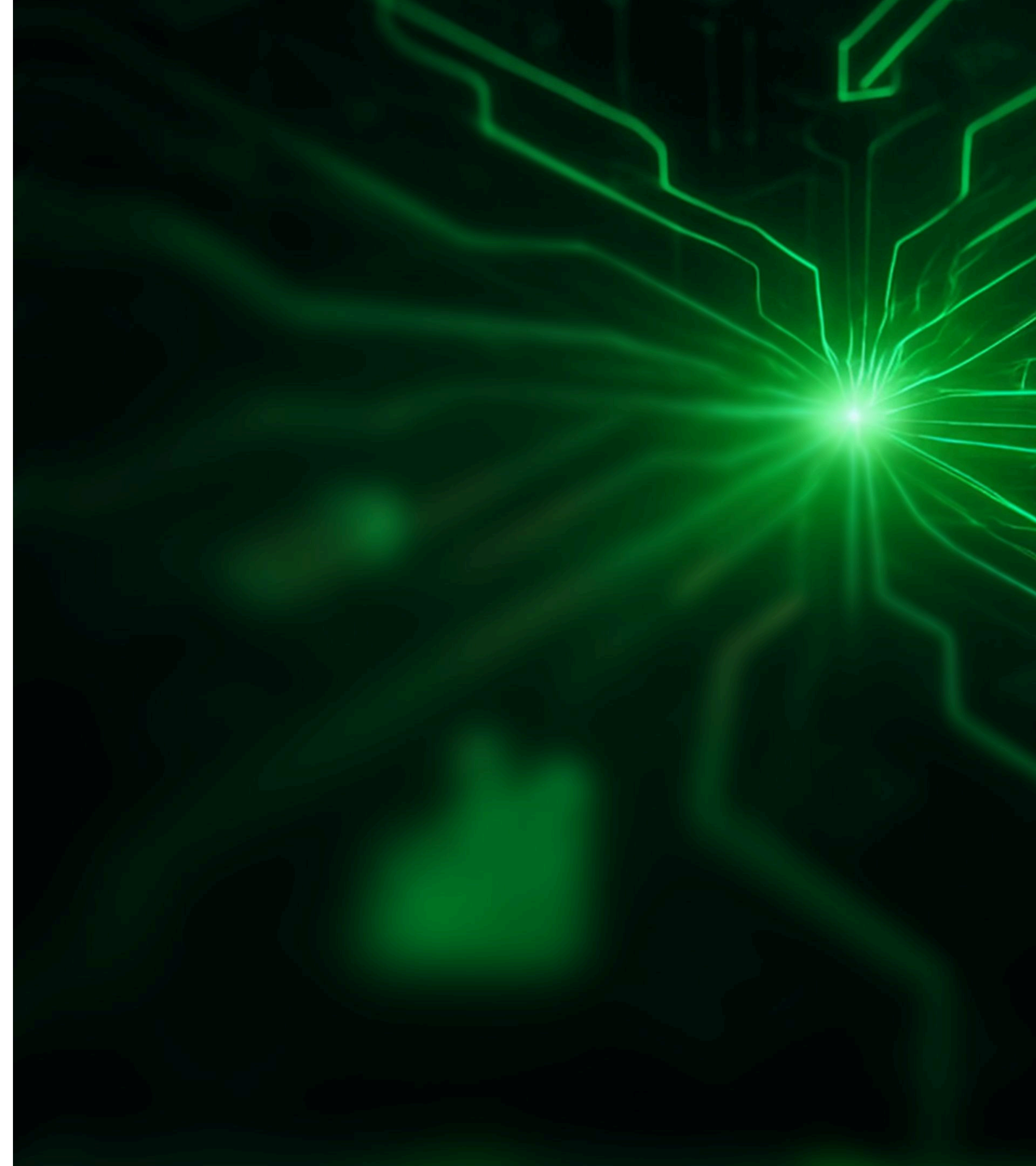


BITBLAZE

BlazeData

Новейшее решение для распределенного хранения
больших и очень больших данных



О компании

Первые в России

2009 год

Год основания компании

BITBLAZE



портировали систему хранения данных с x86 на отечественную платформу E2K (Эльбрус)

СХД + эльбрус



разработали распределенную файловую систему хранения данных для отечественной платформы E2K (Эльбрус)

KFS + эльбрус



разработали кроссплатформенную отказоустойчивую СХД на All-flash NVMe и внесли в реестр МПТ ПП 719/878

СХД + NVMe



x86

bitblaze.ru

Проблемы существующих файловых СХД

Ограниченная производительность с узла (вертикальное масштабирование)

- Сервер на PCIe Gen4 + 24 NVMe SSD способен на 40 ГБ/с и 10 млн. IOPS (по протоколу NVMe-oF)
- XFS локально: 3-4 млн. файловых операций/сек
- NFS: максимум ~10 ГБ/с и 100–200 тыс. IOPS, что снижает эффективность системы
- Lustre, BeeGFS, DAOS: пропускная способность выше 40 ГБ/с, но до несколько сотен тысяч операций/сек

Пример

Несколько лет назад возникла задача обучения ИИ, при которой GPU-ферма требовала для полной загрузки порядка **13 млн. операций** «открыть–прочитать–закрыть» небольших файлов в секунду. Лучшее на тот момент доступное решение для хранения данных обеспечивало лишь около **120 тыс. таких операций с одного узла**. Чтобы удовлетворить требования, потребовалось бы более **100 узлов СХД** – оказалось крайне дорого и невыгодно, учитывая, что сама GPU-ферма стоила значительно меньше.

Файловые системы

Проблемы существующих файловых СХД

Только параллельные файловые системы хранения данных (Lustre, BeeGFS, DAOS и другие) обеспечивают настоящее горизонтальное масштабирование

Однако они отличаются:

- Высокой сложностью настройки и сопровождения
- Повышенным порогом входа и общей стоимостью эксплуатации

BlazeData

Впервые файловое хранилище стало сопоставимо по скорости с блочным

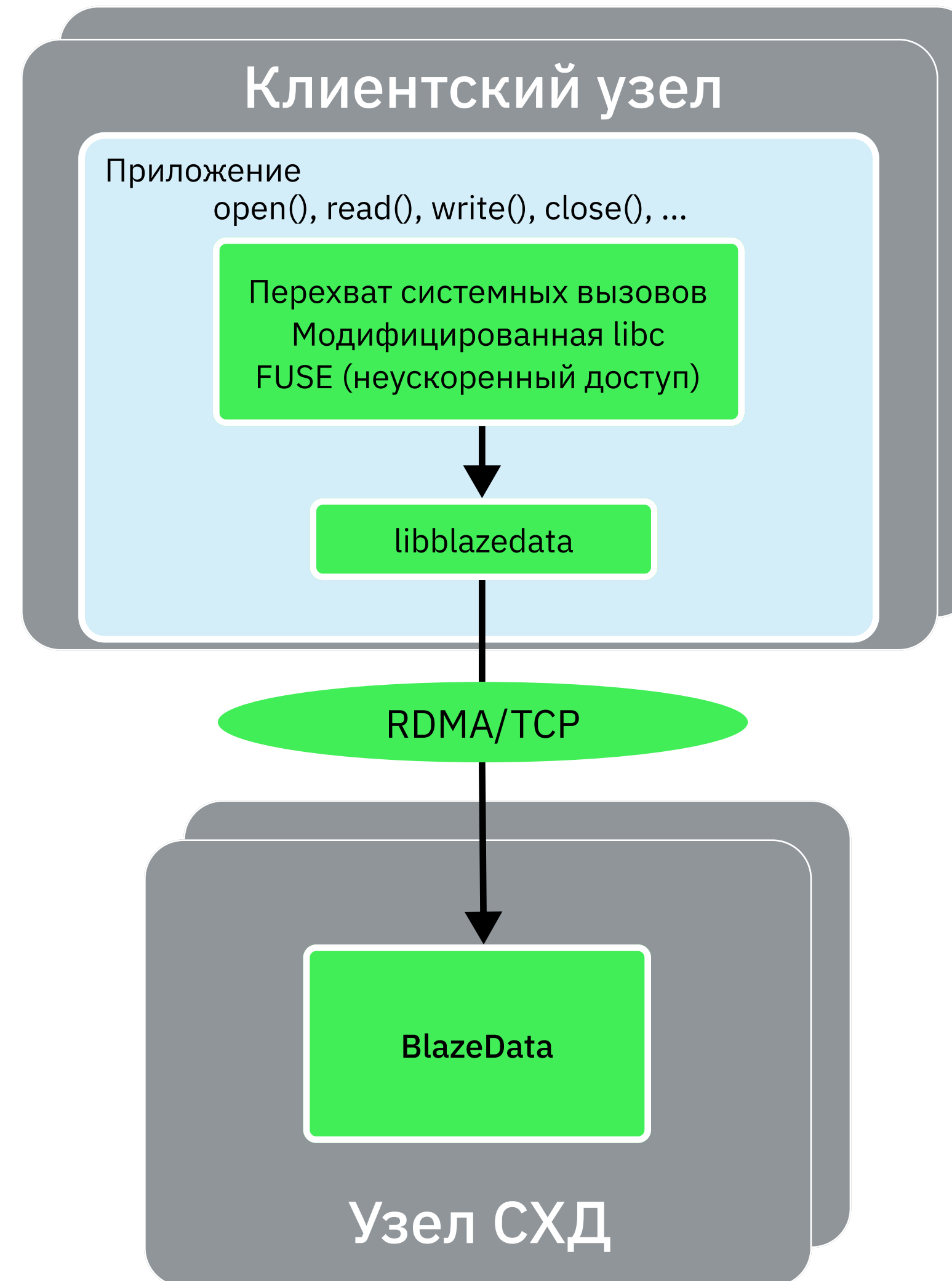
Цель

- Получить производительность локальной файловой системы каждого серверного узла на сетевых клиентах
- До 2 миллионов файловых операций/секунду (открыть—прочитать—заккрыть) с сервера среднего класса и выше
- Максимальная утилизация имеющихся аппаратных ресурсов недорогих серверов
- Минимизация стоимости СХД при требуемом уровне производительности
- Можно приобрести больше GPU ресурсов при том же бюджете
- Масштабирование по производительности и объему
- Максимальная простота использования
- Минимальная стоимость владения

BlazeData

**Впервые файловое
хранилище стало
сопоставимо по
скорости с блочным**

BITBLAZE



💡 Решение

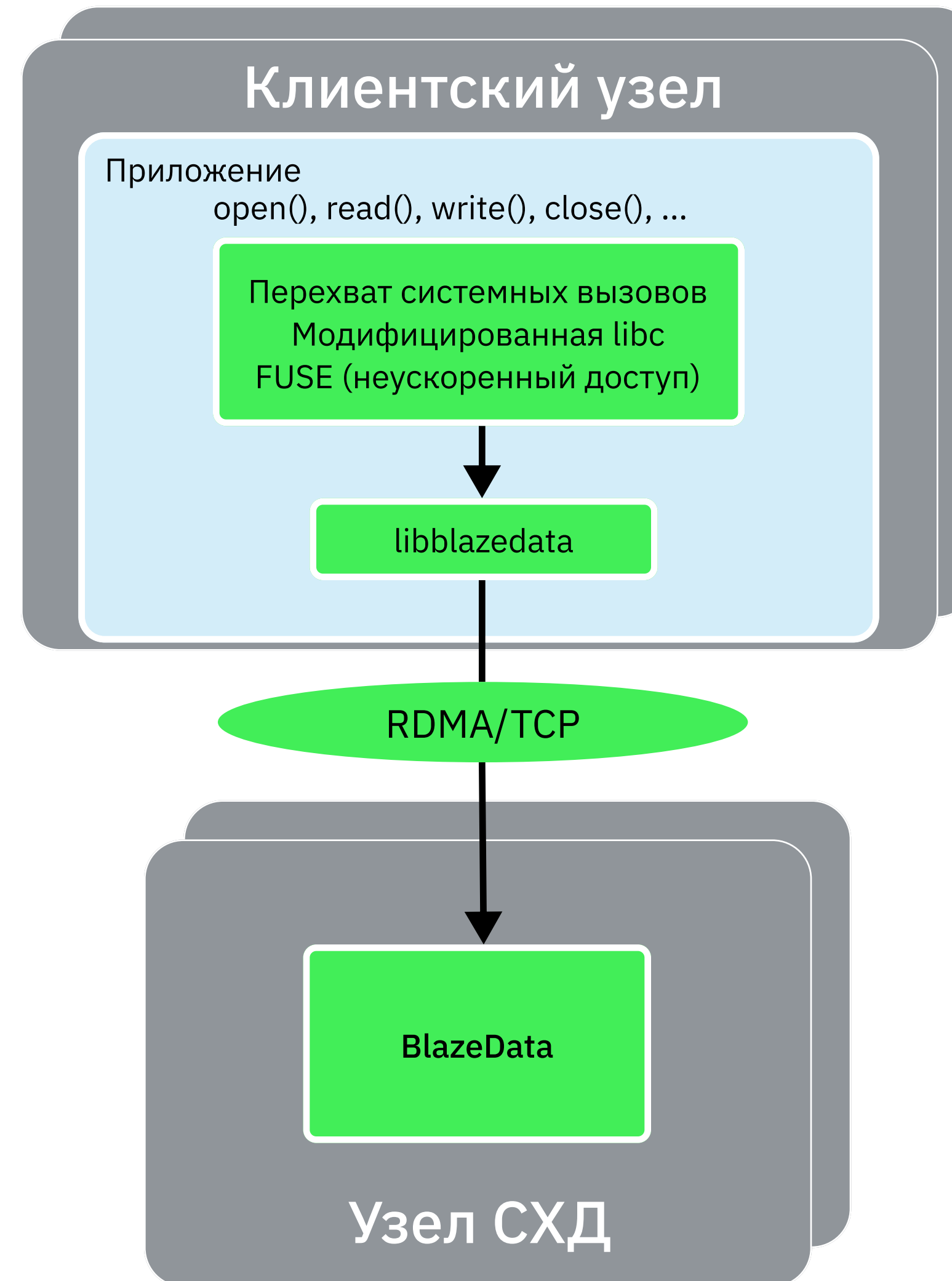
- Новый легковесный высокоэффективный протокол передачи данных
- Клиентское приложение работает в обход всего VFS стека ядра на клиенте
 - Перехватываются системные вызовы (`open`, `close`, `read`, `write`, `create` и т.д.)
 - Вся клиентская обработка происходит в библиотеке `libblazedata`
 - Приложения напрямую связываются с ней

bitblaze.ru

BlazeData

**Впервые файловое
хранилище стало
сопоставимо по
скорости с блочным**

BITBLAZE



💡 Решение

- Файлы разбиваются на части и размещаются на различных серверах
 - Размещение по математическому закону – сервер метаданных не требуется
 - Файл читается сразу с нужного сервера
 - Защита от отказов
- Сервера и клиенты сами находят друг друга в сети и объединяются в кластер
 - Встроенный протокол самоконфигурирования

bitblaze.ru

Доступная скорость и масштабирование

Пример со слайда 2

Для обеспечения 13 млн. операций/сек
Обычная файловая система: 100 серверов
BLAZEDATA: 10 серверов

Результат

- Очень тонкий клиентский стек напрямую внутри приложений
 - Низкая задержка, близкая к аппаратной
 - Полный параллелизм и множество очередей (Multi-Queue)
- Единое файловое пространство имен с множества серверов
- Масштабирование до тысяч серверов и клиентов
- Миллионы файловых операций в секунду и десятки ГБ/с с сервера среднего класса и миллиарды IOPS и десятки ТБ/с с кластера,
- Простота настройки (самоконфигурирование)
 - Сервера и клиенты сами находят друг друга
 - Ручное конфигурирование не требуется: включаешь — работает
 - Прозрачное добавление новых серверов
 - Минимизирует требования к количеству и квалификации обслуживающего персонала
- Минимальная цена владения
 - Поставляется в составе ПАК
 - Минимальная поставка — 1 сервер (2 в отказоустойчивой конфигурации)
 - Планируется включение клиентской части в ведущие российские дистрибутивы Linux (получено предварительное согласие)

Преимущества

ⓘ Доступен стенд для тестирования ваших задач

Дешевле

- Минимальная стоимость производительности – Экономия по сравнению с существующими решениями до 10 раз
- Минимальная квалификация обслуживающего персонала
- Минимальное количество обслуживающего персонала

Быстрее

- Больше производительности как с серверов средней ценовой категории, так и с высокопроизводительных серверов
- Суммарная производительность многих серверов

Проще

- Быстрота развертывания и подключения
- Прозрачное наращивание емкости и производительности
- Минимальная квалификация для обслуживания

Дорожная карта BlazeData

Масштабируемая параллельная файловая СХД
с инновационными возможностями
самоконфигурирования (кроссплатформенное ПО)

Ключевые характеристики

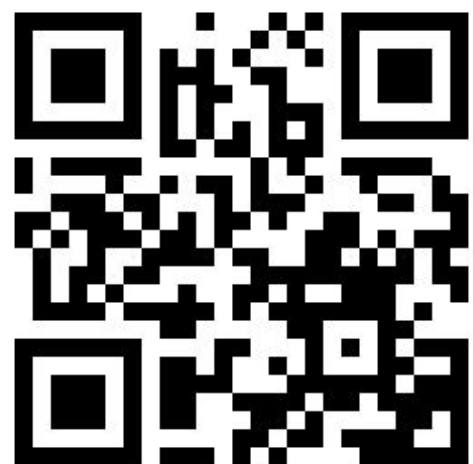


Наши контакты

 info@bitblaze.ru

 bitblaze.ru

 +7 3812 36-11-11



Сайт
BITBLAZE



Телеграм
BITBLAZE

BITBLAZE