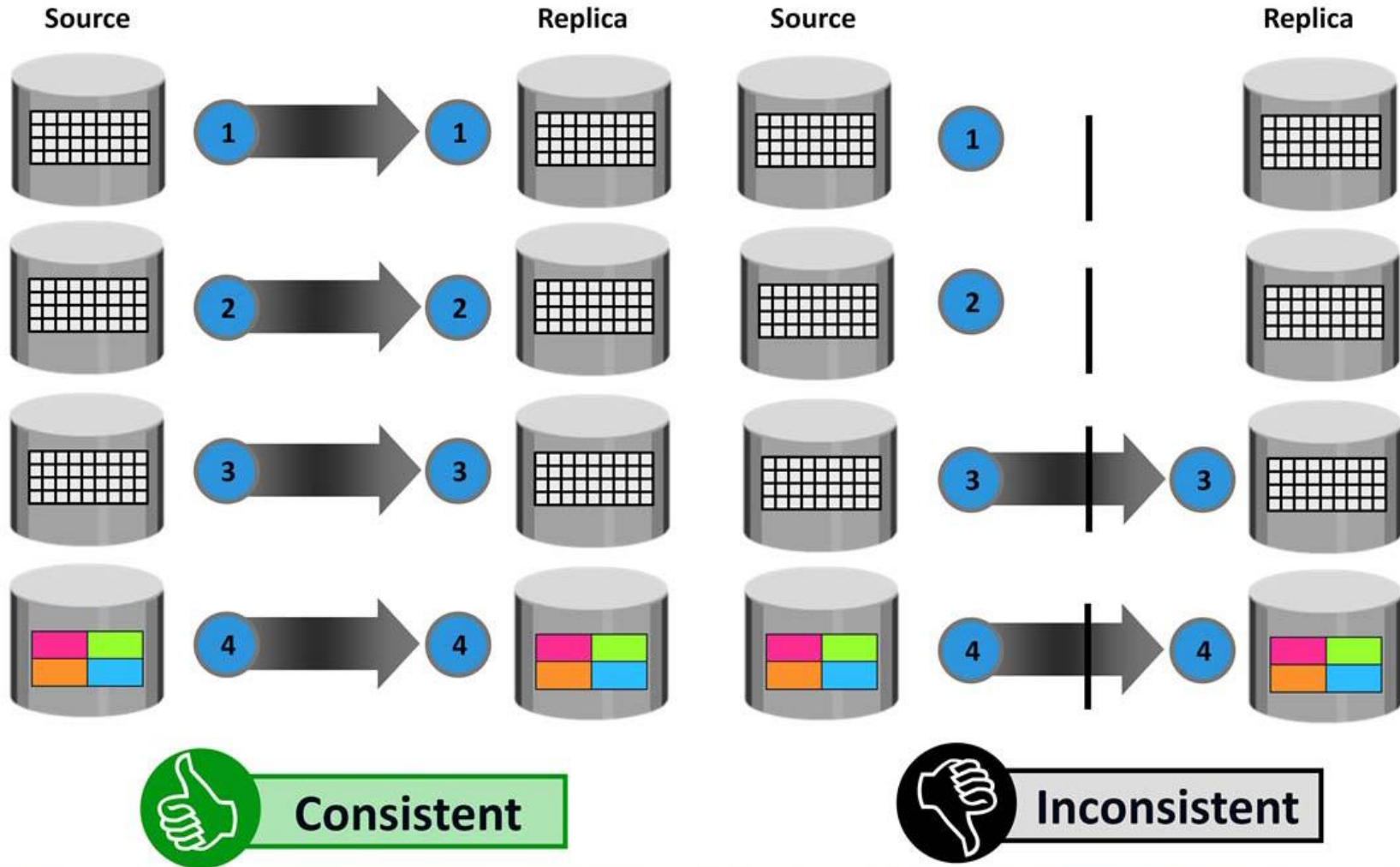
Обеспечение согласованности

	offline	online
Файлы	Отсоединение файловой системы от хоста	Flushing – сброс буферов ОС на диск
Базы данных	Полный останов	<ul style="list-style-type: none">• Зависимая запись• Задержка IO

File System Consistency



Database Consistency

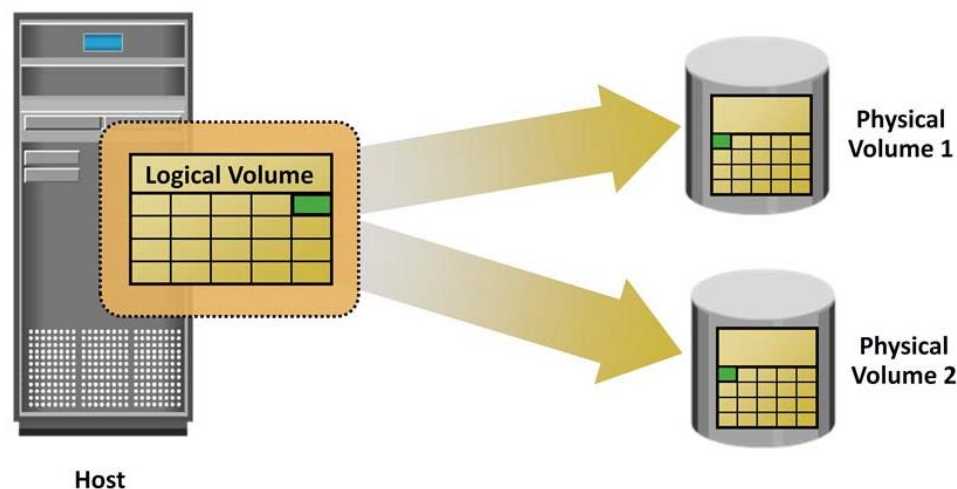


Технологии локальной репликации

- Серверная (host-based)
 - С помощью LVM
 - Снимки файловой системы
- С помощью СХД
 - Зеркальное полное копирование
 - PIT (указатель во времени)
 - Полная репликация с указателем
 - Виртуальная репликация с указателем

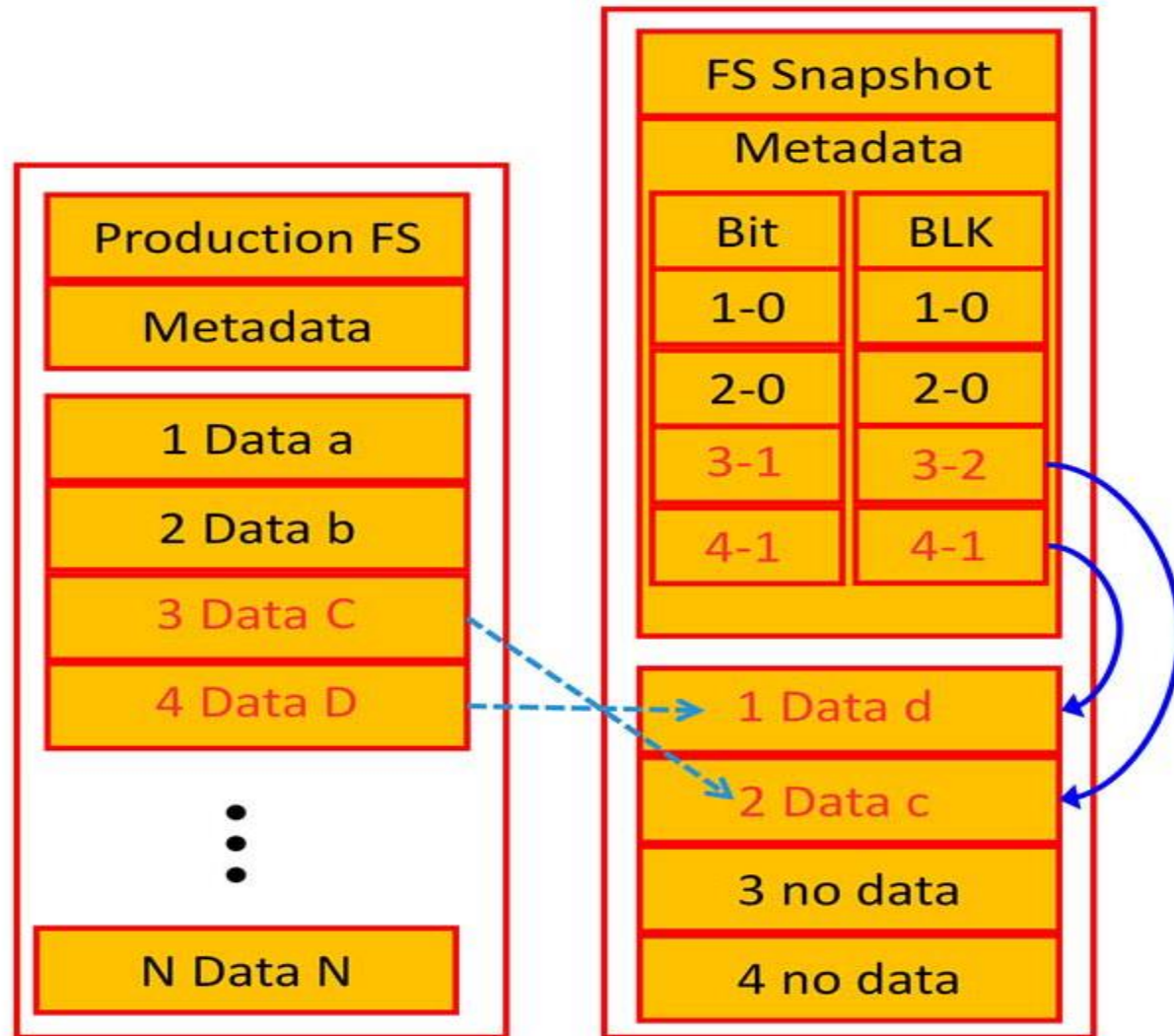
Репликация с помощью LVM

- Запись в два и более тома синхронно
 - LVM зеркалирование
- Не зависит от СХД
- При наличии ошибки копия не приговждается 😊



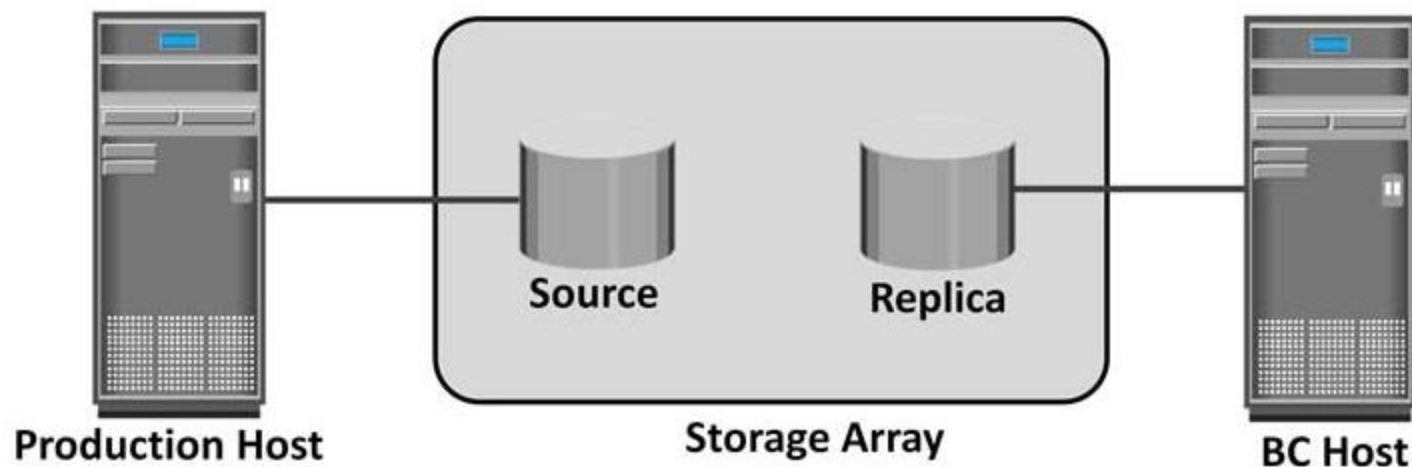
Снимок файловой системы

- PIT
- CoFW
 - copy on first write

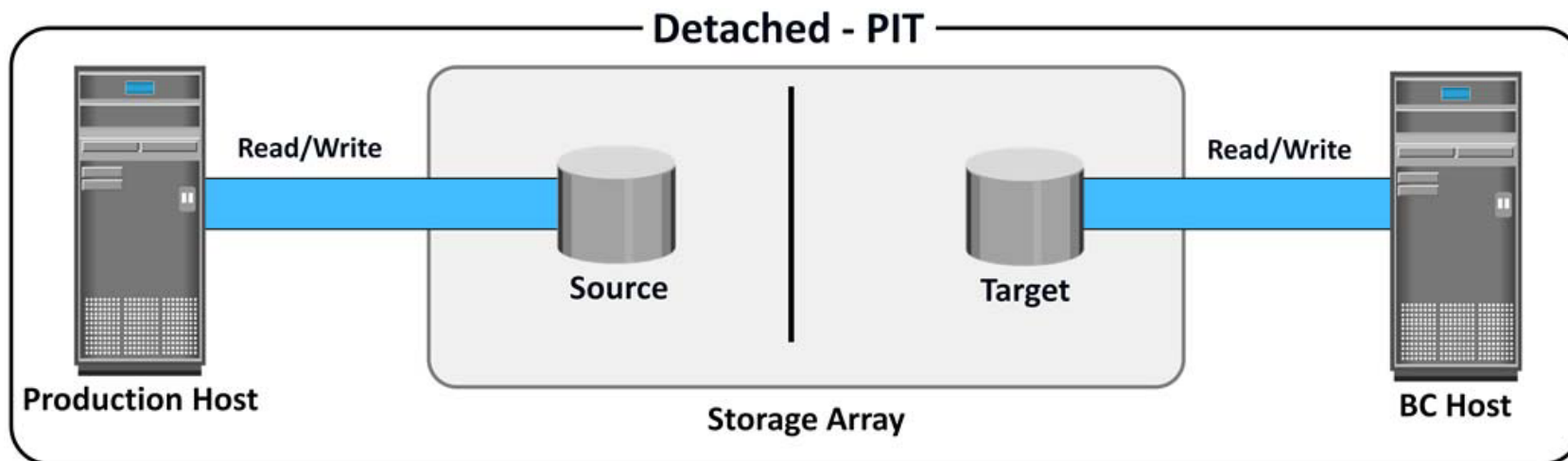
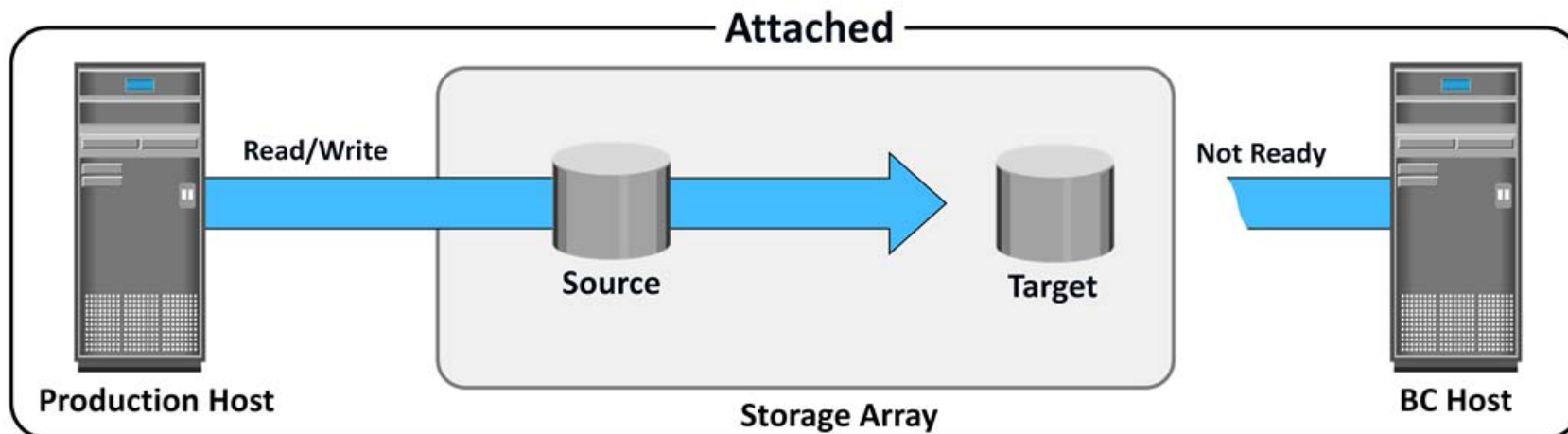


Репликация на стороне СХД

- Ресурсы сервера не используются
- Возможен доступ с другого хоста



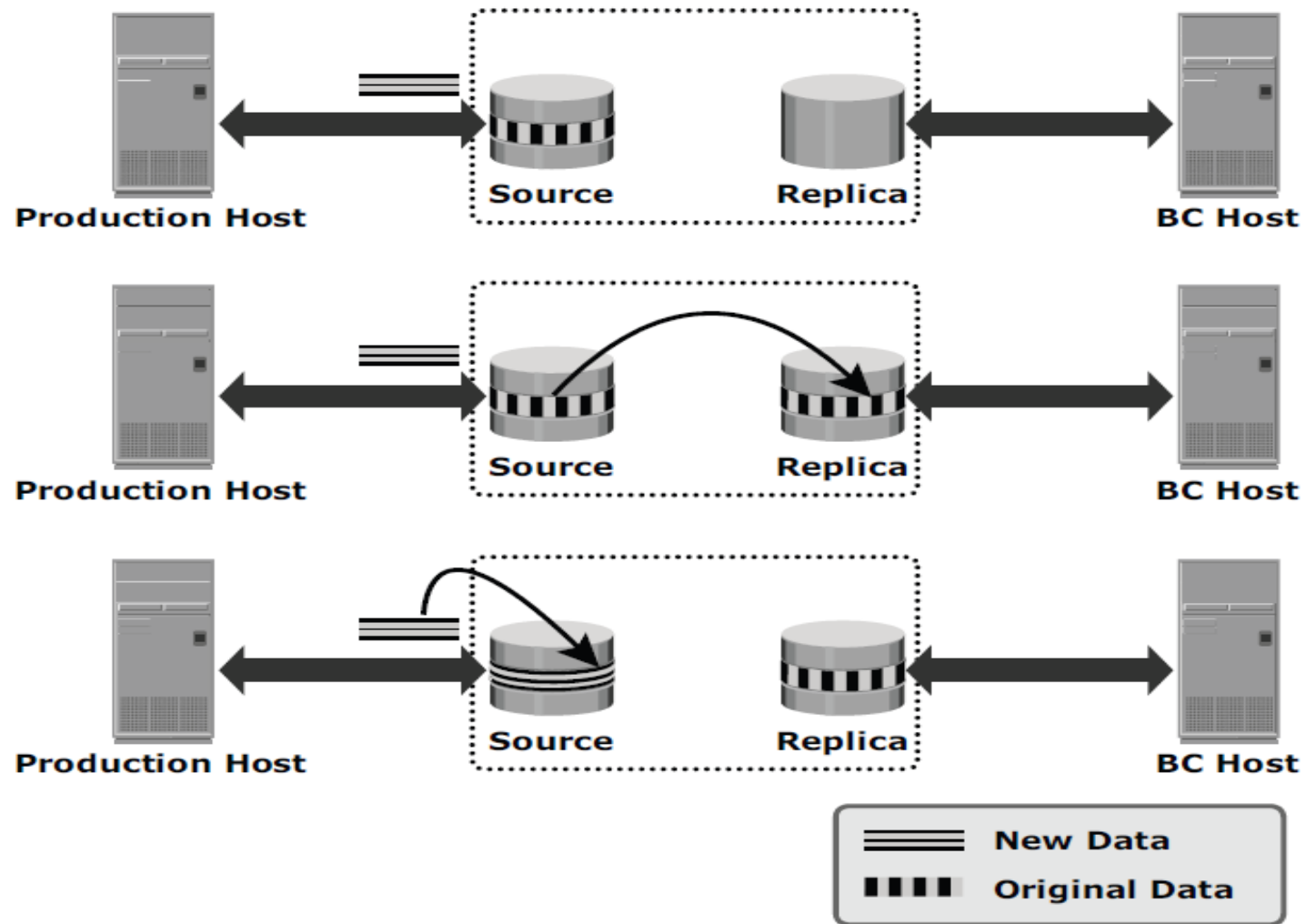
Зеркальное копирование целого тома



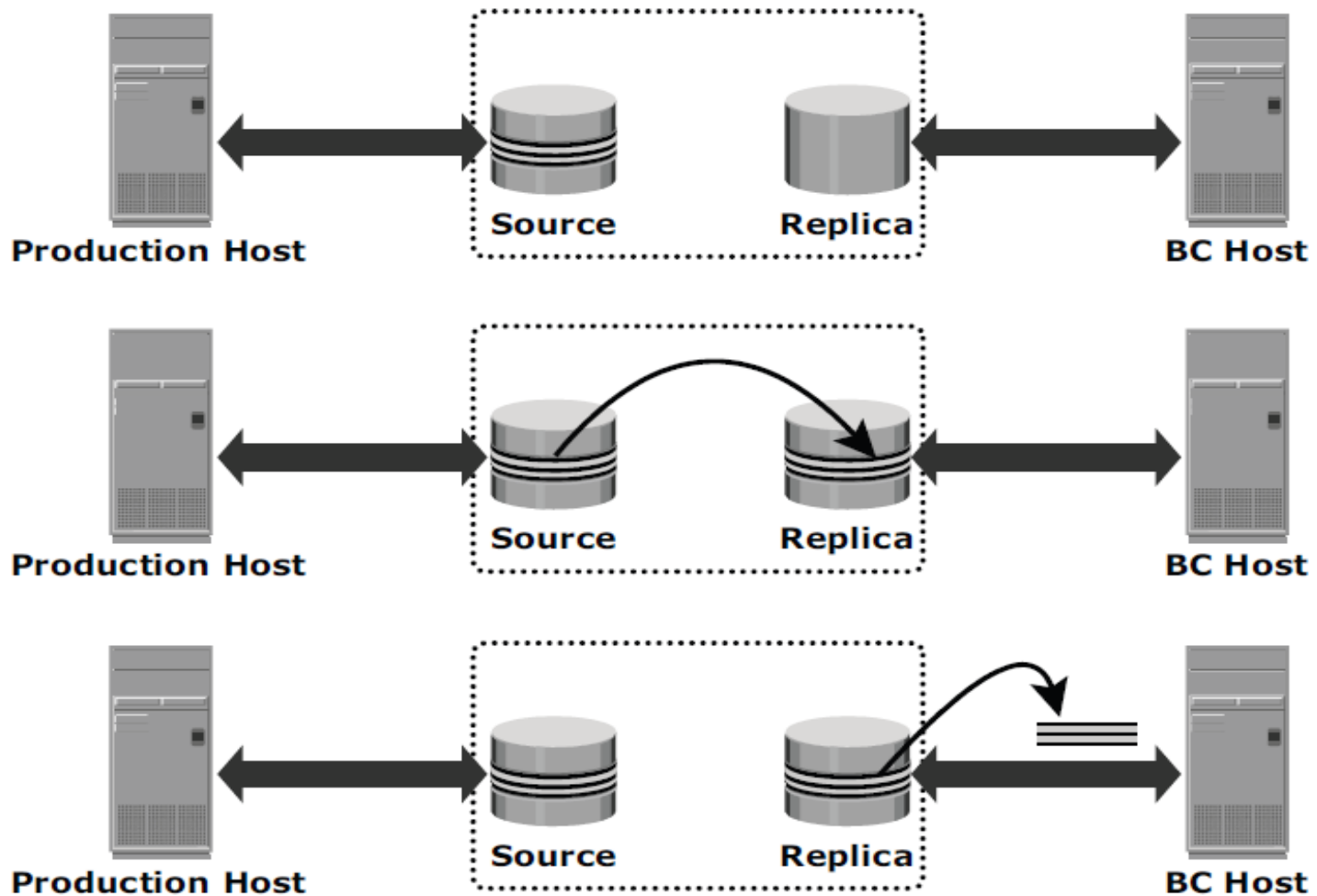
RIT репликация целого тома

- Полная копия источника
- Приемник доступен сразу после активации сессии
- RIT определяется временем активации
- Приемник должен быть не меньше источника
- Два режима
 - Полное копирование в фоновом режиме
 - CoFA (copy on first access)

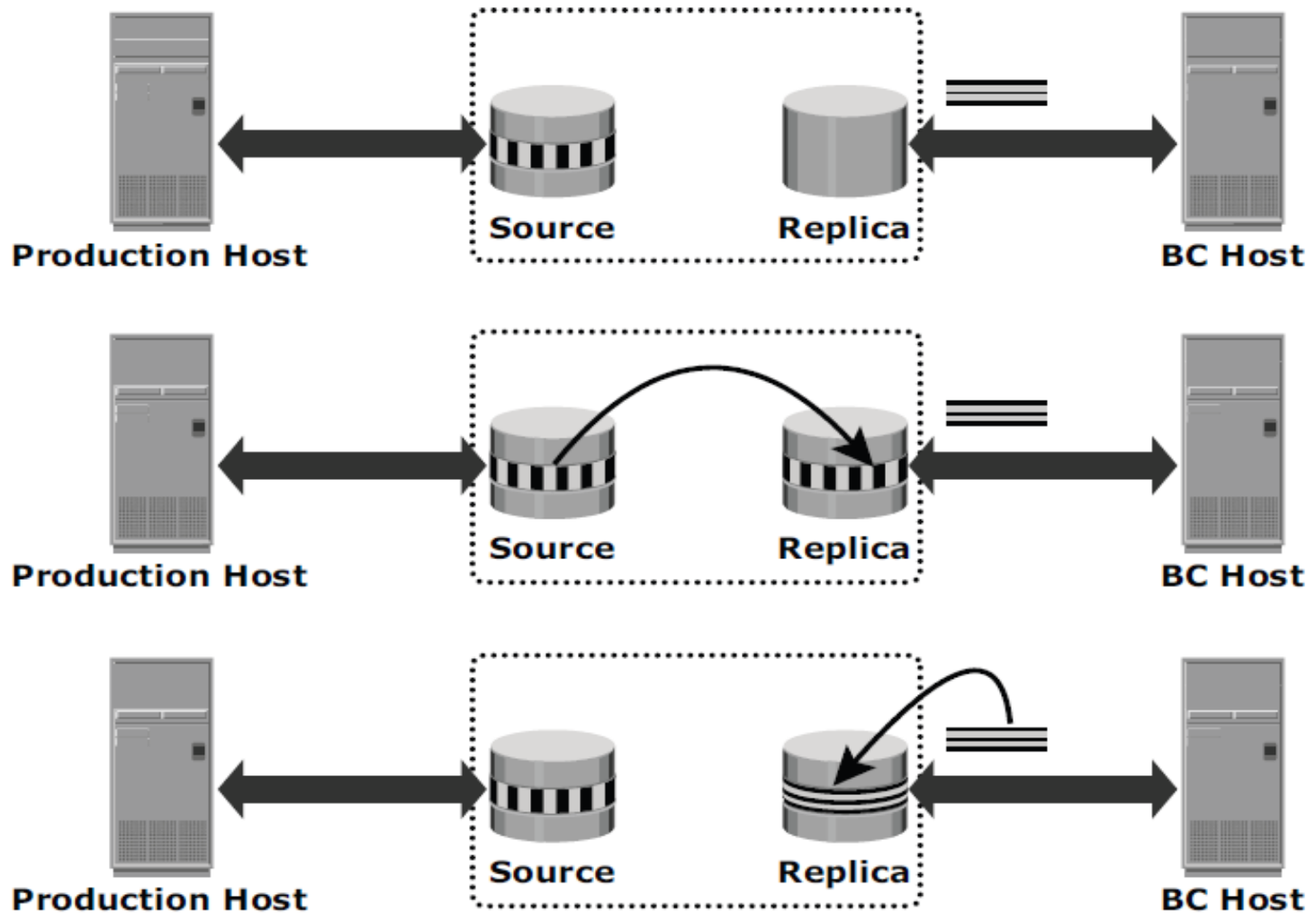
Copy on First Access – Source Write



Copy on First Access – Target Read



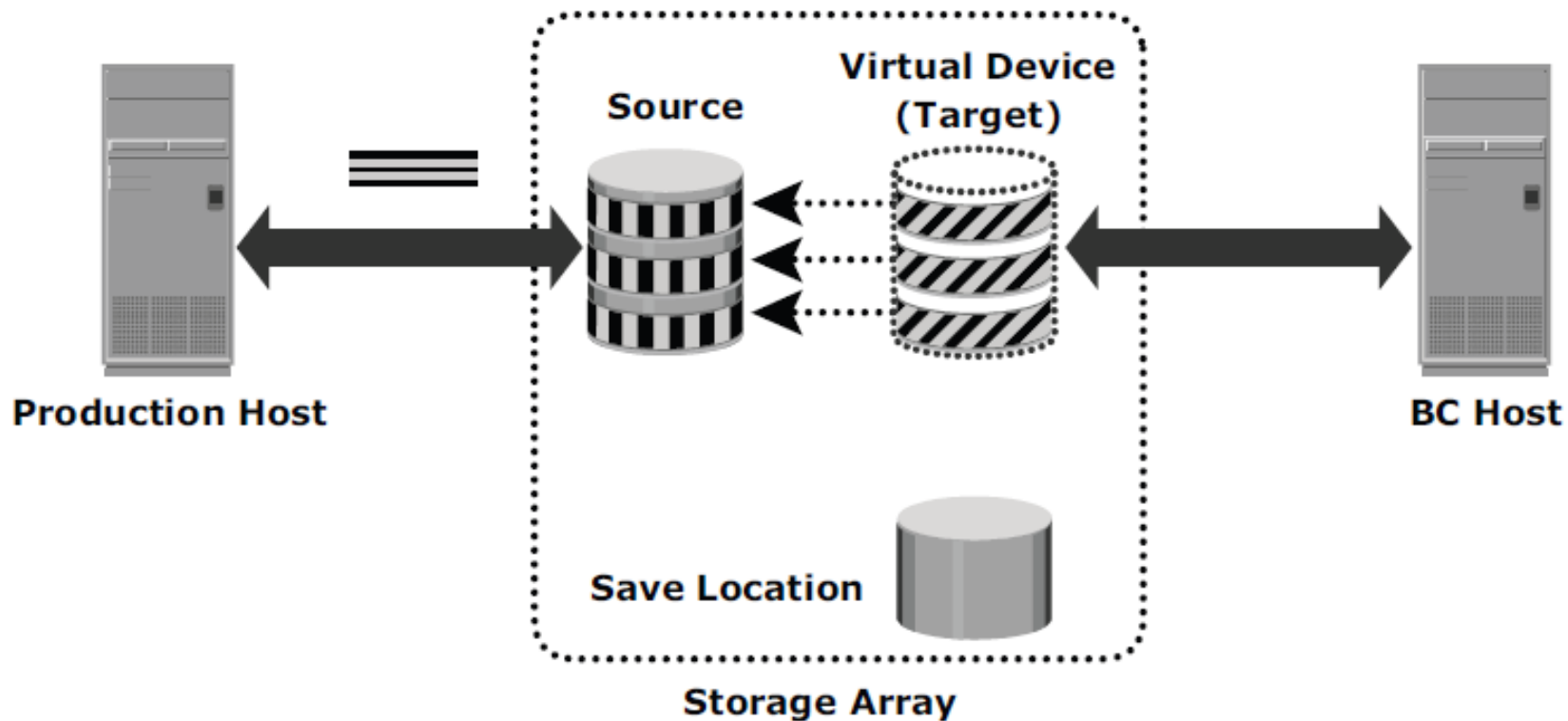
Copy on First Access – Target Write



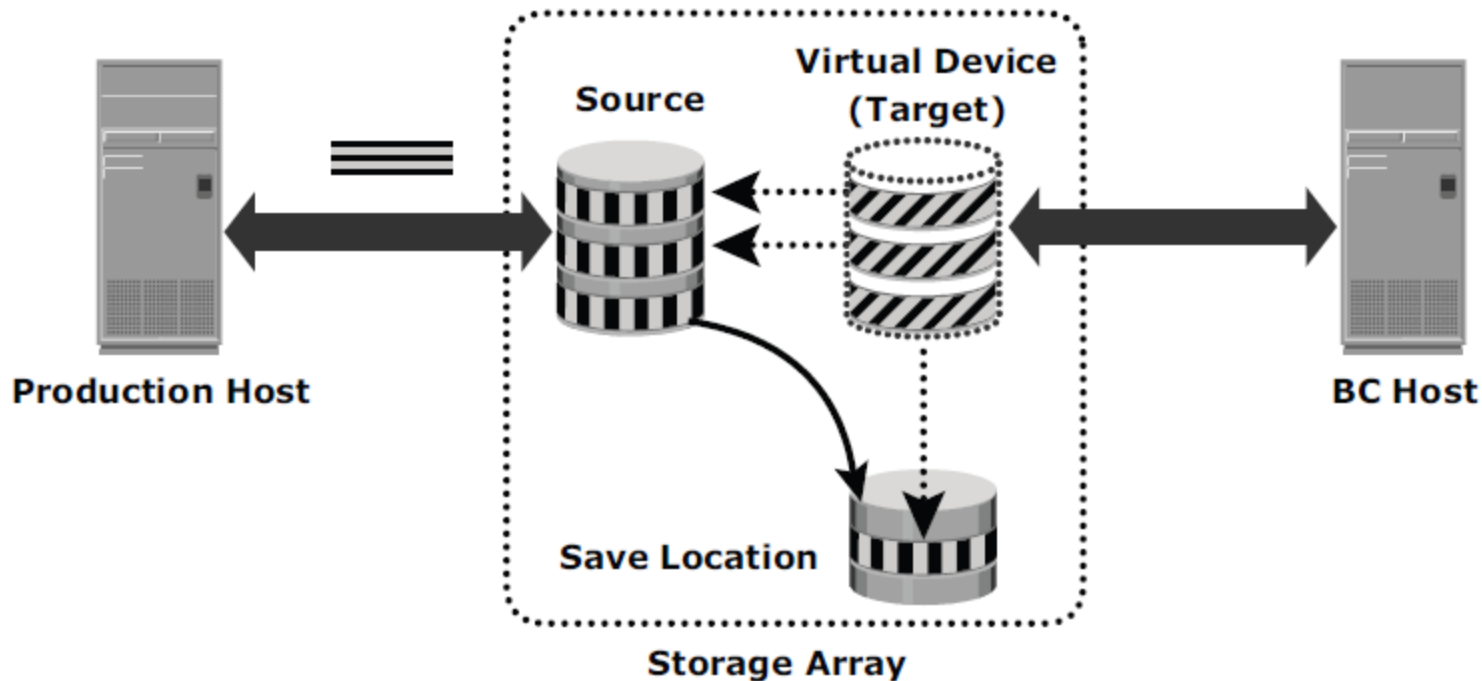
Виртуальная репликация РІТ

- Приемник содержит указатели на изменившиеся данные
- Если нет изменений – указатели на приемник, если есть – на место хранения снимка
- Приемник и место для снимка занимают часть полного объема
- Для изменений данных до 30%
- CoFW – copy on first write

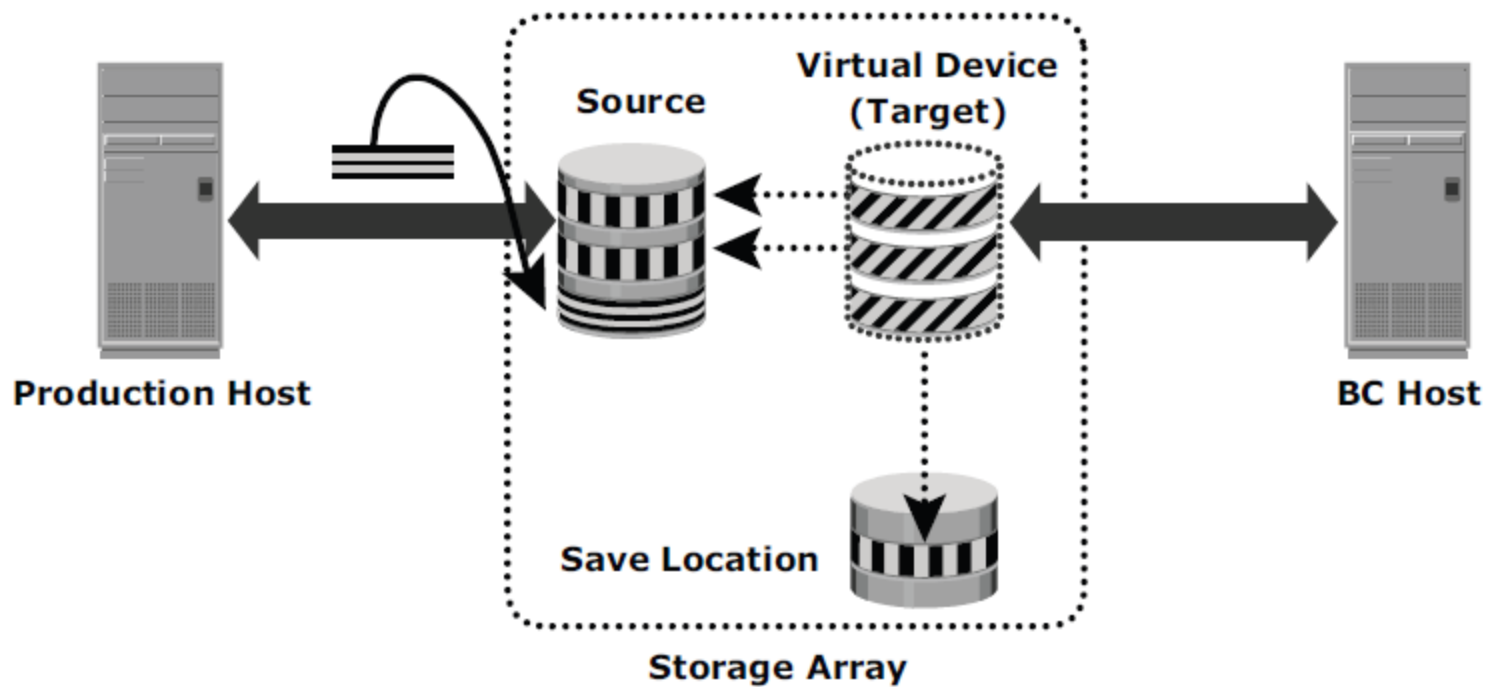
Виртуальная реплика – запись (1)



Виртуальная реплика – запись (2)



Виртуальная реплика – запись (3)



Сравнение технологий локальной репликации

	Зеркало	PIT копия целого тома	Виртуальная копия PIT
Влияние на источник	Не требуется	CoFA - небольшое, полная копия - нет	высокое
Размер приемника	Такой же	Такой же	Малая часть
Доступность источника для восстановления	Не требуется	CoFA – требуется, полная копия - нет	да
Доступность приемника	Только после синхронизации	сразу	сразу

Отслеживание изменений

PIT

	0	1	2	3	4	5	6	7	...
Source	0	0	0	0	0	0	0	0	...

	0	1	2	3	4	5	6	7	...
Target	0	0	0	0	0	0	0	0	...

После PIT

	0	1	2	3	4	5	6	7	...
Source	1	0	0	1	0	1	0	0	...

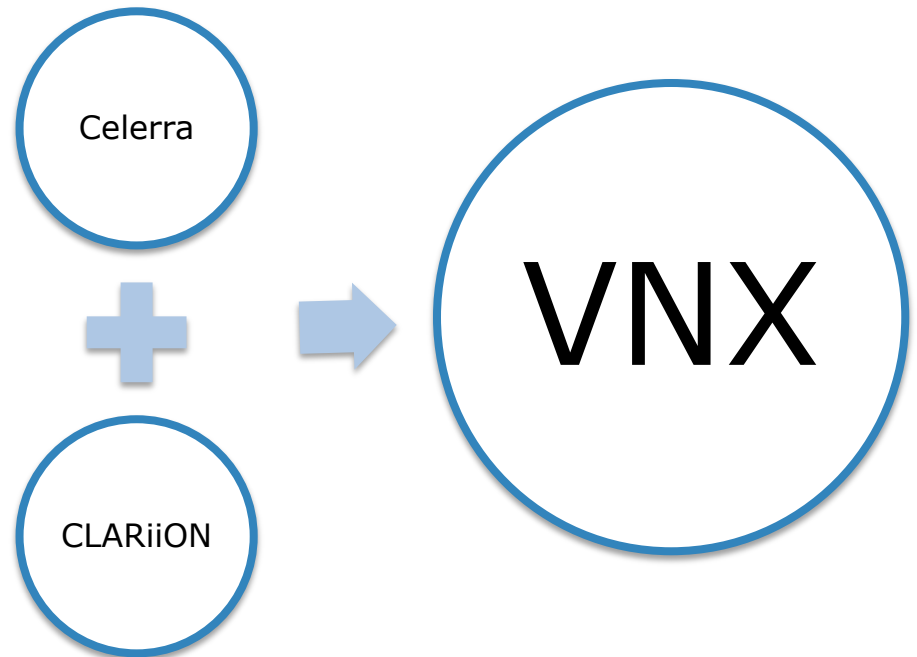
	0	1	2	3	4	5	6	7	...
Target	0	0	1	1	0	0	0	1	...

OR

	0	1	2	3	4	5	6	7	...
Used for Resynchronization/Restore	1	0	1	1	0	1	0	1	...

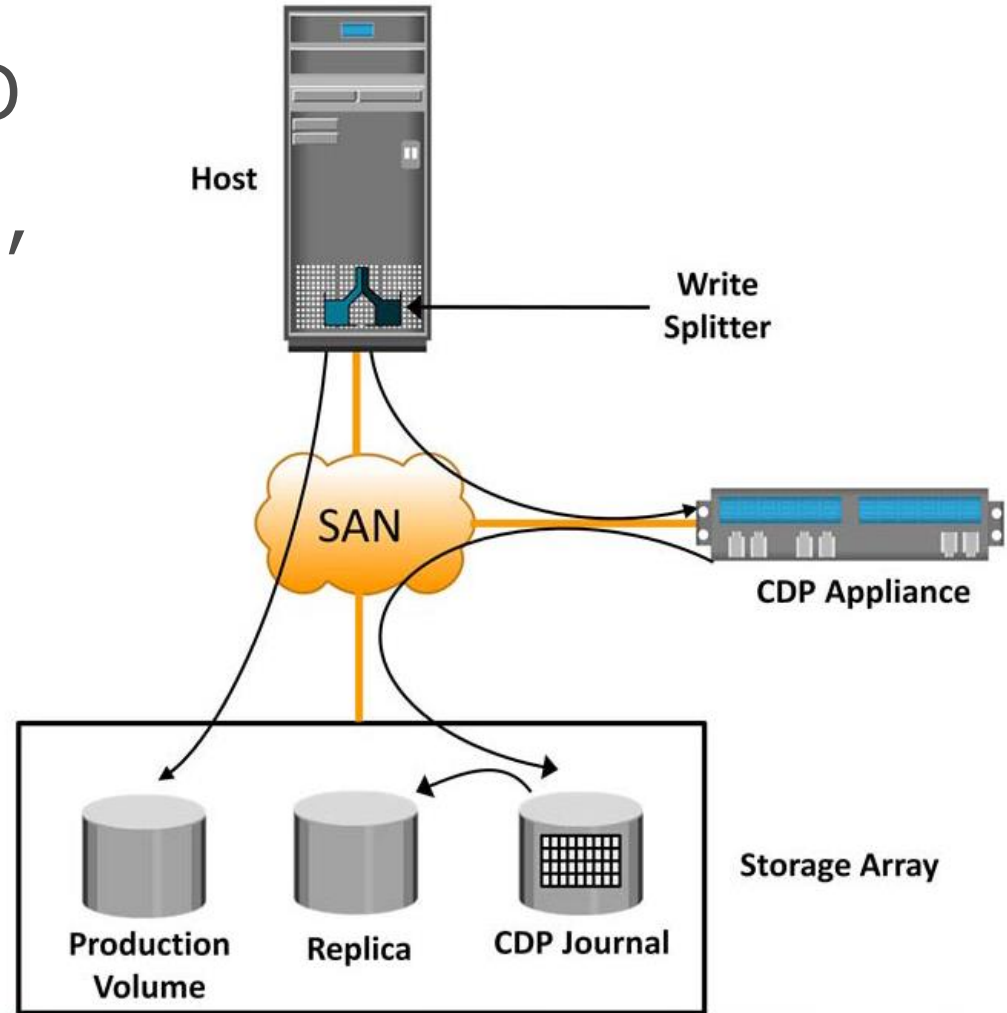
Локальная репликация @ EMC

- Simmetrix – EMC TimeFinder
- CLARiiON – SnapView
- Celerra – SnapSure



Сетевая локальная репликация – continuous data protection (CDP)

- Произвольное RPO
- Сплиттер на хосте, в сети или в СХД
- Сначала – полная синхронизация, затем – журналирование



Удаленная репликация

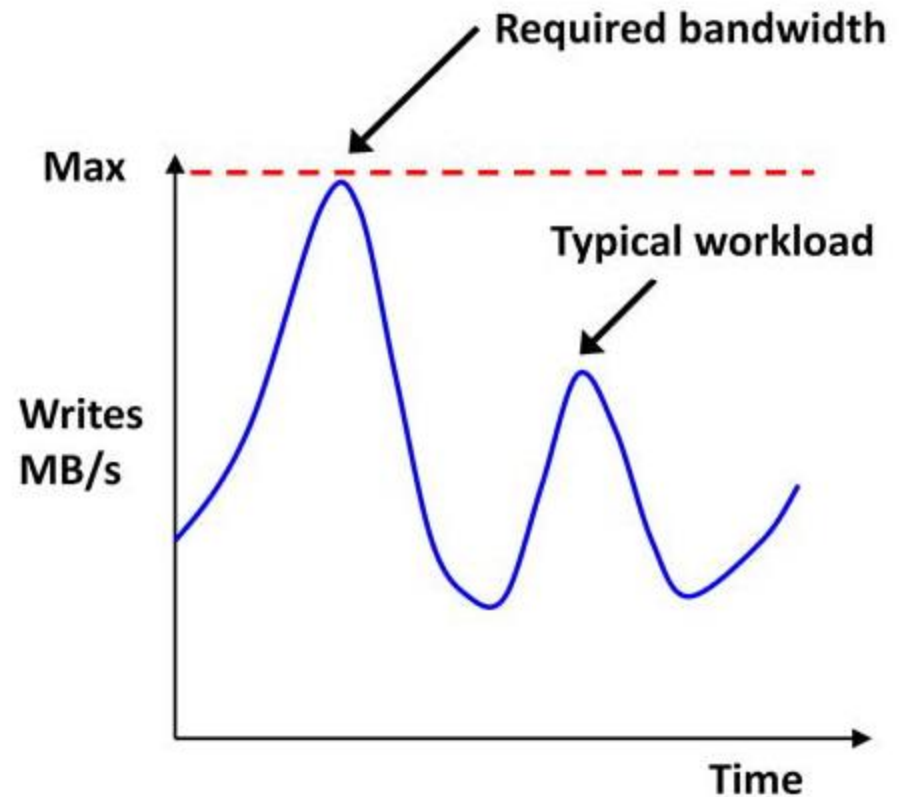
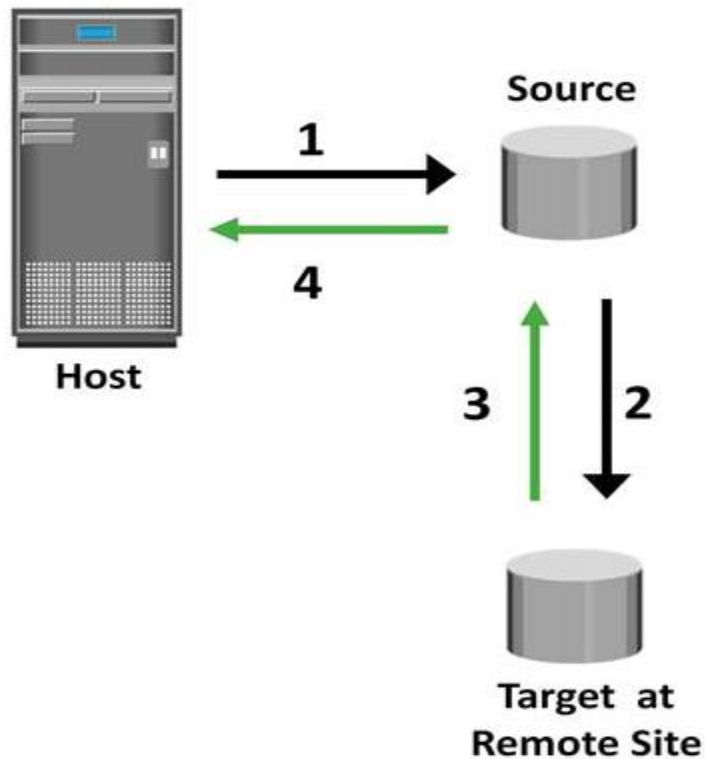
- Процесс создания реплик на удаленных сайтах
- Режимы
 - Синхронный
 - Асинхронный
- Технологии
 - Серверная (LVM, RDBM)
 - СХД



Синхронный режим

- Запись на источник и приемник до подтверждения завершения
- Идентичные копии
- Нулевой RPO, низкий RTO
- Большое время отклика
- До 200 км

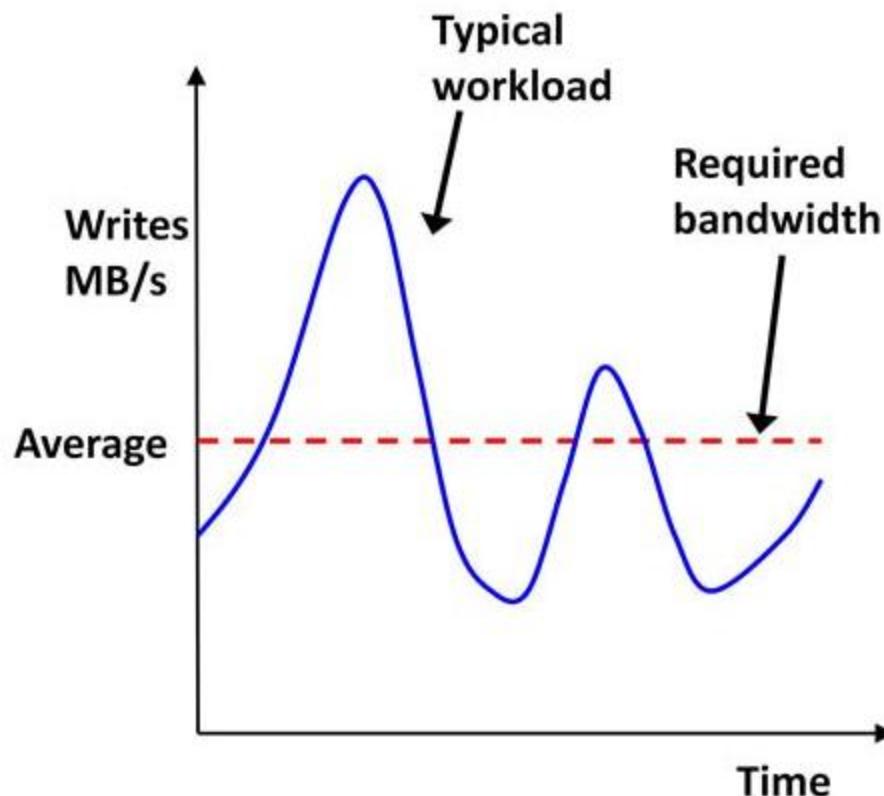
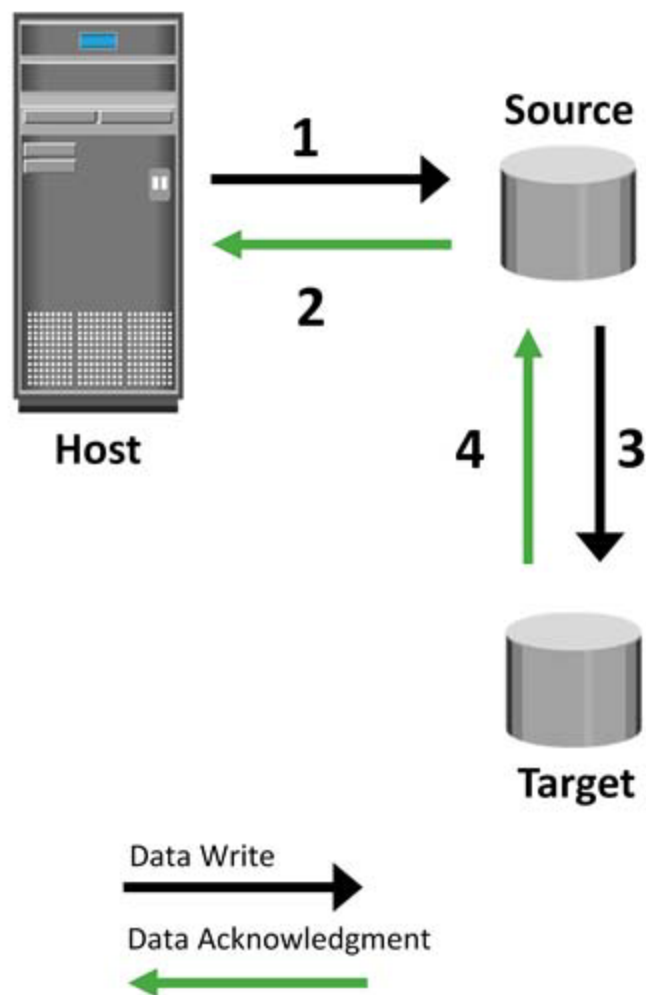
Синхронная репликация в СХД



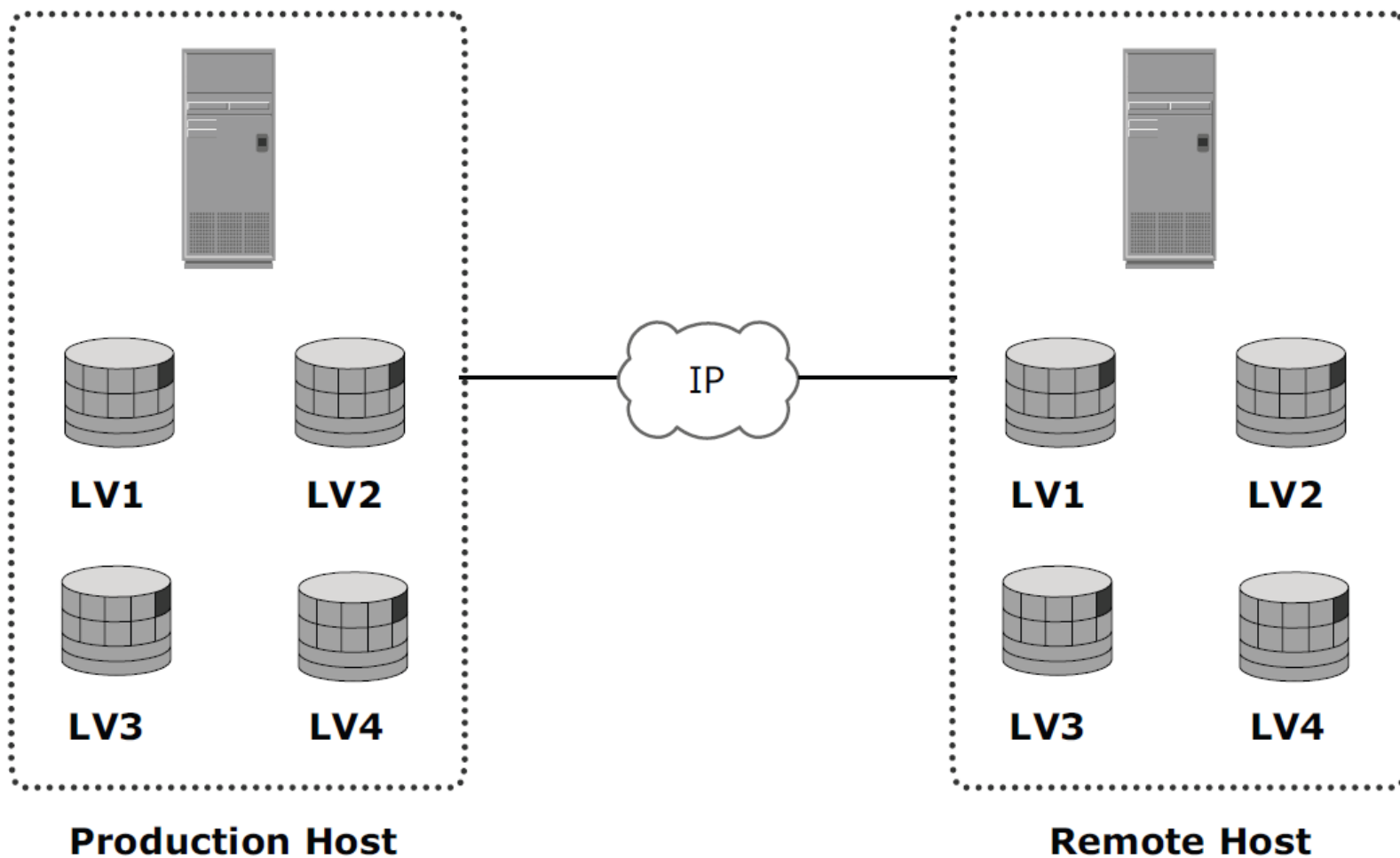
Асинхронный режим

- Немедленный отклик от источника
- Запись на приемник при заполнении буфера или по времени
- Ненулевое RPO – зависит от размера буфера, пропускной способности сети, нагрузки на источник
- Несколько тысяч км

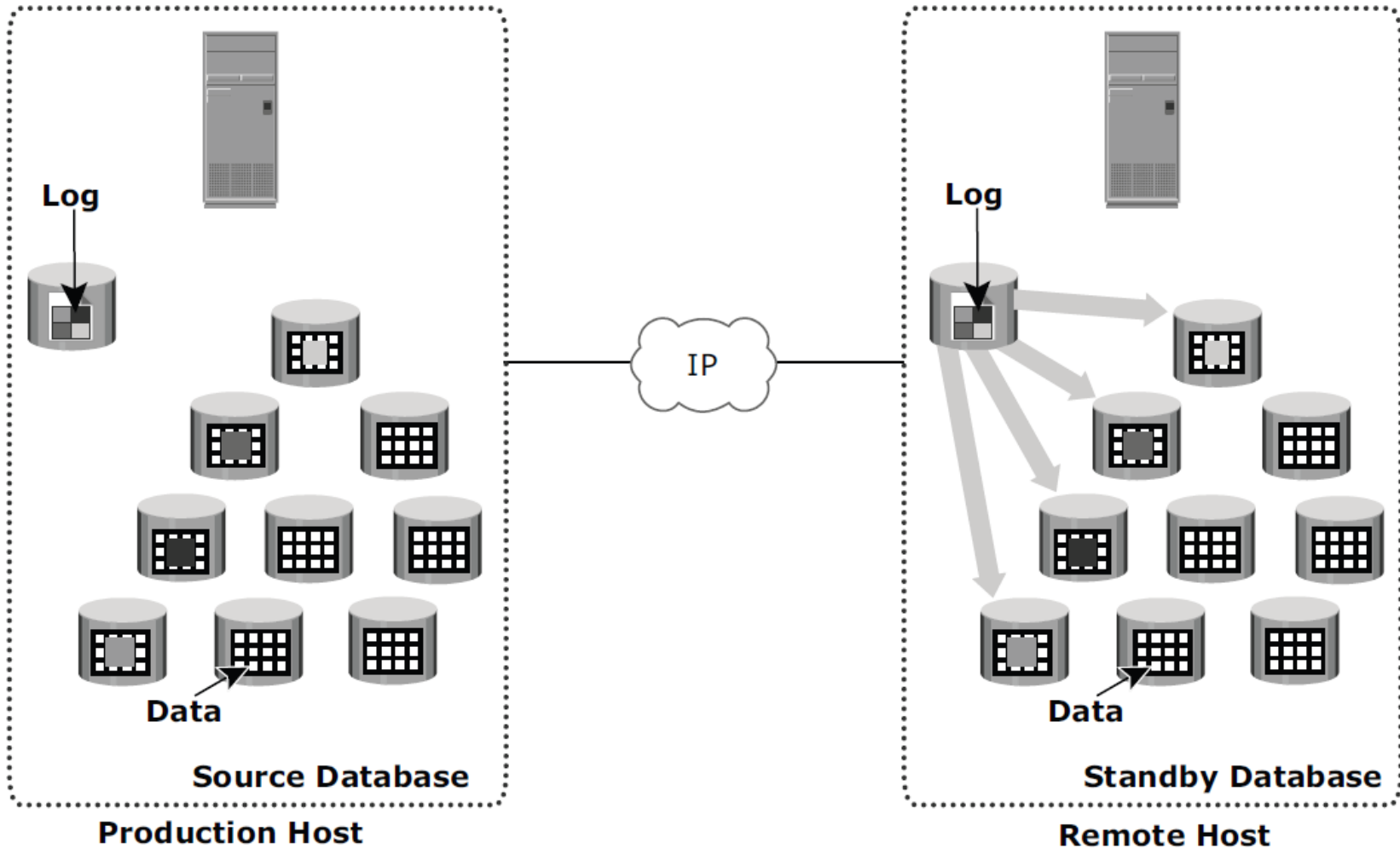
Асинхронная репликация в СХД



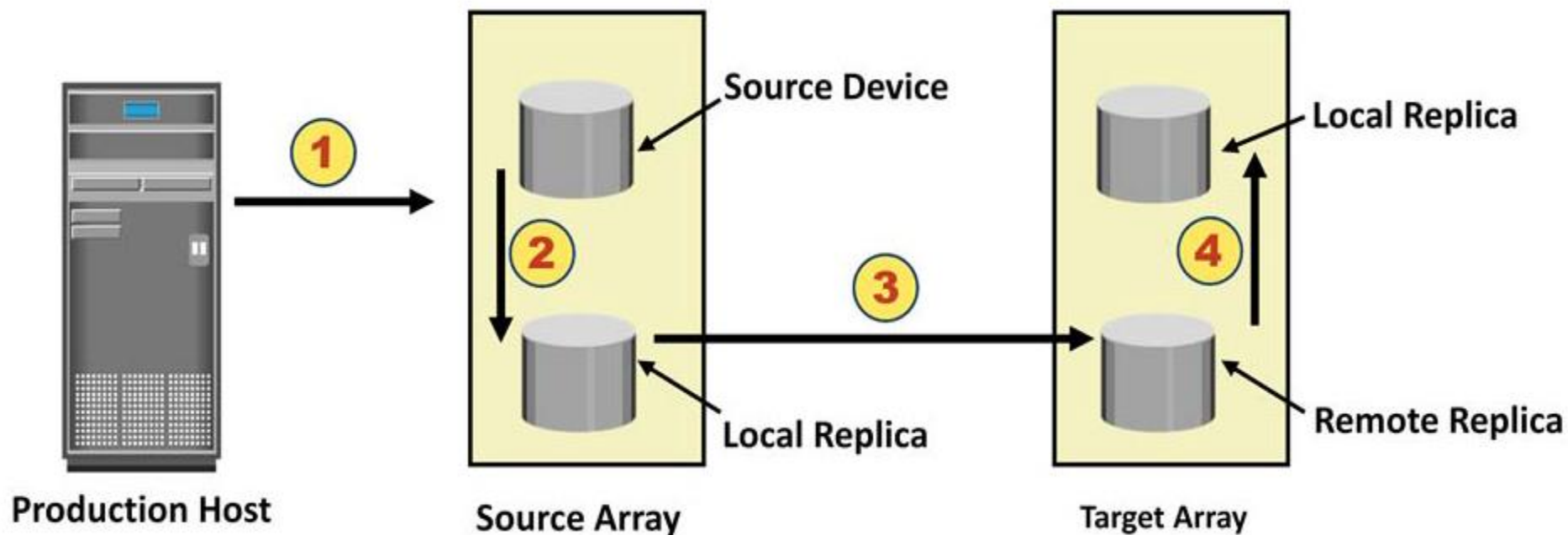
Удаленная репликация с LVM



Журналы транзакций RDBM

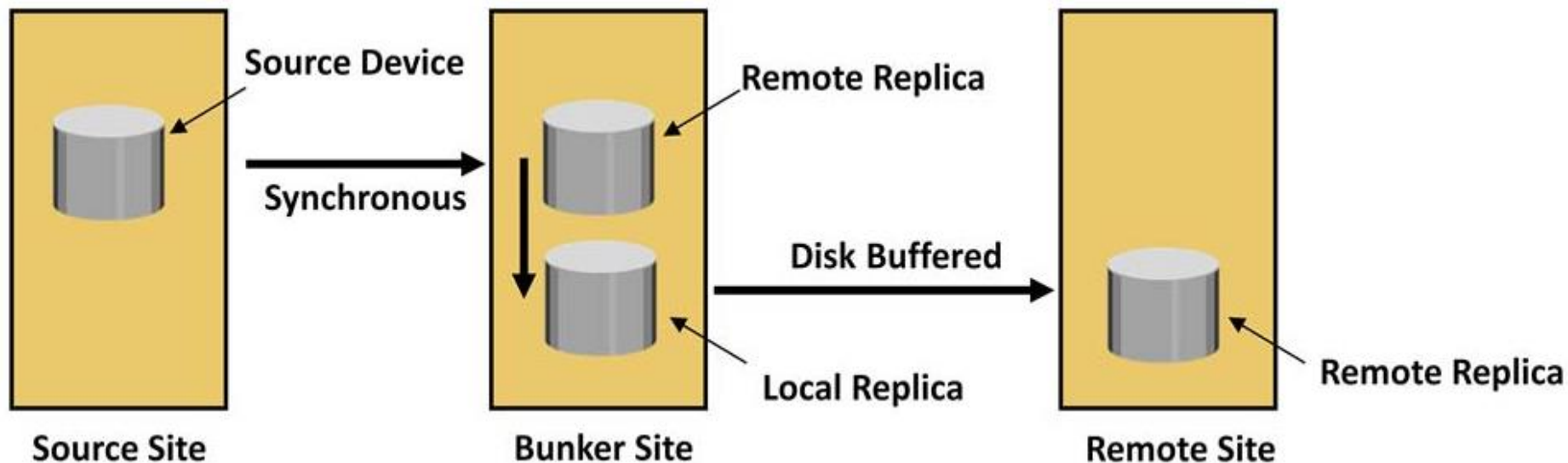


Репликация с дисковой буферизацией



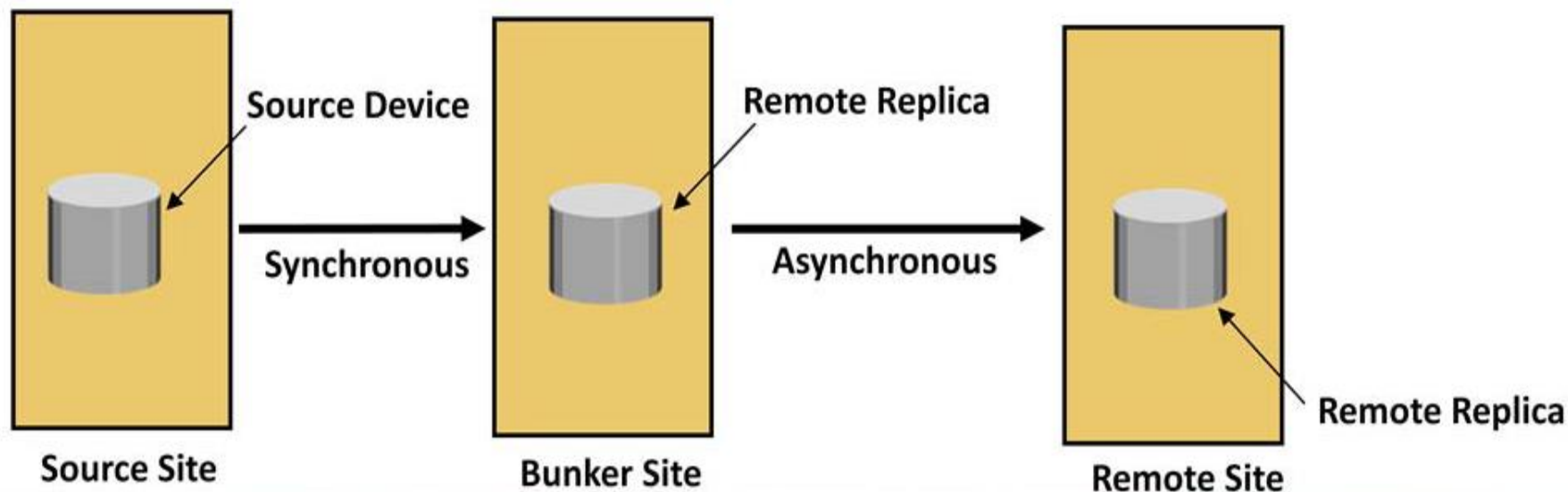
- Network friendly
- Инкрементальная синхронизация

Трехсторонняя каскадная репликация: синхронная + дисковая буферизация



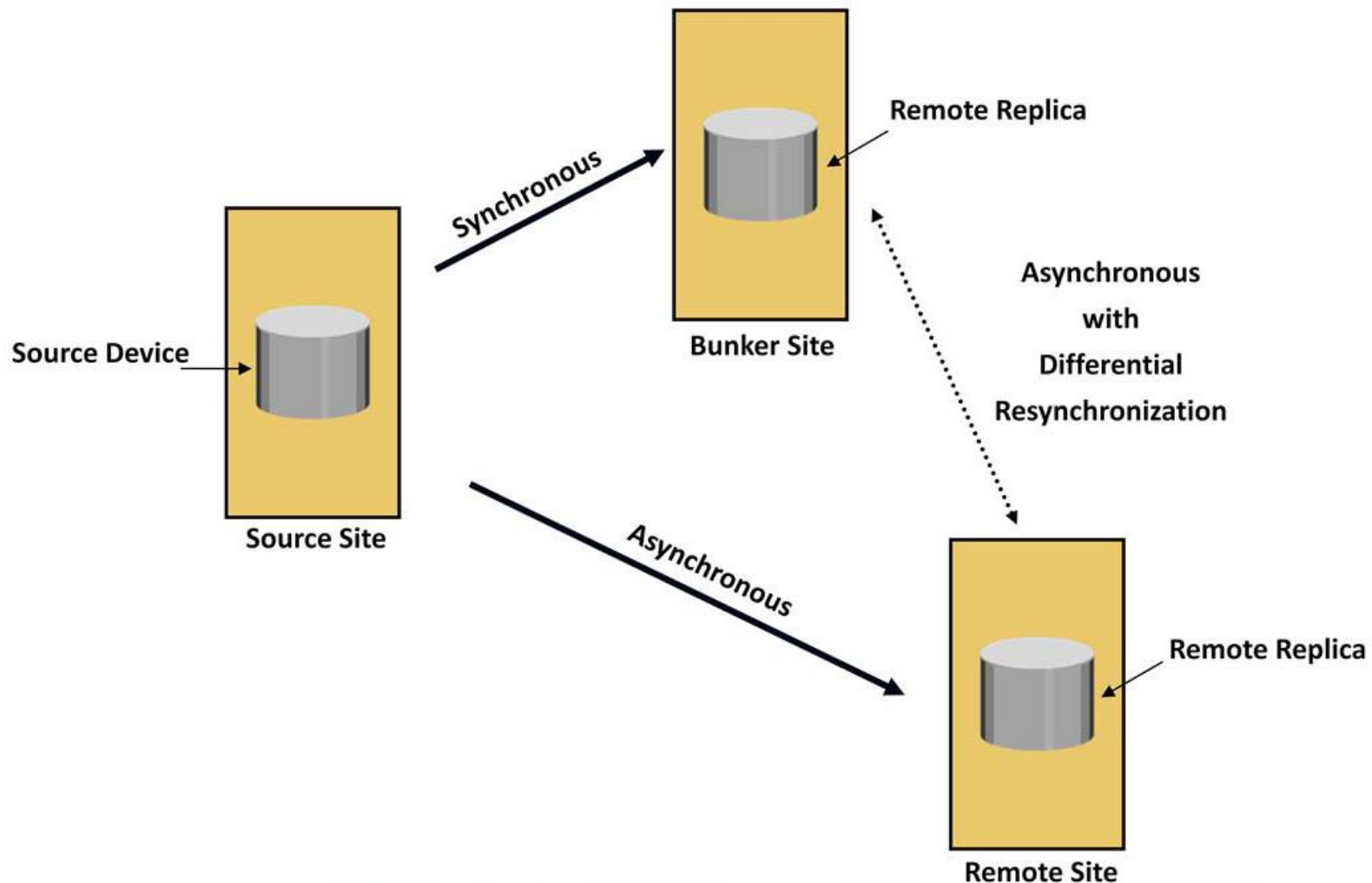
- 4 устройства хранения
- RPO на удаленном сайте – несколько часов

Трехсторонняя каскадная репликация: синхронная + асинхронная

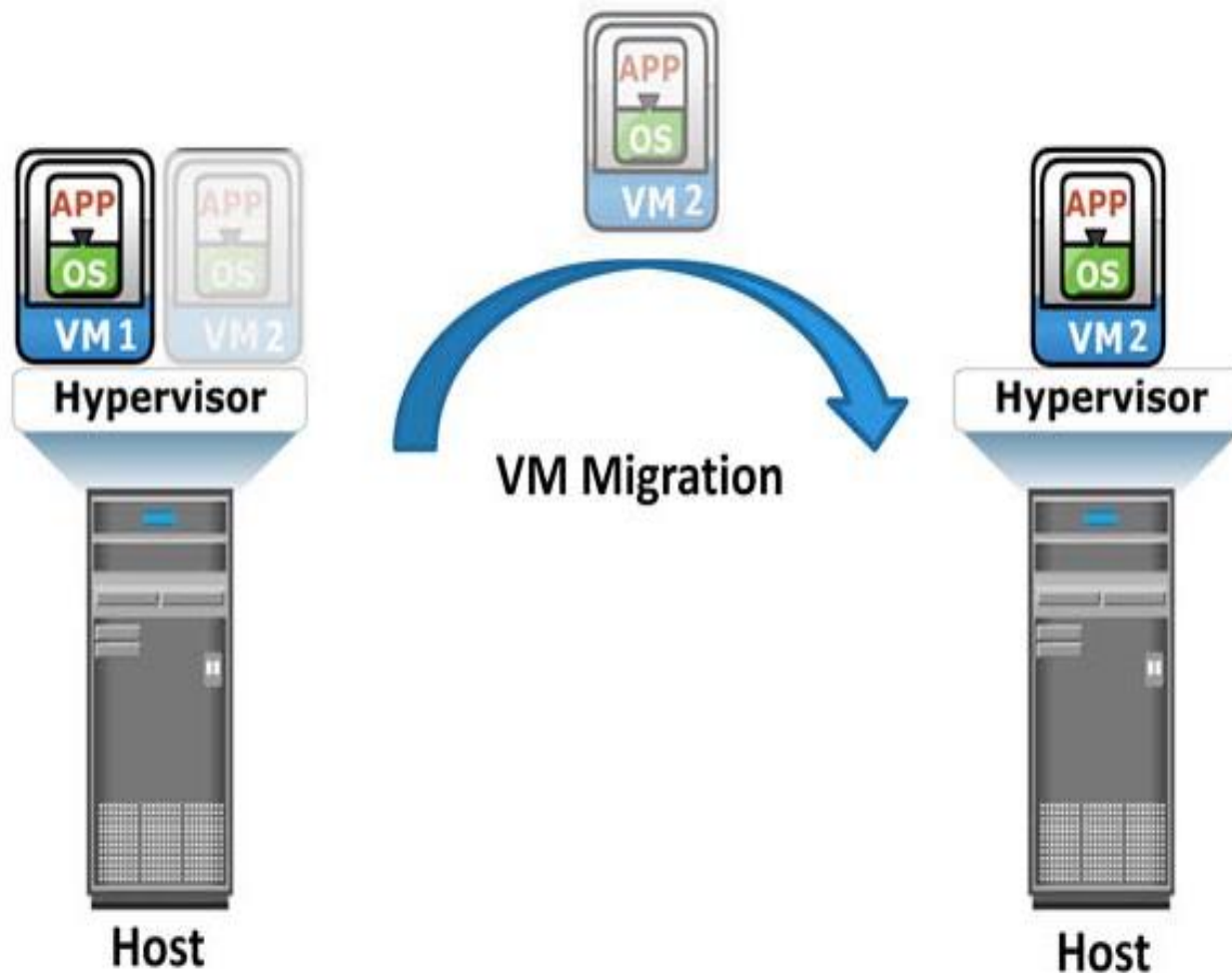


- RPO на удаленном сайте – минуты
- 3 устройства хранения

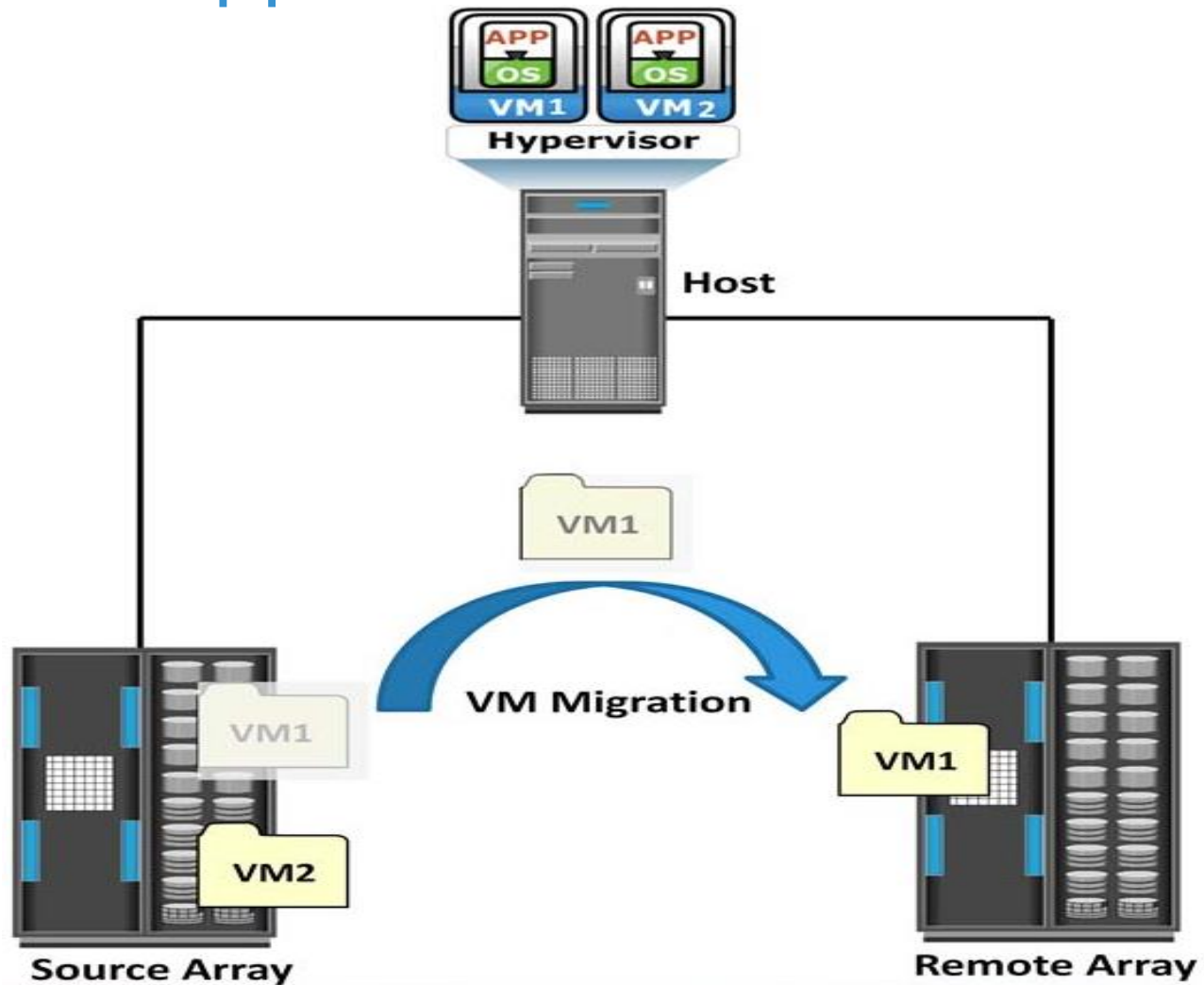
Трехсторонняя треугольная репликация



Репликация в виртуальной среде – с помощью гипервизора



Репликация в виртуальной среде – с помощью СХД



Спасибо!

EMC²®