

# А в «облаке» ли счастье? Сравниваем и разбираем варианты реализации реального и виртуального ЦОД применительно к задачам обработки сейсмических данных в ТЭК России.



*Изображение взято из общедоступного источника сети интернет, все права на него принадлежат автору изображения.*

В рамках международной геолого-геофизической конференции и выставки [ГеоЕвразия-2020](#), которая проходила в Москве 3-6 Февраля 2020 года, состоялся круглый стол «Облачные технологии в обработке сейсмических данных», организованный компаниями «Яндекс.Облако» и ООО «Сейсмотек». Выражаю благодарность организаторам мероприятия за его проведение, учитывая актуальность данной темы в последнее время. Но отведенных на презентации и дискуссию 100 минут было однозначно мало и осталось достаточно много вопросов, которые не вошли в обсуждение. Надеюсь, что похожее мероприятие будет повторено, но с большей длительностью и реальной двусторонней дискуссией. Текущее же событие стало неким катализатором к написанию данной статьи, которая уже давно планировалась.

Следует сразу отметить, что в данном документе, в первую очередь, речь идет о вариантах использования или строительства ЦОД применительно к ресурсоёмким задачам, связанным с обработкой и моделированием сейсмических данных 2D/3D/4C. Но эта информация также окажется полезной при рассмотрении аналогичных задач в других отраслях.

Проанализируем три наиболее вероятных варианта, которые рассматривают компании нефтегазового рынка (нефтяные, газовые, нефтесервисные и т.п.) на начальном этапе выбора направления развития счетных мощностей: строительство собственного центра обработки данных (далее по тексту — ЦОД), аренду физических серверов в существующем ЦОД одного из коммерческих провайдеров или аренду облачных счетных ресурсов у коммерческого провайдера? Помимо этого, в данной статье будет рассмотрено специальное предложение, действующее на Российском рынке услуг, от компаний ООО «Сейсмотек» и ООО «Яндекс.Облако»: сервис обработки сейсмических данных «*Prime.Cloud*» на облачном ЦОД «Яндекс.Облако».

### **Вариант №1: строительство собственного ЦОД**

Данная опция долгое время была доминирующей на Российском рынке (а ранее и на международном) применительно к стратегии развития вычислительных мощностей компаний ТЭК. Что же подразумевает данная опция? Какие особенности следует учитывать при ее рассмотрении? В чем ее плюсы и минусы? Рассмотрим и проанализируем основные моменты, а также приведем сравнение с другими предлагаемыми вариантами.

Начнем с анализа потребностей, то есть с попытки определить каковы будут параметры планируемого ЦОД? Сразу оговоримся, что мы рассматриваем решения сопоставимые с вариантами аренды ЦОД (аппаратного или виртуального). Поэтому варианты размещения стоек с оборудованием в кладовке, подвале, на чердаке без учета необходимых требований к работе такого оборудования и его безопасности для снижения затрат здесь не рассматриваются.

Так как определить потребность в ресурсах? Для одних компаний это может быть минимальное количество счетных ядер CPU (процессора), указываемое в тендерных заданиях, другие располагают более обоснованными потребностями, базирующимися на собственном опыте использования существующего ВЦ, требованием к вычислительным мощностям программного обеспечения (существующего или планируемого к закупке), пониманием объема проектов, которые будут обрабатываться на данном ВЦ или опыта коллег. Третьи собирают статистику потребности в мощностях, привлекая как специалистов отдела обработки сейсмических данных, так и специалистов отдела ИТ, имеющих опыт поддержки жизнедеятельности ВЦ... Поскольку каждая компания заинтересована в том, чтобы инвестиции в ЦОД окупались, возможно привлечение специалистов других отделов (например связанных с геологическим или гидродинамическим моделированием — так как эти процедуры тоже весьма ресурсоемки), которые могли бы обеспечить загрузку ЦОД, ведь простой оборудования в случае его недогрузки, будет приносить прибыль только компании поставщику электроэнергии вашего предприятия. Возможно это станет открытием для руководства предприятия: оборудование ЦОД потребляет значительное количество электроэнергии даже в режиме

минимальной загрузки и это стоит учитывать при планировании ЦОД. Тем более, что тарифы на электроэнергию только растут...

Рассмотрим еще одну важную деталь при проектировании ЦОД. Такой параметр как просто количество ядер CPU не может служить критерием выбора оборудования. Здесь должны учитываться модель процессора, тактовая частота, энергопотребление, адаптация программного обеспечения (некоторые задачи хорошо распараллеливаются и для них оптимальным вариантом будет процессор с большим количеством ядер, но с более низкой частотой. Другие - наоборот: требуют меньшего количества ядер, но более высокой частоты процессора). У большинства производителей вычислительного оборудования в настоящее время ограниченный выбор производителя CPU (основного вычислительного процессора) — это будет либо разработка компании Intel (подавляющее большинство серверов, используемых в ТЭК), либо разработки компании AMD. Обе компании американские. К огромному сожалению, сервера с отечественными процессорами являются большой экзотикой, как и сами отечественные процессоры (Эльбрус, Байкал). А программного обеспечения, адаптированного под эти процессоры и того меньше (по крайней мере в нефтегазовой сфере). Увы... Есть еще ряд другой экзотики, такой как сервера с процессорами архитектуры ARM или RISC, но геофизическое ПО для обработки сейсмических данных сейчас не разрабатывается под этот тип процессоров.

Далее, необходимо учитывать потребность в использовании другого типа счетных модулей: узлов с графическими процессорами (GPU), например, решениями от компании nVidia (например, решениями линейки TESLA K80/P40/P100/V100 и другими моделями этой линейки). Для целого ряда задач существуют алгоритмы, которые адаптированы с использованием графических ядер для ускорения вычислений (причем значительного ускорения счетных задач), но используемое программное обеспечение должно поддерживать такого рода операции, поскольку стандартные процедуры, адаптированные к использованию на классических CPU, не получают прироста производительности от использования дорогостоящих счетных узлов с GPU. Обратной стороной медали здесь будет повышенное энергопотребление счетных узлов с GPU, повышенные требования к системам охлаждения и значительно более высокая стоимость лицензий программных модулей с поддержкой расчетов на GPU.

Продолжим рассматривать конфигурацию счетных узлов. Помимо собственно центральных процессоров, в каждом счетном узле необходимо иметь также: оперативную память (обычно определяемую в количестве гигабайт на вычислительное ядро CPU или как общее количество гигабайт на процессор), наличие дисковых накопителей (классических HDD или более современных и быстрых SSD), сетевых интерфейсов (в зависимости от скорости и типа сети, используемой в ЦОД: от различных типов Ethernet сети до дорогих оптических адаптеров, поддерживающих сети с более высокой пропускной способностью, например InfiniBand EDR). Также необходимо отметить требование к наличию как минимум двух блоков питания с возможностью горячей замены, достаточной мощности для выбранной

конфигурации счетных узлов. Зачем нужно дублирование? Чтобы не пришлось в процессе важного расчета его прерывать из-за перегоревшего блока питания и вышедшего из строя вследствие этого счетного узла ЦОД.

Предположим, что компания определилась с параметрами счетных узлов. Означает ли это, что можно переходить к следующему этапу - выбору поставщиков оборудования? Нет. Остались еще важные вопросы, которые необходимо решить. Ранее упоминались сетевые адаптеры, которые будут использоваться в вашем ЦОД. Допустим вы остановились на одном из вариантов, но для использования и строительства сети, помимо самих сетевых адаптеров в счетных узлах, понадобятся коммутаторы, которые позволят объединить все счетные узлы в кластер. Что необходимо помнить при их выборе? Помимо совместимости (обычно рекомендуется один производитель для сетевых адаптеров и коммутаторов), необходимо опять предусмотреть их резервирование (с использованием схем построения сети с дублированием), ведь вы же не хотите, чтобы в процессе выполнения счета на кластере, который может продолжаться несколько недель, система полностью вышла из строя из-за отказа единственного коммутатора? Такого рода «сюрпризы» точно не обрадуют менеджмент вашей компании, ведь простой ЦОД обходится весьма дорого. В буквальном смысле этого слова.

Продолжим. С параметрами сети ЦОД мы определились, что же еще необходимо учесть? Конечно конфигурацию и тип современной СХД (системы хранения данных) с достаточной емкостью, запасом по пропускной способности сетевых интерфейсов и наличием достаточного дискового пространства, как для хранения данных, так и для обеспечения отказоустойчивости системы в случае выхода из строя одного или нескольких дисков. Существуют различные решения — аппаратные СХД, программные СХД с распределенной файловой системой, смешанные типы СХД — выбор здесь лучше оставить специалистам. Цены также могут сильно различаться в зависимости от типа используемых дисков (классические HDD или быстрые SSD диски), объема памяти, опций СХД (как программных, для которых необходимо приобретать дополнительные лицензии, так и аппаратных), возможностями расширения СХД и т.п. Следует обязательно учитывать, что количество данных в процессе счета только растет, поэтому необходимо предусмотреть возможность ее расширения. Еще одним важным моментом является следующее: поскольку в СХД будет задействовано наибольшее количество дисков по сравнению с другими серверами ЦОД, следует помнить, что диски являются расходным материалом, который требует регулярной замены в следствии износа и выхода из строя. Затраты на их замену лучше сразу заложить в бюджет обслуживания ЦОД.

Все мы рассмотрели? Конечно нет! Мы рассматриваем вариант строительства собственного ЦОД, а это гораздо более «увлекательный» и «затягивающий» процесс, чем могло показаться на первый взгляд. Что же еще осталось учесть? Систему обеспечения бесперебойного питания ЦОД, которая гарантирует его функционирование в случае отказа основного источника питания. В первую очередь речь идет об ИБП (источниках бесперебойного

питания или UPS), которые должны взять на себя обеспечение питания ЦОД в случае аварии на основной линии подачи электроэнергии, а также резервные источники питания (например, альтернативная резервная линия электропитания от другого поставщика электроэнергии или собственный дизель-генератор). Еще раз обращаю ваше внимание на то, что отключение ЦОД из-за проблем с электропитанием точно не будет приятным сюрпризом.

Еще одна немаловажная деталь современного ЦОД – это система кондиционирования и охлаждения помещения ЦОД и счетного оборудования. Здесь существует масса вариантов от классической воздушной системы охлаждения до очень дорогих жидкостных (погружных и полупогружных) систем охлаждения. Последние могут по стоимости превзойти все счетные узлы... Стоит также учитывать, что система охлаждения должна по мощности превосходить потребляемую мощность всех узлов ЦОД и другого оборудования ЦОД, работающего на максимальной мощности потребления электроэнергии. Соответственно получаем «удвоенный» вариант энергопотребления, который должен быть заложен при проектировании ЦОД. И, как говорилось ранее, чтобы не допустить выход из строя счетного оборудования в случае отказа системы кондиционирования, лучше заложить дополнительные резервные устройства кондиционирования, которые можно задействовать в случае прекращения работы части основных систем охлаждения.

На что еще следует обратить внимание при проектировании ЦОД? Это системы резервного копирования данных. Выбор здесь большой — наиболее распространенный вариант: ленточные библиотеки или библиотеки оптических дисков большой емкости. Другим важным элементом любого ЦОД являются системы безопасности и контроля: системы VPN и защиты данных, системы видео наблюдения, системы контроля доступа в ЦОД, система пожаротушения, система освещения.

На этом этапе мы приближаемся к завершению предварительного планирования ЦОД и можем передать собранные данные и пожелания в компанию, которая будет профессионально заниматься разработкой проектов по строительству ЦОД. Почему проектов, а не проекта? Потому что помимо основного проекта будут еще проекты подготовки помещения, проект размещения оборудования, проект пожарной системы, проект системы электропитания и многие другие. На их подготовку может уйти достаточно много времени. Дополнительно необходимо учитывать, что ряд проектов требуют получения согласований в компетентных государственных органах. И эти процедуры далеко не быстрые.

Что еще может повлиять на процесс подготовки проекта ЦОД? Например, выбор материалов для отделки помещения, напольное покрытие (в случае приподнятых полов для размещения кабелей электропитания и сетевых кабелей под поверхностью пола), выбор дизайна и моделей стоек для размещения оборудования — необходимо учитывать удобство расположения счетных узлов, их дизайн с точки зрения оптимального охлаждения, общий вес стойки и общее потребление электричества оборудованием, расположенным в каждой из стоек.

Какие дополнительные подводные камни могут возникнуть при проектировании ЦОД? Если только помещение под ЦОД не было подготовлено и построено специально для этих целей, то стоит отметить следующие важные моменты, значительно усложняющие проект:

- отсутствие требуемого подвода электричества достаточной мощности;
- проблемы с весовой нагрузкой на перекрытия пола помещения ЦОД;
- проблемы с гидроизоляцией помещения от внешних источников воды: систем отопления и водоснабжения зданий, протечек с расположенных выше этажей и крыши здания;
- невозможность размещения внешних блоков кондиционеров на стенах здания;
- отсутствие возможности создания резервных источников электропитания (например, запрет на установку дизель-генераторов в городской среде);
- ограничения по размещению систем бесперебойного питания (ИБП) из-за значительного веса блоков аккумуляторов и превышения допустимой нагрузки на перекрытия;
- проблемы с организацией зон безопасного доступа к инфраструктуре ЦОД;
- Стоит уточнить входит ли монтаж оборудования в стоимость его приобретения, а также существует ли возможность настройки другого оборудования (например, построения сети)? Будут ли эти опции учтены в договоре?
- Не стоит забывать и о программном обеспечении (лицензиях), причем этот вопрос распространяется не только на прикладное специализированное ПО для обработки сейсмических данных или гидродинамического моделирования, но и лицензиях на операционные системы, лицензии на конфигурации модулей оборудования, конфигурацию и опции СХД, сетевого оборудования — об этом многие забывают.
- Забывают также и о необходимости ввести в штат как минимум двух сотрудников поддержки и администрирования своего ЦОД. Почему как минимум двух, а не одного? Ведь вы же не планируете выключать ЦОД на время отпуска или болезни одного из системных администраторов? Причем это должны быть специалисты знакомые с оборудованием, которое вы планируете использовать. Обучение в процессе работы точно не лучший вариант в случае поддержки ЦОД.
- Возможно вам потребуется создать свой склад резервного оборудования: диски, блоки питания серверов, аккумуляторы для ИБП, дополнительные кабели, картриджи для систем архивирования информации — следует учитывать, что в случае выхода из строя оборудования, счет будет вестись на минуты и ожидать поставку необходимой детали на следующей неделе будет не самым лучшим вариантом. Точно.
- Допустим вы приняли решение, оборудование выбрано, проект ЦОД готов, но помимо задержек при подготовке помещения (сроки завершения ремонта всегда трудно предсказуемы, даже если на бумаге все было

идеально распланировано) — далее вас, скорее всего, будет ожидать «сюрприз» в виде возможности получения оборудования только на заказ через 4 - 5 или n-ное количество недель. Получить что-то здесь и сейчас не получится, разве что ваш ЦОД помещается в одну стойку с СХД и ИБП (но мы же не о таком варианте говорим — не правда ли?).

- При выборе вендора оборудования следует учитывать всем знакомую истину: «скупой платит дважды» (или трижды в современных реалиях), правда с поправкой на текущую ситуацию на рынке. Можно выбрать оборудование крупных, известных вендоров, например, Dell, HPE, IBM (и других), это точно будет самый дорогой и проверенный вариант. Он был бы, наверное, идеальным, если бы не «пустяк», а именно: санкции, введенные в отношении Российского ТЭК. То есть может случиться так, что вендор просто не сумеет поставить запчасти, обновить лицензии или сделать что-то еще для вашего оборудования (а возможно и даже поставить его в установленные сроки). Эти риски необходимо учитывать. Альтернативой является приобретение оборудования, произведенного в Китае. В настоящее время подавляющее большинство оборудования производится в Поднебесной. Местные игроки могут предложить различные варианты: начиная от неизвестного местного гаражного производителя, поставляющего оборудование с локализацией интерфейса исключительно на китайском (Оборудование для локального рынка — скорее всего наиболее дешевый вариант, но без какой-либо поддержки. Возможно это хороший повод выучить еще один язык? Или попытаться выучить? Или все же лучше не пытаться?) до весьма солидных и проверенных решений, которые практически ничем не уступают американским решениям. Здесь интерфейс в обязательном порядке будет на английском, как и документация. А возможно даже на русском.
- Следует также уточнить какая гарантия поставляется на оборудование и есть ли у компании производителя дистрибьюторы и представители технической поддержки в России (и в городе, где планируется размещать ЦОД).

Предположим, что все прошло хорошо и ваша компания стала счастливым обладателем собственного ЦОД. Попробуем сравнить все плюсы и минусы этого решения (опять же применимо к нашим задачам по обработке сейсмических данных):

### **Плюсы:**

1. Это ваш собственный ЦОД (все сертификаты и документация на оборудование и ПО у вас на руках и готово к предоставлению в соответствии с тендерными процедурами) и только вы его контролируете (задания запускают только сотрудники вашей компании).
2. Для загрузки данных или выгрузки результатов обработки не нужно ждать (до 1 дня при передаче 1 Тб. данных по внешнему интернет каналу

- 100 Мбит/с) — данные загружаются / выгружаются с локальных носителей или по более быстрой локальной сети.
3. Вы физически контролируете расположение оборудования и данных, которые на нем находятся, то есть теоретически шанс утечки данных ниже.
  4. Вы сами обеспечиваете загрузку ЦОД задачами и обеспечиваете ее распределение между счетными узлами и по времени.
  5. Вы можете пригласить потенциальных клиентов или партнеров и показать им новое, сверкающее помещение ЦОД со множеством мигающих лампочек, красивых стоек, и т.п. То есть с точки зрения пиара — это, конечно, плюс.
  6. Если вы приобрели новейшее оборудование (последние модели процессоров, огромное количество оперативной памяти, очень быструю сеть, быструю систему СХД с огромным объемом быстрых дисков, потратив на все это очень-очень-очень много денег) — то ваш ЦОД точно будет считать быстрее, чем стандартный вариант аренды аппаратного или виртуального ЦОД. Правда возникает вопрос — а оно вам точно было нужно?

### **Минусы:**

1. Строительство собственного ЦОД процесс не быстрый, требующий весьма больших (или даже очень-очень-очень больших) начальных затрат, которые, к огромному сожалению, скорее всего не окупятся в среднесрочной перспективе. Ведь гарантировать правильную загрузку ЦОД, близкую к 100% практически невозможно (особенно если мы говорим об обработке сейсмических данных). А любая недозагрузка ЦОД или простой (авария, сбой, ошибки в конфигурации) только уводят вас в минус.
2. Помимо начальных затрат, не стоит забывать и о постоянных новых затратах, которые появятся у вашей компании: счета за электричество (значительные, поскольку энергопотребление у ЦОД достаточно велико и оно только будет увеличиваться в моменты пиковой загрузки, не стоит списывать и систему охлаждения, работающую в режиме 24x7x365), выплаты заработной платы двум (или более) сотрудникам поддержки и администрации ЦОД, затраты на ремонт ЦОД (ремонт оборудования, замену дисков, сетевого оборудования, блоков питания).
3. К огромному сожалению, компьютерное оборудование имеет тенденцию устаревать и даже если вы приобрели новейшие, на момент создания ЦОД, модели компьютерного оборудования, то через 4-6 лет оно устареет и уже не будет считаться конкурентоспособным на рынке. Потребуется замена. А это новые затраты. А если оборудование изначально было не совсем новым — то конкурентоспособный срок службы снижается еще больше.



4. Два сотрудника службы поддержки и администрирования вашего ЦОД не могут конкурировать со службой поддержки крупных коммерческих ЦОД, с временем их реакции на инциденты, временем решения проблем. Здесь вы будете зависеть от квалификации сотрудников, их обучения, самого оборудования.
5. Представим, что на момент проектирования ЦОД был выбран потолок в 1024 ядра, который рассматривался как максимально необходимый. Но на рынке появились тендеры, в которых прописано, что для участия компании исполнителю требуется 2048 или 4096 или 10240 ядер. Что же делать? Модернизировать ЦОД? Срочно закупать оборудование? Подводить новые электрические мощности? Менять кондиционеры? В виртуальном ЦОДе можно просто указать, что для нового счета потребуются дополнительные счетные узлы: 10, 100, 1000, 10000. А вот с реальным собственным ЦОД такой вариант не пройдет.
6. Аналогично с памятью (оперативной или дисковой) — то что в виртуальном ЦОД потребует пару щелчков мышью (и некоторых дополнительных денежных затрат), в реальном, увы, повторить не получится.
7. В собственном ЦОД модель денежных потоков будет примерно следующей: сначала очень сильный минус (начальные инвестиции в оборудование и помещение), затем попытка восполнить эти затраты за счет выполнения проектов, но не стоит забывать и о постоянных расходах на поддержание ЦОД (электроэнергия, зарплаты ИТ специалистов, ремонт), и в тоже время исполнителям проектов тоже нужно платить зарплату.
8. Реальный ЦОД невозможно получить «сейчас» и даже «завтра». Это длительный, затратный процесс. Это необходимо учитывать при планировании такого проекта.

## **Вывод:**

Было время, когда я отстаивал вариант строительства собственного ЦОД в компании. Но ситуация кардинально изменилась с тех пор. Я более детально познакомился с моделями и сервисами виртуальных ЦОД, предлагаемых на Российском рынке и, на мой взгляд, строительство в настоящее время своего собственного ЦОД компанией, занятой в обработке сейсмических данных, может быть рассмотрено только в трех исключительных случаях:

1. Прямое распоряжение руководства компании о строительстве собственного ЦОД (хотя здесь я бы постарался отговорить руководство от такого решения, представив экономическую модель такого проекта в сравнении с проектом использования стороннего виртуального ЦОД).
2. Наличие не лимитированного количества свободных денежных средств у компании и приобретение ЦОД, как альтернатива покупки, как то шестой яхты или третьего футбольного клуба (в этом случае я полностью

поддерживаю такое решение и готов предоставить свои услуги такой компании в реализации данного проекта ;-))

3. Наличие в компании очень строгой службы безопасности, которая имеет 100% влияние на руководство компании. В этом случае, люди, которые могут быть не до конца знакомыми с современными способами защиты информации (или вообще не знакомыми с ними и работающими по старинке) могут убедить руководство, что использование стороннего ЦОД является угрозой прямой умышленной утечки информации (тезисно: интернет — это зло, компьютер не должен иметь возможность связываться с другими компьютерами по сети, флешки и внешние диски запрещены, все входящие и выходящие из ЦОД сотрудники должны досматриваться). Печально, но такие ограничения, действительно существуют в ряде компаний. Жаль...

### **Вариант №2: аренда физических серверов в ЦОД**

Наиболее старая опция на рынке ЦОД. В этом случае, компании нет необходимости строить свой ЦОД на территории предприятия — вместо этого приобретаются права на временное использование серверов, сетевого оборудования и СХД, приобретенных провайдером услуг (или самой компанией) и размещенных у него в ЦОД. Для доступа используются каналы связи — чаще всего это будет комбинация собственной сети предприятия, каналы связи интернет провайдера и сетевая инфраструктура ЦОД. Именно это обстоятельство и вызывает опасения служб безопасности крупных нефтегазовых компаний: как передавать данные компании через интернет? А вдруг кто-то их перехватит? Здесь необходимо добавить, что некоторые сетевые провайдеры предлагают опцию выделенной аренды сети (в первую очередь оптоволоконной) от здания компании до одного из крупных сетевых узлов, например, М9 или даже непосредственно до ЦОД (если провайдер представлен там). В этом случае компания получает свой собственный канал связи, который повышает безопасность передаваемой и получаемой информации. Но опция эта не из дешевых. Опасения распространяются и на сами сервера — они же ведь расположены за пределами вашей компании — а что если системный администратор или кто-то еще из сотрудников, работающих в ЦОД скопируют данные? О том, что данные в таком случае в обязательном порядке шифруются, как при передаче, так и при хранении (опционально) они забывают. Есть ли здесь повод для беспокойства? Теоретически — да, поскольку много чего может произойти на территории, не подконтрольной заказчику. Но такие же проблемы могут возникнуть и из-за недобросовестных сотрудников в самой компании или в случае передачи информации подрядной организации. Так что же? Не пользоваться услугами подрядчиков? Ах, да! Подрядчики подписывают соглашение о неразглашении. Стоп, но подписывая договор аренды вы также подписываете соглашение о конфиденциальности данных. Так в чем же дело? Продолжать можно долго. Каждая из сторон будет приводить свои доводы, но все будет зависеть от провайдера, с которым вы работаете. Если это крупная

компания, дорожающая своей репутацией, с большим опытом работы, собственной службой безопасности, имеющей опыт работы с крупными клиентами рынка — думаю, что проблем в этом случае не возникнет. Если же вы планируете работать с провайдером, только что вышедшем на рынок и имеющем юридический адрес, зарегистрированный в подвале пятиэтажки на окраине маленького городка далеко от Москвы — в этом случае, я соглашусь с мнением службы безопасности. Качественные и надежные услуги никогда не будут самыми дешевыми на рынке.

Посмотрим в чем же отличие от первого варианта? Во-первых, нет необходимости создавать собственное помещение для размещения ЦОД, проводить планирование, ремонт, отделку, закупать и монтировать оборудование, обеспечивать бесперебойное электропитание... То есть значительная экономия средств. Во-вторых, нет дополнительных ежемесячных платежей за потребленную электроэнергию, нет необходимости держать в штате ИТ специалистов по поддержке оборудования ЦОД, не нужен ремонт и обновление оборудования — это все покрывается договором аренды оборудования в ЦОД. Опять, еще одна важная деталь — нет очень (или как говорилось ранее очень-очень-очень) высоких первоначальных затрат на создание ЦОД (оборудование, помещение, охрана, администрирование).

Но появляются новые затраты: ежемесячный счет на аренду оборудования (обычно контракты на аренду заключаются на длительный срок с ежемесячными платежами), аренда канала связи для доступа к вашему ЦОД, аренда дискового пространства в СХД, доступной для ваших счетных узлов.

Попробуем опять рассмотреть недостатки и преимущества такого варианта:

### **Плюсы:**

1. Нет необходимости строить собственный ЦОД (подготавливать помещение, закупать и монтировать оборудование, заниматься поддержкой и ремонтом). Нет огромных начальных затрат.
2. Более низкая начальная стоимость доступа к вычислительным ресурсам на правах аренды и гораздо более сжатые сроки начала работы такого ЦОД по сравнению с первым вариантом.
3. Арендуются физические сервера — тем самым гарантируется, что только сотрудники вашей компании их будут использовать и не придется делиться ресурсами с другими компаниями.
4. Скорость работы таких серверов может быть выше по сравнению с виртуальными серверами (конфигурация согласовывается с заказчиком, хотя и зависит от возможностей провайдера ЦОД — типа оборудования которое было закуплено).
5. Работа идет практически как на собственном ЦОД, только расположенном где-то еще. Важно наличие достаточно быстрого канала связи с ЦОД.
6. Возможность аренды серверов с GPU (графическими процессорами — ускорителями расчетов) — но обычно это значительно более дорогая

опция по сравнению с арендой обычных CPU серверов. Важно помнить, что просто графические ускорители не ускорят работу системы, если нет лицензий на специализированное программное обеспечение с поддержкой расчетов на таком оборудовании!

### **Минусы:**

1. Необходимо будет выбрать из возможных вариантов оборудования, имеющегося в наличии у провайдера ЦОД (возможно оно не совсем совпадает с тем, что вам хотелось бы использовать, если бы закупили оборудование сами).
2. Теоретически возникают риски по безопасности передачи и хранения данных, поскольку они будут покидать периметр компании.
3. Необходимо согласовывать с провайдером специализированного программного обеспечения возможность его установки и лицензирования на арендованном ЦОД (аренда лицензий на время проекта или предоставление неким образом доступа к уже закупленным в компании лицензиям). Важно: в подавляющем большинстве лицензионных договоров, особенно с иностранными производителями программного обеспечения, в лицензии указано, что она не может быть использована в арендованном ЦОД. Так что проверяйте и согласовывайте данную опцию заранее!
4. Скорость загрузки исходных данных и выгрузка результатов расчетов: многие забывают, что объемы сейсмических данных достаточны велики и объем в 1 Тб. полевых сейсмических данных по 3D проекту никого не удивляет. В локальном ЦОД — данные могут быть загружены по быстрой локальной сети, а в удаленном ЦОД скорость загрузки будет регламентирована скоростью канала данных провайдера и при загрузке такого объема через канал 100 Мбит/с (в идеальном случае, если ничего дополнительного не передается) придется подождать около 24 часов, чтобы данные попали из офиса в ЦОД. Выгрузка будет занимать также достаточно много времени, которое будет считаться временем простоя ЦОД.
5. Поскольку речь идет об аренде на определенный временной интервал, в данной практике обычно кратный месяцу, то крайне вероятна ситуация когда-либо не хватит времени аренды для выполнения проекта и ее придется продлевать, либо наоборот ЦОД будет недозагружен. Ведь в данной модели работы оборудование предоставляется вам в аренду, и вы должны каким-то образом обеспечить его загрузку. Если серверов мало — обработка может длиться долго, если много, то хотя счет будет идти быстрее, но придется платить и за время простоя. Ведь не все процедуры требуют полной загрузки процессоров и памяти.
6. Вы арендовали оборудование, работаете над проектами и вдруг понимаете, что мощностей текущей конфигурации недостаточно. Нужно больше счетных узлов, больше процессоров, больше памяти. Конечно

можно оставить заявку провайдеру ЦОД, возможно через некоторое время они сумеют добавить несколько серверов или закупят дополнительную память для установки в существующие. Но для модификации оборудования придется остановить, произвести установку дополнительных модулей, переконфигурировать сервера и перенастроить кластер в целом. То есть опять речь идет о простом, дополнительных затратах и, как следствие, потери части прибыли.

## **Выводы:**

Данная альтернатива решения проблемы с заменой строительства собственного ЦОД, на аренду физических серверов у одного из коммерческих провайдеров на рынке может быть интересна, но, в последнее время, такой вид аренды вытесняется третьим вариантом — арендой виртуальных серверов и созданием настоящего виртуального ЦОД. Об этом можно будет почитать в следующем разделе документа. Здесь остается только добавить, что, хотя данный вариант более интересен и прост по сравнению со строительством собственного ЦОД, он может быть рекомендован только в случае если компании необходимы какие-то особенности серверов, которые не могут быть предоставлены виртуальными серверами или требуется работа непосредственно с физическими, а не виртуальными вычислительными узлами. Стоит перед принятием решения еще раз проанализировать экономику проекта и посмотреть сумеет ли такой проект себя окупить и начать приносить прибыль. Или он уйдет в минус из-за недозагрузки оборудования.

## **Вариант №3: аренда виртуальных серверов в ЦОД**

Наиболее востребованная и активно развивающаяся опция на рынке ЦОД. Под виртуальными серверами подразумевается, что компания будет арендовать не конкретный физический сервер, расположенный в ЦОД провайдера, а виртуальные машины, реализованные на базе различных средств виртуализации, когда заказчику предоставляются конфигурации серверов, параметры которых могут быть изменены в течении 5 минут (а чаще и еще быстрее самим пользователем). Нужно вам не 10, а 100 серверов — пожалуйста! Необходимо увеличить память каждой машины с 4 до 16 Гб. на ядро процессора — нет проблем. Заканчивается место на диске — два клика мышкой и размер увеличен в два, три, десять — практически сколько угодно раз. А раз все это виртуально — то удобно хранить снимки и резервные копии машин, легко подстраивать счетные узлы под требуемые задачи. Если мы говорим о крупных игроках на рынке, таких как [Yandex.Cloud](#), [Mail.ru Cloud Solutions](#) или ... достаточно крупных игроков на рынке сейчас много (я говорю только о Российском рынке, хотя есть конечно и международные гиганты: *Google, Oracle, Amazon, Microsoft* – все они сейчас имеют сервисы предоставления виртуальных серверов, но расположены они за пределами России). Так что каждый может выбрать вариант, который лучше подходит

именно ему. Расценки тоже варьируются, как и опыт взаимодействия с отраслью ТЭК. Но поскольку это крупные игроки, в других областях их ресурсы практически безграничны. Что может быть очень полезным для рынка обработки сейсмических данных — поскольку ваш виртуальный ЦОД может быть, как маленьким, так и огромным и все это только зависит от ваших потребностей. Добавим сюда дополнительный бонус — многие провайдеры предлагают сейчас поминутную оплату сервиса аренды — а это может означать сокращение простоя и отсутствие оплаты не используемых сервисов.

Проведем сравнение с ранее рассмотренными вариантами создания ЦОД. Здесь стоит отметить, что сохраняются все преимущества варианта №2 — то есть отсутствие необходимости закупки дорогостоящего оборудования (огромных первоначальных затрат), инфраструктуры, строительства помещения для ЦОД, отсутствие ежемесячных платежей за электричество, потребляемое вашим ЦОД и системой охлаждения ЦОД (~50% от затрат на электричество в классическом варианте). Также отсутствует необходимость введения в штат двух системных администраторов, отвечающих за работу ЦОД, отсутствие затрат на ремонт и обслуживание оборудования. В дополнение к варианту №2 появляется возможность гибко распределить затраты: оплата идет только за то время, пока сервисы используются и есть возможность варьировать мощность ЦОД в зависимости от потребностей (причем это происходит практически мгновенно — нет необходимости ждать поставок оборудования, его ввода в эксплуатацию, настройки обновленной конфигурации ЦОД).

Конечно есть и некоторые минусы, они практически идентичны тем, которые указаны для варианта №2: это ежемесячные платежи за аренду выбранной конфигурации (правда в этом случае размер платежа может варьироваться в зависимости от используемых ресурсов), аренда канала связи для доступа к вашему ЦОД, задержка при передаче большого объема информации (обычно на начальном этапе загрузки исходных данных для проведения обработки и на этапе выгрузки результатов, хотя объемы здесь обычно значительно меньше). Также, в некоторых редких случаях, вам может потребоваться доступ непосредственно к «железным» серверам — поскольку по сравнению с ними виртуальные все же немного медленнее (однако это может быть компенсировано путем задействования большего количества виртуальных серверов).

Рассмотрим недостатки и преимущества этого варианта:

### **Плюсы:**

1. Нет необходимости строить собственный ЦОД (подготавливать помещение, закупать и монтировать оборудование и инфраструктуру, заниматься поддержкой и ремонтом).
2. Значительно более низкая начальная стоимость доступа к вычислительным ресурсам на правах аренды и гораздо более сжатые

сроки начала работы такого ЦОД по сравнению с первым и даже со вторым вариантом.

3. Арендуются виртуальные сервера — заказчик сам регулирует требуемую в конкретный момент времени конфигурацию, которая может быть легко изменена в зависимости от текущих потребностей. Также возможно изменять количество серверов, объемы дисковой и оперативной памяти.
4. Виртуальные сервера могут предоставляться с поддержкой GPU (графических ускорителей), аналогично физическим серверам если ваше программное обеспечение поддерживает опцию расчета с использованием графических процессоров.
5. Тарификация идет с поминутной (а иногда и с посекундной) оплатой. То есть заказчик действительно платит именно за те ресурсы, которые он использует в отличие от первых двух вариантов, когда оборудование может простаивать при отсутствии загрузки.
6. Существуют дополнительные опции, которые могут снизить затраты на эксплуатацию виртуальных серверов, например, специальный вид серверов с ограниченным жизненным циклом (не более 24 часов). Стоимость аренды таких серверов значительно ниже обычных (данная опция доступна не у всех провайдеров).

### **Минусы:**

1. Виртуальные сервера немного медленнее реальных «железных» серверов, но этот недостаток компенсируется возможностью легко и быстро менять их конфигурацию, что недоступно для обычных серверов.
2. Теоретические риски по безопасности передачи и хранения данных, поскольку они будут покидать периметр компании. Здесь стоит отметить, что такие же риски всегда существуют при заказе подрядных работ, но от последних компании не отказываются. Также требования безопасности на ЦОД крупных провайдеров могут быть даже более строгими, по сравнению с регламентами, принятыми в вашей компании.
3. Необходимо согласовывать с провайдером специализированного программного обеспечения возможность его установки и лицензирования на арендованном ЦОД (аренда лицензий на время проекта или предоставление неким образом доступа к уже закупленным в компании лицензиям). Важно: в подавляющем большинстве лицензионных договоров, особенно с иностранными производителями программного обеспечения, в лицензии указано, что она не может быть использована в арендованном ЦОД. Так что проверяйте и согласовывайте данную опцию заранее!
4. Скорость загрузки исходных данных и выгрузка результатов расчетов: многие забывают, что объемы сейсмических данных достаточно велики и объем в 1 Тб. полевых сейсмических данных по 3D проекту никого не удивляет. В локальном ЦОД — данные могут быть загружены по быстрой

локальной сети, а в удаленном ЦОД скорость загрузки будет регламентирована скоростью канала данных провайдера и при загрузке такого объема через канал 100 Мбит/с (в идеальном случае, если ничего дополнительного не передается) придется подождать около 24 часов, чтобы данные попали из офиса в ЦОД. Выгрузка будет занимать также достаточно много времени, которое будет считаться временем простоя ЦОД. Но, справедливо будет заметить, что на время выполнения проекта, которое может измеряться месяцами, данное ограничение скажется незначительно.

## **Выводы:**

Вариант аренды виртуальных серверов является наиболее предпочтительным по сравнению с вариантом строительства собственного ЦОД. В мировой практике все больше нефтегазовых и нефтесервисных компаний выбирают именно вариант аренды виртуальных серверов, поскольку по соотношению цена/качество он однозначно является наиболее привлекательным.

## **Практическое применение**

Рассмотрим вариант практического применения технологий облачных виртуальных ЦОД для проведения обработки сейсмических данных. Поскольку одно из основных направлений проекта «[AGB Corp.](#)» связано с вопросами [импортозамещения](#) — проанализируем отечественные решения в этой области.

С одной стороны, в данном сегменте не так уж много (хотя правильнее было бы сказать мало) отечественных решений, обеспечивающих полный цикл обработки сейсмических данных 2D/3D применительно к нефтегазовому сектору и доступных для приобретения коммерческими заказчиками в России.

Точнее сказать, пока на рынке представлен только один отечественный пакет\*, который может составить конкуренцию зарубежным решениям с адаптацией для использования в облачном ЦОД. Это собственная разработка компании [ООО «Сейсмотек»](#), пакет обработки сейсмических данных 2D/3D/4D/3C/4C «[Prime](#)» и его версия «[Prime.Cloud](#)», специально адаптированная для развертывания и работы в облачном ЦОД.

Кратко о данном пакете можно сказать, что «[Prime](#)» - это полная интерактивная система обработки сейсмических данных, включающая все необходимые процедуры от чтения исходных данных до построения окончательного мигрированного изображения среды и вывода его на печать. Особенностью обработки, реализованной в «[Prime](#)» является собственный алгоритм построения глубинно-скоростной модели среды, являющийся основой для многих других процедур обработки, например, миграции. Концепция «[Prime](#)» разработана доктором физ.-мат. наук В.М. Глоговским. Дополнительная



информация о системе «*Prime*» доступна на сайте компании ООО «*Сейсмотек*». Следует отметить, что система разработана для использования на платформе Linux x64 (RHEL / CentOS 6.x/7.x) с поддержкой алгоритмов распределенных (кластерных) вычислений. Пакет внесен в реестр отечественного программного обеспечения Минкомсвязи: [№2561 от 23 Декабря 2016 года](#).

В чем же особенность системы применительно к использованию в виртуальном ЦОД? В отличие от большинства конкурентов, которые допускают возможность применения своего ПО на облачных ЦОД и компаний провайдеров технологий, который адаптируют и оптимизируют настройки ЦОД к конкретному бинарному пакету, компания ООО «*Сейсмотек*» использовала многолетний опыт сотрудничества с одной из крупнейших ИТ компаний в России — компанией *Яндекс*, а начиная с 2018 года с компаний ООО «*Яндекс.Облако*». В результате этого сотрудничества в исходные коды системы были внедрены программные решения поддерживающие оптимизированные алгоритмы для работы системы на облачных ресурсах компании ООО «*Яндекс.Облако*». Среди прочих можно упомянуть возможность настройки и создания счетных узлов виртуального центра обработки данных прямо из системы «*Prime.Cloud*» под конкретные счетные задачи. Для конечного пользователя данной системы это означает значительную экономию времени и средств при использовании «*Prime.Cloud*» для обработки сейсмических данных на виртуальных серверах провайдера ООО «*Яндекс.Облако*». Ведь эти сервера будут созданы именно на время расчетов и по окончании последних и сохранении полученных результатов они будут автоматически остановлены (а при необходимости и удалены), что подтверждает предложенную концепцию оптимизации средств — клиент платит только за ресурсы, которые использует и простой арендованных ресурсов сводится практически к нулю.

Опыт сотрудничества с компанией ООО «*Яндекс.Облако*» позволил провести действительно глубокую адаптацию пакета «*Prime.Cloud*» к использованию на облачных вычислительных ресурсах провайдера ООО «*Яндекс.Облако*».

Возможно ли перенести полученный опыт для адаптации системы «*Prime*» под решения других облачных провайдеров — да, возможно. Конечно это займет некоторое время, но, в случае заинтересованности клиентов, уверен это можно будет сделать.

То есть вышеприведенный пример подтверждает, что в России могут быть реализованы высокотехнологичные проекты, не уступающие западным решениям основанные на собственных разработках. А кооперация между разработчиками программного обеспечения для обработки сейсмических данных и провайдерами облачных вычислительных ресурсов могут приводить к значительной оптимизации использования ресурсов конечными пользователями этих систем. Следует отметить, что данное преимущество распространяется и на сервисные работы, выполняемые компанией ООО «*Сейсмотек*», поскольку оптимизация использования счетных ресурсов (которые составляют значительную часть статьи расходов любого сервисного контракта на обработку сейсмических данных) здесь также будет применяться.

Ознакомится с географией применения программного обеспечения «Prime» можно на [этой странице](#) сайта. Если вас заинтересовала возможность использования ПО «Prime», «Prime.Cloud» на ресурсах «Яндекс.Облако», или вы хотите заказать сервисные услуги по обработке сейсмических данных в компании ООО «Сейсмотек» - вы можете обратиться к [автору](#) статьи (контактная информация представлена ниже) или обратиться непосредственно в компанию [ООО «Сейсмотек»](#) (сославшись на материал данной статьи).

### **Использованные ресурсы**

При подготовке данной статьи были использованы материалы сайтов: проекта [«AGB Corp.»](#), компании [ООО «Сейсмотек»](#), компании [ООО «Яндекс.Облако»](#) и ряда других общедоступных ресурсов интернет по тематике виртуальных облачных вычислений.

### **Контактная информация**

Ваши комментарии, исправления, предложения о сотрудничестве вы можете направлять автору статьи, независимому бизнес-консультанту Богданову Алексею Георгиевичу по адресу: [support@agbcorp.ru](mailto:support@agbcorp.ru) или через форму обратной связи <https://www.agbcorp.ru/contacts.html> на сайте проекта [«AGB Corp.»](#).

Контактная информация компании ООО «Сейсмотек», правообладателя и разработчика ПО «Prime» и «Prime.Cloud» представлена на их сайте <https://seismotech.ru/kontakty/>, также вы можете отправить запрос по электронной почте [mail@seismotech.ru](mailto:mail@seismotech.ru) или через форму обратной связи <https://seismotech.ru/kontakty/feedback/>.

Контактная информация компании ООО «Яндекс.Облако» представлена на их сайте: <https://cloud.yandex.ru>.