

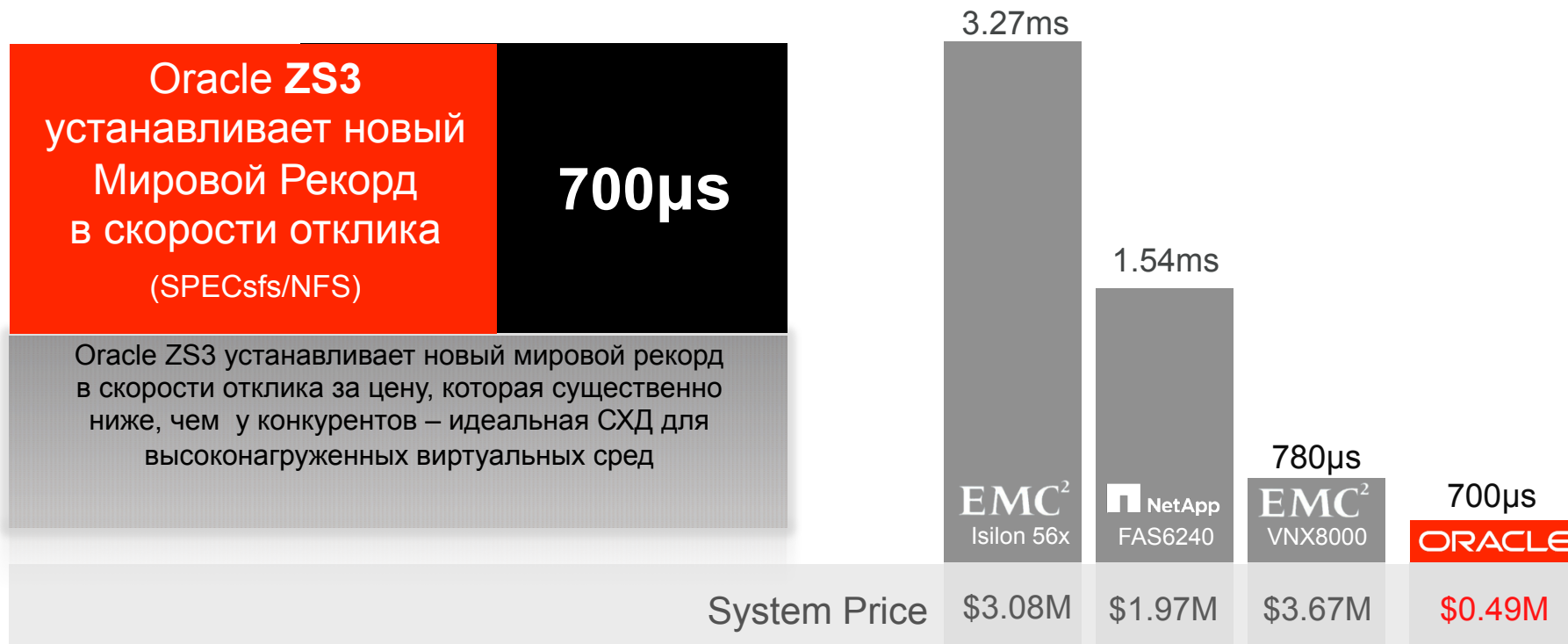
Системы хранения Oracle – все, что Вам нужно знать о них

Басов Андрей
Ведущий технический специалист
направления Oracle
basov@marvel.ru

ORACLE®

Certified Specialist

Рекорд производительности



Results from www.spec.org as of 09/10/2013. Oracle ZFS Storage ZS3-4 delivers 450,702 SPEC SFS2008_nfs.v3 ops/sec, 0.70 msec ORT. Configuration Bill of Materials total list price for the Oracle ZFS Storage ZS3-4 is \$487,738 from oracle.com. NetApp 8-node FAS6240 cluster delivers 512,667 SPEC SFS2008_nfs.v3 ops/sec, 1.54msec ORT. Configuration Bill of Materials total list price is \$1,973,880 from www.netapp.com/us/media/nalist-usd-netapp-custom-state-new-discounts.pdf. EMC Isilon 56 node cluster delivers 456,223 SPECsfs2008_nfs.v3 ops/sec, 3.27 msec ORT. Configuration Bill of Materials total list price is \$3,080,280 from www.emc.com/sales/stateoffl/index.htm. EMC VNX 8000 8 node cluster delivers 580,796 SPECsfs2008_nfs.v3 ops/sec, 0.78 msec ORT. Configuration Bill of Materials total list price is \$3,676,106 from www.emc.com/sales/stateoffl/index.htm

Zettabyte File System



128 битная файловая система

В $18,4 \times 10^{18}$ раз больше данных, чем в 64 битные ФС

все пределы – теоретические,

например:

2^{48} - количество снимков (2×10^{14});

16 эксабайт (2^{64} byte) - максимальный размер файла;

2^{56} — количество файлов в директории;

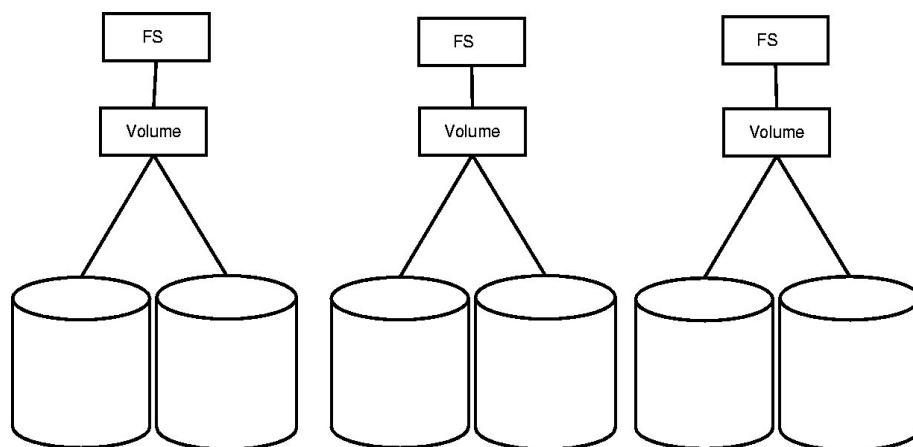
2^{64} — количество устройств в пуле;

2^{64} — количество пулов в системе;

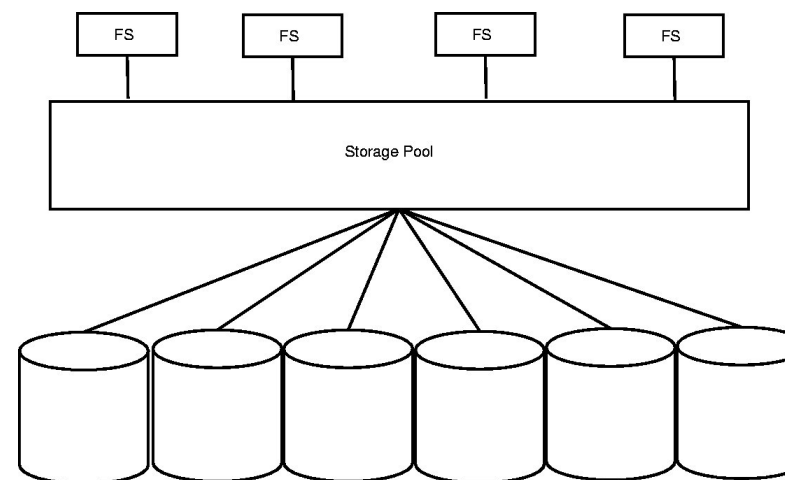


ZFS. Пулы устройств

Традиционный подход

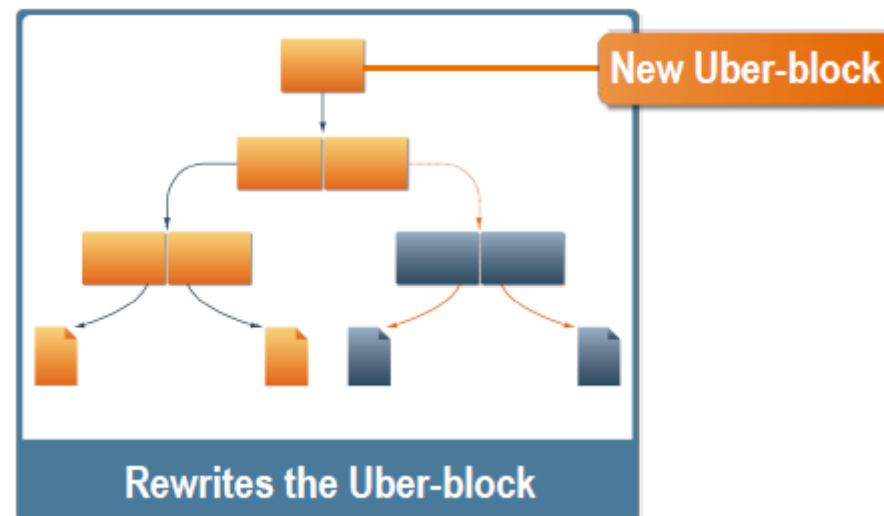
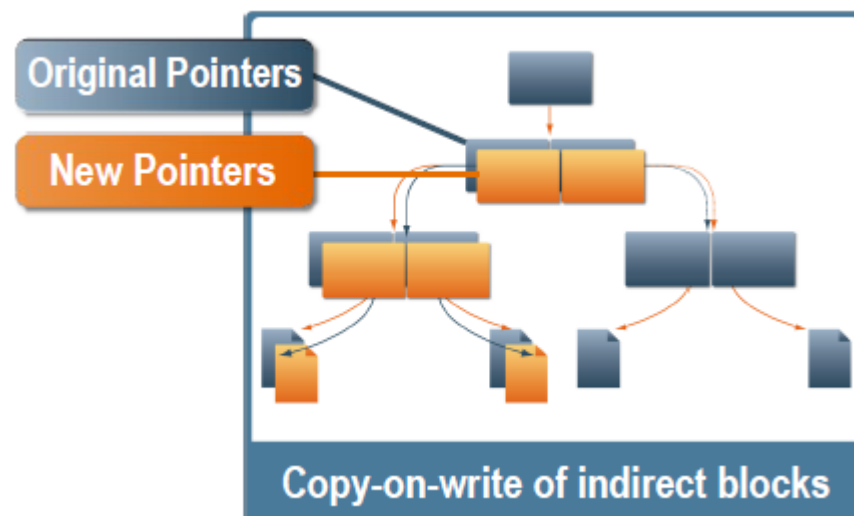
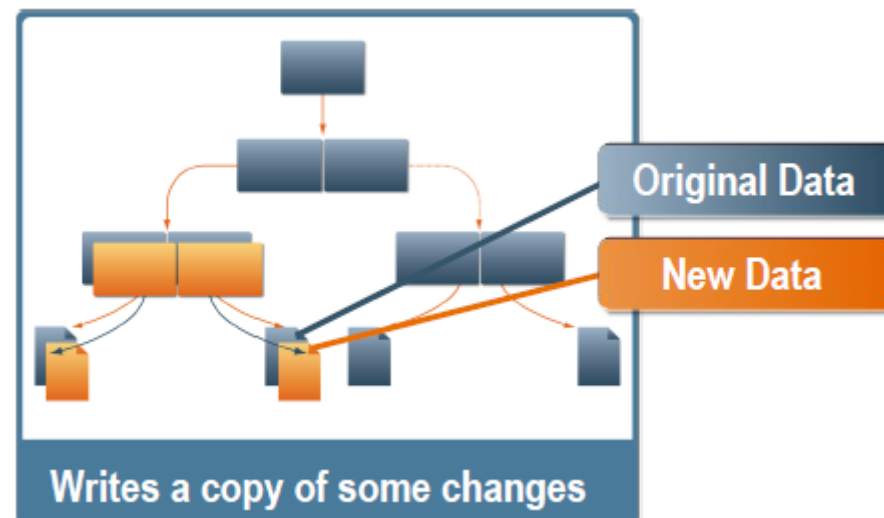
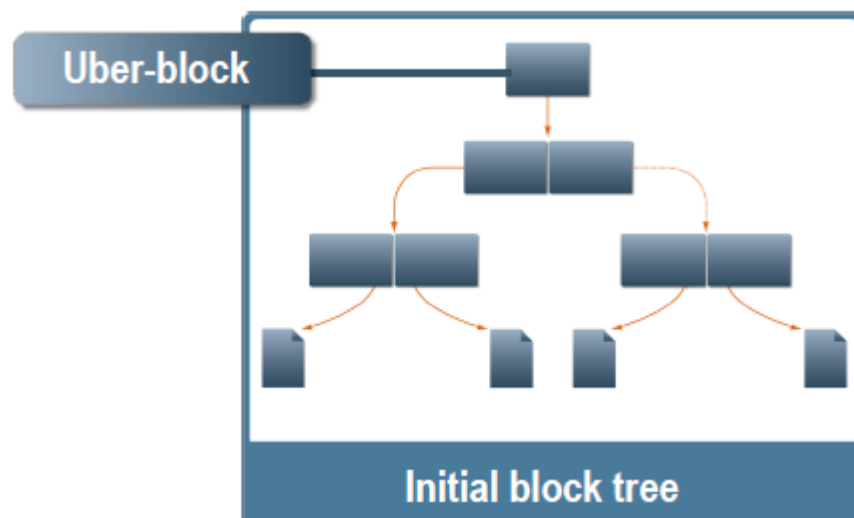


Пулы устройств

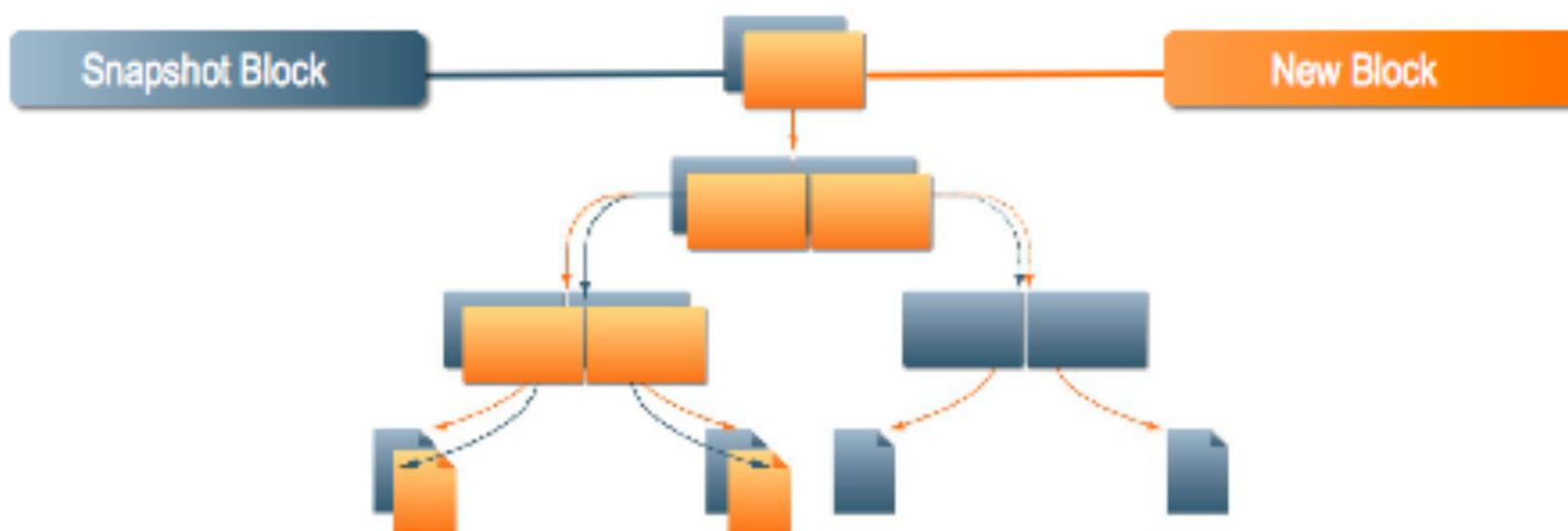


- **общий поток**

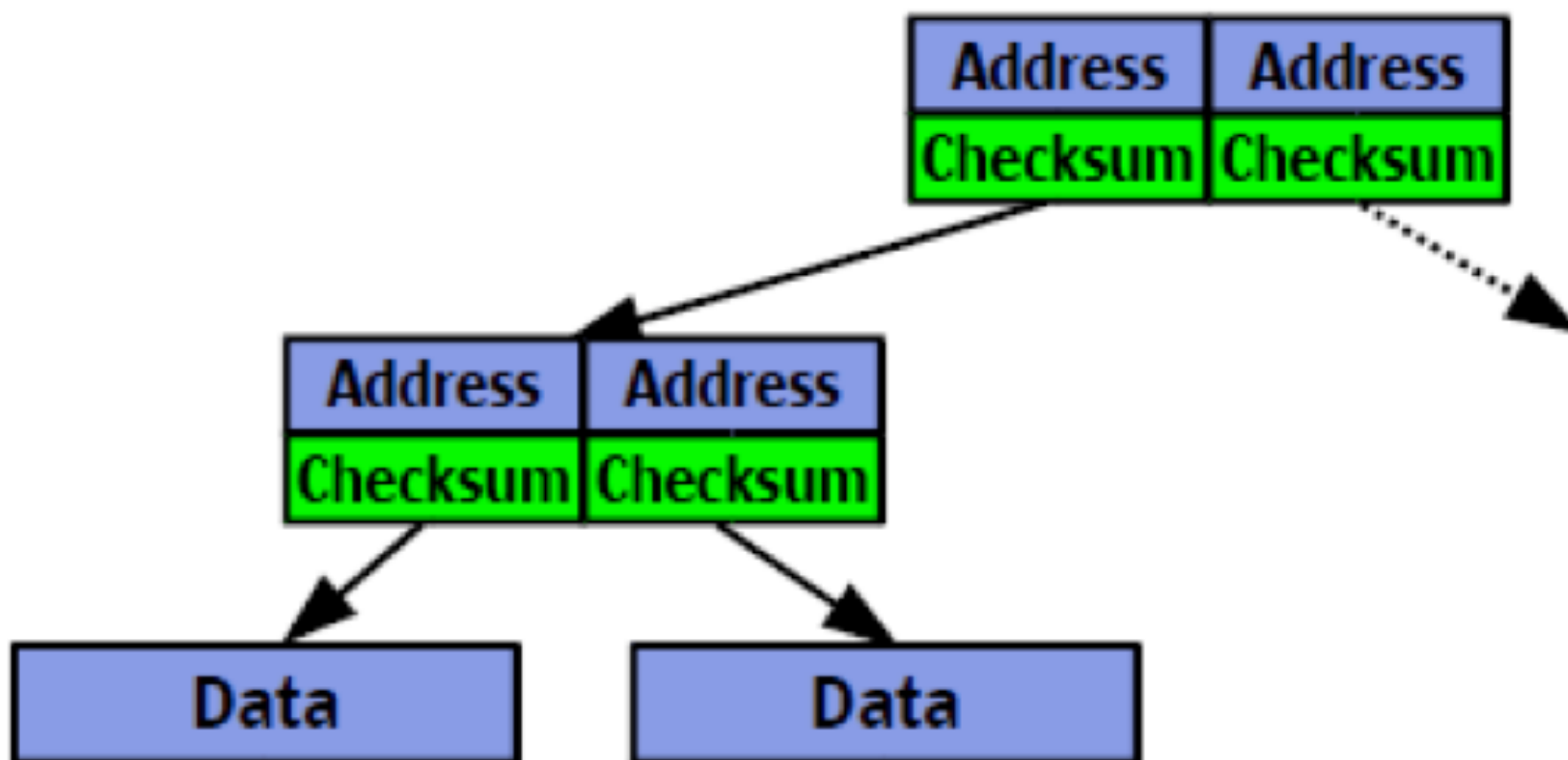
ZFS. Copy-on-write



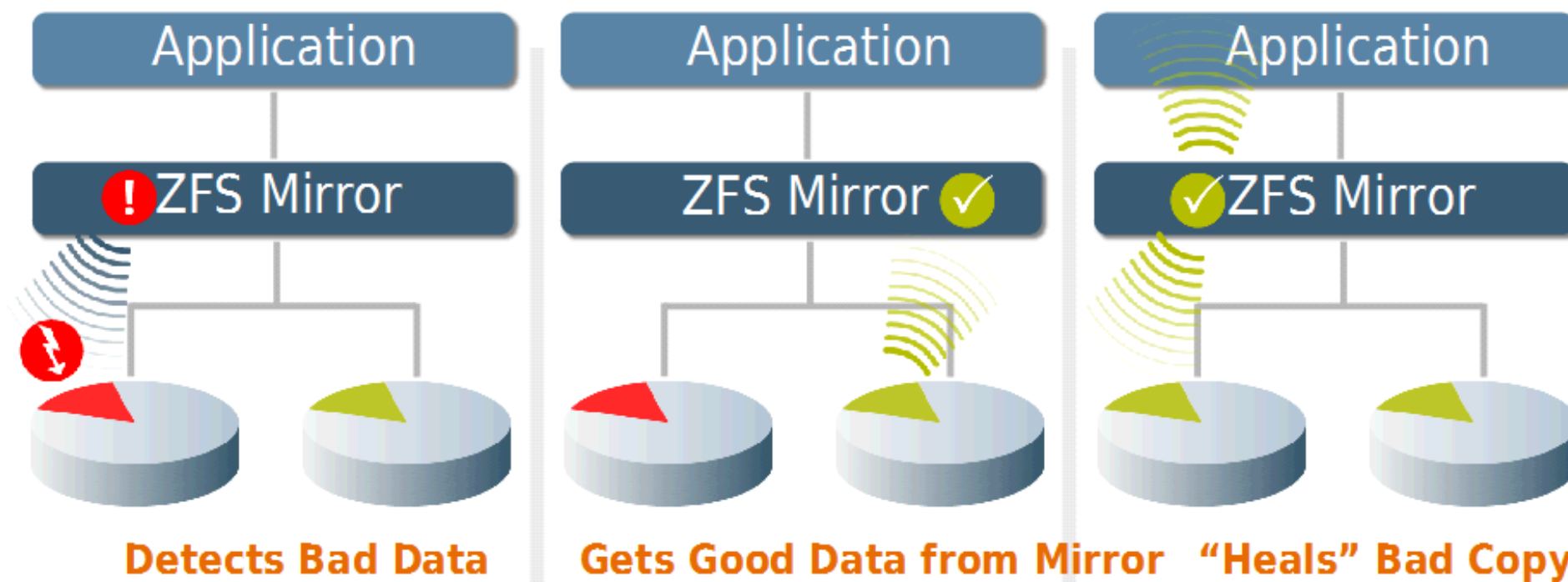
ZFS. Моментальные снимки



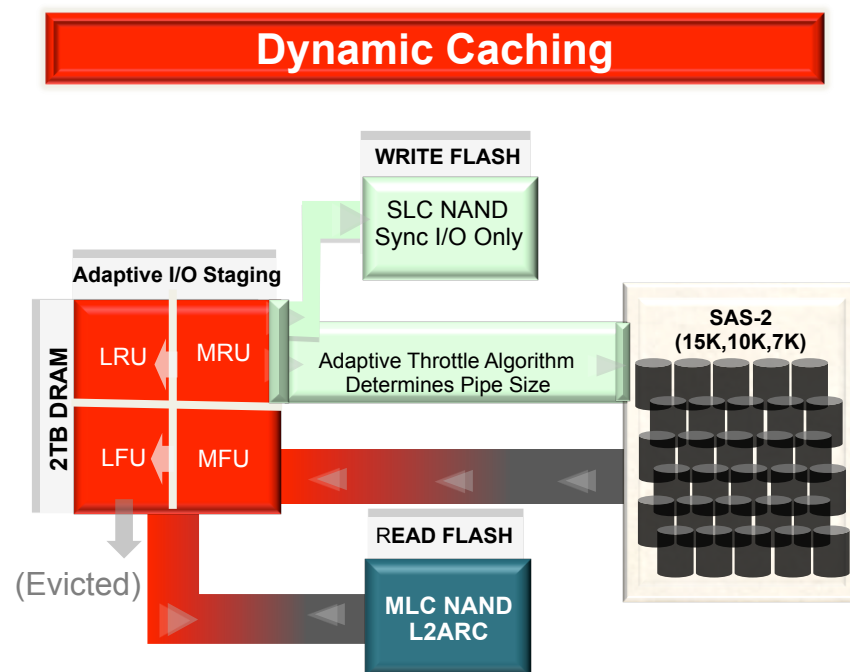
ZFS. Сквозной контроль целостности



ZFS. Самовосстановление



ZFS. Адаптивный кэш



Автоматизированная миграция данных в реальном времени между DRAM, Flash-памятью и дисками.

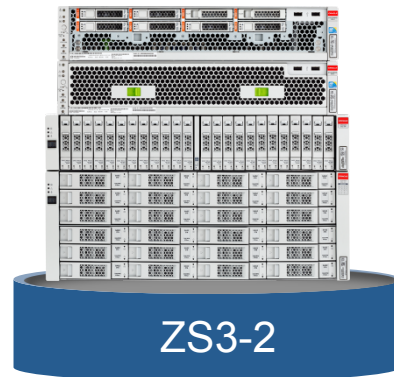
Новая линейка Oracle ZS3



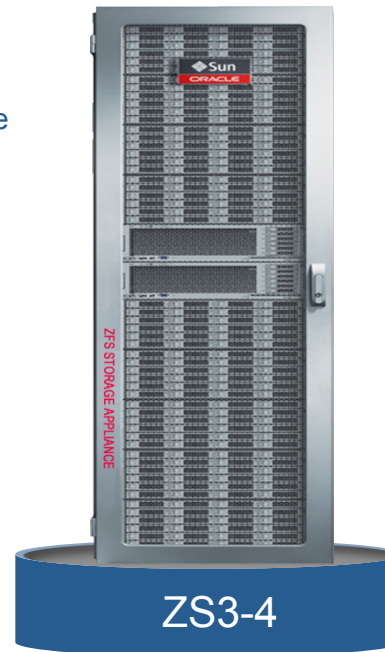
Операционная система ZFS Storage

Наиболее мощный программный пакет для СХД

Создана вместе с Базой Данных и приложениями Oracle



ZS3-2



ZS3-4



ZS3-BA

В 2 раза быстрее
В 3 раза гибче

- Один/два контроллера
- 512GB DRAM кеш
- 8 PCIe слотов
- 8 дисковых полок/750TB
- 12TB SSD кеш на чтение
- 2TB SSD кеш на запись

- Один/два контроллер
- 2TB DRAM кеш
- 14 PCIe слотов
- 36 дисковых полок/3.4PB
- 12TB SSD кеш на чтение
- 10TB SSD кеш на запись

- Настроена для резервного копирования Инженерных Систем
- Смонтирована & оптимизирована
- Наличие Configuration Utility

В каждом контроллере: **2 X 8 core**

4 X 10 core

СХД разработана вместе с приложением



База Данных и
приложения
взаимодействуют с
СХД

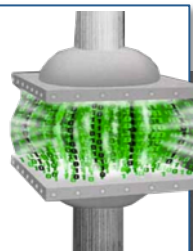


СХД взаимодействует с
Базой Данных и
приложениями

Гибридное колоночное сжатие



База Данных
Oracle с гибридным
поколоночным
сжатием



Сжатие данных
в 10 – 15 раз

Идеально для: Хранилищ данных,
Database ILM, DR, тестирования/
разработки

Sun ZFS
Storage
Appliance



3–5x меньше дисков

3–5x ниже энергопотребление

3–5x ниже стоимость

Oracle Database 11g R2 EE на любом сервере, без дополнительной стоимости

Выше эффективность и производительность



С НСС данных сжаты:

- На дисках
- В кэш-памяти СХД
- При передаче в сети
- В оперативной памяти
- В системной кэш памяти

База данных (retail)	Начальный размер	Таблицы после НСС	Улучшение
Размер	735 GB	38GB	19.3x
Запрос 1	52:06	10:52	4.7x
Запрос 2	2:22:44	16:50	8.4x
Запрос 3	7:06	1:06	6.4x

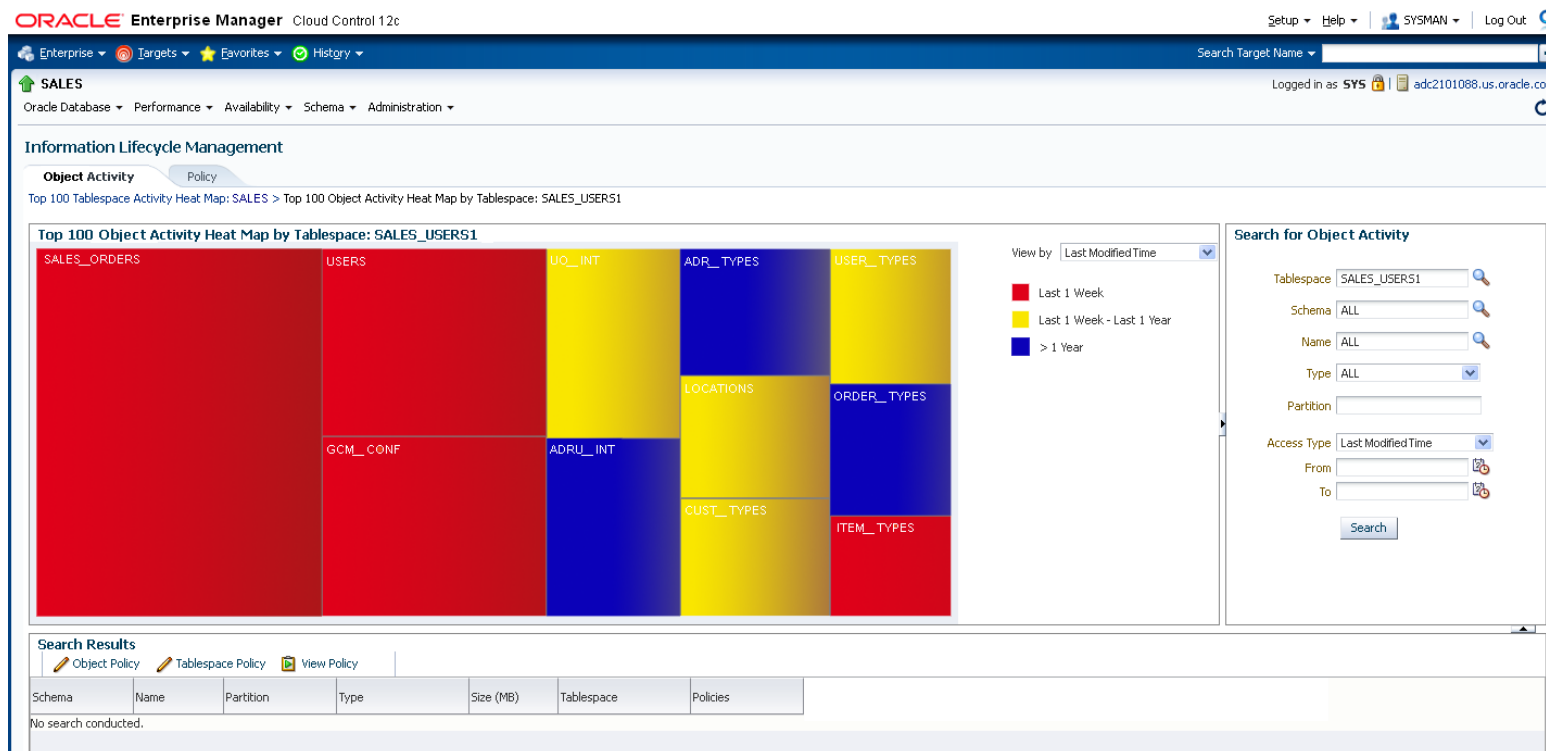
Больше данных размещаются в памяти

Короче окна резервного копирования

Oracle Database 12c Automatic Data Optimization



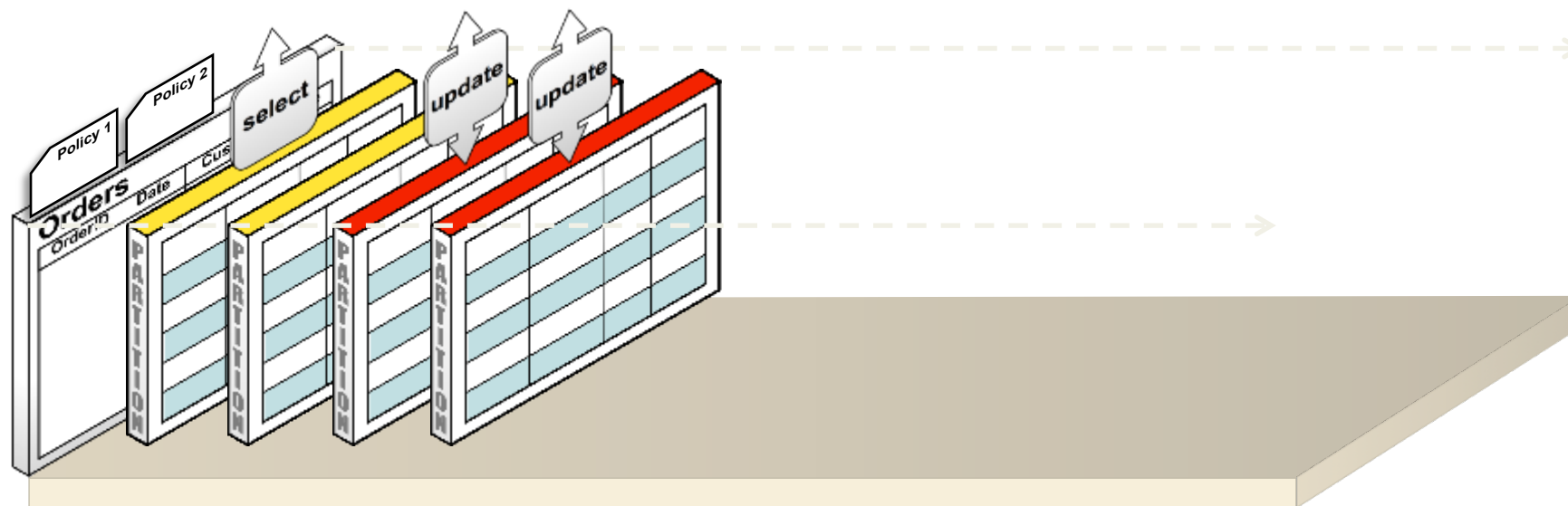
Сжатие и перемещение данных на основе температурной карты



Автоматическая Оптимизация Данных



Температурная карта отслеживает активность сегментов и блоков



Partitioned By Time

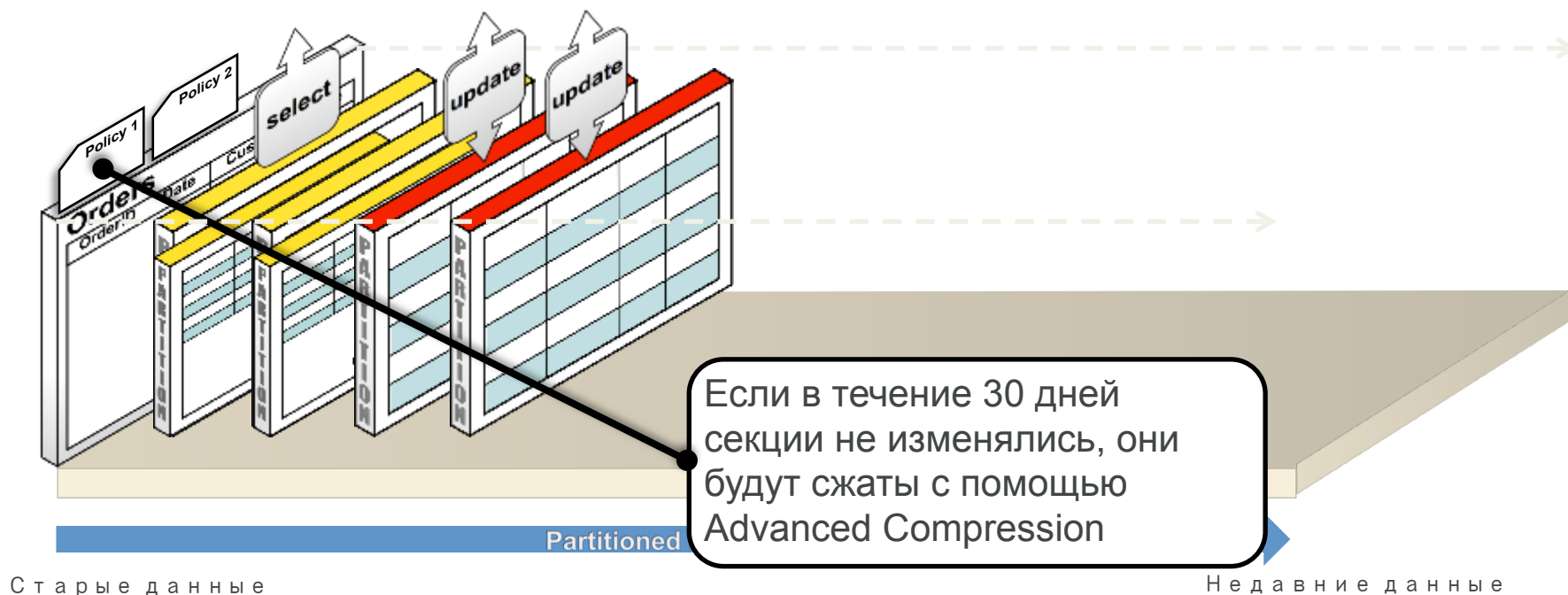
Старые данные

Недавние данные

Автоматическая Оптимизация Данных



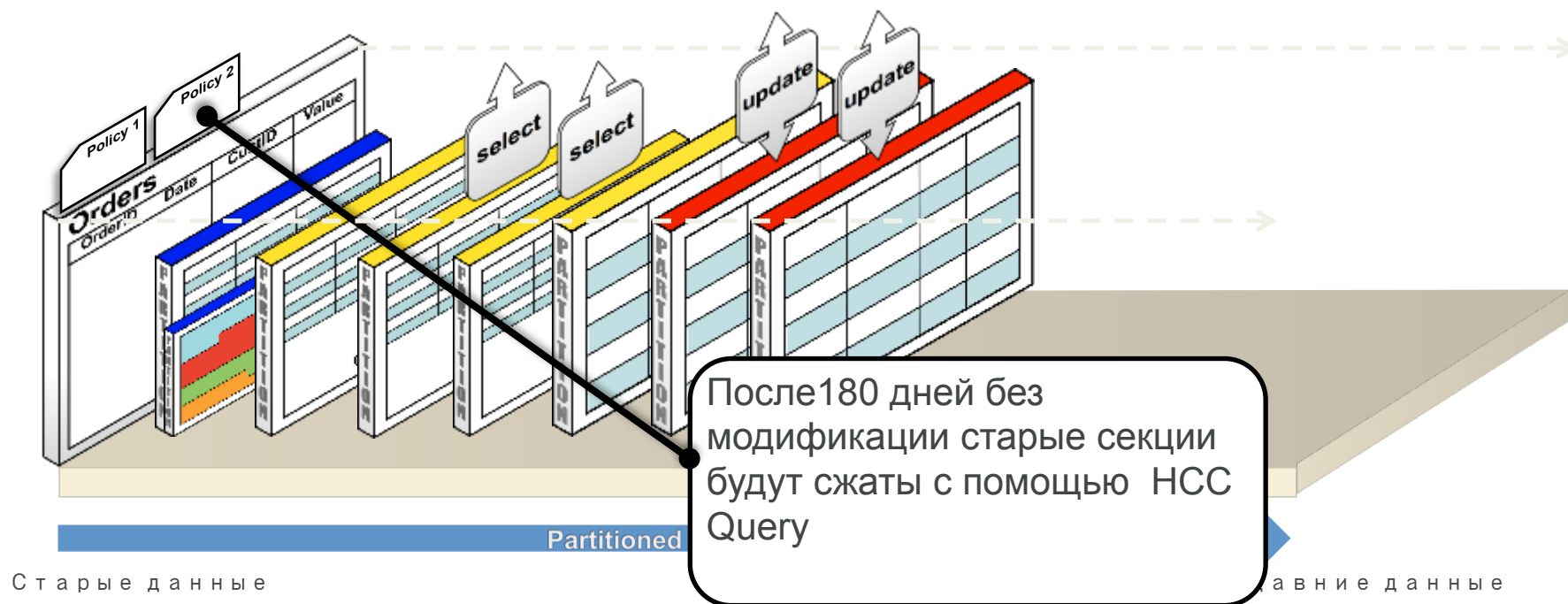
Политики автоматически применяются к таблицам



Автоматическая Оптимизация Данных



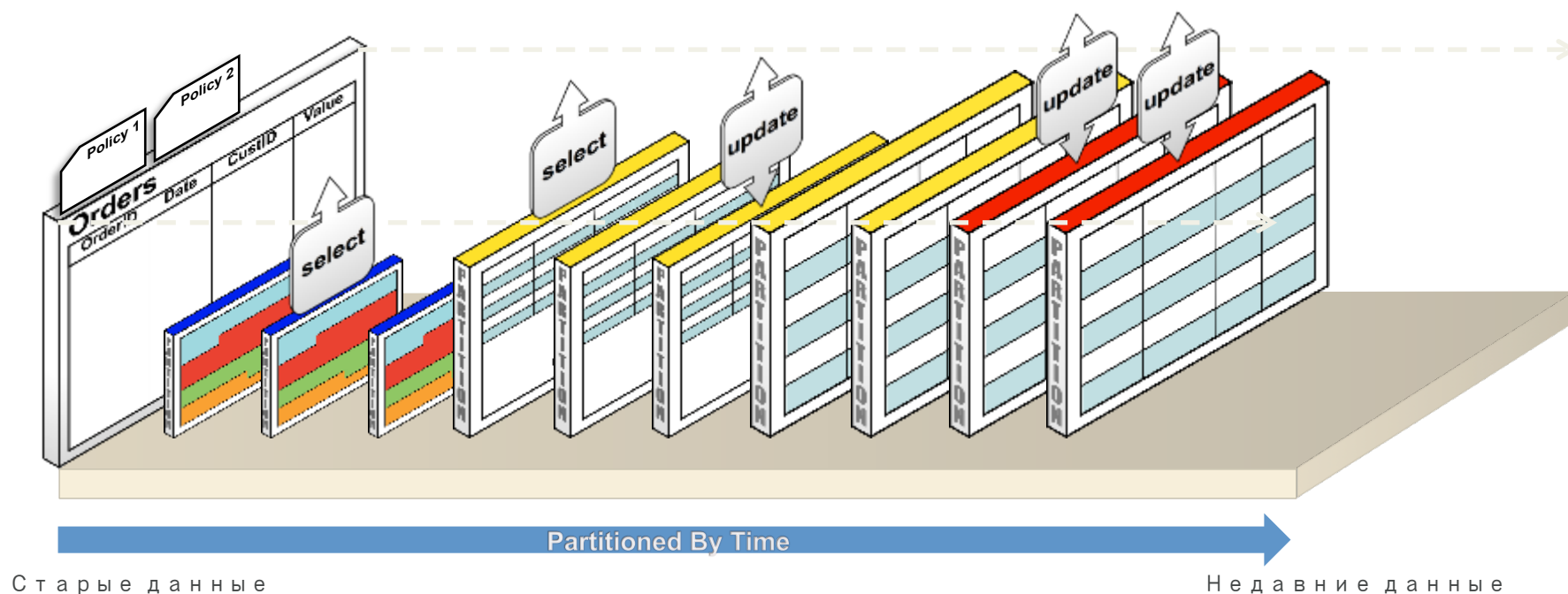
Политики автоматически применяются к таблицам



Автоматическая Оптимизация Данных



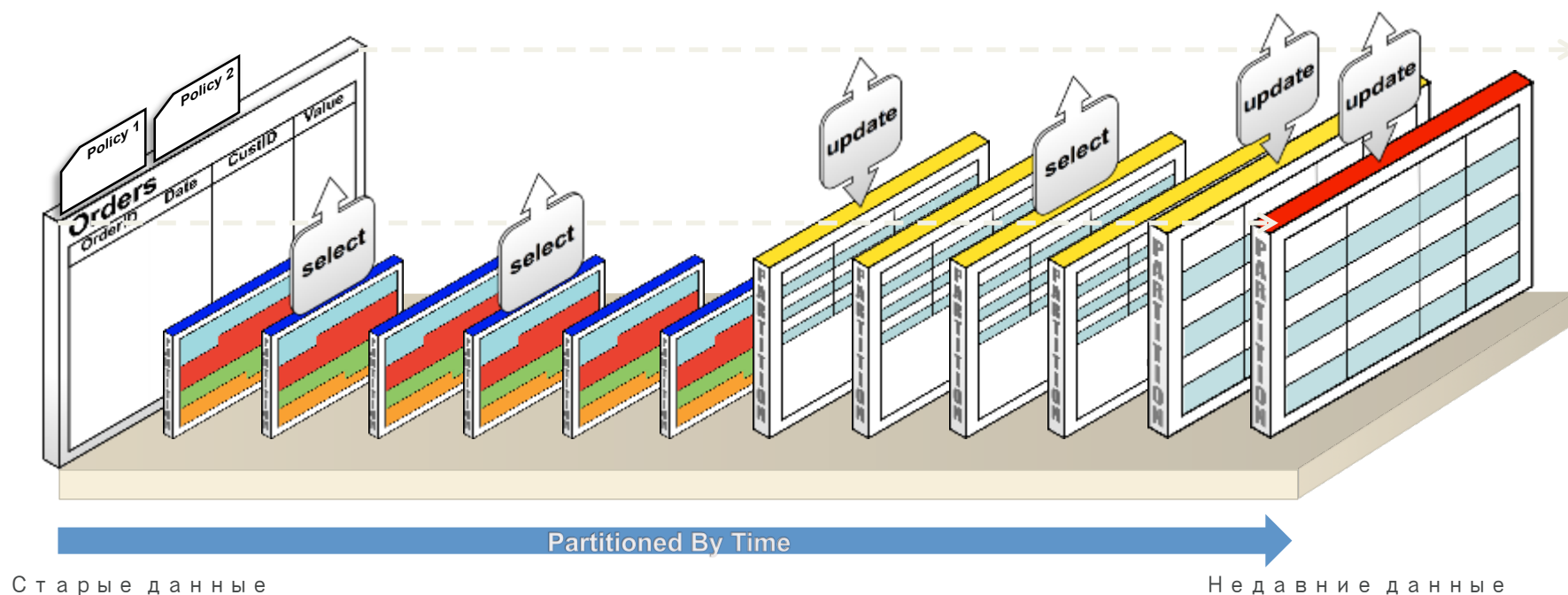
Политики автоматически применяются к таблицам



Автоматическая Оптимизация Данных



Уменьшается объем хранения, чтение сжатых данных выполняется быстрее



Автоматическая Оптимизация Данных



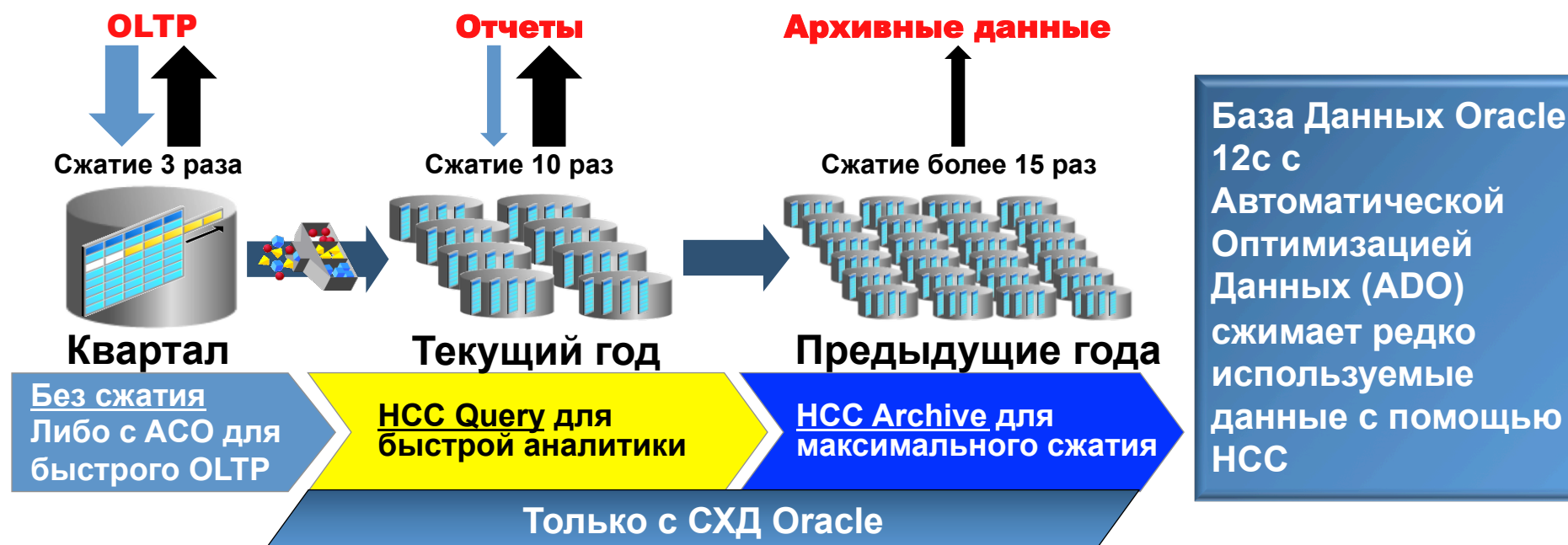
Автоматическое перемещение данных для снижения стоимости хранения



Только с СХД Oracle



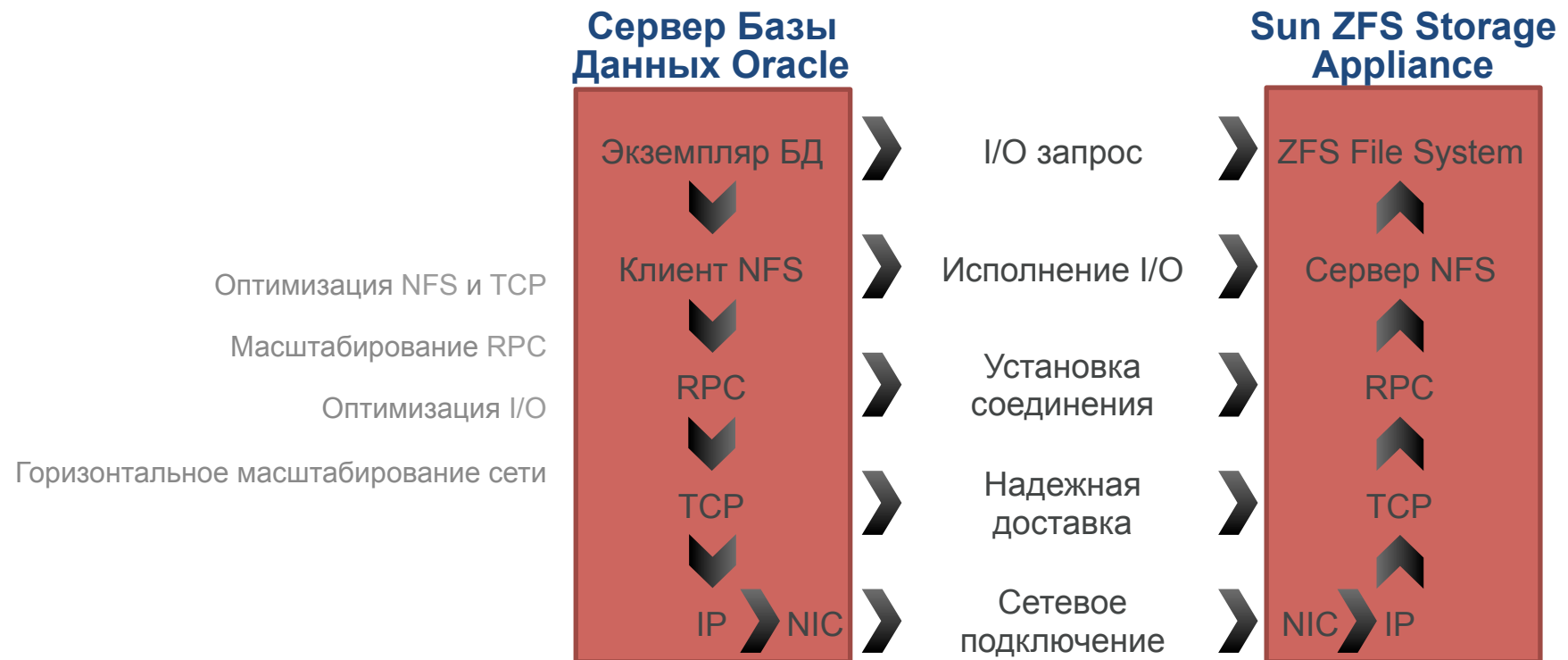
Автоматическая миграция строк и таблиц исходя из использования



Kernel NFS



Oracle dNFS: оптимизация I/O и повышение производительности



Oracle Direct NFS



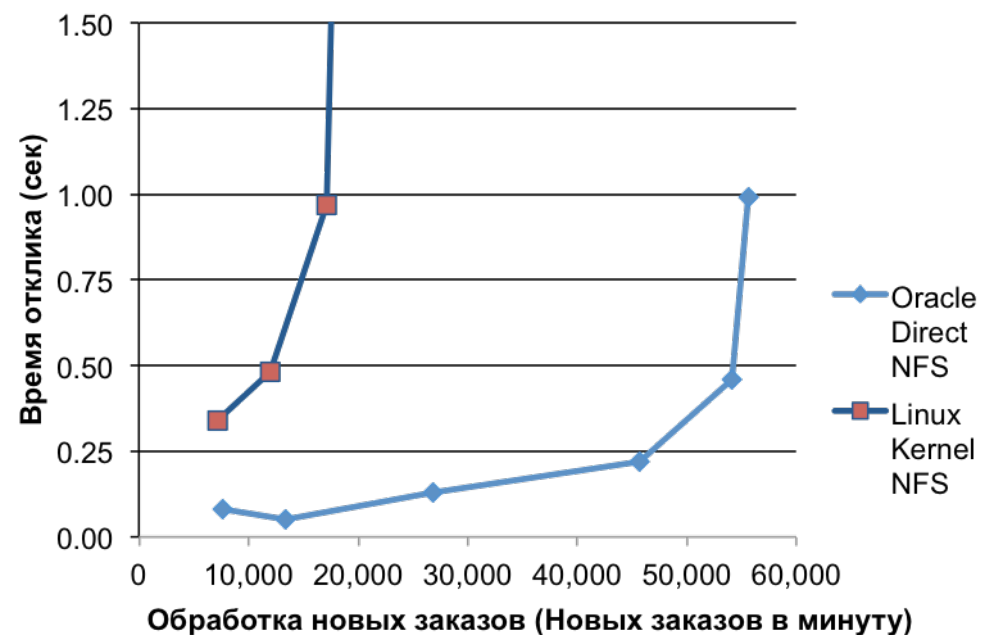
Oracle dNFS: оптимизация I/O и повышение производительности



Oracle Direct NFS: обработка OLTP



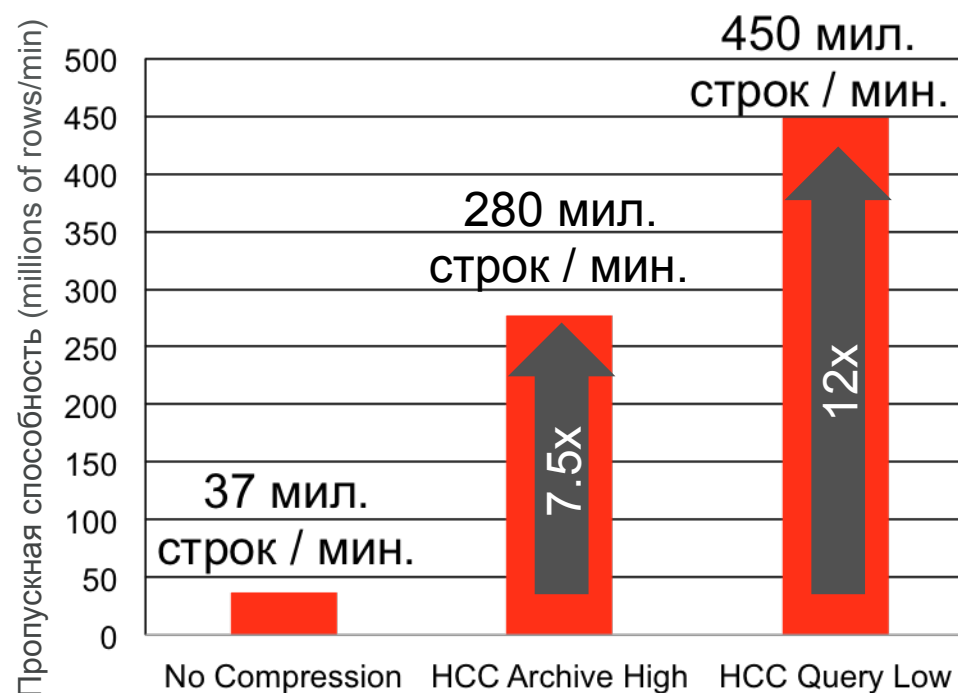
- В 3.2 раза выше пропускная способность при одинаковом времени отклика
- В 4.3 раза ниже время отклика при одинаковой пропускной способности
- В 2 раза выше эффективность использования CPU
- Узким местом становится приложение



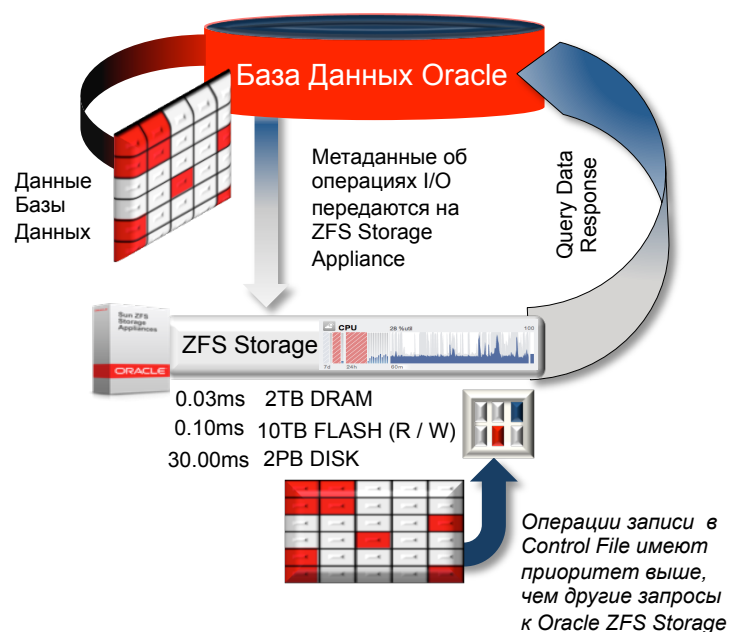
Пропускная способность с Oracle dNFS и HCC



- Без сжатия
 - Узкое место 1GbE
 - Базовый уровень
- HCC Query Low
 - Узкое место 1GbE
 - В 12х раз выше пропускная способность
- HCC Archive High
 - Узкое место CPU
 - В 7.5х раз выше пропускная способность



Oracle Intelligent Storage Protocol



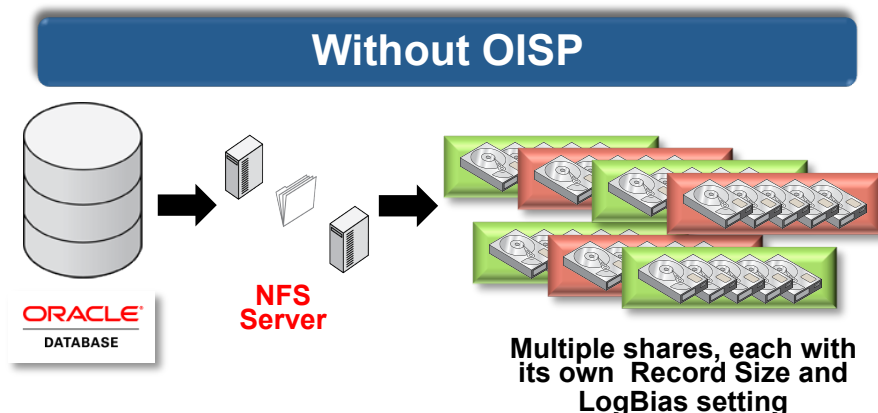
Oracle Intelligent Storage Protocol: Уникальный протокол, который активирует динамический обмен данными между Базой Данных Oracle и Oracle ZFS Storage.

Пример пересылаемых данных:

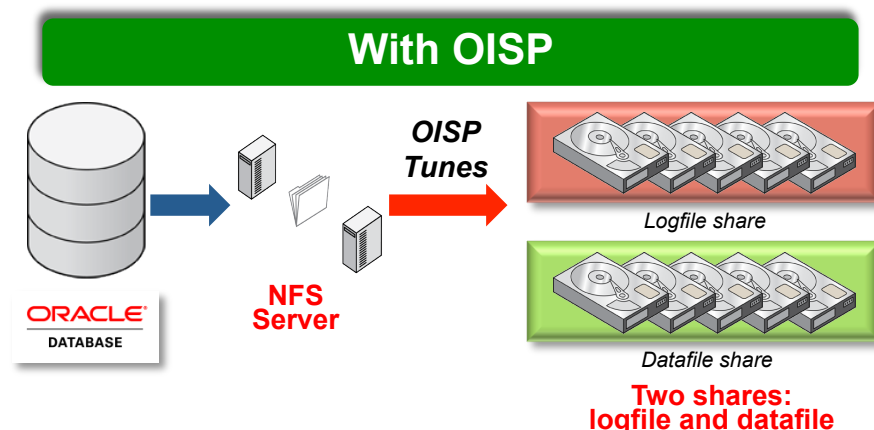
- Тип данных: таблицы, redo логи, RMAN копия, и т.д.
- Приоритет I/O

Oracle ZFS Storage динамически выделяет системные ресурсы для оптимизации производительности Базы Данных Oracle

Oracle Intelligent Storage Protocol



/mnt/dbname/redo (**Record Size, LogBias**)
/mnt/dbname/control (**Record Size, LogBias**)
/mnt/dbname/pfile (**Record Size, LogBias**)
/mnt/dbname/datafile (**Record Size, LogBias**)
/mnt/dbname/tempfile (**Record Size, LogBias**)
/mnt/dbname/chgtrack (**Record Size, LogBias**)
/mnt/dbname/backup (**Record Size, LogBias**)



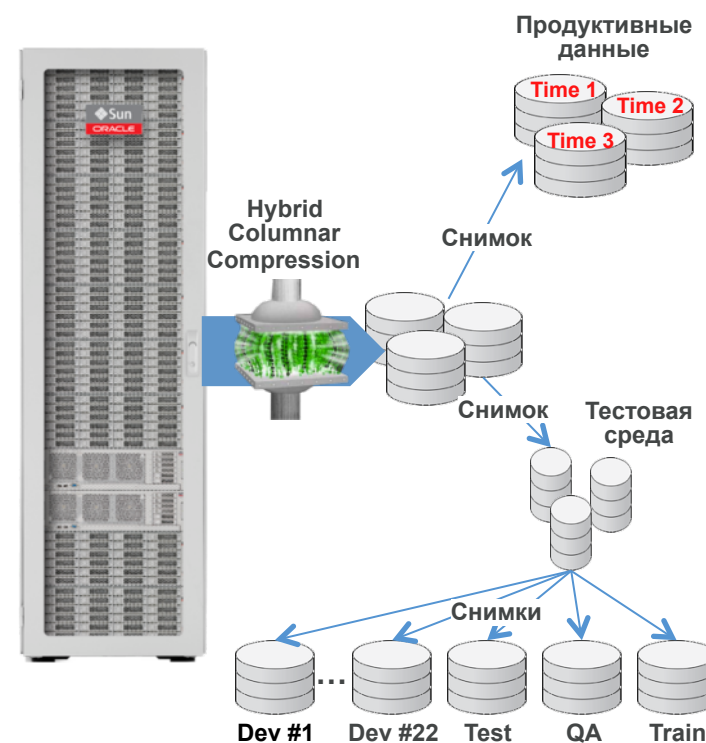
/mnt/dbname/logfile (**OISP sets Record Size, LogBias**)
 →redo
/mnt/dbname/datafile (**OISP sets Record Xize, LogBias**)
 →control
 →pfile
 →datafile
 →tempfile
 →chgtrack
 →backup

Снимки и клоны Oracle ZS SA



Быстрое, эффективное управление данными для Test/Dev

- **Быстро создавать** моментальные для защиты данных и разработки
- Создание резервных копий продуктивных данных и клонов для тестирования и разработки **на лету**
 - Создавать **больше тестовых** сред меньшем количестве систем
 - Тестовые среды **работают с высокой** производительностью благодаря Гибридным Пулам Хранения HSP
- Легко и быстро тестировать новые конфигурации и патчи **на полном объёме продуктивных данных**

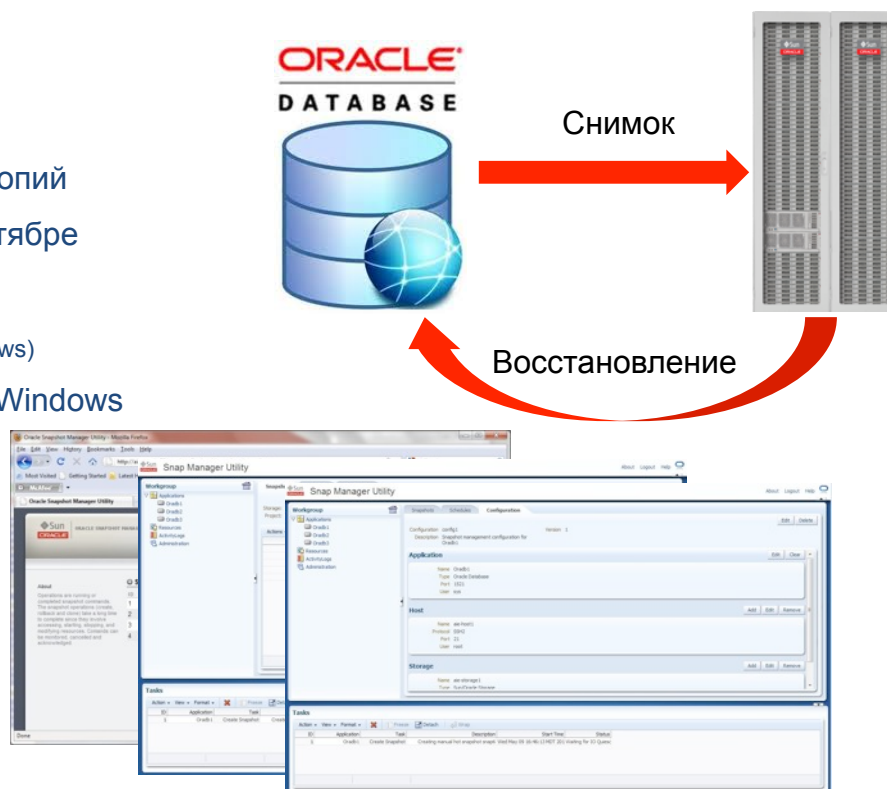


Snap Manager для БД Oracle



Позволяет администраторам легко создавать и управлять копиями Базы Данных Oracle с Oracle ZFS Storage

- Позволяет быстро и эффективно копировать и восстанавливать экземпляры Базы Данных
- “Application Aware” → обеспечивает целостность копий
- Предварительная версия, General Availability в Октябре
- Поддержка Oracle 10g, 11g, RAC
 - Поддержка протоколов NFS и CIFS (FC & iSCSI в будущем)
 - CLI версия для большинства платформ (Solaris, Linux, Windows)
- Поддержка серверов Баз Данных с Solaris, Linux, Windows
- Типы снимков
 - Горячие: Online
 - Холодные: Offline
 - Standby: DataGuard standby clone
- Управление снимками
 - Автоматизированный планировщик
 - Политики хранения
 - Хранение метаданных о снимке
- Восстановление снимка



Аналитика DTrace для ZS SA



Простое управление СХД для Базы Данных Oracle 12c

Автоматическая визуализация в реальном времени с глубокой детализацией

Понять специфику нагрузки, создаваемой Базой Данных и приложениями

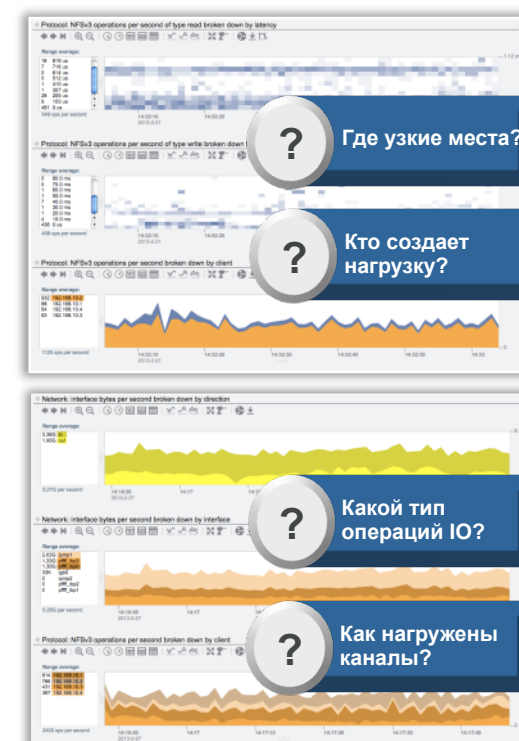
Примеры:

Утилизация системы: Легко обнаружить причину низкой производительности (бутылочное горлышко)

Производительность системы: Легко определить какие пользователи создают максимальную нагрузку

Настройка системы: Легко определить некорректные настройки (напр.: неправильный размер блока)

Балансировка нагрузки: Легко оптимизировать нагрузку благодаря визуализации нагрузки в реальном времени

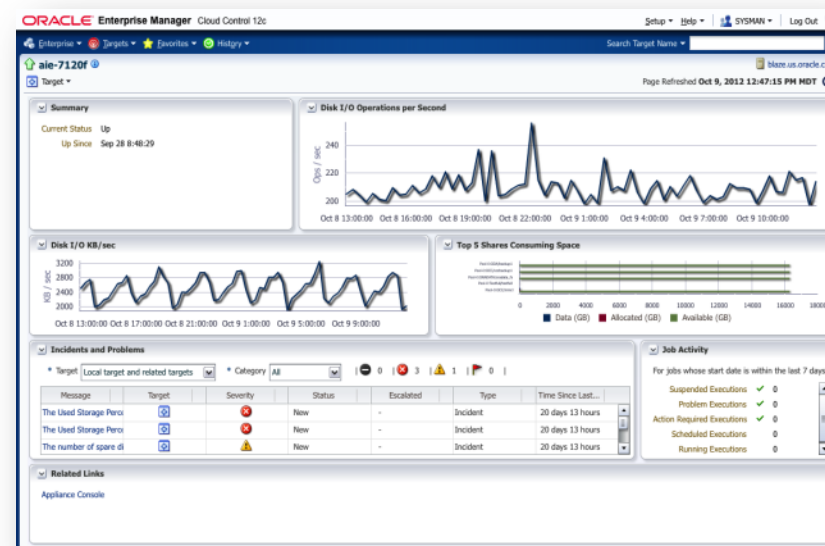


Интеграция с Oracle Enterprise Manager



Существенно снизьте сложность администрирования

- Plug-in для мониторинга и выявления проблем СХД
- Отображает предупреждения, метрики/аналитику, возникшие проблемы, отчеты
- Единая панель управления благодаря функции EM Group для множества систем и приложений
- Настраиваемый вид панели
- Поддержка EM 10g/11g & 12c
- Активно используется Oracle IT



Быстрая настройка



Быстрая настройка



Настройка СХД за минуты

Нет скрытых лицензий

Встроенный функционал копирования

Масштабирование до 2.6 PB

Интеграция с ПО Oracle

Упрощение управления



Простое управление



Автоматический тюнинг СХД

Аналитика производительности Базы Данных

Богатый программный функционал

Приложения для iPad / iPhone

Создана вместе с инженерными системами

Единая консоль управления

Эффективная разработка



Эффективные test/dev



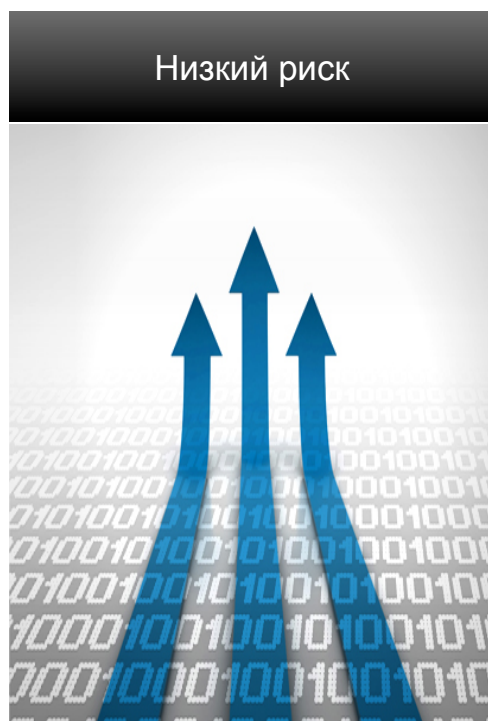
Быстрое создание неограниченного количества снимков и клонов продуктивных данных

Снижение стоимости дискового пространства для разработки

Увеличить производительность с HSP

Перенесите test/dev с Exadata на ZFS SA

Снижение рисков



Решения разрабатываются и тестируются вместе

Ниже риск при меньшем количестве систем

Совместимость с Oracle Exadata

Стабильная производительность

Ниже операционные расходы

Простое внедрение и развертывание решений



Быстрая настройка



Простое управление



Эффективные test/dev



Низкий риск



Спасибо за внимание! Вопросы?



Басов Андрей

Ведущий технический специалист, направление Oracle,

- Телефон: +7 (812) 326-3232 доб. 1077
- E-mail: basov@marvel.ru

Общий внешний адрес по вопросам Oracle: oracle@marvel.ru



ZS SA controller options



	ZS3-2	ZS3-4	ZS3-BA
CPU Cores	2 X 8 core	4x10 core	4x10 core
DRAM	256 GB	1 TB	1 TB
Max Read Flash	4 SSDs	4 SSDs	Base units ship with 1 per controller; allow up to 4 through x-options
Max Write Flash	64 SSDs*	144 SSDs	144 (HC) or 132 (HP)
HA Option	Yes	Yes	HC and HP systems come clustered
# Data HDDs in-head	0	0	0
Max # Disk Enclosures	16*	36	36
Max # Drives	384*	864	860
# Open PCIe Slots	4	7	7
PCIe Version	3.0	2.0	2.0
Form Factor	2U (4U HA)	3U (6U HA)	6U

* ZS3-2 specs at release are half that indicated (8 DE, 32 Write Flash, etc.)

Oracle ZS SA



Data protocols

- Fibre channel
- iSCSI
- InfiniBand over IP/RDMA
- iSER
- SRP
- NFS V3 and V4
- CIFS
- HTTP
- WebDAV
- FTP/SFTP/FTPS
- ZFS NDMP V4

Data services

- OISP
- Hybrid Columnar Compression
- Hybrid Storage Pools
- Single, double and triple-parity RAID (RAIDZ, RAIDZ2, RAIDZ3)
- Mirroring and triple mirroring
- End-to-end data integrity
- Local and remote replication
- Snapshots and clones
- Quota(s)
- In-line deduplication
- Compression
- Thin provisioning
- Antivirus via ICAP protocol
- Online data migration
- Clustering

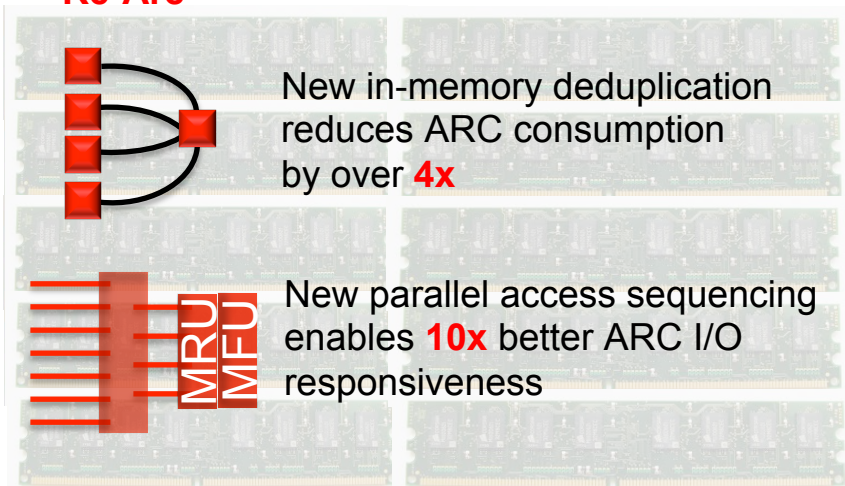
Management

- Browser and CLI interface
- Management dashboard
- Hardware/component view
- Role-based access control
- Phone home
- Event and threshold based alerting
- Storage analytics
- Scripting
- Workflow automation
- Advanced networking
- Snap Management Utility
- Source aware routing

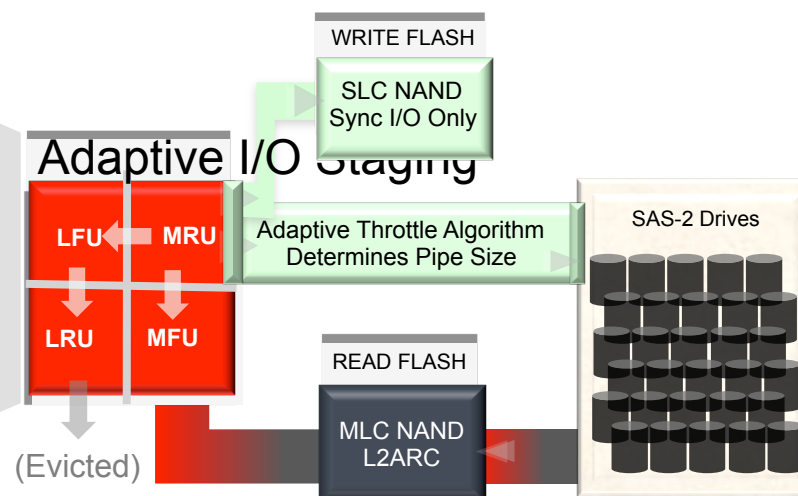
Oracle ZS SA



Re-Arc



Increases system performance by up to 2x
(using the same hardware)



Automated, real-time (sub-LUN or share) data migration from
DRAM to multi-class flash, to multi-class disk storage

Oracle Intelligent Storage Protocol

- Oracle Disk Manager (ODM) hints are passed ...
- ... via NFS through Oracle Direct NFS ...
- ... to NFS Server and ZFS File System

