



## Сервер для интенсивных вычислений в задачах искусственного интеллекта и машинного обучения

### Ключевые характеристики

- 2 × Intel® Xeon® Scalable 4/5 поколения, TDP до 350 Вт
- До 8 GPU PCIe 5.0 x16 двойной ширины
- До 4 PCIe 5.0 x8 для карт расширения или NVMe
- До 8 ТБ DDR5-5600, RDIMM/LRDIMM ECC
- До 12 накопителей (8 × SAS/SATA, 4 × SAS/SATA/NVMe)

### Ключевые сферы применения

- Искусственный интеллект (AI)
- Машинное обучение (training)
- Аналитика и прогнозирование данных (inference)
- Облачный рендеринг и визуализация

### Откройте новые горизонты для вашего бизнеса, используя силу искусственного интеллекта

Поддержка GPU двойной ширины с интерфейсом PCIe 5.0 x16 открывает доступ к передовым технологиям ИИ, машинного обучения, анализа данных и рендеринга, обеспечивая решение задач с повышенными требованиями к вычислительным ресурсам, в частности, тренинг и инференс для моделей разной величины. Возможность развертывания специализированных ИИ-приложений максимально раскрывает потенциал графических ускорителей и позволяет внедрять новейшие алгоритмы прогнозирования, создавать аналитические модели для широкого спектра задач и работать со сложным мультимедийным контентом в реальном времени.

### Максимальная производительность при пиковых нагрузках

Сервер с двумя процессорами Intel® Xeon® Scalable 4-го/5-го поколения создает мощный фундамент для ресурсоемких задач и обработки больших объемов данных. Оптимальная архитектура минимизирует задержки, ускоряет выполнение сложных операций и гарантирует стабильную высокую производительность в любых условиях.

### Гибкость адаптации под различные задачи

Масштабируемость до 8 GPU и поддержка технологии NVLink™ Bridge позволяют начать с базовой конфигурации и постепенно наращивать производительность по мере увеличения потребностей бизнеса. Широкий выбор поддерживаемых графических карт (включая H100 и RTX4090) делает платформу универсальной для тренинга, инференса и других нагрузок с гибкой адаптацией под изменяющиеся задачи.

### Создание распределенного GPU-кластера

GPU-серверы можно объединить в распределенный кластер через сетевые интерфейсы, а высокоскоростные адаптеры PCIe 5.0 минимизируют задержки передачи данных, что является ключевым фактором для распределенного обучения моделей и работы с глубокими нейронными сетями. Интеграция с системами хранения увеличивает объем хранилища, ускоряет обработку обучающих выборок и обеспечивает стабильный доступ к данным для машинного обучения и анализа.

### Масштабируемая память и гибкое хранение данных

До 8 ТБ оперативной памяти DDR5 обеспечивает быструю обработку больших объемов данных и минимизирует задержки в задачах обучения нейронных сетей и анализа, а 12 отсеков для накопителей SAS, SATA и NVMe добавляют гибкость и масштабируемость хранилищу.



## ОБЩИЕ ХАРАКТЕРИСТИКИ СИСТЕМЫ

<b>Форм-фактор</b>	4U, стандартный 19-дюймовый шкаф	
<b>Модуль расширения</b>	<ul style="list-style-type: none"><li>• 2 × Intel® Xeon® Processor Scalable 4/5 поколения</li><li>• TDP до 350 Вт</li></ul>	
<b>Поддерживаемые GPU</b>	<ul style="list-style-type: none"><li>• NVIDIA® H200 NVL*, RTX PRO 6000 Blackwell*, H100 NVL, H100, A100, L40S, L40, A40, A30, A10, A16, 6000 ADA, 4000 ADA, L4, T4, RTX4090</li><li>• TDP до 600 Вт*</li><li>• Поддержка NVLink Bridge</li></ul> <p><b>* Не более 4-х карт в сервере при температуре окружающей среды не более 25 °C</b></p>	
<b>Возможности расширения</b>	Спереди: <ul style="list-style-type: none"><li>• 1 × PCIe 5.0 x16 FHFL (активно x8)</li></ul>	Сзади: <ul style="list-style-type: none"><li>• 8 × PCIe 5.0 x16 FHFL</li><li>• 2 × PCIe 5.0 x16 FHFL (активно x8 в каждом разьеме)</li><li>• 1 × OCP 3.0 (PCIe 5.0 x8)</li></ul>
<b>Память</b>	<ul style="list-style-type: none"><li>• До 8 ТБ (32 × 256 ГБ)</li><li>• 32 × DDR5-5600</li><li>• Поддержка RDIMM/LRDIMM ECC (2S × 8 каналов × 2DPC, 16 × DIMM на процессор)</li></ul>	
<b>Локальное хранение данных</b>	12 × LFF/SFF (спереди): <ul style="list-style-type: none"><li>• 8 × SAS/SATA</li><li>• 4 × SAS/SATA/NVMe</li></ul>	
<b>Поддержка RAID</b>	<ul style="list-style-type: none"><li>• RAID-контроллеры с поддержкой RAID 0, 1, 10, 5, 50, 6, 60, памятью до 8 ГБ</li><li>• Поддержка VROC для NVMe</li><li>• Опционально – суперконденсатор для защиты содержимого кэш-памяти при отключении питания</li></ul>	
<b>Интерфейсы ввода-вывода</b>	Передние: <ul style="list-style-type: none"><li>• 1 × USB 3.0</li><li>• 1 × USB 2.0</li><li>• 1 × VGA</li></ul>	Задние: <ul style="list-style-type: none"><li>• 2 × USB 3.0</li><li>• 1 × VGA</li><li>• 1 × serial BMC (MicroUSB, RS-232)</li></ul>
<b>Подключение к сети</b>	1 × 1GbE RJ-45 (выделенный порт управления BMC)	
<b>Управление и мониторинг</b>	<ul style="list-style-type: none"><li>• ПО на основе OpenBMC с корпоративным функционалом от YADRO®</li><li>• Доступ через IPMI 2.0, WebUI, CLI, SNMP, Redfish API</li><li>• ПО управления YADRO® СУПРИМ и набор инструментов VEGMAN® Satellites</li></ul>	
<b>Целевые ОС (сертификации и тесты в работе)</b>	<ul style="list-style-type: none"><li>• РЕД ОС 7.3, 8.0</li><li>• ALT Linux 10, 11</li><li>• Astra Linux 1.7, 1.8</li><li>• Astra Linux SE 1.8</li><li>• SberLinux 9.0</li><li>• Windows Server 2022, 2025, 2025 (Hyper-V)</li></ul>	<ul style="list-style-type: none"><li>• ESXi 7.3, 8.x</li><li>• Ubuntu 24.04 LTS</li><li>• Moncloud 1.5</li><li>• Базис 4.3</li><li>• zVirt 4.4</li><li>• Брест 3.3</li><li>• Альт Виртуализация 11</li></ul>
<b>Электропитание</b>	3+1 CRPS Platinum до 2700 Вт	



## ОБЩИЕ ХАРАКТЕРИСТИКИ СИСТЕМЫ

<b>Охлаждение</b>	12 двухроторных вентиляторов размера 60×56 мм с резервированием N+2 и поддержкой горячей замены
<b>Климатический режим</b>	От 5 до 40 °С при относительной влажности от 5% до 90% (отдельные конфигурации могут иметь иные ограничения)
<b>Монтажные размеры (Ш × Д × В)</b>	844,3 × 447 × 174,8 мм
<b>Установочный комплект</b>	Рельсы (опционально – направляющая для укладки кабелей)
<b>Гарантия и поддержка</b>	1 год (9 × 5), доступны расширенные программы и опции поддержки
<b>Вес</b>	До 65 кг