ЦОД-2026: агенты, сегрегация и фотоны

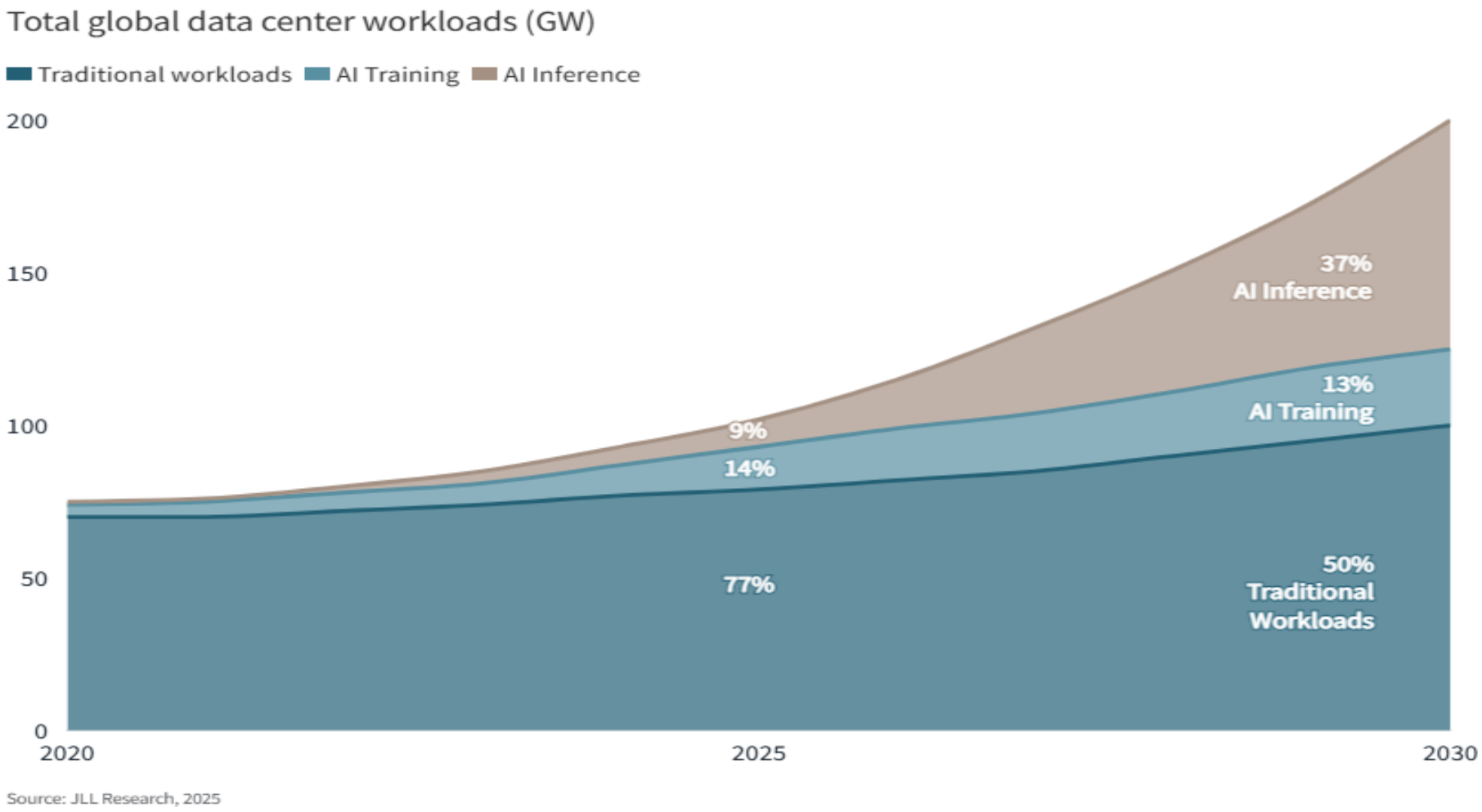


Александр Барсков,
директор по контенту, Аналитический Центр ИКС

© АЦ ИКС, 2026

Типы нагрузок в ЦОД

В 2027 г. инференс обгонит обучение ИИ-моделей



AI-нагрузки формируют рост спроса на мощности ЦОД

Что такое AI-ЦОД

Типовое определение

AI-ЦОД — это дата-центр, специально спроектированный и построенный для выполнения задач искусственного интеллекта

Классический ЦОД	ЦОД для AI
В основном CPU	GPU/AI-ускорители
5–15 кВт на стойку	30–150+ кВт на стойку
Воздушное охлаждение	Жидкостное охлаждение
Ethernet	InfiniBand / Ethernet (RDMAoCE) 400–800G

AI: обучение и инференс

С точки зрения требований к инфраструктуре ОБУЧЕНИЕ и ИНФЕРЕНС – вещи очень разные

Параметр	Обучение	Классический инференс
GPU	*****	**
CPU	**	***
Память	*****	***
Сеть в ЦОДе	*****	***
Мощность на стойку	**80+ кВт**	**15-40 кВт**
Длительность задачи	**Недели/месяцы**	**Секунды**

AI-агенты в деле

AI-агент — программный помощник, использующий AI-модель для самостоятельного выполнения задач через взаимодействие с внешними системами и данными.

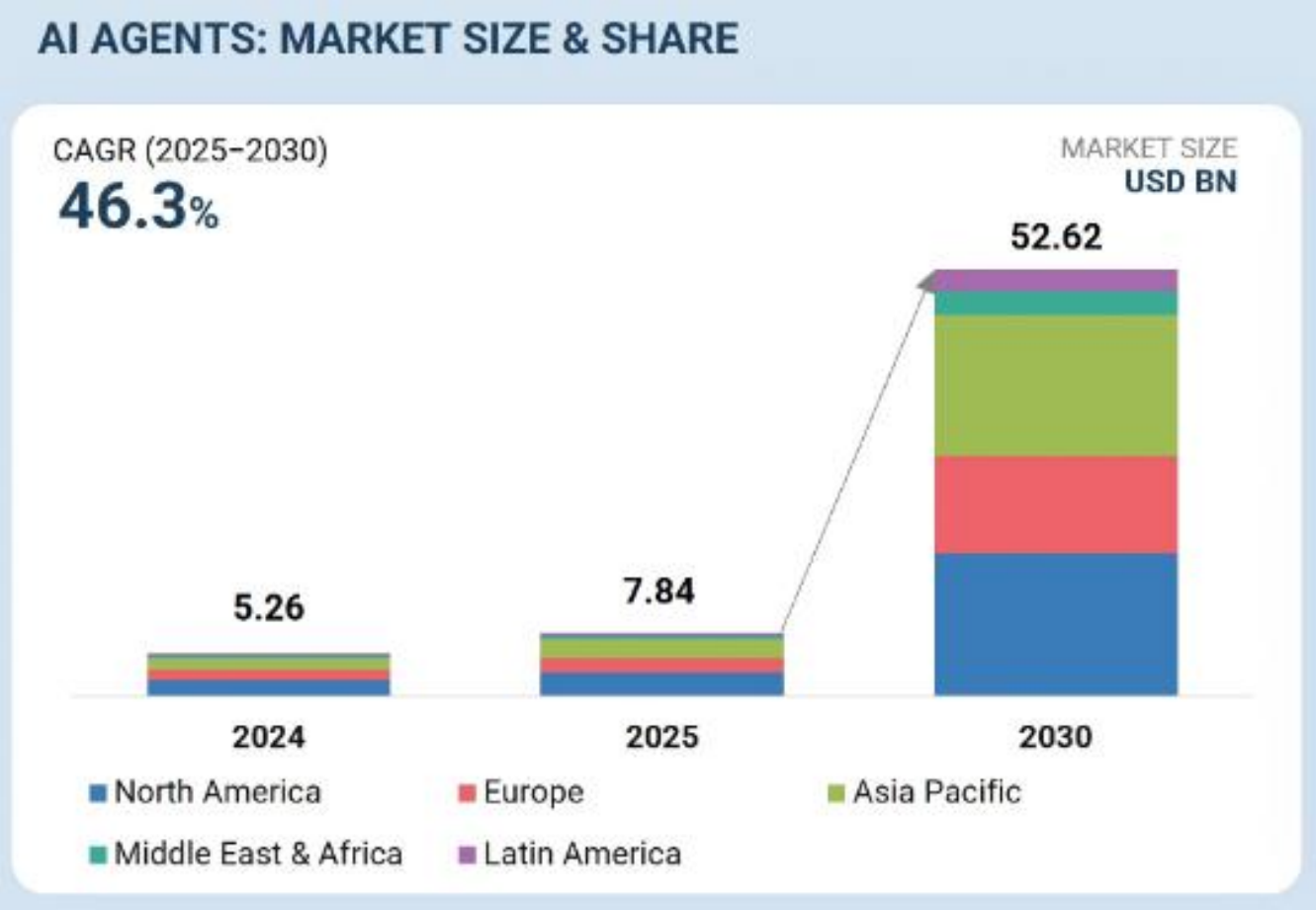
Работа агентов — это в основном инференс

Параметр	Обучение	Классический инференс	AI-агенты
GPU	*****	**	***
CPU	**	***	****
Память	*****	***	****
Сеть	*****	***	****
Мощность на стойку	**80+ кВт**	**15-40 кВт**	**20-60 кВт**
Длительность задачи	**Недели/месяцы**	**Секунды**	**Минуты**

AI-агенты. Массовое внедрение

Массовое применение

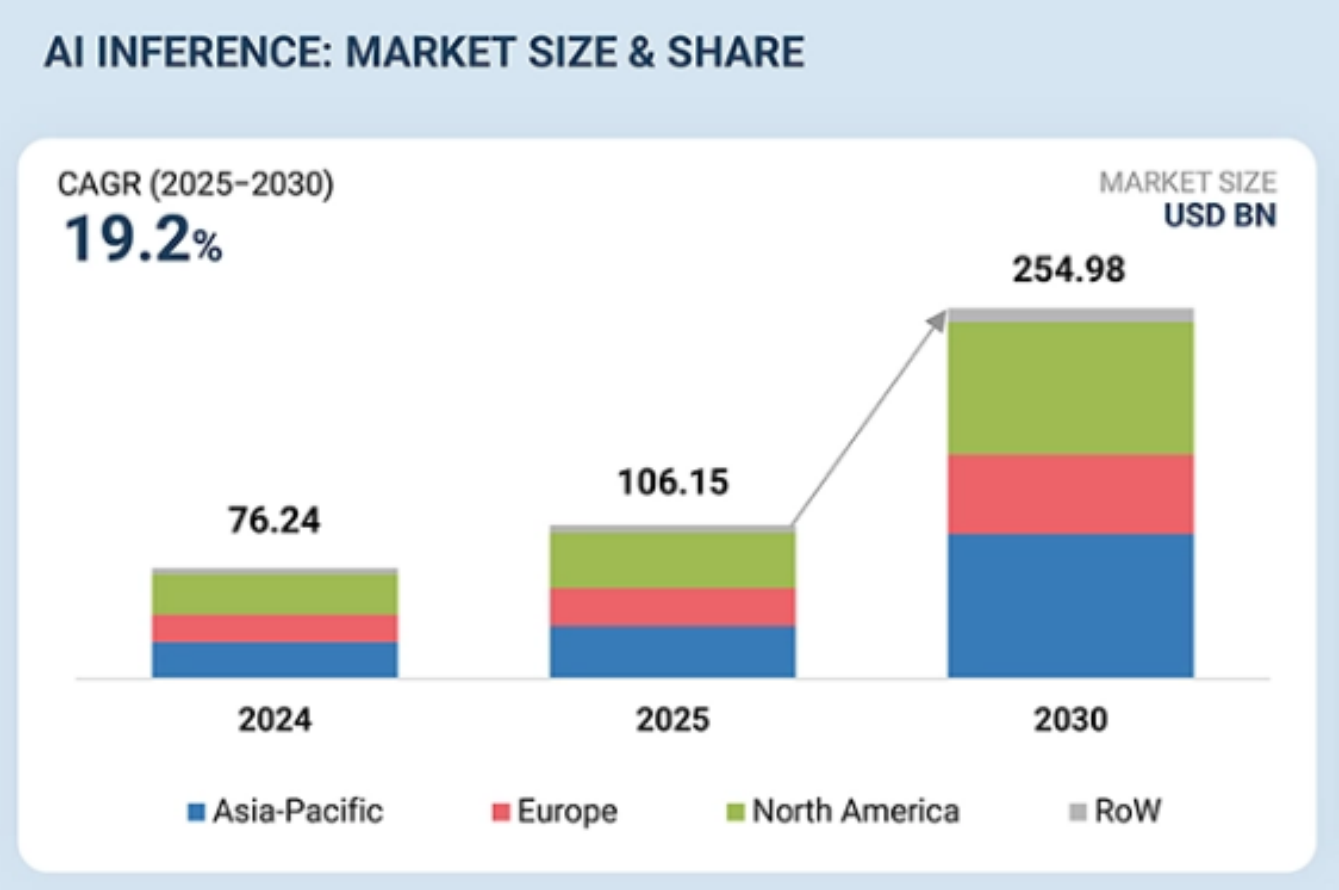
В 2026–2028 гг. AI-агенты начнут массово встраиваться в корпоративные приложения. Соответствующий рынок может вырасти до $50 млрд к 2030 г.



Источник: MarketsandMarkets

AI-агенты. Инфраструктурный эффект

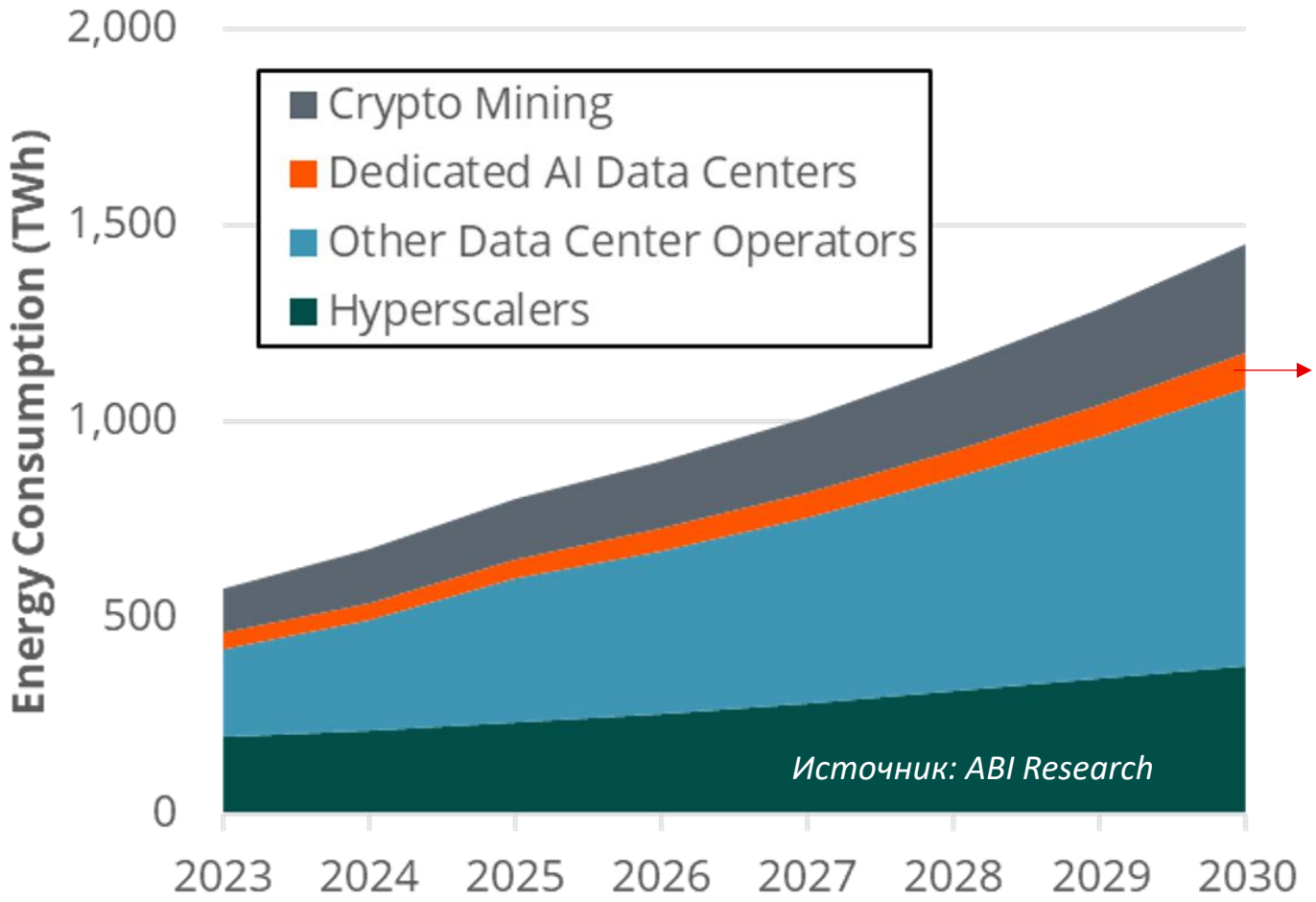
AI-агенты станут одним из главных генераторов спроса на инференс, рынок которого к 2030 г. может вырасти до $255 млрд



Источник: MarketsandMarkets

Агентные платформы формируют новый класс AI-ориентированных ЦОД с плотностью 20–60 кВт на стойку

Сколько будет AI-ЦОДов?



Доля спец. AI-ЦОД

6% по мощности

~1% по числу стойко-мест

Энергия: разрыв между спросом и предложением



4%

рост генерации электроэнергии

15%

рост потребления электроэнергии ЦОДами

Источник: МЭА

Где брать энергию? Локальная генерация

Солнечная и ветровая генерация – в составе гибридных систем с применением систем накопление энергии (СНЭ)

Сегодня – газовая генерация



В перспективе – атомные станции малой мощности (АСММ)

Сценарий использования	2025 г.	2030 г.
Резервное электроснабжение	**~80%**	Доля снизится*
Совместное использование локальной генерации и централизованного электроснабжения	**10-15%**	**50-70%**
Полностью автономное электроснабжение	**~1%**	**25-30%**

Прогноз АЦ ИКС и «Парус электро»

**из-за резервирования средствами ИТ и перевода систем резервного питания в режим гибридного электроснабжения*

Где брать энергию?
Новая архитектура вычислений

Проект XFRA — это попытка использовать свободную электроэнергию частных домов и небольших зданий для создания распределённого ЦОДа для инференса

SPAN совместно с NVIDIA планируют развернуть до 8 тыс. микроЦОД в частных домах. Каждый узел -- отдельный вычислительный модуль с ПЖО + АКБ + система управление энергией

Для инференса не требуется тесная синхронизация тысяч GPU (как для обучения), поэтому нагрузку можно распределять по множеству географически разнесённых узлов.

Установка 8000 блоков XFRA:

в 6 раз быстрее

в 5 раз дешевле

строительства ЦОД сопоставимой мощности (~100 МВт)



Источник: https://www.xfra.ai/

Фотоника и энергодефицит

Оптический вычислитель использует не электронику, а свет. Чип берёт на себя самую энергозатратную часть вычислений в ИИ — умножение матриц. Свет выполняет эти операции почти без выделения тепла.

Благодаря фотонике возможно КРАТНОЕ снижение энергопотребления крупных AI-ЦОД

Прогноз промышленного внедрения: 10-15 лет



Фотонный вычислитель Lightelligence PACE



Фотонный процессор Lightmatter

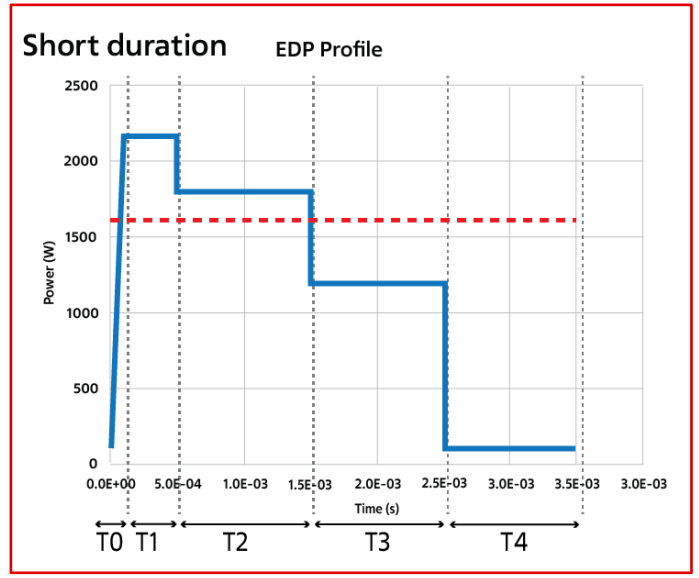
В июне 2026 г. Сбер представил оптический вычислитель для ИИ. В его основе — фотонная интегральная схема, архитектуру которой полностью разработали исследователи компании.

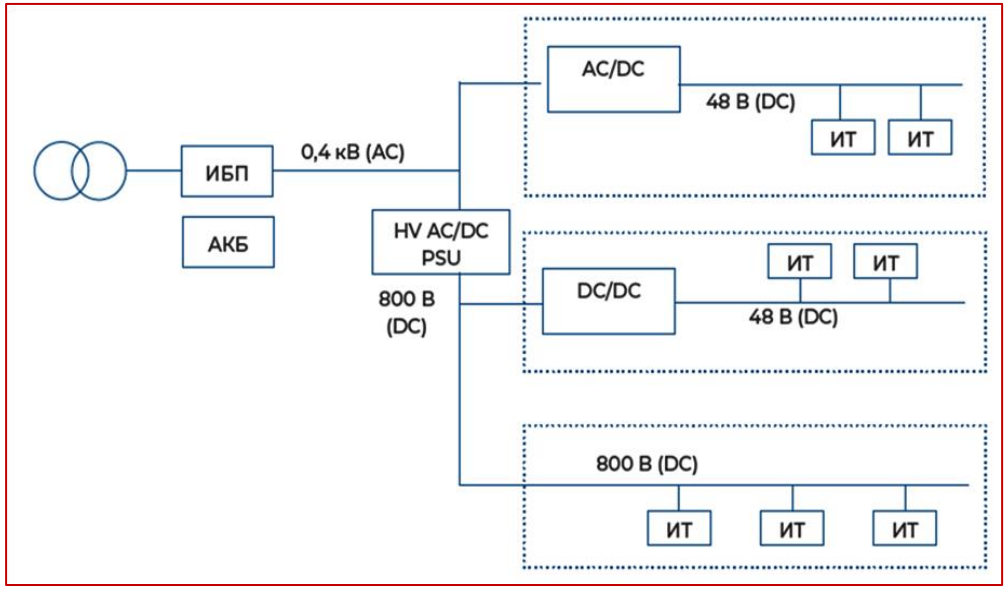
Изменения в инфраструктуре

Рост плотности мощности, дефицит электроэнергии, изменение характера нагрузки…

- Внедрение жидкостного охлаждения
- Переход на гибридные ИБП (АКБ + суперконденсаторы)
- Отказ от PDU в стойках
- Переход на схемы 400/800 В DC
- Внедрение СНЭ
- Превращение фальшпола в небольшой технический этаж с трубопроводами

…





О формальном определении «ИИ-ЦОД»

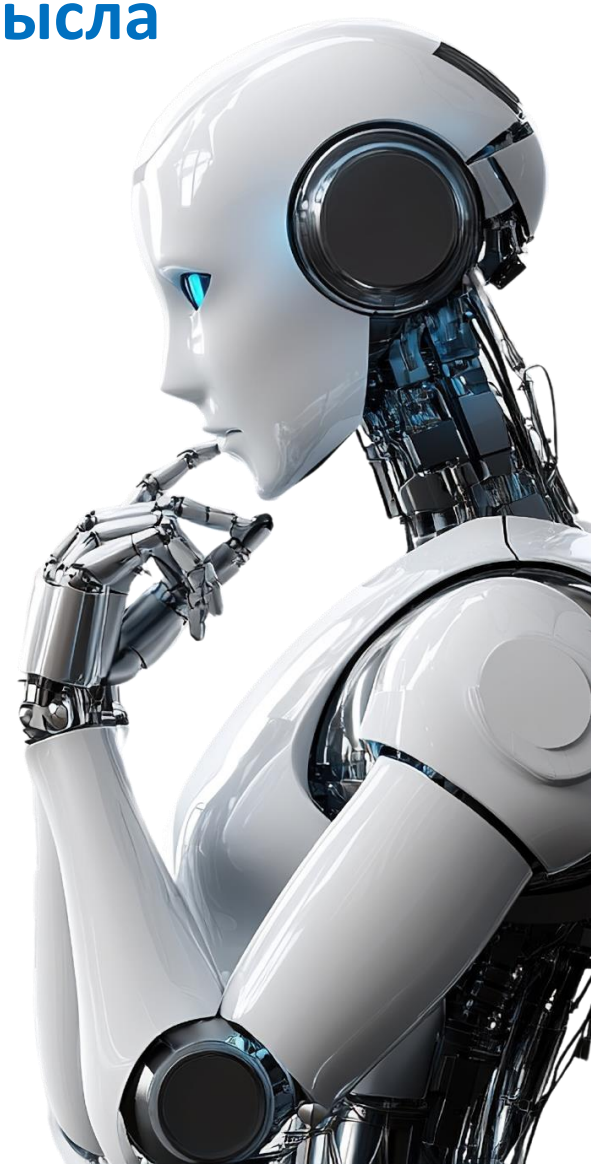
Сегрегация ЦОДов по функционалу не имеет смысла

ЦОДы – **универсальные** инженерные комплексы, предназначенные для размещения оборудования, обеспечивающие хранение и обработку **ЛЮБЫХ** данных, а не только данных **ОТДЕЛЬНЫХ** приложений, в том числе использующих ИИ.

ИИ ≠ GPU – в агентской экономике CPU будут играть все большую роль

GPU ≠ ИИ – высоконагруженные GPU-кластеры нужны не только для ИИ, это:

- научные и инженерные расчёты
- моделирование в промышленности
- рендеринг и компьютерная графика
- игровые и VR/AR-системы…



Вопросы?... ответы в материалах АЦ ИКС











© АЦ ИКС, 2026