ИНФЕРИТ

Национальная премия Бренд года в России | 2024

Высокопроизводительные решения Инферит для систем ИИ с жидкостным охлаждением

ПРОДУКТЫ, РЕШЕНИЯ И КОМПЕТЕНЦИИ

Группа компаний Инферит

Тех ИБ ОС ИТ Фин Бил Об

Экосистема продуктов вендора Инферит

Техника

Тех

Компьютерное и серверное оборудование

Безопасность

ИБ

Надёжная защита инфраструктуры от несанкционированного доступа, утечки и кибератак

Операционная ситема

ОС

Семейство ОС «МСВСфера» для серверов и рабочих станций

ИТМен

ИТ

ПО для инвентаризации, учёта и контроля ИТ-инфраструктуры

FinOps

Фин

FinOps-платформа и консалтинг для эффективного управления затратами на облака

Биллинг

Бил

Billogic Platform — биллинговая платформа для автоматизации современной подписочной модели продаж

Облако

Об

Российский облачный провайдер инфраструктурных и ИТ-сервисов

«Инферит Техника»

подразделение вендора «Инферит» по производству компьютерной техники и серверного оборудования.

Гибкость и кастомизация

Подстраиваем аппаратную часть и ПО под задачи бизнеса

Индивидуальный подход

Разрабатываем конфигурации IT-оборудования под конкретные запросы

Полный цикл

Проектируем, производим и интегрируем IT-решения

Отгружено в 2024 году

500+
серверов

50 000+
компьютеров

2 000+
компаний приобрели ПК

Преимущества

2 000+
клиентов

7 000 м
производство

300 000
устройств в год

Линии поверхностного монтажа

Свой R&D
Прототипирование изделий

Собственный промышленный дизайн

Собственный склад

Сборочный конвейер

Собственное производство в наукограде Фрязино

Предпосылки к созданию

Бурный рост систем Искусственного Интеллекта

- Потребность в специализированной ИТ инфраструктуре для систем ИИ
- Параллельные вычисления на базе GPU значительно эффективнее традиционной архитектуры x86.
- Решения с несколькими GPU в одном корпусе — один из самых быстрорастущих сегментов рынка.

Воздушное охлаждение — технология из прошлого, не отвечает современным вызовам.

- Эффективность воздушного охлаждения подошло к своему пределу и становится сдерживающим фактором для развития ИТ инфраструктуры систем ИИ.
- Топовые конфигурации GPU серверов, не могут работать в долгосрочном режиме без троттлинга

Рост требований к электропитанию и охлаждению инфраструктуры для систем ИИ

- Современные чипы становятся энергоэффективнее, но выделяют всё больше тепла.
- Рост энергопотребления серверов с GPU в несколько раз по сравнению с серверами общего назначения
- Рост требований к эффективному отведения тепла

Жидкостное охлаждение — это решение позволяющие выйти за рамки ограничений

- Использование жидкостного охлаждение позволяет существенно повысить плотность размещения GPU карт
- Топовые конфигурации GPU серверов, функционируют на полную мощность
- Возможна долговременная работа на турбо частотах

Рабочая станция Инферит

Tex ИНФЕРИТ

ЛАУРЕАТ
НАЦИОНАЛЬНАЯ ПРЕМИЯ
БРЕНД ГОДА
В РОССИИ | 2024

Рабочая станция Инферит

Рабочая станция Инферит для задач AI deep learning

Типовая конфигурация

4x H10094GB или 2x H200 141GB

Совместимо со следующими GPU картами:
NVIDIA: 3090, 4090, 5090, RTX A6000, RTX 6000 ADA, A40, L40, L40S, A100, H100, H200, RTX PRO 6000.
AMD: W7800, W7900.

Типовые конфигурации для различных сценариев

Основное направление:

- **AI Inference**
  (запуск обученных ИИ моделей):
  4-8x RTX 5090 (32GB), L40S и W7900 (48GB) или RTX PRO 6000 (96 GB)

- **AI Training**
  (обучение ИИ моделей):
  4-6x H100 (80GB и 94GB), H200 (141 GB) или RTX PRO 6000 (96 GB)

Другие отрасли:

- **Криминалистика** (восстановление цифровых паролей) - 4-8x RTX 5090
- **Виртуальное продакшн —** 2x RTX 6000 ADA или 2x RTX PRO 6000
- **Рендеринг** - 4-8x RTX 5090, RTX 6000 ADA или RTX PRO 6000
- **Научные исследования в области биологии и медицины** - 4-8x RTX 5090
- **Высокочастотная торговля** - разогнанный Threadripper 7000 (до 5.1 ГГц)

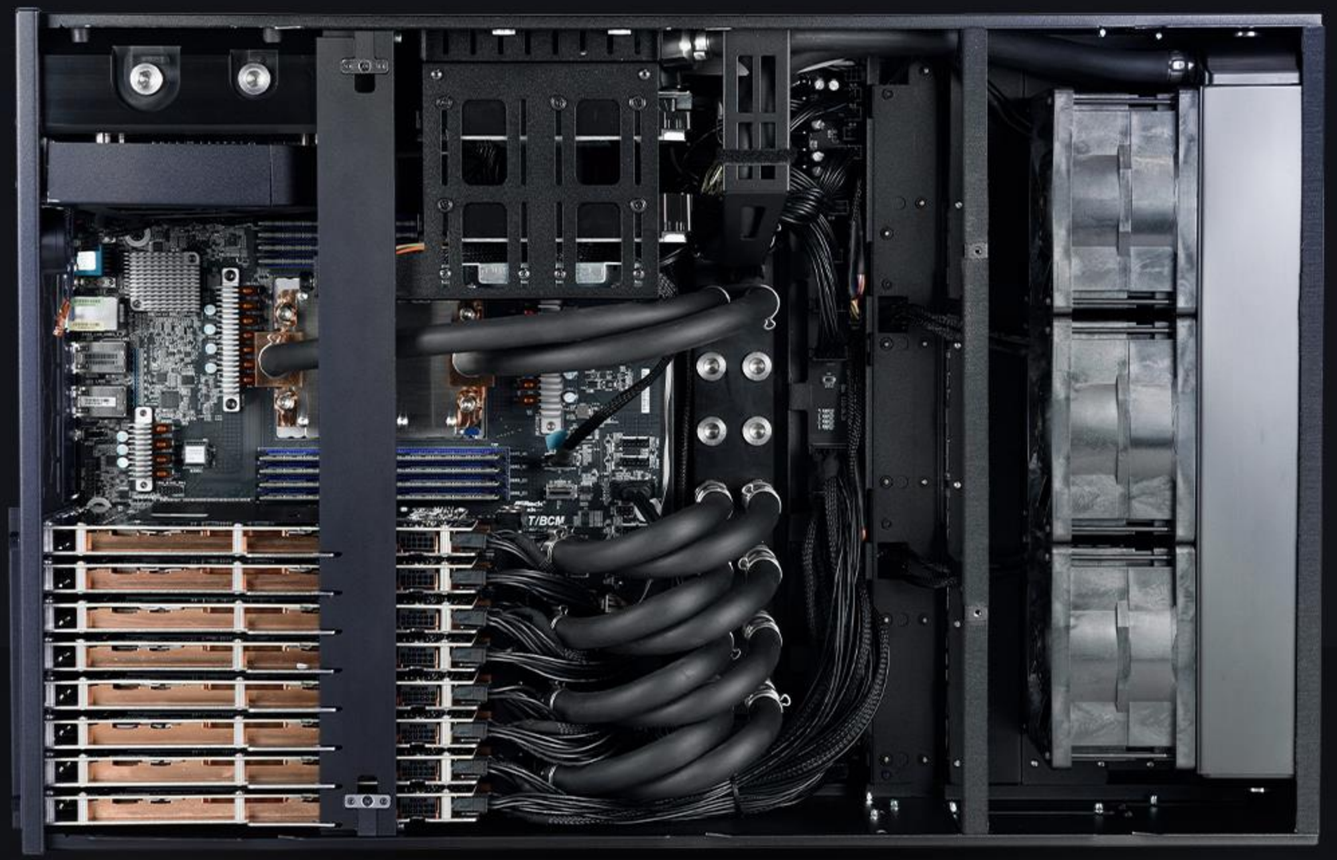
Уникальность решений

Новый тип решений на рынке, идеально подходящий для рабочих групп и автономного использования.

До 8 графических процессоров и 2 центральных процессоров (до 5,5 кВт).

- Уникальное решение, изначально спроектированное под жидкостное охлаждение.
- Экстремальная производительность.
- Компактный корпус и высокая плотность вычислений.
- Низкий уровень акустического шума.
- Система резервирования питания.

Сервер для систем ИИ



Tex ИНФЕРИТ

ЛАУРЕАТ
НАЦИОНАЛЬНАЯ ПРЕМИЯ
БРЕНД ГОДА
В РОССИИ | 2024

Сервер Инферит для задач AI inference

4x H200 - сервер CRPS 4U

4x Nvidia H200 141GB с NVLink =
564GB видеопамяти (VRAM)

Потребление H200 до 600W –
только жидкостное охлаждение

До 4 карт H200 можно объединить с помощью NVLink

Оптимальное решение
для обучения ИИ и запуск тяжеловесных моделей с
высоким квантом (DeepSeek, GPT-4, LLaMA-4, Mixtral)

Преимущества наших решений

Специально разработан для:

- Круглосуточной работы в сложных условиях при температуре до 38ºС.
- Эффективного охлаждения и предотвращения троттлинга современного высокомощного оборудования (GPU и CPU мощностью 600 Вт и более).

Создан на основе реальных сценариев использования:

- Максимальная производительность в конфигурации — до 8 GPU.
- Предварительно настроен, адаптирован и протестирован под программный стек клиента.
- Возможны сборки по требованию Заказчиков

Пример готового решения

Ключевые бизнес-применения AI Inference

Диагностика в здравоохранении
Анализ медицинских изображений (рентген, МРТ, КТ) в реальном времени.

Обнаружение мошенничества, восстановление паролей, криминалистика
Мониторинг финансовых транзакций для выявления мошеннической активности.

Применение обработки естественного языка (NLP)
Чат-боты, виртуальные ассистенты или расшифровка речи в реальном времени.

Автономные системы
Принятие решений в реальном времени для автономных автомобилей или дронов.

Персонализированные рекомендации
Онлайн-платформы, использующие ИИ для предложения товаров, контента или услуг.

Умное производство
ИИ-модели, оптимизирующие процессы на производстве или предсказывающие поломки оборудования.

Ключевые бизнес-применения в Life Sciences

Разработка лекарств
Прогнозирование молекулярных взаимодействий, поиск кандидатов на лекарства, моделирование результатов клинических испытаний.

Персонализированная медицина
Разработка индивидуальных планов лечения на основе генетических данных, биомаркеров или истории болезни пациента.

Эпидемиология и общественное здравоохранение
Моделирование вспышек заболеваний и прогнозирование потребностей в здравоохранении.

Медицинская визуализация и патология
ИИ-модели для анализа образцов тканей или изображений на наличие аномалий.

Оптимизация клинических испытаний
ИИ, помогающий в подборе участников и проектировании клинических исследований.

Геномика
Секвенирование и анализ геномных данных для изучения генетических заболеваний или эволюции.

Биопроизводство
Оптимизация условий в биореакторах и производственных процессов с помощью ИИ.

ИНФЕРИТ Тех

НАЦИОНАЛЬНАЯ ПРЕМИЯ БРЕНД ГОДА В РОССИИ | 2024

Московская обл, г.о. Фрязино, г Фрязино, тер. Восточная Заводская промышленная, д. 3, стр. 5, помещ. 1030, Россия, 141190

info@inferit.com

8 800 707-85-53

inferit.com