

ЦОДы в космосе

Александр Герасимов, N4A Analytics

Сергей Мальцев, независимый эксперт по космическим технологиям

«Жить просто – нельзя! Жить надо с увлечением!» - С.П. Королев.



В детстве взахлеб прочитал¹ случайно найденную в домашней библиотеке фантастическую книжку Андрэ (Эндрю) Нортон «Саргассы в космосе»², повествовавшую о том, как в далеком будущем прорвавшиеся к звездам люди стали находить миры, выжженные дотла в невообразимой по масштабам войне, случившейся за миллионы лет до выхода человека в космос. Мысль Нортон очевидна: чем совершеннее технологии, тем выше ответственность человека за их использование во благо всей цивилизации, а не во вред, и развитие технологий должно идти в ногу с совершенствованием социальных отношений, их гуманизацией, иначе технологическое развитие приводит к катастрофе.

Тут вспоминается другая цитата Сергея Павловича Королева про гармонию природы и технологий: «...Мы должны понять и изучить весь этот гигантский, постоянно изменяющийся механизм и, не ломая его, каким-то образом подключить к нему машину нашей цивилизации...».

Космические ЦОДы – зачем?

Но перейдем все же к технологиям. Центры обработки данных (ЦОДы) в космосе. Выглядит как первоапрельская шутка, и, одновременно, как очередная безумная фантазия компаний с триллионной капитализацией в попытке задрать ее еще выше. Хотя, казалось бы, куда выше то?

Публичная аргументация техногигантов слабовата. Космические ЦОДы как решение проблемы временного энергодефицита в отдельных регионах США, внезапно случившегося ввиду интенсификации строительства гигантских ЦОД под обучение LLM? Это как у Михаила Жванецкого в одной из миниатур: «...Вроде колхоз не настоящий... И тем, кто никогда не были в колхозе, что-то не верится, а те, кто живут в нем, жутко ругаются и матерно кроют Голливуд...». Те, кто хоть немного разбирается в экономике ЦОД, тоже в полном недоумении – как такое вообще можно «скормить» публике.

Если реальные цели проектов создания космических группировок ЦОД действительно соответствуют заявленным, то единственно верная стратегия непричастных – наблюдать со стороны, когда лопнет очередной технологический пузырь, находясь подальше от эпицентра события. Тем более что тот же SpaceX в рамках подготовки своего IPO честно предупреждает своих

¹ Автор вводного текста – Александр Герасимов.

² <https://libcat.ru/knigi/fantastika-i-fjentezi/28809-andre-norton-sargassy-v-kosmose.html>

потенциальных инвесторов, что затея с мега-ЦОДами в космосе имеет высокие риски не реализоваться в виду технологической новизны и непонятной экономики.

А если нет? Что если ЦОДы в космосе – это все-таки реальная потребность экономики будущего, причем в определенной конфигурации технически реализуемая за разумные деньги и потому будущего недалекого?

Давайте начнем с экономической потребности, но с реальной, а не выдуманной ради шумных заголовков в СМИ и вдохновения инвесторов, чтобы те не обращали внимание на фантастичность мультипликаторов капитализации к выручке и прибыли. И рассмотрим ее в связке с техническими аспектами, причем не космическими, а исключительно ИТшными, то есть абстрагируемся пока от слова «космический», и сосредоточимся на аббревиатуре «ЦОД».

Экономика будущего – это цифровая экономика, технологически реализуемая через «приход AI в физический мир», в реальный сектор экономики, как средство автоматизации производственных и бизнес-процессов до уровня полной автономности от непосредственного участия человека, причем не только на уровне исполнения этих процессов, но и на уровне их модификации. Именно в этом, а не в написании курсовых и сочинений студентам и школьникам, может состоять основной экономический смысл, потенциально дающий окупаемость масштабным AI-проектам.

Еще пять лет назад мы в N4A Analytics предприняли попытку просчитать соотношение между затратами и возможным экономическим эффектом от полномасштабной цифровизации реального сектора экономики³. Сделали это не для абстрактной «глобальной экономики», образованной экономиками совершенно разных стран, а на примере двух вполне конкретных экономик: России и США. России как примера экономики с низким подушевым ВВП - такие же удельные показатели характерны для большинства стран БРИКС+, включая Китай, а США – как примера «богатой» экономики с высоким подушевым ВВП.

Результаты оказались схожи: налицо два общих аспекта.

Аспект №1: «структурный перекося». Гросс экономический эффект от цифровизации крайне неравномерно распределен по отраслям, его распределение совершенно не соответствует отраслевому распределению ВВП. Львиная доля экономического эффекта от цифровизации приходится на логистику и транспорт, где основной задачей цифровизации является динамическое оптимизационное управление транспортными средствами, то есть подвижными объектами. Причем на всех уровнях: от непосредственно управления транспортным средством до управления парками транспортных средств и даже производственно-торгово-логистическими кооперационными цепочками. Следует отметить, что чем ниже подушевой ВВП, тем выше доля логистики и транспорта в общем гросс-эффекте от цифровизации.

Аспект №2: затратность цифровизации. Наши расчеты показывают, что цифровизация – чрезвычайно затратный проект. Не зря первый «заход на цифровизацию», сделанный в СССР в 1960-х годы на пике успехов научно-технической революции в ВПК, оценивали как более дорогостоящий, чем ракетно-космический, атомный и противовоздушный / противоракетный проекты СССР вместе взятые. И ничего с тех пор фундаментально не изменилось – затраты действительно космическо-атомные.

Разумеется, большие затраты или маленькие – понятие относительное. В данном случае относительное в сравнении с экономической выгодой. И проблема как раз в этом соотношении: экономическая выгода от цифровизации в подавляющем большинстве отраслей реальной экономики оказывается либо очень близка к затратам, либо даже меньше их. Одно «но» – это если

³ <https://www.n4a.info/autonomous-systems-ru>

брать удельные затраты «как есть». Отсюда простой вывод: надо кардинально снижать удельные затраты на цифровизацию.

У затрат на цифровизацию две составляющие: ИКТ-инфраструктура и труд. Начнем с труда, но ограничимся кратким описанием, поскольку наша статья все-таки про ИКТ-инфраструктуру.

Цифровизация – чрезвычайно трудоемкий процесс. Причем даже в варианте, который просчитывали мы, когда все необходимые для корпоративной цифровизации программные приложения предоставляются провайдерами корпоративным потребителям в готовом виде, аaaS, причем с требуемым SLA. Основная трудоемкость – в необходимости глубокой перестройки (реинжиниринга) производственных и бизнес-процессов, и создании над ними еще одного уровня процессов – координационных, охватывающих сквозным образом всю кооперационную цепочку. В основном эти затраты ложатся на сами цифровизирующиеся бизнесы, то есть это их внутренние затраты. За примерами далеко ходить не надо – это маркетплейсы, фактически сами являющиеся технологическими компаниями. Причем чем ниже подушевой ВВП, тем выше перекоп в сторону inhouse, то есть внутренней составляющей затрат на цифровизацию.

Как снизить удельные внутренние затраты? Выход здесь начинает прослеживаться – это AI агенты. В теории их применение сокращает или даже полностью исключает необходимость длительного и чрезвычайно трудоемкого процесса реинжиниринга процессов. В случае с AI-агентами нужна лишь начальная точка, пригодная для старта, - минимально работоспособная конфигурация процесса, и корректно выстроенный процесс самообучения AI-агентов. Аналогичный подход применим и к задаче снижения трудоемкости разработки тиражного софта для процессной автоматизации любого уровня: от АСУТП до ERP/BI/BPM.

А вот ИКТ-инфраструктура? У нее, в свою очередь, две составляющие – вычислительная и телекоммуникационная.

Начнем с вычислительной. Да, затраты на ее развитие уже сейчас гигантские. Только в США суммарный объем запланированного строительства новых ЦОДов под AI-нагрузки по состоянию на конец 2025 года превысил 800 млрд. долл., причем это исключительно частные инвестиции. То есть ожидается что они должны окупиться. Предполагается, что за счет дальнейшего роста размера ЦОД удастся существенно снизить удельную стоимость вычислений. Речь идет о строительстве даже не гиперскейл, а слово то подобрать сложно каких размеров ЦОДов – целых городов.

Однако есть еще одна критически важная составляющая, определяющая удельную себестоимость вычислений – это уровень средней загрузки вычислительных мощностей, который может «гулять» в широком диапазоне значений, от единиц процентов загрузки CPU до выше 60%. Понятно, что уровень загрузки в единицы процентов полностью «убивает» экономику любого ЦОД, где бы он ни располагался и какого размера он бы ни был.

Проблема повышения средней загрузки вычислительных мощностей под AI-нагрузками сейчас стоит чрезвычайно остро и пока очень далека от своего решения. Так, SpaceX готов заплатить 60 млрд. долл. за маленький стартап Cursor⁴, чтобы тот помог повысить загрузку AI-суперкластера SpaceX Colossus (крупнейший в мире суперкомпьютер для AI-нагрузок) с провальных 11% до хотя бы 30-35% при обучении LLM⁵. Про привычные для других видов нагрузок уровень утилизации мощностей гиперскейлов в 60% загрузки CPU и более пока речи даже не идет.

На средний уровень загрузки вычислительных мощностей основное влияние оказывает качество работы системы управления ЦОД (домена) и кросс-доменной системы оркестрации. Но пока решать проблему пытаются не на уровне систем управления, а в основном инфраструктурно, путем

⁴ <https://x.com/SpaceX/status/2046713419978453374>

⁵ <https://cursor.com/blog/spacex-model-training>

дальнейшего укрупнения датацентров. Так проще мультиплексировать нагрузки различных tenants, комбинируя крупные и малые полезные нагрузки. Но, как видно из примера со SpaceX, это далеко не универсальный рецепт, скорее даже тупиковый. Другая крайность – это оптимизация процессов обучения LLM, чем занимается, в частности Cursor, но без вмешательства в уровень управления распределенной сетью-вычислительной инфраструктурой. Очевидно, что нужна оптимизация на всех трех уровнях: от AI-моделей до инфраструктуры через промежуточный уровень управления инфраструктурой.

Вернемся к вычислительной инфраструктуре. Обратная сторона медали централизованной схемы с гиперскейлами применительно к задаче использования AI для управления физическими объектами – это интенсивный трафик между объектом управления и гиперскейл-ЦОД, причем с жесткими требованиями по доступности канала и уровню задержки сигнала (latency). Но сетевой ресурс ограничен, особенно если это радиочастотный ресурс, ведь мы говорим в первую очередь о задаче управления подвижными объектами как формирующей основной вклад в экономический эффект от цифровизации.

Идеальная конфигурация ИКТ-инфраструктуры – это гибкая система из корневых и краевых вычислительных мощностей с управляемой сетевой связностью между ними, способная динамически, в моменте, подстраиваться под текущую нагрузку различных приложений, в том числе AI. В теории такая система позволяет минимизировать нагрузку на сетевую компоненту и снизить требования по сквозной задержке сигнала, перенеся наиболее критичную к задержке вычислительную нагрузку ближе к объекту управления. Так, если нужна задержка в 10 мс и менее между объектом управления и вычислительным центром, то без краевых вычислений на практике это сложно реализовать. И даже в этом случае часть вычислений должна выполняться непосредственно на борту объекта управления.

Чтобы такая конфигурация была экономически жизнеспособна, требуется обеспечить высокую среднюю загрузку не только корневых, но и краевых вычислительных мощностей, а также сетей связи между ними и объектом управления и вычислительных мощностей на борту объекта управления. Это чрезвычайно сложная задача, которая до сих пор находится, мягко говоря, не в фокусе внимания техногигантов. Пытаясь в меру своих скоромных сил заполнить данный пробел, мы посвятили этой теме множество публикаций, наиболее свежая из которых – в журнале Connect за март-апрель 2026 года⁶. Наиболее проблематичный аспект задачи – балансировка вычислительной и сетевой нагрузки между доменами, управляемыми различными хозяйствующими субъектами.

Вернемся к краевым вычислениям и зададимся вопросом: а «краевые» — это, собственно, где? Где этот край физически должен находиться? Вообще говоря, краевые вычисления обсуждаются в основном в контексте сетей 5G под термином MEC. Физически это серверные шкафы на сайтах базовых станций 5G.

Почему только 5G? Почему не LTE? Дело в том, что за счет разделения контуров управления и передачи пользовательских данных, то есть отдельной передачи сигнального и пользовательского трафика, в 5G сетях стало возможным локально выводить трафик на краевые ЦОД. В сетях LTE это невозможно – весь пользовательский трафик надо «гнать» в ядро сети, где реализована функция его коммутации (аналог UPF в 5G), а потом обратно в краевой ЦОД, что полностью обесценивает главное преимущество краевых вычислений – низкую задержку сигнала.

Несмотря на пять лет активного развертывания сетей 5G в мире, даже для богатых стран характерна проблема ограниченного покрытия сетями 5G, особенно сетями stand alone, SA (см. нашу публикацию в апрельском Connect) – только такие сети позволяют реализовать все преимущества

⁶ <https://www.connect-wit.ru/izdaniya-connect.html>

сетей 5G. Покрытие 5G SA до сих пор сосредоточено преимущественно в крупных городах с наиболее платежеспособными абонентами. И, похоже, эта ситуация сохранится на многие годы вперед.

Однако выход AI в физический мир, основу которого формирует транспорт и логистика, уже сейчас требует тотального покрытия широкополосными IP-сетями поверхности земного шара, включая океаны. Если не принимать во внимание совсем уж экзотические технические решения вроде гигантских дирижаблей в стратосфере⁷, опытно-конструкторские работы по которым были свернуты еще пять лет назад, то остается только один вариант – это низкоорбитальные спутниковые группировки как основная сеть и наземные сети 5G как резервное очаговое покрытие регионов с высокой плотностью платежеспособного населения.

Соответственно, в случае со спутниками краевые вычисления должны располагаться там же – в космосе, причем тоже на низких орбитах. Опять же, с точки зрения задержки: если для наземных сетей 5G теоретически при определенных условиях возможно достигать требуемого для AI-управления в режиме реального времени уровня задержки сигнала без использования краевых вычислений, то для спутниковых сетей это в принципе невозможно.

Но может спутник вообще не подходит для этих задач? Ведь в режиме ретрансляции даже для низких орбит задержка (round trip delay) составляет не менее 20 мс⁸? А речь то про 10 мс и ниже, вплоть до единиц мс⁹?

Тут важно понимать архитектуру решения в целом. Как отмечено выше, наиболее критичные к задержке вычисления вообще должны производиться на борту. Но не полностью – это экономически неэффективно. То есть сам AI-агент должен быть распределенным: располагаться на борту и в краевом ЦОД. В такой конфигурации задержки между бортовым и краевым вычислителем до 10 мс (до 5 мс в одну сторону) вполне допустимы, а передаваемые на наземный гипер-ЦОД со сквозной задержкой более 10 мс данные будут использоваться только для обучения LLM, то есть не будут являться данными, используемыми в режиме реального времени. Для таких кейсов допустима суммарная задержка до 50 мс и более (см. Рис. 1).

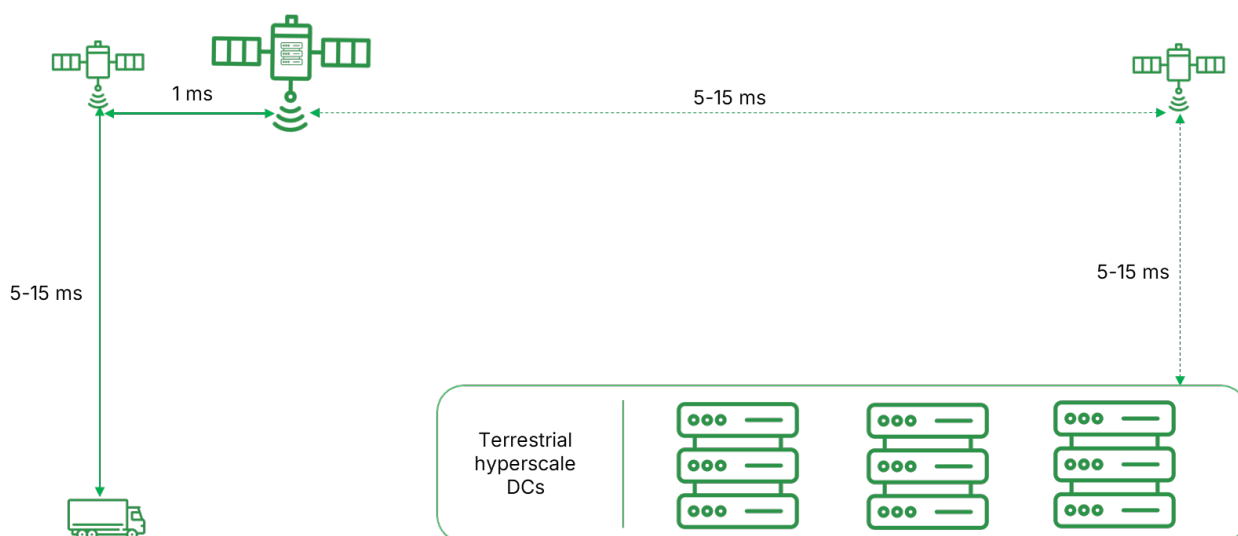


Рисунок 1. Схема ИКТ-системы с элементами в космосе для AI-управления транспортом

⁷ <https://pro-arctic.ru/10/04/2015/expert/15168>

⁸ <https://starlink.com/map?view=latency>

⁹ 3-10 мс для автономных автомобилей, стандарт 5G для сетевого слоя V2X.

Но вот режим ретрансляции для рассматриваемой задачи действительно невозможен. Чтобы обеспечить задержку до 10 мс, краевые вычислители должны находиться в космосе, в составе группировки спутников связи. Иначе всю самообучаемую систему управления придется располагать непосредственно на объекте управления, то есть делать его автономным не только с точки зрения неучастия человека в процессе управления, но и с точки зрения исключения интенсивного внешнего информационного обмена. А это кардинально ограничит сферу применения AI только очень крупными, а, следовательно, малочисленными, объектами типа крупных кораблей, что неприемлемо.

Технически в добавлении спутников-вычислителей в низкоорбитальные группировки спутников-ретрансляторов нет больших проблем. Например, начиная с версии 1.5., то есть с 2022 года, все новые спутники Starlink оснащены оптической (лазерной) линией связи с другими спутниками группировки, и могут ретранслировать сигнал не только на Землю, но и на другой спутник, который может быть вычислителем, а не ретранслятором. Да, спутник – краевой вычислитель должен быть крупнее ретранслятора, но это отнюдь не мега-ЦОД в космосе, а объект размером с гагаринский «Восток».

Есть еще один большой плюс такой конфигурации: в случае единой комбинированной спутниковой группировки (спутники-ретрансляторы и спутники-вычислители) можно избежать наиболее сложной и не решенной до сих пор проблемы кросс-доменной оркестрации – оркестрации в гетерогенной среде, образованной доменами различных хозяйствующих субъектов, ведь оба этих домена, а также домен наземных гиперскейлов будут под управлением одной компании-оператора.

Но технические проблемы конструирования спутника-вычислителя все же существуют, причем весьма критичные. В основном они связаны с необходимостью использования инновационных систем охлаждения и радиационной защиты крупногабаритного вычислительного оборудования.

Технологические инновации для спутников-вычислителей.

Очевидно, что космический ЦОД кардинально дороже наземного. Это и стоимость спутника, и стоимость его вывода на орбиту, и стоимость эксплуатации. Казалось бы, если речь о кардинальном удешевлении удельной стоимости вычислений, то вывод ЦОД в космос - направление обратное от требуемого.

Какова сейчас разница в стоимости вычислений на Земле и в космосе и каковы возможности по ее сокращению?

1 ГВт вычислительной мощности на орбите может стоить \$42,4 млрд (практически в три раза дороже, чем на Земле¹⁰) и еще только предстоит преодолеть ряд принципиальных технических проблем. По расчетам, для приближения стоимости орбитальных вычислений к наземным ЦОД нужно снизить стоимость выведения в 18 раз, с текущих \$3600/кг у Falcon 9 до \$200/кг для Starship в 2030-е гг. При этом рынок пусковых услуг является не вполне эластичным - сложная регуляторика, ограниченный масштаб, B2B/B2G-клиенты с длинными циклами продаж и оптовыми контрактами, вертикальная интеграция у перспективных мегагруппировок: до 75% пусков Falcon 9 приходится на КА Starlink, по этому же пути идут Amazon Leo/Blue Origin и китайские Guowang/Thousand Sails. То есть ответ – да, возможно, но не для всех, а только для вертикально-интегрированных гигантов.

Второе необходимое условие для приближения стоимости орбитальных ЦОД к наземным – радикальное снижение стоимости самих вычислительных КА. Стоимость 1 кВт электроэнергии наземных ЦОД колеблется от \$570 до \$3000 в год, в зависимости от местных условий. Да, на орбите солнечная энергия бесплатна, но нужно еще закупить компоненты, разработать КА и запустить его,

¹⁰ <https://andrewmccalip.com/space-datacenters>

а также обеспечить нормальное функционирование. В итоге стоимость 1 кВт КА Starlink, по расчетам ведущих сотрудников Project Suncatcher, составляет \$14 700 в год¹¹. Снижение стоимости в 5 раз возможно, учитывая массовое производство и ожидаемый для массового производства КА production learning curve на уровне 95% (снижение стоимости производства единицы продукции на 5% при каждом удвоении тиража)¹². В целом, кратное снижение и даже падение стоимости пусковой услуги и производства КА на порядок теоретически возможно.

Однако, остается решить еще крайне сложную задачу отвода тепла в вакууме. Да, солнечные батареи работают на орбите в 5-8 раз эффективнее, чем на поверхности Земли. Но отсутствие атмосферы препятствует и эффективному отводу тепла. По предварительным расчетам, на каждые 20 кВт отводимой мощности требуется 20 м² излучающей площади при стандартной температуре работы процессоров в 80°C¹³. Таким образом площадь одних только радиаторов КА SpaceX AI Mini Sat мощностью 100 кВт может приблизиться к 100 м². Что вполне возможно технически, хотя и по-прежнему амбициозно. Маск в своем стиле назвал такой КА Mini, предполагая еще более масштабные изделия. Для сравнения - электрическая мощность всех солнечных батарей МКС составляет 120 кВт при площади радиаторов в 150 м². Планы Маска по разработке кастомизированных радиационно-стойких AI-чипов, способных работать при более высокой температуре 97°C как раз призваны повысить температуру теплоносителя и общую эффективность радиаторов при снижении площади рассеивания.

Однако помимо теплоотведения, предстоит еще борьба с радиацией, чтобы избежать битовых ошибок (bit flip). Традиционный подход с использованием радиационно стойкой ЭКБ space grade неприемлем из-за своей дороговизны (один RAD750 от BAE Systems стоит, к примеру, \$200 тыс.¹⁴). Для КА вычислительных мегагруппировок нужно использовать обычную ЭКБ auto/industrial grade, но с защитной оболочкой. И лабораторные испытания Google с такими покрытиями плотностью 5-10 г/см² в ускорителе частиц внушают оптимизм¹⁵, однако предстоит провести еще реальные летные испытания. Поэтому все с нетерпением ждут запуска двух тестовых КА Google и Planet по проекту Starcatcher в 2027 г. с развертыванием на борту ИИ-моделей.

Удорожание одного элемента системы не обязательно приводит к росту стоимости системы в целом

Если все перечисленные выше инновации будут успешно доведены до широкого коммерческого применения, все равно космический ЦОД всегда будет проигрывать в экономике наземному, даже если сравнивать сопоставимое, а именно граничные ЦОДы небольшого размера – наземные и космические.

Есть только одно «но», и оно полностью переворачивает взгляд на эту проблему: следует рассматривать экономику всей системы, а не отдельного ее элемента. Как описано выше (Рис. 1), под системой мы понимаем как минимум два вида ЦОД – граничный и корневой, а также каналы связи между объектом управления и граничным ЦОД, и между корневым ЦОД и граничным ЦОД. Причем сравниваем не гипотетические конфигурации такой системы, а реалистичные, с учетом требований по пропускной способности, бюджету на задержку и уровню доступности системы.

Как отмечено выше, для подвижных объектов, представляющих наибольшее значение с точки зрения вклада в экономику цифровизации, с учетом требований по задержке сигнала подходят только два варианта: 5G SA и низкоорбитальная спутниковая связь типа Starlink. Причем обе только

¹¹ <https://arxiv.org/pdf/2511.19468>

¹² <https://www.cbo.gov/publication/58794>

¹³ <https://www.youtube.com/watch?v=FIQYU3m1e80>

¹⁴ <https://www.useluminix.com/reports/industry-analysis/data-centers-in-space>

¹⁵ <https://research.google/blog/exploring-a-space-based-scalable-ai-infrastructure-system-design/>

в конфигурации с MEC. С учетом ограничений покрытия 5G SA, даже на суше базовым вариантом обеспечения сетевой связности может быть только спутниковая группировка. Устойчивое покрытие 5G SA с MEC в малонаселенных районах – это гораздо большая фантастика нежели ЦОДы в космосе, а еще ведь остаются воздушный и водный виды транспорта, для них спутник вообще не имеет альтернативы.

Таким образом, мы располагаем только одной возможной конфигурацией системы, и имеем в этой системе один экстремально дорогостоящий, но необходимый элемент – спутники-вычислители. Парадоксально, но именно его наличие открывает возможности для удешевления остальных элементов системы. Это и существенно меньшие при наличии устойчивой связи требования к аппаратной платформе бортовой системы управления автономного транспортного средства, что критически важно ввиду массовости таких объектов управления, и экономия на пропускной способности каналов связи, как радио, так и наземных проводных (оптических), и практически полное отсутствие ограничений на выбор мест размещения корневых дата-центров – можно не строить ЦОДы в энергодефицитных регионах, что улучшает их экономику в разы.

Еще один фактор, возможно даже более критичный чем стоимость спутников-вычислителей – это эффективность системы сквозной оркестрации столь сложной коммуникационно-вычислительной системы. Как отмечено выше, вариант конфигурации с космическими ЦОД не усложняет, а наоборот упрощает задачу оркестрации, поскольку в этом случае сводится к минимуму либо полностью исключается необходимость использования наземных сетей широкополосной передачи данных и все элементы ИКТ-инфраструктуры могут быть развернуты одним хозяйствующим субъектом в рамках единой технической политики.

То есть теоретически кейс может сойтись, учитывая высокий удельный экономический эффект цифровизации именно в транспортно-логистической сфере. Нельзя сбрасывать со счетов и роль личности в истории, ведь все-таки есть личности, которым, как Королеву, важно не просто жить, а жить увлеченно.

Попытки реализации: SpaceX, xAI, TeraFab

В феврале 2026 г. SpaceX запросил у американского регулятора FCC разрешение на развертывание распределенного орбитального ЦОДа для AI-нагрузок общей мощностью 100 ГВт, состоящего из миллиона КА. Каждый такой КА AI Mini Sat массой около 1 тонны будет иметь мощность 100 кВт, что в 4-5 раз выше текущего поколения Starlink V2 Mini. Аппараты будут связаны лазерными каналами Starlink.

Теоретически это позволит объединить между собой как огромное число КА, так и существующий наземный вычислительный сегмент корневых дата-центров недавно поглощенного xAI с орбитальным. Во многом именно для этой вычислительной мегагруппировки SpaceX/xAI и Tesla анонсировали запуск проекта TeraFab – вертикально интегрированной фабрики по производству 2 нм GPU общей мощностью 1 ТВт/год – это 1 млрд AI-чипов (дизайн процессоров, литография, фабрикация, производство чипов памяти, корпусирование и тестирование). Стоимость первой фазы строительства оценивается в \$25 млрд, а первые серийные чипы появятся через 3-5 лет. В будущем предполагается масштабирование производства для покрытия потребностей не столько на Земле (на Tesla FSD/Dojo/robotaxi, Optimus, xAI/Grok прогнозируется только 20%), сколько в космосе (80% производственных мощностей на орбитальные ЦОДы, и освоение Луны и Марса).

Объединение SpaceX и xAI, а также замах Маска на миллионную вычислительную группировку (напомним, у Starlink сейчас около 10 тыс. активных КА и это считается пока самой масштабной орбитальной группировкой в мире) может быть PR-ходом для повышения рыночной оценки

объединенной компании до \$1,75 трлн в преддверии выхода Starlink на IPO¹⁶. При этом глобальные производственные возможности всех ведущих фабрик AI-процессоров не превышают 20 ГВт/год, лишь 2% от требуемого Маску 1 ТВт¹⁷. Очень амбициозные планы, даже с учетом того, что индустриальным партнером проекта выступит Intel (техпроцесс 1,8 нм, опыт в организации массового контрактного производства)¹⁸.

Тем не менее, на волне всеобщего интереса к AI, к орбитальным вычислениям проявляют живой интерес и другие техногиганты. Так Google совместно с Planet Lab анонсировал Project Suncatcher – орбитальный кластер из 81 аппарата на основе Tensor Processing Units (TPU)¹⁹, специально разработанных тензорных процессоров для ускорения машинного обучения и вычислений на основе нейронных сетей. Первый прототип такого КА планируется запустить в 2027 г.

В то же время Google инвестировал в стартап Starcloud²⁰, который подал заявку в FCC на развертывание группировки орбитальных ЦОДов из 88 тыс. КА. Один такой 5 ГВт кластер Starcloud-4 на орбите, в основе которого будут лежать вычислительные модули NVIDIA Space-1 Vera Rubin, будет иметь размеры 4x4 км - в жизнеспособность таких монолитных мегакопструкций на НОО верится с трудом, учитывая актуальную проблему космического мусора²¹.

Blue Origin Джеффа Безоса также планирует вычислительную мегагруппировку на орбите Project Sunrise²² из 51 600 КА с полной вертикальной интеграцией – в качестве платформы для AI КА выступит, очевидно, тяжелый орбитальный буксир Blue Ring, запуски будут осуществляться PN New Glenn, а передача данных на Землю - по лазерным каналам перспективной группировки TeraWave из 5 408 КА (общая пропускная способность 6 Тбит/с).

Китайский Adaspace также анонсировал AI-мегагруппировку из 2800 КА с вычислительными мощностями.

Очевидно, что часть заявленных проектов будет свернута, а оставшаяся – не достигнет изначальных сверхзавышенных целей, по-другому и быть не может. Но объективная потребность в относительно небольших спутниках-вычислителях в космосе существует – не для обучения LLM, а для управления физическими объектами с помощью AI-агентов, и вполне возможно будет реализована в не столь отдаленном будущем.

В завершении о будущем

AI, опирающийся на распределенную сете-вычислительную инфраструктуру с краевыми вычислителями в космосе – это уже вполне себе Skynet. Есть и достаточный уровень инвестиций, и реальная потребность и технические возможности для ее разработки и развертывания.

Чем плоха история со Skynet, мы знаем по фильму «Терминатор». Поэтому возвращаясь к началу данной публикации, еще раз отметим, что инновации в социально-экономических отношениях не менее, а более важны для того, чтобы технологические новшества не уничтожили человечество, а служили ему во благо. Но это тема отдельной серии публикаций, размером с несколько выпусков журнала Connect...

¹⁶ <https://finance.yahoo.com/markets/stocks/articles/spacex-going-public-1-75-202500656.html>

¹⁷ <https://finance.yahoo.com/sectors/technology/articles/heres-real-reason-spacex-teaming-131200500.html>

¹⁸ <https://www.forbes.com/sites/jonmarkman/2026/04/10/intel-joins-terafab-to-build-elon-musk-25b-ai-chip-project/>

¹⁹ <https://cloud.google.com/tpu>

²⁰ <https://starcloudinc.github.io/wp.pdf>

²¹ <https://habr.com/ru/companies/ruvds/articles/595695/>

²² <https://datacentremagazine.com/news/project-sunrise-blue-origin-data-centre-space-race>