

Построение контура хранения данных в инфраструктуре суперкомпьютера

на отечественной аппаратной
и программной платформе

TROK × Спинтроника × Байкал Электроникс



Почему результаты вычислений нужно хранить надежно

Потеря результата означает повторное использование вычислительных ресурсов. Иногда исходные условия или эксперимент невозможно воспроизвести

1

задача запускается
на сотнях или тысячах
ядер

1000+

часов могут стоить
подготовка модели
и расчет

0

допустимых потерь для
уникального результата



Поэтому система хранения для НРС должна сохранять данные при отказах и давать предсказуемый доступ к ним

Четыре слоя формируют рабочий контур хранения

В исследовании оценивается
совместная работа стека, включая
передачу данных по NFS
и хранение данных в SDS TROK

Вычислительная
архитектура

Baikal-S

ARM-процессор



Проектирование и
сборка

Спинтроника

Серверная
платформа



Драйверы, сервисы,
управление

Astra Linux

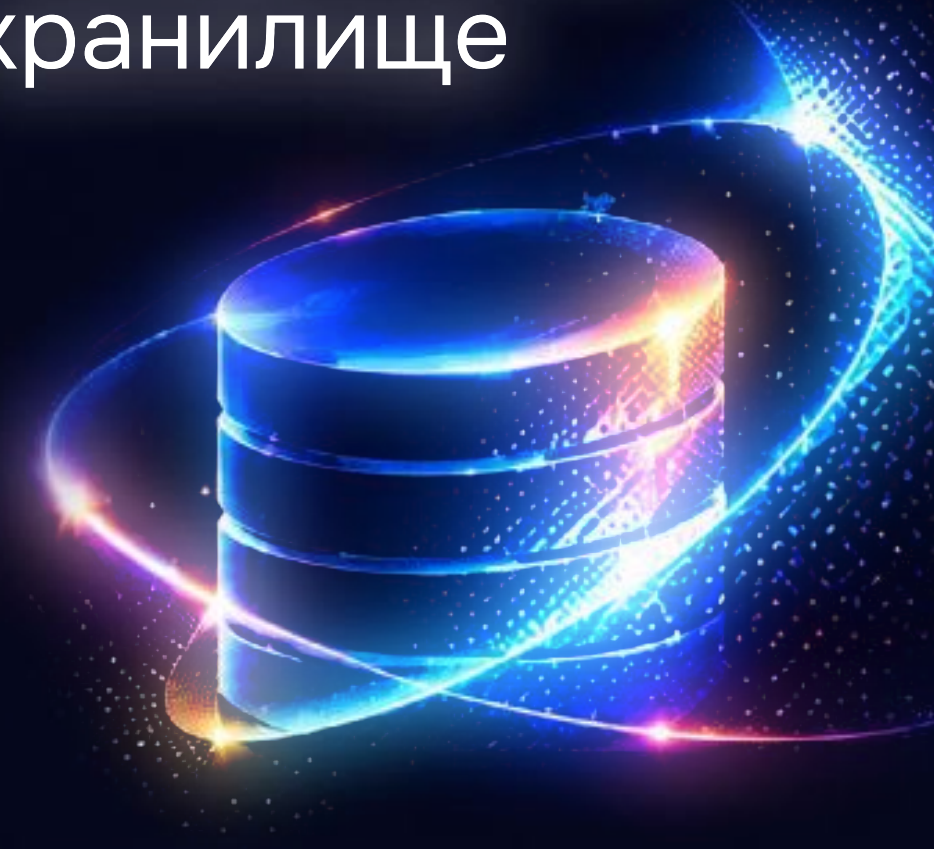
ОС для ARM



Данные, доступ,
отказоустойчивость

TROK

Программно
определяемое
хранилище



Программная архитектура

Вычислительный кластер

Метеорологические расчеты

Геодезические расчеты

Узлы хранения TROK

TROK-worker

Drbd-reactor
resource agent

targetcli

nfs-server

NVME-oF

TROK S3 GW

Блочное устройство DRBD

Drbd-utils

LVM/LVMThin/ZFS

Kmod-drbd9

TROK-harbor



Слой доступа к данным

DRBD Diskless

iSCSI

NFS

NVME-oF

S3

Узлы управления TROK (HA)

TROK-controller-api

TROK-server

TROK-common

TROK-server DB

TROK Metadata

TROK WEBUI

TROK-harbor-api



Аппаратное обеспечение



Процессоры
Baikal-S



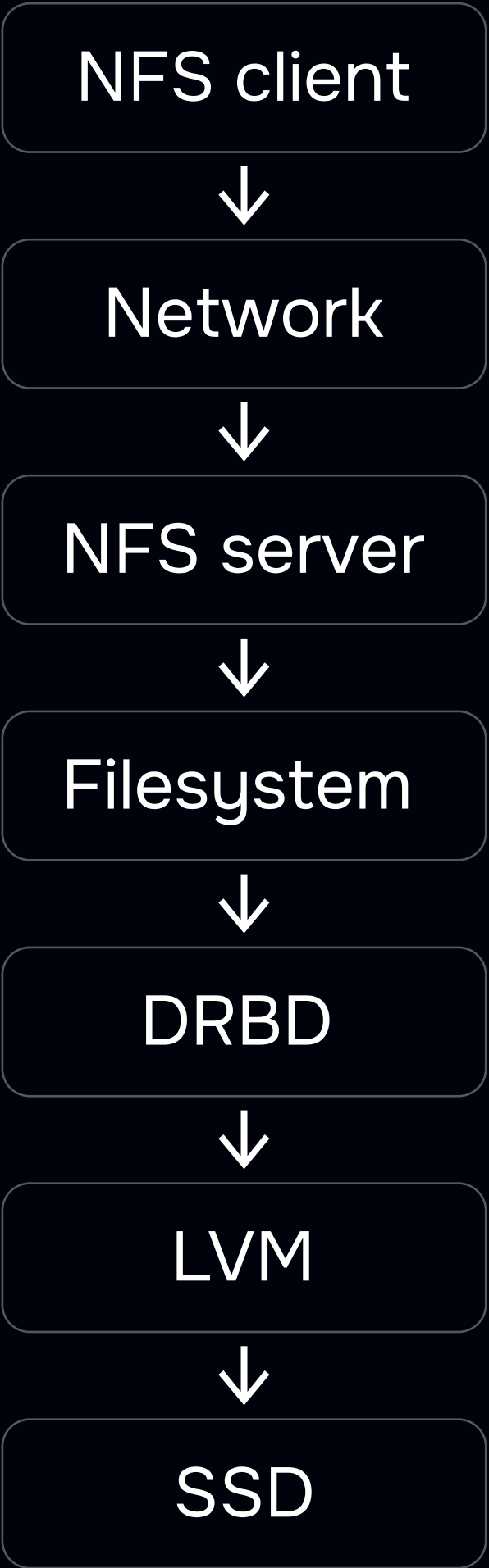
Серверы
Спинтроника



Тестовый стенд

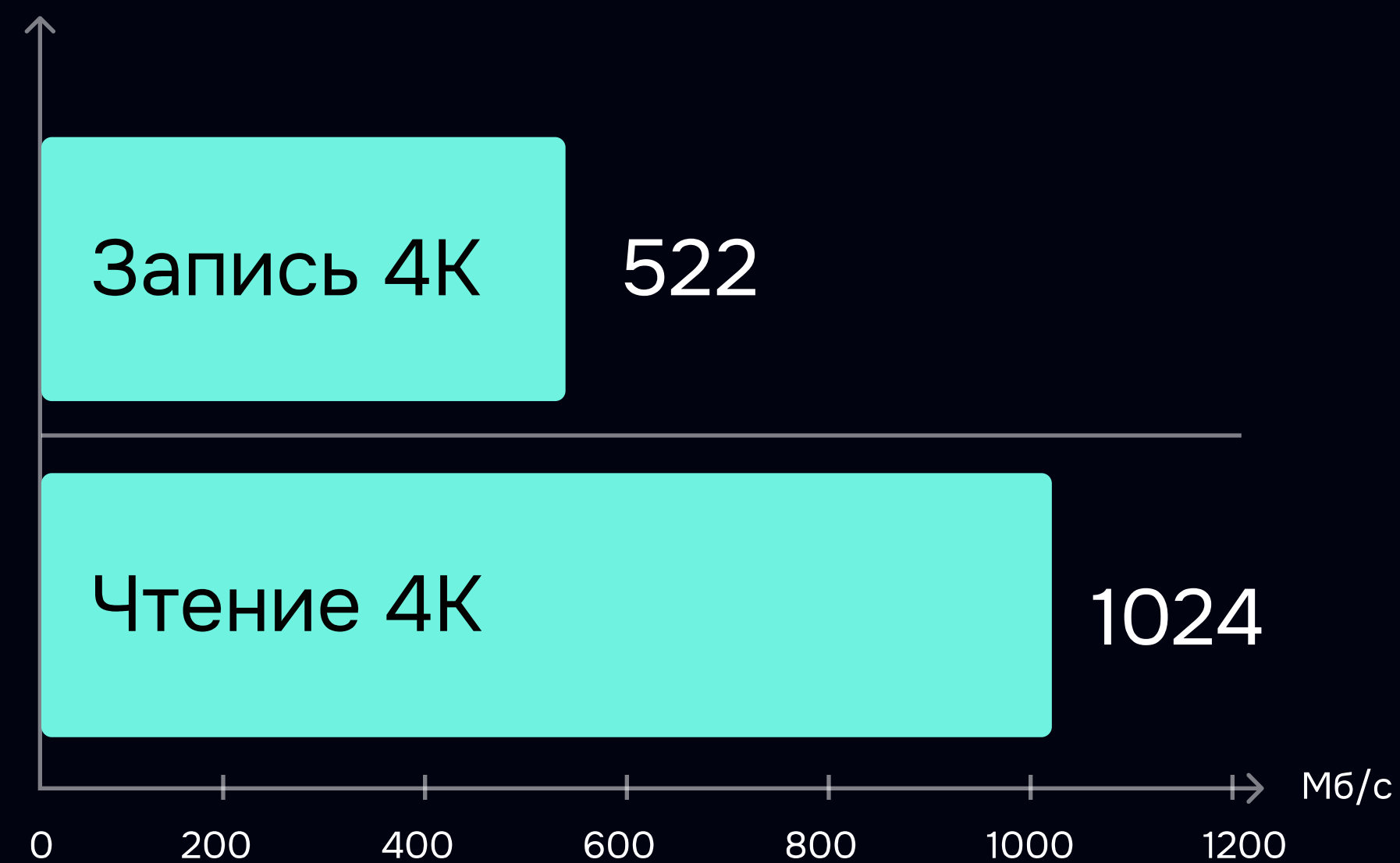
Параметр	Описание
Узлы хранения	6
Протокол	NFS 4.2 / TCP
Сеть	25 Gbit/s
Backend	12 × enterprise SSD, 15 TB каждый
Агрегация	LVM linear, около 28 TB, без striping
Генератор	FIO 3.28, Direct I/O
Тестовый файл	100 GiB

Путь данных



NFS обеспечивает около 1 GiB/s на чтение и до 37,3 тыс. IOPS

Последовательные операции



0,121 ms

Random Read 4K, QD1

Случайные операции 4к,
4 × qd32



0,165 ms

Random Write 4K, QD1



Полный набор
измерений
показывает профиль
производительности

Queue	Профиль	Результат	p99
Backend SSD, Seq Read 1M	QD32	≈ 1,08 GiB/s	—
NFS Seq Read 1M	QD32	≈ 1,0 GiB/s*	28 ms
NFS Seq Write 1M	QD32	522 MiB/s	87 ms
NFS Random Read 4K	4 × QD32	37,3k IOPS	4,3 ms
NFS Random Write 4K	4 × QD32	23,3k IOPS	37 ms
NFS Mixed 4K 70/30	4 × QD32	≈ 31k IOPS	6 ms
NFS Random Read 4K	QD1	9,28k IOPS	0,121 ms
NFS Random Write 4K	QD1	8,12k IOPS	0,165 ms

Отечественный стек технологий хранения показал следующие результаты

≈ 1 GiB/s

последовательное
чтение по NFS

37,3k

IOPS на случайном
чтении 4K

0,121 ms

базовая latency чтения
при QD1

Хранилище для СУБД

Высокая производительность и отказоустойчивость делают TROK надёжной платформой для работы с базами данных в критичных приложениях



Хранилище для высоконагруженных приложений и виртуализации

Поддержка NVMe-oF и быстрая репликация данных обеспечивают мгновенный отклик даже при высоких нагрузках



S3-совместимое хранилище для облачных сервисов

Используйте TROK как объектное хранилище S3 для облачных платформ и внутренних сервисов



Хранение резервных копий

Поддержка моментальных снимков и отказоустойчивость обеспечивают надёжную защиту данных и быстрый доступ к резервным копиям



Сценарии применения TROK

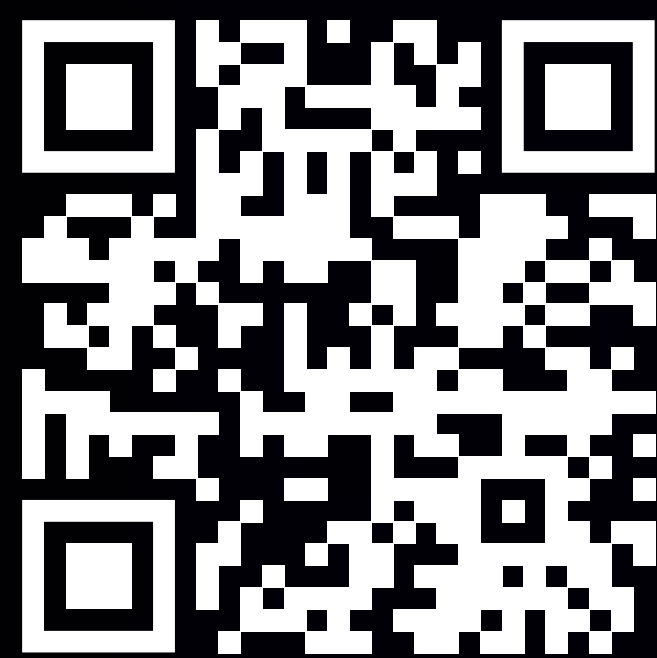
Географически распределённые системы

Встроенная репликация между узлами в разных дата-центрах позволяет строить катастрофоустойчивые решения

Спасибо за внимание



trok-sds.ru



spintronica.ru



baikalelectronics.ru

TROK × Спинтроника × Байкал Электроникс

