Высокопроизводительные программные СХД

Сергей Платонов
«Рэйдикс»

RAIDIX

Компания основана в 2009 году экспертами по хранению данных и учеными-математиками

Решения для медиаиндустрии, корпоративного сектора, видеонаблюдения, HPC и других отраслей

Более **10 технологических патентов**, собственные алгоритмы помехоустойчивого кодирования и модели RAID

Продукты RAIDIX **включены в реестр Минкомсвязи** и рекомендованы к закупке государственными структурами

Собственная исследовательская лаборатория, отвечающая за инновации и развитие технологии

Лидирующие в отрасли показатели производительности, пропускной способности и отказоустойчивости

Стратегическое партнерство
с мировыми лидерами индустрии

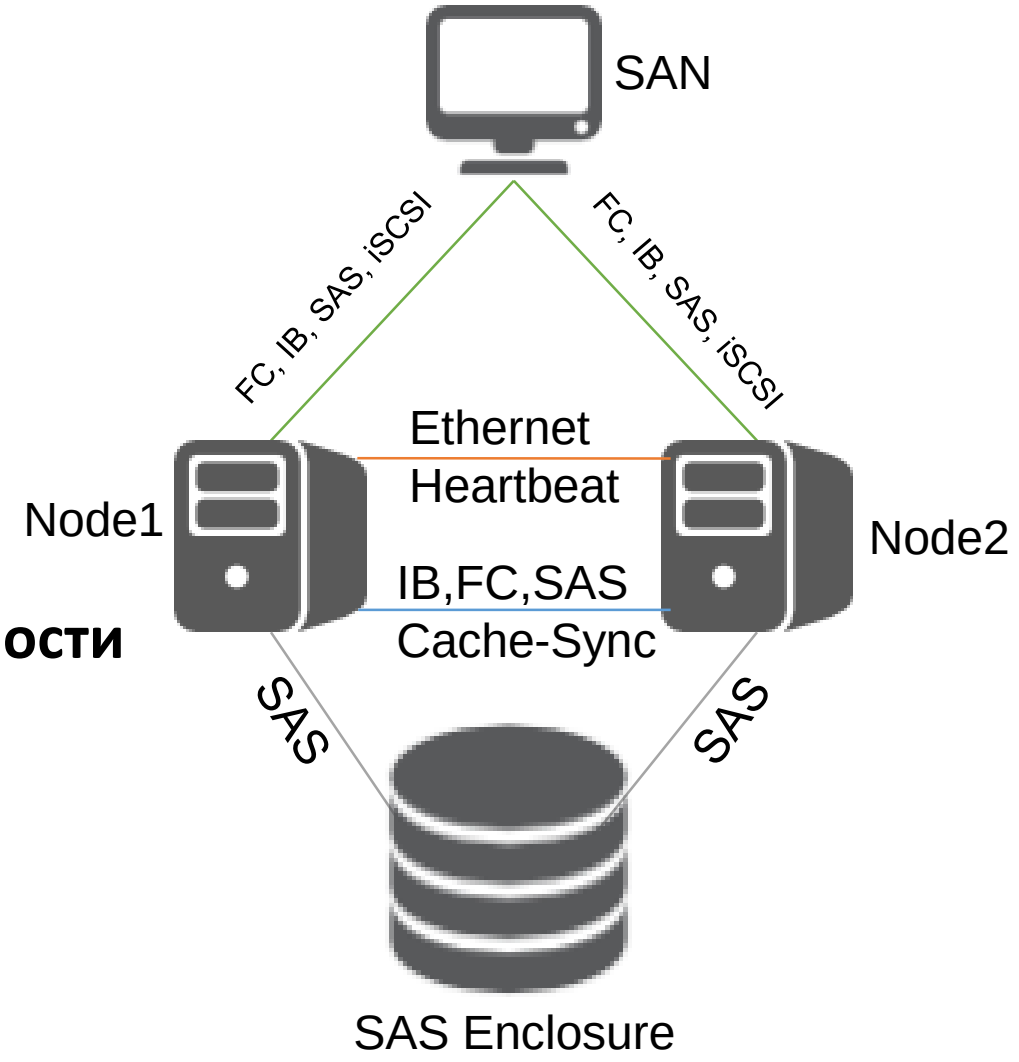
Что же такое RAIDIX?

RAIDIX

Программное обеспечение для СХД

RAIDIX

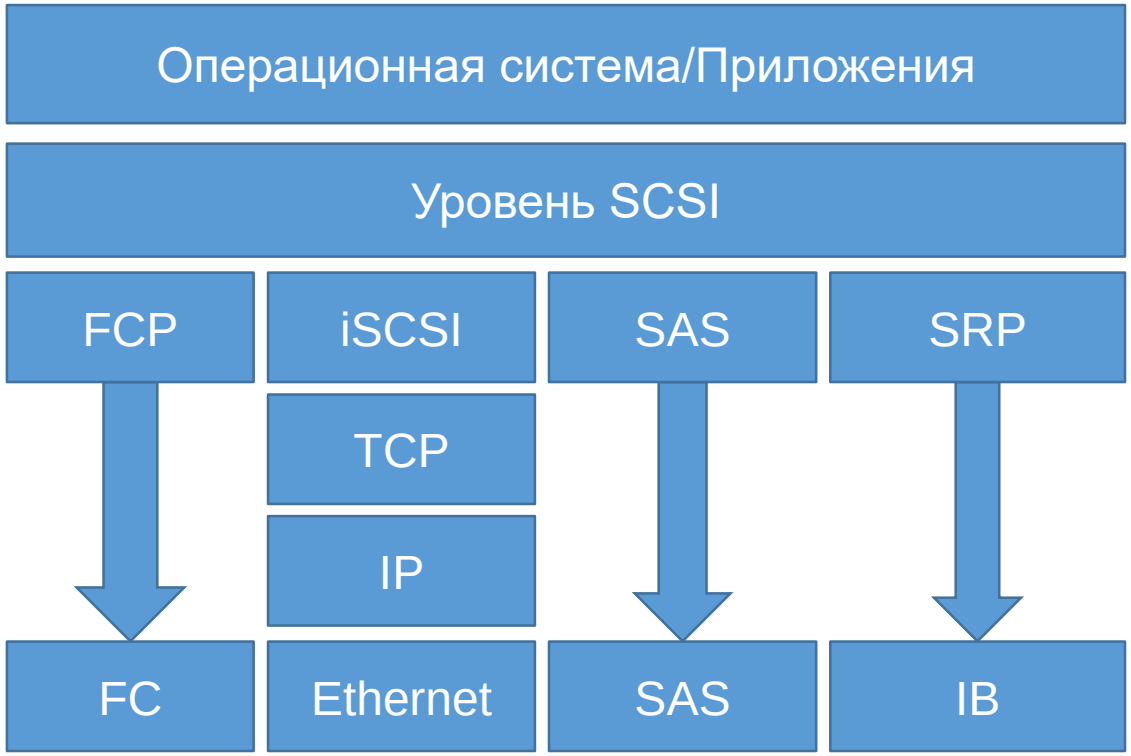
- ✓ **Файловый и блочный типы доступа**
- ✓ **Без единой точки отказа**
- ✓ **Отличный баланс стоимости и производительности**



RAIDIX

Блочный доступ:

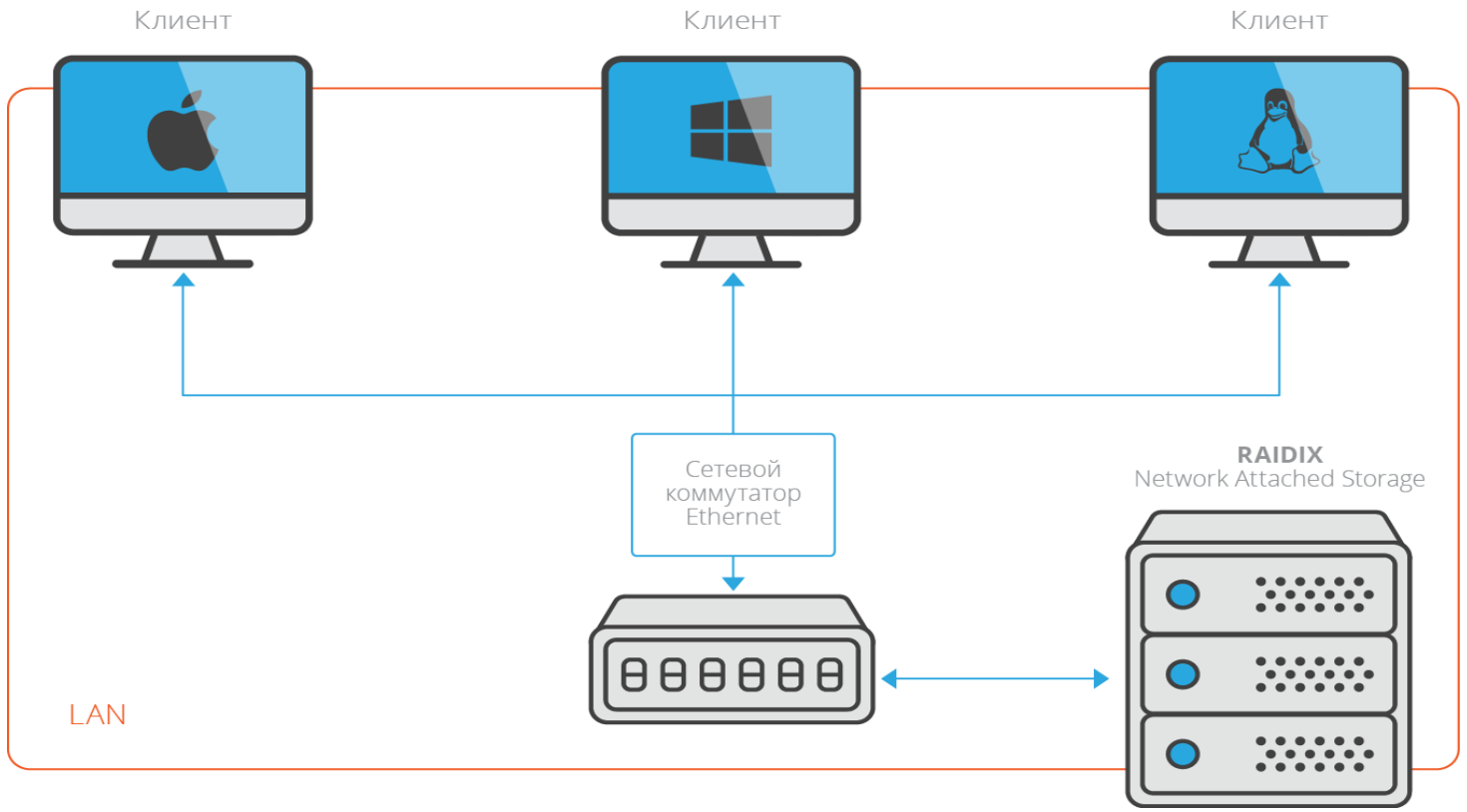
- Fibre Channel – 8Gbit, 16Gbit
- InfiniBand SRP – 20Gbit, 40Gbit, 56Gbit, 100Gbit
- iSCSI – 10Gbit, 25Gbit, 40Gbit
- SAS 12Gbit



RAIDIX

Файловый доступ:

- SMB v2.0
- NFS v3.1
- AFP
- FTP



Для кого?

RAIDIX

Назначение системы

- ✓Высокоинтенсивные задачи, для которых требуется гарантированная пропускная способность
- ✓Высокая стоимость данных и большие объёмы хранения, которые необходимо разместить на ограниченной площади

- ✓Производительность
  - ✓18GBps на узел
    - ✓ Ext4
  - ✓300K IOps на узел
    - ✓ Блочный том
    - ✓ 100% random, 16QD, 8T
    - ✓ 70/30 rw, 4k
- ✓Отказоустойчивость
  - ✓До 32 накопителей в одной RAID-группе могут быть потеряны
  - ✓Потеря производительности при отказе накопителей не превышает 10%

RAIDIX

Основные конкурентные преимущества

✓Самый быстрый в индустрии RAID
Максимальная утилизация оборудования и гарантированный уровень производительности

✓Оптимизация под конкретные задачи
Возможность создания конвергентных и уникальных для решения задачи СХД

✓Высочайшая отказоустойчивость

✓Технологии сжатия

RAIDIX

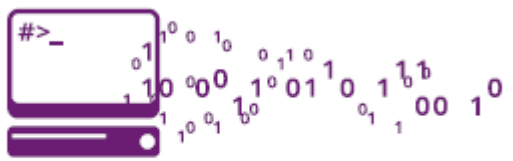
Типовые сценарии

RAIDIX

ENTERPRISE

MEDIA

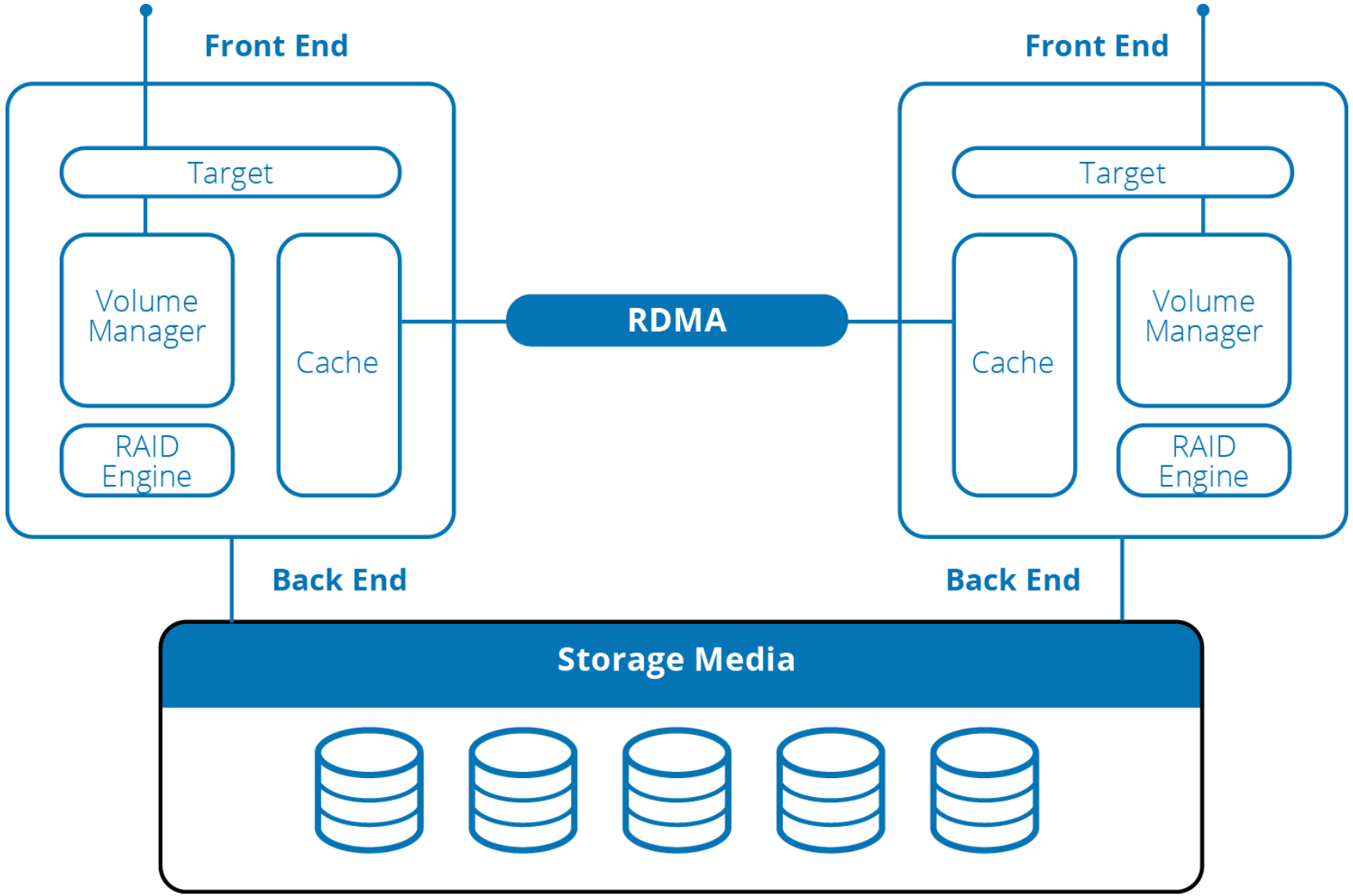
VIDEO SURVEILLANCE



HPC

Как это работает?

RAIDIX



RAIDIX

RAID

RAIDIX

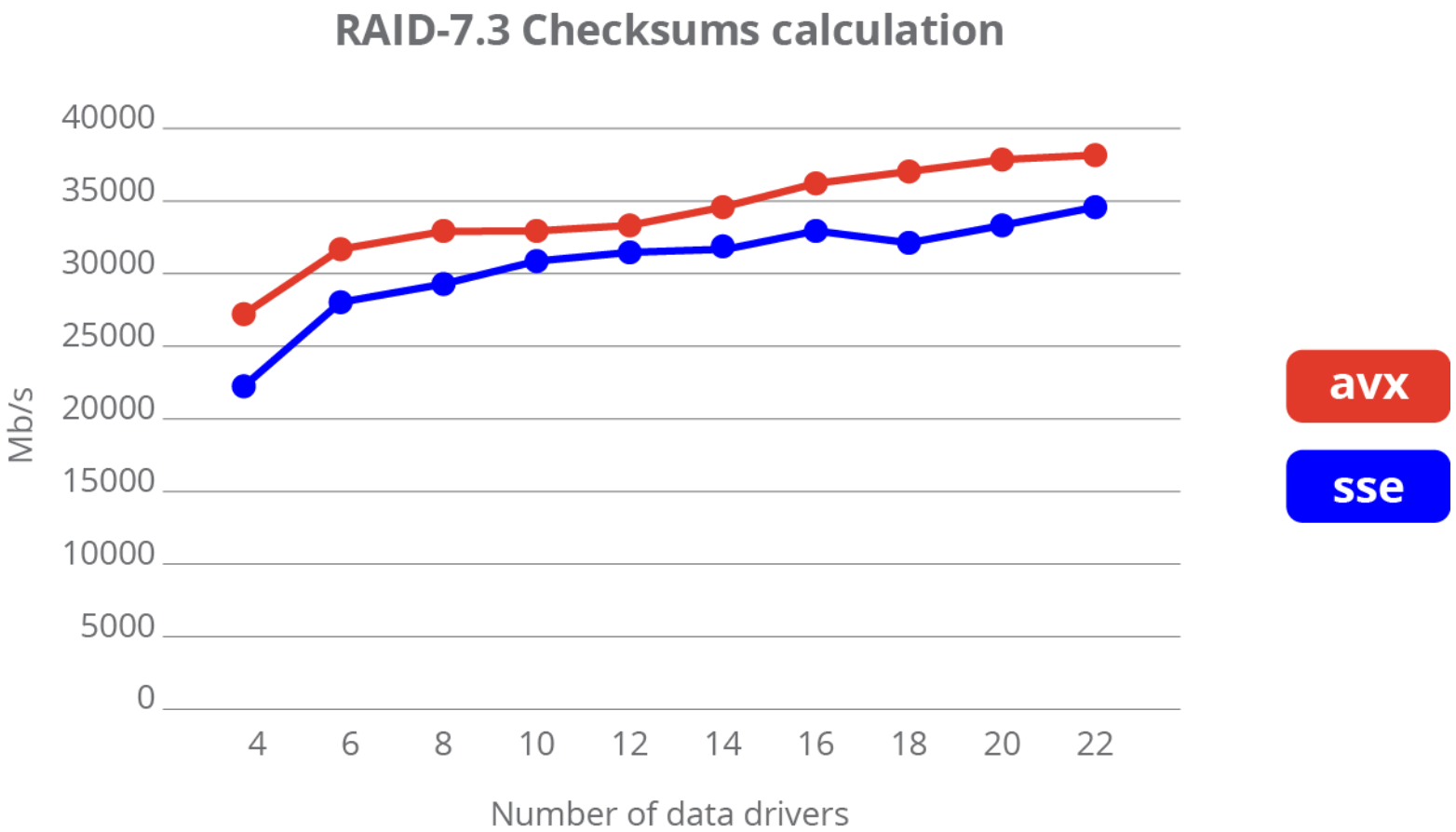
RAIDIX 4.x поддерживает RAID 0, RAID10, RAID 5, RAID 6, RAID 7.3 и RAID N+M

RAID 6 — массив с двойной чётностью

RAID 7.3 —массив с тройной чётностью, обеспечивает **сохранность** данных при выходе из строя **до 3**-х дисков

RAID N+M — массив с количеством дисков чётности **до 32-х**

RAIDIX



RAIDIX

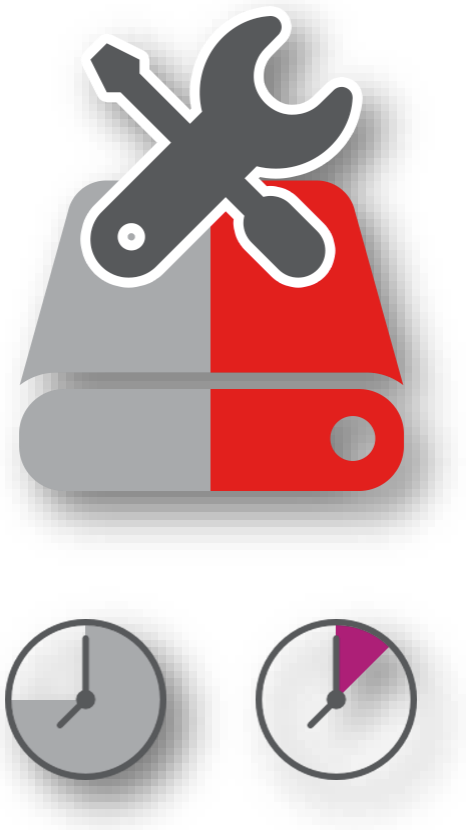
Возможность оптимизации скорости чтения

Из процесса исключается до двух дисков (в случае RAID 7.3 - до трех дисков), скорость чтения с которых ниже, чем у остальных.

RAIDIX

- **Возможность восстанавливать только часть диска** за счет использования собственный алгоритма расчёта RAID-массива. Это позволяет существенно уменьшить время реконструкции при случайном извлечении диска из массива

- Частичная реконструкция очень эффективна для массивов большого объёма



RAIDIX

- RAIDIX использует уникальный алгоритм, позволяющий обнаружить и устранить **скрытые повреждения данных**
- Анализ метаданных без провала производительности
- Обнаружения и исправления ошибок идёт в фоновом процессе

RAIDIX

Кэш

RAIDIX

Двухуровневый кэш – механизм SSD-кэширования позволяет обеспечить высокие показатели производительности при случайном чтении.

Сквозная запись – хост получает подтверждение записи только тогда, когда данные были записаны в основную память (на диски). Сквозная запись существенно уменьшает риск потери данных и улучшает производительность в двухконтроллерной конфигурации, т.к. не требует синхронизации кэшей.

Настраиваемые алгоритмы кэширования

RAIDIX

Кэш второго уровня

- ✓Лог-структурированный подход
- ✓Адаптивное выделение пространства RWC/RRC
- ✓ACD

RAIDIX

Дедупликация

RAIDIX

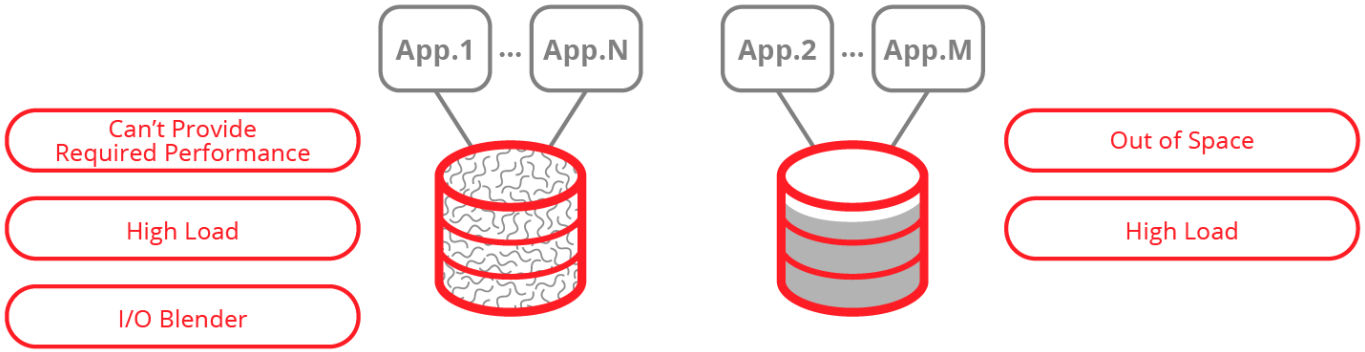
- ✓ В режиме реального времени
- ✓ Постоянный размер блока 4k
- ✓ Высокая производительность без просадок
- ✓ “Умный” индекс

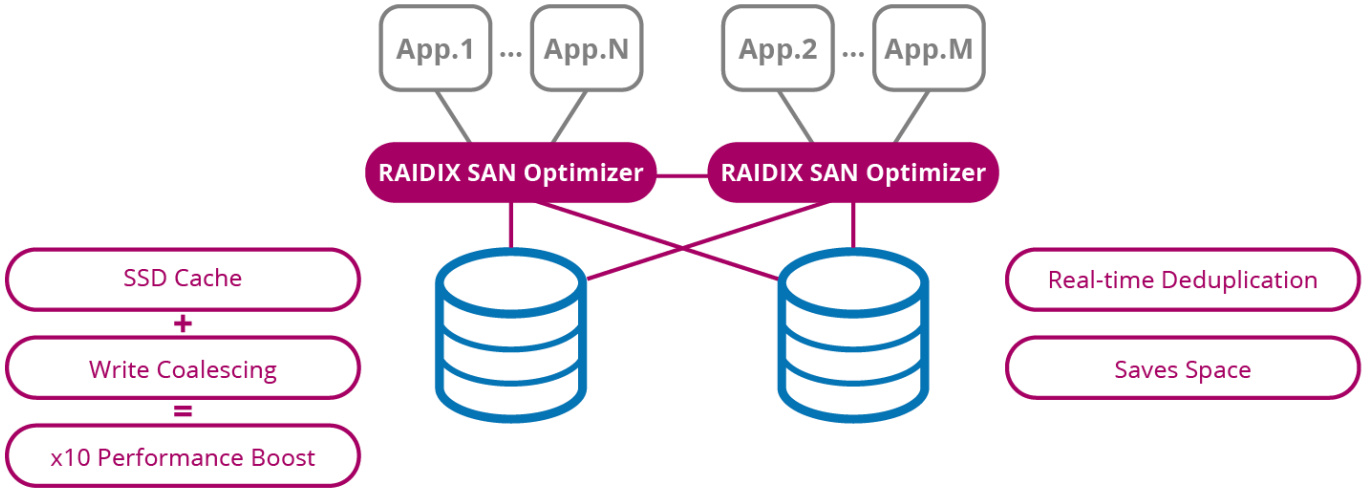
RAIDIX

	RAIDIX	Comp 1	Comp 2	Comp 3	Comp 4
Inline Deduplication	Yes	No	Yes	Yes	Yes
Dedupe Granularity	4 KB	4 KB	Dedupes based on IO size 512 B – 32 KB	4 KB to 128 KB (64 KB default)	8 KB
Inline Compression	Yes	Yes	Partial	Yes	Yes
Dedupe before Compression	Yes	No	No	No	Yes
4 K Random IO performance RW80/20	650K	N/A*	150K	<100K	150K per node
Memory Usage	0.1 GB/TB	N/A	4 GB/TB	20 GB/TB	25GB/TB
Scalability	256 TB per pool	100 TB per node	70 TB	Limited by memory requirements	20 TB per node

RAIDIX

- В RAIDIX 4.x реализована возможность мониторинга параметров работы RAID-массивов и системы в целом – в режиме реального времени
- Вся информация представлена в виде графиков в веб-интерфейсе





Спасибо за внимание!

Сергей Платонов
platonov.s@raidix.com

RAIDIX

Дедупликация. Индекс

Высокопроизводительный индекс
200,000 IOPS/сек/ядро

Масштабируемость
640 млрд. объектов/сервер

Низкие требования к RAM
0.1 байт/записей в RAM

RAIDIX

Виртуальный том

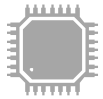
Производительность

1.3M (4 КБ) IOPS

Без «сбора мусора»

Масштабируемость

До 4 ПБ физически/1 ЭБ логически



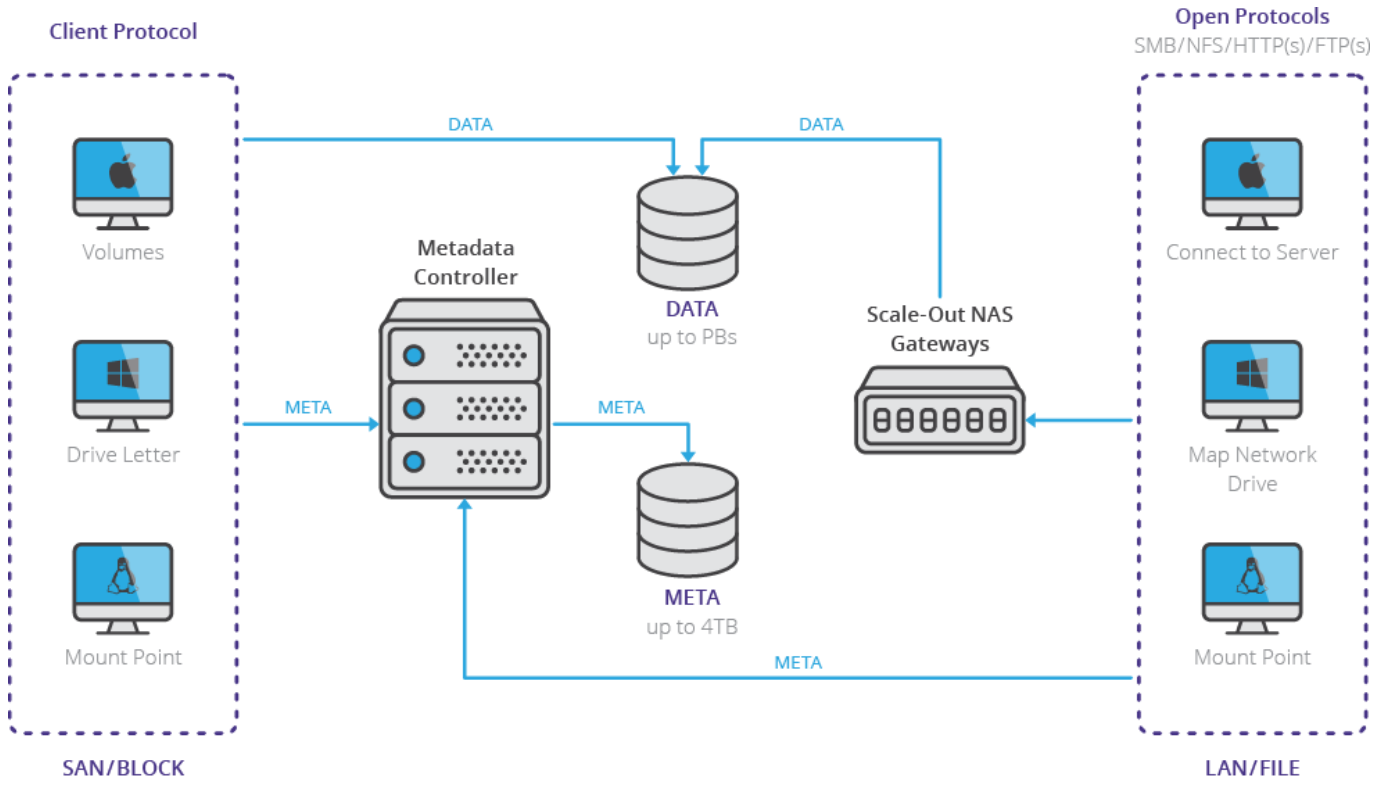
Низкие накладные расходы

280 МБ RAM/1 ТБ физического хранения с блоками по 4 КБ

RAIDIX

RAIDIX EXASPHERE. ГОРИЗОНТАЛЬНО-МАСШТАБИРУЕМЫЙ NAS И SAN ОБЩЕГО ДОСТУПА

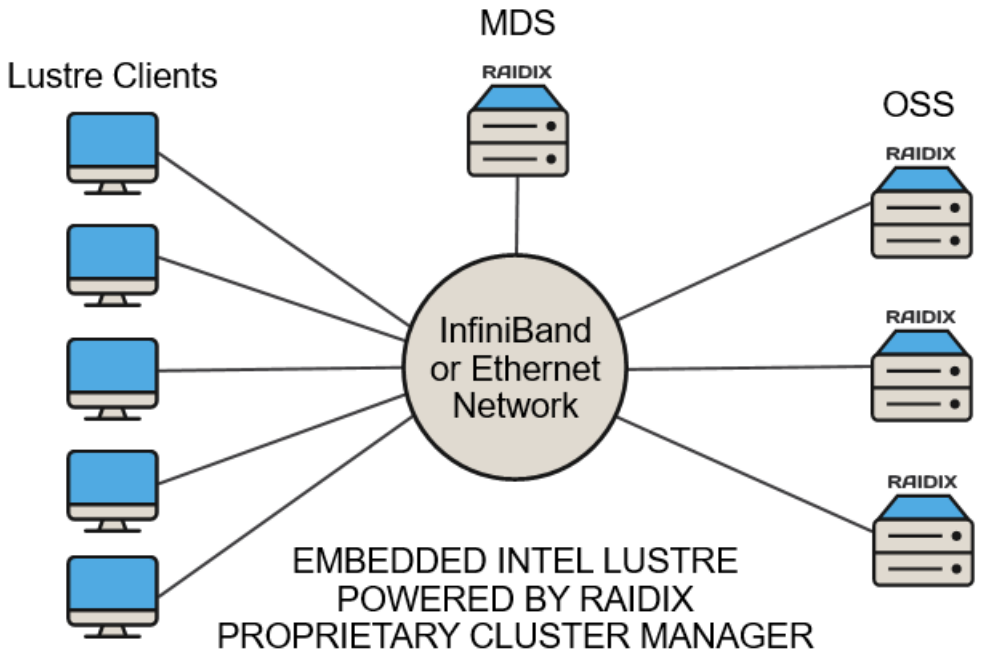
- Узлы хранения представлены как единый пул
- Масштабируемость: возможность увеличить емкость по мере роста объемов данных
- Высокопроизводительные блочные интерфейсы
- Встроенный сервер метаданных
- Отказоустойчивая конфигурация



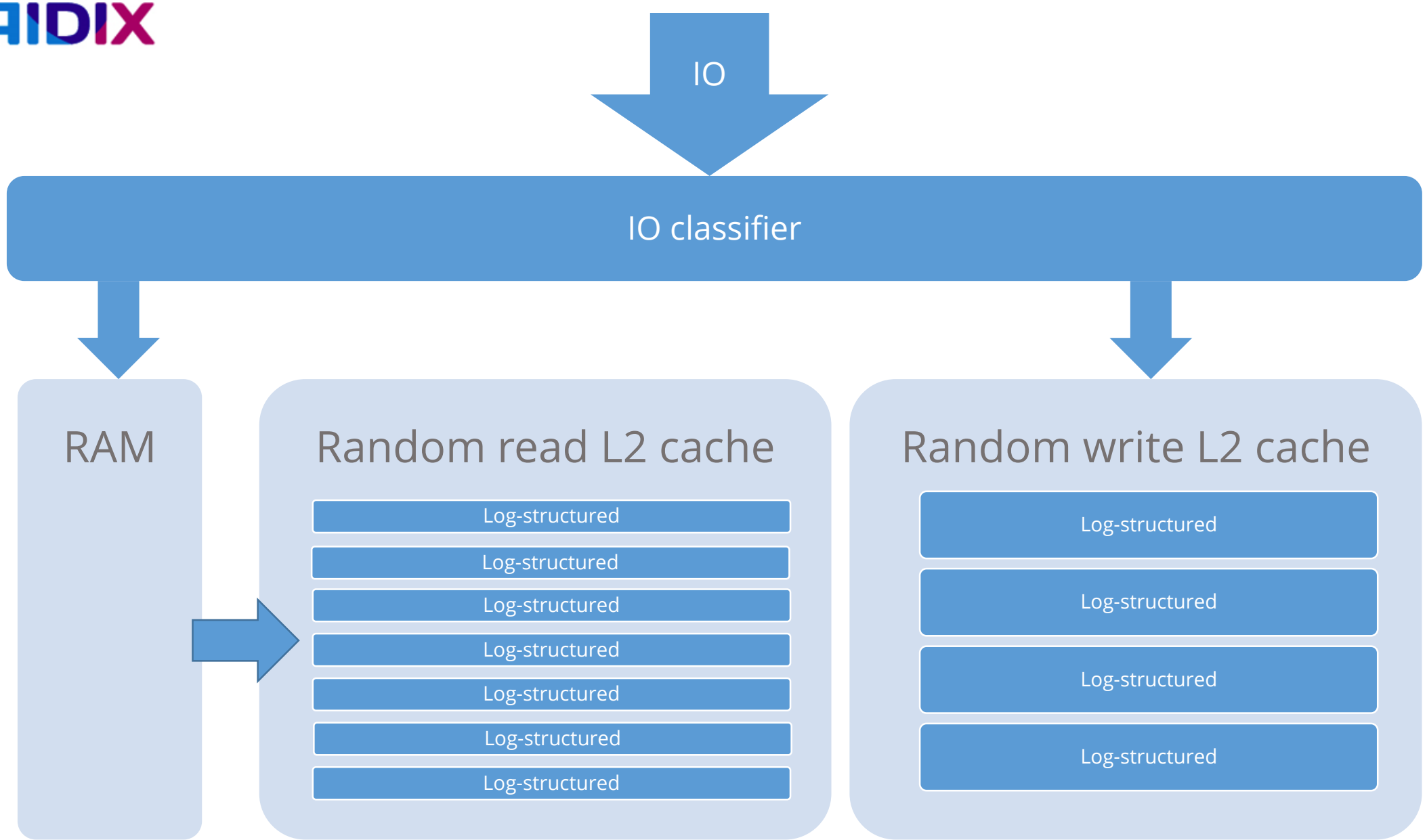
RAIDIX

RAIDIX EXASPHERE. ГОРИЗОНТАЛЬНО-МАСШТАБИРУЕМЫЙ NAS И SAN ОБЩЕГО ДОСТУПА

- Технология RAIDIX HPC/cluster-in-a-box («готовый кластер») + Intel® Enterprise Edition for Lustre*
- Построение высокопроизводительных кластеров хранения данных
- Масштабируемость до 512 ПБ, пропускная способность до 2 ТБ/с.
- Высокая производительность, отказоустойчивость и быстрый ввод резерва
- Асимметричная двухконтроллерная архитектура
- Прямая установка Intel® Enterprise Edition for Lustre*, OSS+OST (Active-Active) и MDS+MDT (Active-Passive) на узлы хранения
- Сервисы Lustre* могут управлять томами локально, отсюда более низкие задержки и улучшенная производительность.



RAIDIX

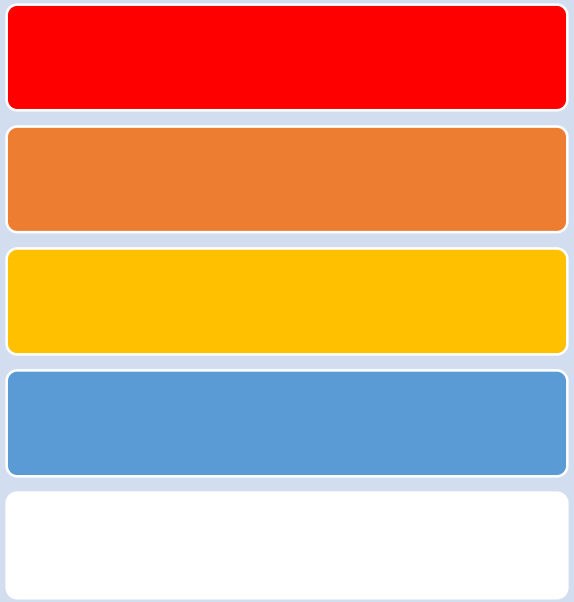


RAIDIX

ТЕХНИКА ВЫТЕСНЕНИЯ RRC

«Призрачная очередь» (Ghost queue)

Выбор самого «холодного» буфера



Вытесняются только «холодные» блоки, «горячие» и «теплые» блоки перемещаются в свободный буфер