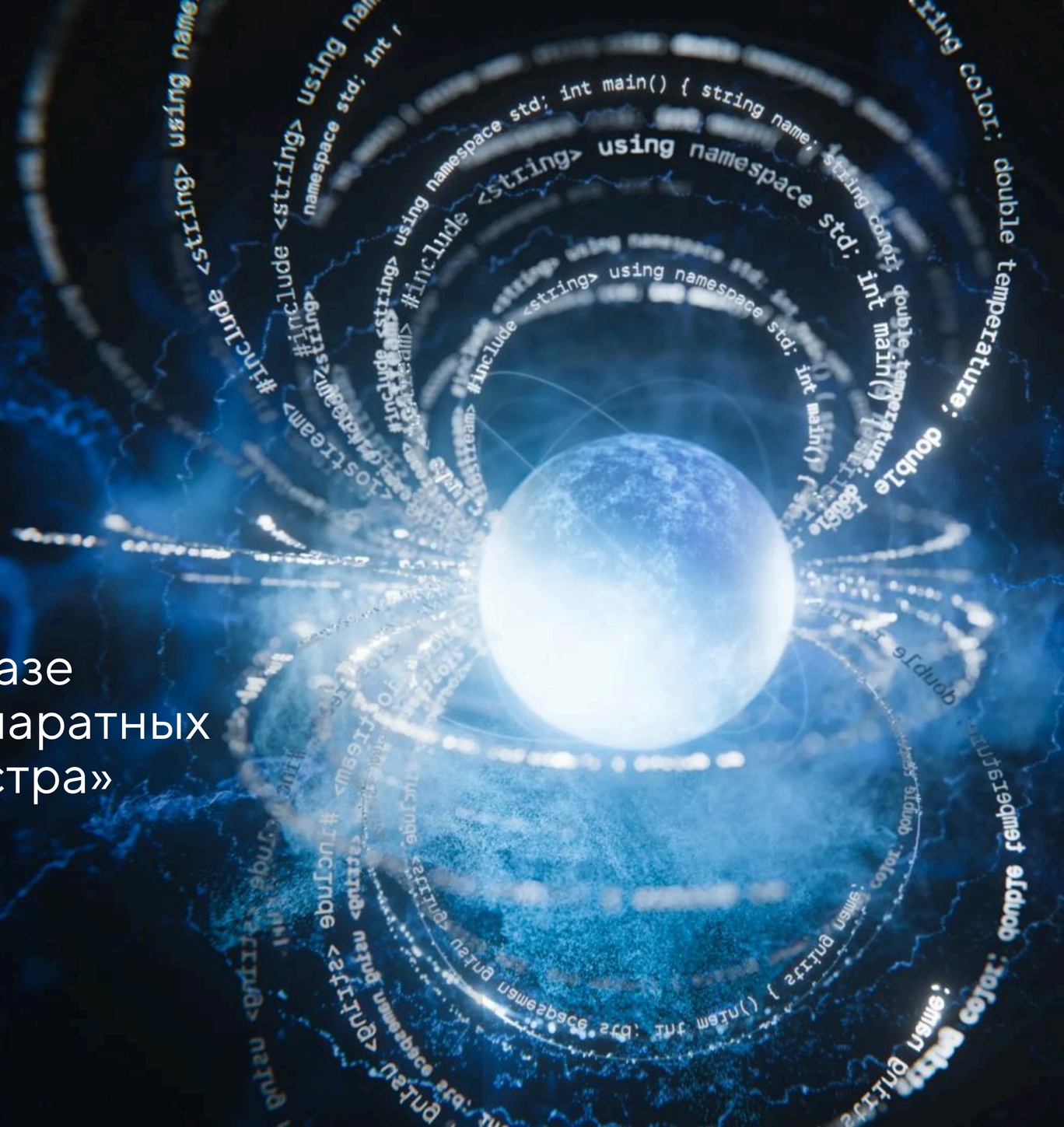
АСТРА

Инфраструктура «под ключ» на базе единой линейки программно-аппаратных комплексов ПАКов от «Группы Астра»

Руководитель направления XPlatform
Никита Кораблев



АСТРАКОНФ

СВЕРХНОВАЯ

Продуктом владеет тот,
кто контролирует развитие
его основного функционала

Что является ядром ПАК?

Лидеры рынка — софтверные компании

IBM
NUTANIX
ORACLE

Развивают собственную программную функциональность и используют аппаратную платформу как средство ее поставки

Программное обеспечение

Задает архитектуру, логику и ключевую функциональность решения

Оборудование

Создает базовый слой функционирования, но не определяет функции и предельные возможности

Софт определяет жизненный цикл ПАК

Интегратор - не владеет софтом и жизненным циклом

Производитель, который не контролирует ключевое ПО и его R&D, интегрирует чужие компоненты, а не определяет эволюцию решения


АСТРА
ASTRA
XPLATFORM

ПАК: IT Appliance

Appliance —
прибор, устройство
для конкретной задачи

Пример

- телефон
- NGFW
- телевизор
- СХД



Свойства

- Гарантированный результат
- Простота эксплуатации
- Не обязательно разбираться в деталях функционирования

СОФТ
+
СЕРВЕР
< ПАК

IT Appliance

ПАК

Целостная система
для критичных нагрузок

- Повышенная производительность
- Снижение TCO
- Единый SLA

ASTRA
XPLATFORM

АСТРА

Единая линейка ПАКов

01

Appliance —
устройства
с конкретной узкой
задачей

02

Комплексный ПАК
для построения
ИТ-инфраструктуры



Appliance — устройства с конкретной узкой задачей

XConnect

Аппаратный контроллер доставки приложений



XTime

Архивное хранение данных



XGenAI

Платформа AI-инфраструктуры



XData

Машина баз данных

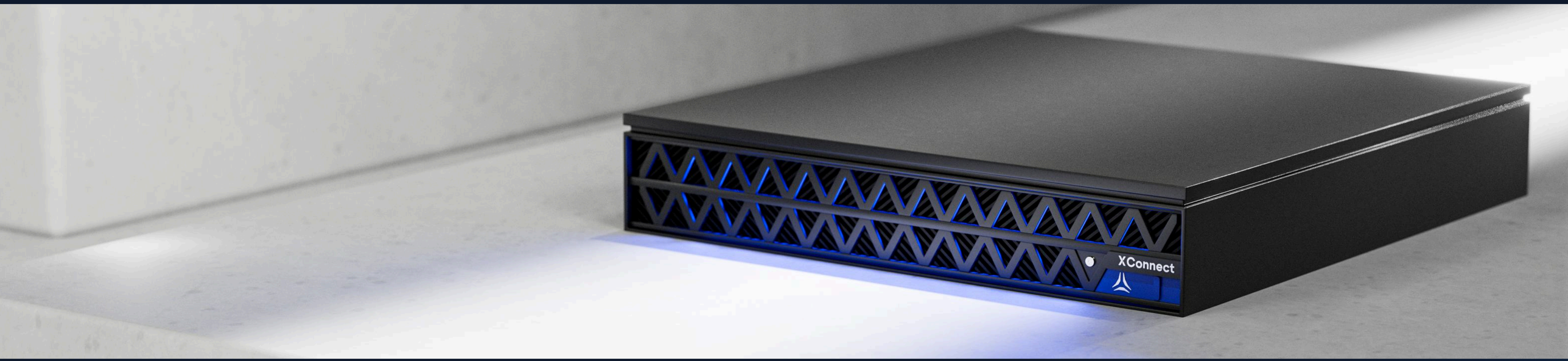


XPlatform заменяет западные appliance

Защищает вашу инфраструктуру от рисков остановки

Аппаратный контроллер доставки приложений

XConnect | АСТРА



Примеры применения

- Балансировка нагрузки и оптимизация доставки контента
- Построение точки балансировки трафика для обеспечения Disaster Recovery
- Распределение нагрузки между серверами приложений для быстрого и надежного подключения пользователей
- Снижение сетевой нагрузки за счет перенаправления пользователей на ближайшую площадку

Преимущества

Аппаратное SSL-ускорение

Глобальная балансировка нагрузки (GSLB)

По 4 интерфейса 100, 25 и 1 Гбит

Поддержка отказоустойчивовго кластера

Работает в сложных сетевых топологиях

благодаря Virtual Routing Forwarding, VLAN-протоколам динамической маршрутизации и агрегации каналов

ПАК архивного хранения данных



Примеры применения

Дешевое архивное хранение данных
с защитой от удаления и шифрования

Обеспечение возможности отчуждения
архивных носителей от основной системы

Преимущества

Высокоэффективная дедупликация и компрессия ZSTD

(Коэффициент сжатия 6:1 — 12:1)

Высокая производительность

Технология BR-Boost ускоряет
передачу данных до 3 раз

Защита от киберугроз

Система WORM с защитой
от ransomware

Отчуждаемость архива

Возможность хранения архивных
данных в другом ЦОДе

ПАК для ИИ и машинного обучения

XGenAI | АСТРА



Комплексное решение задачи

По внедрению ИИ от инфраструктуры до инференса

Устранение зависимости

Масштабирования от объёма GPU-памяти

Предсказуемая производительность

При росте нагрузки

АСТРА ИИ

Управляющий Software-слой

Hardware-слой

Инференс и прикладные задачи

Биллинг с ролевой моделью + управление ресурсами GPU

4× Пользователей на том же железе

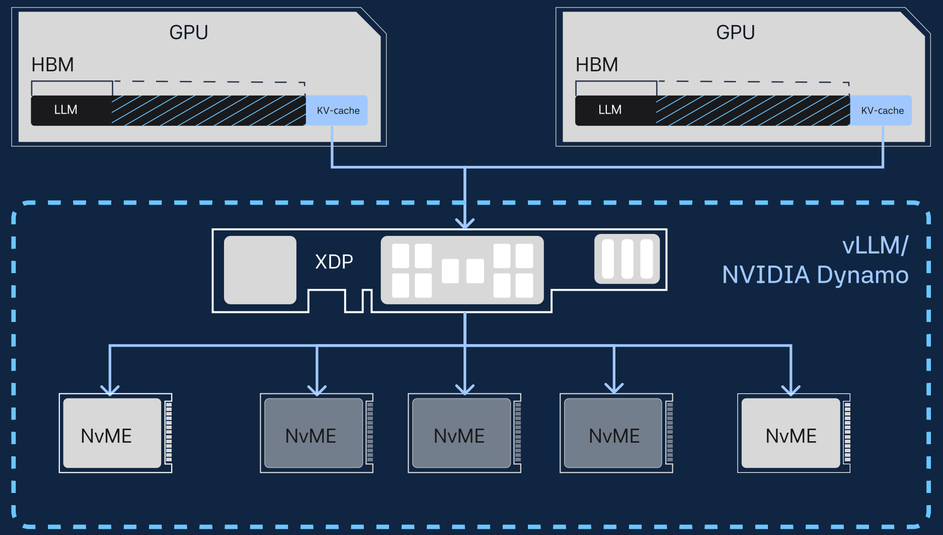
Рост пользователей в ИИ-инфраструктуре

ПРОБЛЕМА

ИИ-вычисления упираются в память,
а не в производительность ядер GPU

РЕШЕНИЕ

Перенос контекстного кеша из памяти
видеокарты на NvME-диски



x4

пользователей
на том же кластере

x4 Больше успешных запросов в секунду (Throughput req/s)

x4 Средняя скорость выдачи последующих токенов (TPOT)

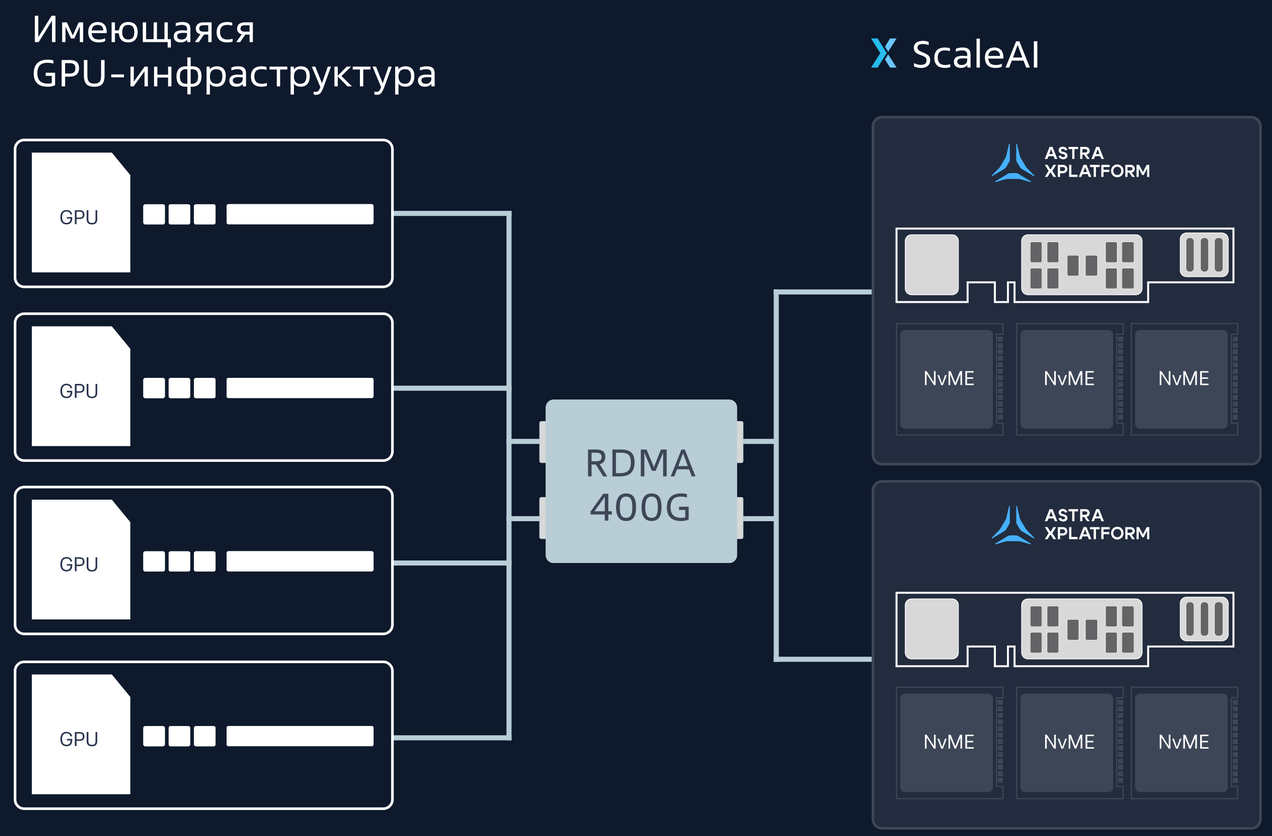
x3 Быстрее первый токен (TTFT)

50-75%

снижение затрат

Как увеличить количество пользователей для текущей GPU-инфраструктуры?

X ScaleAI —
x4 пользователей к вашему текущему инференсу без изменения SLA



Что еще может ПАК ИИ от XPlatform:

Оптимизация за счет gateway

Перераспределение запросов в модели с прогретым кешем

Биллинг по токенам/времени

Гибкая настройка ограничений доступа по времени, типам моделей при помощи ролевой модели

Мониторинг

SSO

Доступ к внешним LLM через биллинг

Безопасность обеспечивает подключение к внешним моделям через Guardrails и DLP

Магазин LLM с готовыми параметрами под аппаратные составы ПАКа

Не нужно подбирать параметры развертывания для каждой модели

Хранение истории запросов

Нужно для

1. Расследования инцидентов
2. Обучения модели на датасете из отработанных запросов

Комплексный ПАК для построения ИТ-инфраструктуры

ПАК инфраструктурного сервиса

ПАК частного облака

ПАК корпоративного бизнес-приложения

Комплексный ПАК инфраструктуры и его модули

Модули



Модуль хранения

Базовый модуль

ASTRA LINUX®

dcimanager

Модуль коммутации

Инфраструктурные ПАКи (планы на 2026-27 гг)

- X Worker
- X Directory
- X Monitoring
- X Trok
- X Backup
- X Post
- X Start

Платформа для построения частного облака на базе Astra Cloud Platform



Astra Cloud Platform включает в себя:

IAM

API Gateway

Мониторинг

Служба каталога

Биллинг

Система хранения данных

Средство резервного копирования

Программно-определяемые сети (SDN)

Платформа контейнеризации

Средства виртуализации

Создание и управление виртуальными машинами и инфраструктурой

Примеры применения

Быстрое развертывание сложных ИС, особенно при импортозамещении

Готовность ПАК к аттестации: Сертификаты ФСТЭК на программные компоненты XCloud** и сертифицированные СЗИ*

*(Средство мониторинга событий безопасности;Средство доверенной загрузки;Средство межсетевого экранирования / обнаружения(предотвращение) вторжений;Средство межсетевого экранирования/ обнаружения;(предотвращение) вторжений уровня веб-сервера;Средство антивирусной защиты;Средство контроля (анализа) защищенности информации;Средства криптографической защиты (VPN, TLS))

**(Astra Linux No2557 (обновление 1.8, 8.1, 4.7, 3.8); ПК СВ «Брест» No4864;RuBackup No4879;ALD Pro No4830)

ПАК корпоративного бизнес-приложения

для комфортного импортозамещения с единой точкой ответственности за результат



Примеры систем

ERP-системы

Bi-платформы

PLM

Варианты поставки:

XCloud

ПАК корпоративной ИС
с оказанием сервиса
(SLA на функционирование
корпоративной ИС
с привлечением производителя
ИС/ проф. партнера)

XCloud + XData

Для самых
высоконагруженных ИС

Инфраструктура на базе ПАКов XPlatform



ЦОД 1

Резервный ЦОД 2

ASTRA
XPLATFORM

Целевая платформа для критичных сервисов

**От архитектуры —
к подтверждённому результату**

Запись на архитектурную
сессию