

УДК 004.5

Р. М. Алыгулиев, д-р техн. наук, зав. отделом, e-mail: r.aliguliyev@gmail.com,
И. Я. Алекперова, зав. сектором, e-mail: airada.09@gmail.com,
Институт информационных технологий НАН Азербайджана, г. Баку

Проблема BIG DATA в нефтегазовой отрасли: состояние и перспективы

Тема больших данных (BIG DATA) в нефтегазовой отрасли многогранна. Крупнейшие нефтегазовые компании мира уже давно имеют дело с большими объемами данных для принятия решения. Чтобы увеличить производство и устоять перед конкурентами, нефтегазовые компании внедряют эффективный анализ больших данных.

Выявлены проблемы, связанные с большими данными, дана классификация источников данных в нефтегазовой промышленности. Также исследованы подходы, используемые крупнейшими нефтегазовыми компаниями для решения проблем, связанных с хранением и обработкой больших данных.

Ключевые слова: нефтегазовые компании, большие данные, информационные технологии, анализ данных, аналитика, сложные аналитические запросы

Введение

В то время как быстро развиваются информационные технологии, возможность доступа, анализа и управления огромными объемами данных приобретает все большее значение для разведки и доставки нефти и газа. Проблемы, связанные с большими данными в нефтегазовой отрасли, сегодня как никогда актуальны. Процессы разведки, разработки и производства, связанные с добычей нефти и природного газа, генерируют большие данные. Компании соперничают друг с другом и используют самые новейшие информационные технологии для хранения и обработки больших данных.

В современном мире нефтегазовая промышленность является высокотехнологичной, капиталоемкой отраслью экономики, и в мировой нефтеперерабатывающей промышленности происходят заметные концептуальные, структурные сдвиги. В больших нефтегазовых компаниях, таких как Saudi Aramco (Саудовская Аравия), ExxonMobil (США), Royal Dutch Shell (Голландия — Великобритания), PetroChina (Китай), Chevron (США), Kuwait Petroleum Company (Кувейт), BP (Великобритания), Total (Франция), Petrobras (Бразилия), Qatar Petroleum (Катар), "Роснефть" (Россия), "Газпром" (Россия), Национальная нефтяная компания Ирана (Исламская Республика Иран) и т. д., среднесуточная добыча нефти составляет несколько миллионов баррелей. Эти компании хорошо осознают ценность анализа больших данных и используют их в новых проектах. Это не только коммерческие проекты. Основные из них сконцентрированы на нефтегазоносных пластах, другие — на бурении скважин.

Чтобы увеличить добычу нефти и газа, крупнейшие нефтегазовые компании все чаще прибегают к исследованиям, связанным с большими данными. Неудивительно, что в погоне за нефтью некоторые из них решили попытаться сократить свои невероятные расходы и риски, связанные с добычей природного углеводородного топлива. Было определено, что если большие данные помогут хоть одной крупной нефтедобывающей компании добыть 1 % дополнительной нефти, этого процента хватит, чтобы три года обеспечивать продуктами, полученными из этого объема, целый мир [1].

Нефтяная промышленность является отраслью экономики, которая занимается добычей, переработкой, транспортировкой, хранением и продажей нефти и сопутствующих нефтепродуктов. Современные компании, работающие в сфере нефтегазового бизнеса, являются либо вертикально интегрированными, либо специализированными. Интегрированные компании включают в свой состав всю производственную цепочку — от скважины до продажи бензина, добычу и переработку газа, а также нефтехимическую отрасль и производство альтернативных топлив. Специализированные компании работают в более узкой сфере — это геологоразведочные работы, добыча, переработка, транспортировка нефти и газа. Но эти компании тесно взаимодействуют с интегрированными компаниями. И интегрированные, и специализированные компании связывает одна цель — дополнительная прибыль и конкурентоустойчивость.

Погоня за дополнительной нефтью подталкивает компании всерьез заняться пересмотром производственных процессов, обращая основное внимание на повышение эффективности и внедрение инноваций,

так как сейчас одинаково важно снизить затраты и найти новые направления роста производительности. В этих условиях большое значение приобретают большие данные. В нефтегазовых компаниях используются огромные объемы различной информации. Речь идет не только об анализе больших данных геологоразведки, но также важен анализ всех данных, полученных при управлении технологическими процессами, начиная от перекачки нефти и газа и заканчивая механизмами контроля всего производственного процесса — от добычи до доставки потребителю.

Проблемы больших данных

В современном мире нефтегазовые компании работают в быстро меняющемся рынке. Новые методы извлечения сырья и появление альтернативных видов энергии создают новые проблемы. Разные виды энергии выходят на рынок и при этом создают избыточное предложение. Политические события в мире также создают дефицит информации о нефтегазовых компаниях. Понимание направления рыночного ценообразования и спроса имеет решающее значение. Например, когда местная поставка начнет обгонять спрос, нефтеперерабатывающим предприятиям, возможно, потребуется изучать другие рынки и политические отношения соседних стран.

Крупнейшие нефтегазовые компании стремятся повысить эффективность бизнеса и производительности, но при этом сталкиваются с целым рядом проблем, основные из них перечислены ниже:

- нефть трудно найти. Для обнаружения и описания нефтяных залежей требуются новейшая технология и дорогое оборудование;
- производство нефти стоит дорого. Для производства одного барреля нефти необходимо большое количество техники и трудовых ресурсов;
- добыча нефти представляет потенциальные проблемы для охраны окружающей среды и безопасности человека, которые необходимо безотлагательно решать.

Неопределенность и изменчивость цен на нефть и газ, изменение политики в области энергетики, экологические проблемы (например, от глобального потепления, цунами и борьбы с последствиями добычи сланцевой нефти), возникающие вследствие конкуренции со стороны новых источников энергии, а также постоянные операционные затраты на управление и т.д. создают большие проблемы, которые требуют срочного и эффективного решения [2].

Аналитики утверждают, что начало новой эры нефти и газа будет играть огромную роль в будущем энергетических компаний. Несмотря на развитие возобновляемых технологий мировой спрос на нефть по-прежнему будет расти до 2035 г., поскольку все больше людей используют автомобили и другие средства передвижения, которые работают на бензине. А так как добыча становится все более трудным делом, необходимы для удовлетворения будущих потребностей [3] инвестиции на \$20 трлн.

Многие компании сталкиваются с проблемами и неопределенностью, но продолжают делать упор на более эффективное управление активами и контроль.

Также остается необходимость сосредоточиться на эффективной и действенной разведке нефти и газа. Не только поддержание благоприятной экологической репутации, общественные и государственные отношения являются приоритетом для хорошего бизнеса, но и получение одобрения правительства для новой разведки и добычи, когда это необходимо. Типы данных, используемых аналитиками для этих целей, могут варьироваться в широких пределах. Одной из главных задач остается осмысление этих данных способами, которые окажут влияние на бизнес.

Источники больших данных (BIG Data)

Столкнувшись с разными проблемами, многие нефтегазовые компании видят решение проблем в исследовании больших данных, непрерывно растущих и хранящихся в различных базах данных компаний. Они могут предоставить такие данные в качестве важных источников информации, необходимых для оптимизации геологоразведочных работ, бурения, добычи и доставки. Нельзя не отметить новые источники данных, такие как данные из социальных медиа, мобильных устройств и личных архивов. Они могут обеспечить понимание настроения местного населения, недовольного работой промышленных объектов и трубопроводов или пострадавшего в авариях или при несчастных случаях. Хорошо управляемые нефтяные объекты имеют важное значение для поддержания положительной репутации в оценке финансовых активов.

В своем стремлении узнать, что лежит ниже поверхности земли и как это добыть, энергетические компании на протяжении многих лет инвестируют в сейсмические исследования программное обеспечение, средства визуализации и другие цифровые технологии. На сегодняшний день нефтяные и газовые компании анализируют данные из различных источников. Эти источники данных могут включать (рис. 1, см. четвертую сторону обложки):

- данные, полученные во время разведки нефти и газа, бурения, добычи, транспортировки и переработки нефти;
- традиционные корпоративные данные операционных систем;
- социальные медиа;
- информационные веб-проекты;
- данные из объявлений, в том числе о работе;
- демографические данные;
- история разведки нефти и газа;
- данные о доставке, хранении, а также о цене нефти и газа.

Основываясь на опыте ведущих нефтегазовых компаний, можно проследить несколько основных тенденций использования данных. Одна из них — разведка новых месторождений нефти и газа. Для решения этой важной задачи необходимо проводить сбор и анализ данных картирования океанского и морского дна и на его основе выявлять особенности и признаки, которые могут указывать на наличие месторождений. Однако огромные размеры файлов сканирования и резкий ежегодный рост совокупных объемов хранимых данных требуют наличия горизон-

тально масштабируемой платформы больших данных для быстрого добавления емкости.

Если говорить о сейсморазведке, информация при сканировании может дать отдельные файлы размером в несколько терабайтов, а это значит, что нефтегазовые компании должны каждый год увеличивать емкость на несколько петабайт, чтобы справиться с объемами поступающих данных. Кроме того, по мере совершенствования алгоритмов анализа и поиска новых решений особое значение приобретает ретроспективный анализ архивированных данных. В таких данных могут быть скрыты непредсказуемая информация, геологические тайны. Для эффективного анализа этих данных важно размещение всех архивных данных на более доступном уровне хранения. Некоторые нефтяные компании уже работают в этом направлении, развертывая файловые системы для хранения и извлечения огромных объемов данных, относящихся к разведке, добыче и переработке нефти и газа [3, 4].

Нефтегазовые компании все чаще осознают, что их огромные объемы данных хранятся в разных репозиториях. Это сильно усложняет аналитику и снижает производительность. Объединение этих ресурсов в единой системе станет стимулом для разработки нового поколения приложений для аналитики. Разведка новых месторождений нефти и газа и объединение данных формируют информационную основу для третьей важной задачи, которая заключается в использовании аналитики как стимула для других драйверов эффективности. Это, прежде всего, относится к использованию более сложных алгоритмов поиска при анализе данных сейсморазведки, но нефтегазовые компании хранят огромные объемы других полезных данных. К ним относятся:

- данные видеонаблюдения промышленных объектов, которые можно использовать для повышения безопасности или оптимизации процессов;
- данные работы оборудования, входящего в сложные цепочки взаимосвязанных процессов обработки, добычи и поставки, которые можно использовать для прогнозирования техобслуживания и снижения убытков в случае непредвиденного отключения или выхода из строя оборудования.

Правильное и полное использование этих данных становится приоритетной задачей для компаний. Скорость, с которой генерируются эти данные, быстро растет и приводит к более высоким темпам потребления аналитики, которую жаждут получить производители. Такие увеличение скорости передачи данных и рост источников, естественно, приводят к увеличению совокупных объемов данных. Аналитики хотят получать больше данных при более высоких скоростях, хранить их дольше и анализировать быстрее. Но при решении этих задач возникают проблемы, связанные с большими данными. Технологии BIG Data помогают нефтегазовым компаниям решать эти задачи. BIG Data — это подход, используемый для описания теории и практики применения передовых компьютерных технологий при анализе постоянно растущей информации [4]. Аналитика осуществляет открытие и связь значимых закономерностей в данных (см. рис. 1 на четвертой стороне обложки), как правило, с использованием сложных методов моде-

лирования. Существуют различные аналитические процессы, основные среди них — приобретение, очистка, визуализация, моделирование и оценка данных. Примерно 80 % времени анализа тратится на приобретение и очистку данных, поэтому аналитика всегда остается мишенью для улучшения процесса анализа.

Нефтяные и газовые компании уже давно ориентированы на повышение эффективности геологоразведочных работ за счет использования передовых аналитических средств, применяемых к разнообразным данным [5]. Сейсмическое отображение дает подсказки о том, где бурить. Эти данные играют большую роль при выборе систем наблюдений, изучении условий возбуждения сейсмических колебаний и способов их регистрации [6].

В сложных современных условиях нефтяным компаниям важно, чтобы стоимость новых исследований не росла, а показатели добычи улучшались, чтобы поддерживать рентабельность. Аналогичным образом повышение пропускной способности пласта без превышения затрат является основным фактором сохранения прибыльности. Такие данные традиционно анализируются в хранилищах данных, состоящих из реляционных баз данных. Сегодня некоторые из этих действий перемещаются в кластеры Hadoop — отчасти из-за относительно низкой цены кластеров Hadoop, отчасти из-за его "schema-less" файловой системы, которая является идеальной для прогнозного анализа рабочих нагрузок [7]. Необходимо хорошо понимать важность повышения эксплуатационной эффективности на буровых площадках, трубопроводах и нефтеперерабатывающих заводах. Это поможет лучше понять, когда обслуживание должно произойти раньше, и путем мониторинга изменять состояние ключевых компонентов, предоставляя еще большую эффективность и экономию.

По большей части информация идет от датчиков и других потоковых источников данных с внедрением системы управления BIG Data, включающей традиционные хранилища данных и новые резервуары данных (с участием Hadoop и/или NoSQL базы данных). Более широкие типы данных могут быть проанализированы, чтобы гарантировать то, что бизнес станет более гибким. Прогнозные аналитические решения, развернутые в системах управления BIG Data (включая Hadoop), становятся обычной практикой для повышения общей безопасности, надежности и снижения стоимости. Подобный анализ также указывает на потенциальные проблемы безопасности и выявляет экологические факторы риска до возникновения дорогостоящих аварий. Прогностическая аналитика широко используется при создании хранилищ данных, оптимизации маршрутизации транспортных средств и расходных материалов для разведки и добычи нефти и газа [7]. Такая маршрутизация имеет основополагающее значение для обеспечения минимально возможной стоимости доставки при сохранении рентабельности и решении основных вопросов технического обслуживания и безопасности.

Цены на газ и нефть могут быть весьма неустойчивыми, но их стоимость в течение срока реализации проекта может определить, сколько финансов требуют разведка, добыча и переработка. Здесь прогностическая аналитика может сыграть решающую

роль в понимании вероятного направления ценообразования и определения правильности разведки и выбора уровня производства. Такие данные могут также указывать на относительность вложения правильных инвестиций в недвижимость.

Возможности больших данных

Многие существующие бизнес-возможности расширяются, когда большое количество данных и их разнообразие становятся частью информационной архитектуры. Информационные технологии нефтегазовых компаний, как правило, используются для решения задач, которые обеспечивают следующее:

- улучшение разведки рентабельности инвестиций. Разведка нефтегазовых месторождений требует обширных сейсмического и экологического анализов, а также анализа стоимости производства, с тем чтобы определить, можно ли получить прибыль. Также надо учитывать меняющиеся рыночные условия;
- повышение эффективности производства. Эффективность производства обеспечивается за счет постоянного анализа буровых работ и своевременного технического обслуживания оборудования. Дистанционный мониторинг и анализ данных также имеют особое значение для определения состояний трубопроводов и нефтеперерабатывающих заводов;
- решение вопросов, связанных с окружающей средой и безопасностью. Анализ больших данных играет важную роль в защите окружающей среды, прогнозировании природных явлений и решении вопросов, связанных с безопасностью, которые могут быть побочным продуктом неоптимального управления операциями;
- улучшение отношений с общественностью и укрепление связи с государственными органами. Понимание общественного мнения и оперативное реагирование на запросы со стороны государственных и правительственных учреждений чрезвычайно важны при установлении доверительных отношений с более широкими сообществами;
- повышение рентабельности и своевременной цепочки поставок при материально-техническом обеспечении. Своевременная поставка запчастей, расходных материалов, оборудования и персонала является ключевым фактором в работе предприятия. Он обеспечивает оптимальное сохранение производственных мощностей и минимальный простой времени. Эффективное управление логистики также может обеспечить более экономически эффективную цепочку поставок;
- анализ рынка текущих цен, являющийся решающим фактором в определении того, когда надо проводить исследования, расширять производство, покупать и продавать недвижимость и менять инвестиционную стратегию;
- эффективное использование современных информационных технологий. Эффективность информационных технологий при определении больших объемов данных часто бывает трудно доказуемой, но иногда ИТ-организации склонны к первоначальному обоснованию при развертывании этих типов решений.

Технологии BIG Data дают возможность оптимизировать отображение геологической среды, указывать, каким образом и где бурить, определять производительность труда, а также обеспечивать точность в работе грузовиков и железнодорожных операций. Используя аналитику больших данных, можно увеличить время бесперебойной работы, сократить стоимость обслуживания техники, повысить производительность труда, уменьшить число ошибок и переделок. Области функционирования, трудности и возможности решения вышеупомянутых задач с помощью технологий BIG Data показаны ниже.

Исследование. В рыночных условиях трудности являются одной из наиболее "жизнеспособных" областей. Анализ больших данных при этом дает возможности:

- улучшения геологического анализа с помощью моделирования неисправностей и сейсмического анализа, что приводит к большим находкам, более предсказуемым результатам, а также к лучшему планированию скважин;
- улучшение прогнозирования рынка, чтобы определять, когда возможны разведка и добыча.

Бурение и добыча. Основной проблемой является повышение эффективности работы при одновременном соблюдении экологических стандартов. Анализ больших данных дает возможности:

- лучшей диагностики обслуживания, способствующей более высокому уровню доступности оборудования;
- предсказывать и избегать риска катастрофического сбоя датчиков на сверлах на основании показателей;
- улучшения управления энергопотреблением, что позволяет снижать стоимость бурения;
- эффективного управления активами и персоналом, снижающего стоимость производства;
- дистанционного управления операциями и мониторинга трубопроводов и оборудования для более низкой стоимости, снижения риска для окружающей среды, а также повышения безопасности.

Цепи поставок и материально-техническое обеспечение. Оптимальное время для доставки частей, расходных материалов и персонала, необходимых при разведке и добыче на участках, является основной проблемой. Анализ больших данных дает возможности:

- прогнозировать время поставки важнейших частей, оборудования и квалифицированного персонала;
- минимализировать время простоя;
- оптимально использовать топливо и стоимость доставки.

Управление недвижимостью. Продажа и вывод из эксплуатации недвижимости не считается финансово привлекательным. Анализ больших данных дает возможности:

- продажи недвижимости, находящейся в надлежащем состоянии, за приемлемую цену;
- продажи недвижимости для финансирования новых исследований на более прибыльных местах;
- слежения за соблюдением экологических норм на местных участках и несением потенциальных расходов по очистке.

Отношения с общественностью и государственными органами. Основой этой задачи является поддержание хороших рабочих отношений с обществом и

правительственными чиновниками. Анализ больших данных дает возможности:

- быстрого определения общественных настроений по отношению к разведке, бурению и добыче;
- более активного реагирования в случае внезапного изменения настроений персонала или руководства;
- быстрого реагирования на требования по воздействию на окружающую среду и другие производственные процессы бурения.

Прогнозирование цен на рынке. Необходимо для понимания направления рыночного ценообразования и будущего влияния на рентабельность участков, которые в настоящее время изучаются и в производстве. Анализ больших данных дает возможности:

- принятия более взвешенного решения о том, где и когда начать разведку нефти и газа;
- лучшего распределения кадров, оборудования и расходных материалов по требованиям;
- более эффективного управления недвижимостью.

Опыт нефтегазовых компаний в области больших данных

В большинстве нефтегазовых компаний исследования вводятся по основным компонентам. Менеджмент, безопасность и руководство имеют решающее значение и всегда стоят на первом месте. Типичная информационная архитектура для обработки больших данных показана на рис. 2 (см. четвертую сторону обложки).

Метод используется многими нефтегазовыми компаниями, где данные сохраняются и структурируются в зависимости от обстоятельств, а затем анализируются для принятия значимых решений [7]. Разнообразие базовых платформ обеспечивает решающую роль в нефтегазовых компаниях.

Некоторые компании используют методологии на основе поэтапных подходов. Эти подходы почти всегда наиболее успешны там, где для обработки больших данных необходимо определить текущее состояние и пробелы, чтобы лучше понять, в каком направлении двигаться дальше.

Компания **Royal Dutch Shell** уже некоторое время разрабатывает идею "нефтяного месторождения с управляемыми данными" в попытке снизить стоимость бурения, так как это главный расход в отрасли. **Royal Dutch Shell** является одной из крупнейших нефтегазовых компаний и четвертой в мире по объему выручки после **BP**, **Chevron**, **Total** и **ExxonMobil** [8].

Shell использует волоконно-оптические кабели, созданные в партнерстве с **Hewlett-Packard** для этих датчиков, и данные, передаваемые на свои частные серверы, поддерживаемые **Amazon Web Services** [8, 9]. Это дает гораздо более точную картину больших данных. Чтобы справиться с циклопическим массивом данных, **Shell** начала работать с **Nadoor** в **Amazon Virtual Private Cloud** [9]. Новые датчики могут сохранить большее количество данных, которые затем анализируются с использованием искусственного интеллекта и создаются 3D- и 4D-карты нефтяных скважин и резервуаров.

Использование интеллектуального анализа данных также помогает обнаружить нефть в казалось бы уже иссякших скважинах. Что именно там происхо-

дит и как проводится анализ данных — секрет фирмы и делиться опытом **Shell** категорически отказывается. Стало известно, что с помощью больших данных ожидается построение подробнейшей модели структуры Земли до глубин от 1 до 12 км. Если учесть, что в будущих планах компании предполагается развертывание датчиков примерно на 10 тысячах скважин, то возможно, что скоро одна только **Shell** будет владеть десятками эксабайтами данных.

По словам президента энергетической технологии компании **Chevron** Пола Сигла (**Paul Siegele**) данные, участвующие в разведке нефти и газа, колеблются. Тем не менее, без технологии, которая включает в себя распределенные датчики, высокоскоростные средства связи, массовые операции передачи данных по добыче и управлению данными дистанционных буровых работ, решение задач поиска и использования новых природных ресурсов было бы почти невозможными. Сегодня внутренний ИТ-трафик компании **Chevron** превышает 1,5 тбайт в день [10].

Chevron инвестирует значительные средства в технологии визуализации, которые позволяют создавать карты энергетических залежей, расположенных на десятки тысяч футов ниже поверхности дна океана. Инновации в области информационных технологий, в том числе возможности визуализации больших данных и прогнозного анализа, дают возможность таким компаниям, как **Shell**, **BP**, **PLC** и **Chevron**, используя поток данных, выявить новые месторождения нефти и газа.

BP объявила в декабре прошлого года, что она строит самый большой в мире коммерческих исследований суперкомпьютер в Хьюстоне. Используемые компанией аналитические системы значительно продлили жизнь крупнейших нефтяных месторождений, принадлежащих компании [11].

Компания **Schlumberger** подписала с американской компанией **Chevron** контракт по обеспечению доступа к программной платформе **Petrel E & P**. Программа предназначена для сбора и обработки больших данных. По данным **Schlumberger**, программная платформа **Petrel E & P** позволяет нефтяным и газовым компаниям стандартизировать рабочие процессы от разведки до производства и принимать более обоснованные решения с четким пониманием возможностей и рисков [12].

Компания **Halliburton** имеет в своем арсенале пакеты, помогающие предсказывать простои и незаметные потери времени за счет анализа в режиме реального времени на платформе **Decision Space**. В сфере сейсмических исследований внедрены системы автоматического обнаружения отказов на базе машинного обучения. Кроме того, для анализа вращения при каротаже во время бурения применяются такие алгоритмы, как метод опорных векторов и метод регрессии. В течение многих лет некоторые специалисты выступают за использование данных алгоритмов в нефтегазовой отрасли. Однако никогда ранее нефтегазовая отрасль не генерировала такие большие объемы данных, как сегодня. На сегодняшний день алгоритмы машинного обучения и визуализация данных реализуются в виде моделей предсказания или в режиме реального времени. Из проектов однократ-

ного применения анализ превратился в непрерывный процесс совершенствования [13].

Компания Halliburton предлагает пакет решений для анализа данных в сфере моделирования скорости проходки при бурении, например, с использованием обучения анализу данных на базе нескольких машин. На основании полученных в системе неявных знаний модуль анализа незаметных потерь времени (Decision Space Invisible Lost Time, DSILT) определяет возможные неполадки в процессе бурения и дает рекомендации, какие действия следует предпринять. Для построения моделей DSILT требуются компьютеры высокой производительности. Системы реального времени для моделирования DSILT могут работать на настольных компьютерах и малых устройствах непосредственно на месторождении. В соответствии с другими подходами также широко используется кластерный анализ больших массивов данных [7, 13].

Харольд Уолтерс (Harold Walters), главный технический консультант компании Halliburton, заявил, что буквально через десять лет машинное обучение и анализ больших массивов данных получат повсеместное распространение и будут внедряться всеми специалистами, отвечающими за принятие решений [14].

Для компании **Saudi Aramco** анализ больших данных проводится несколькими организациями. Данные собираются в хранилище данных, которое обеспечивает объединение, очистку и представление данных. После того как данные поступили в хранилище данных, становится намного легче обнаруживать аномалии. Для этого используются сложные аналитические запросы. Преимущество подхода в том, что в то время, как можно сканировать в режиме реального времени хранилища данных, базы являются гораздо более управляемыми, и это облегчает сканирование недостающих данных и шаблонов для поиска, где данные не вяжутся с остальной частью показаний [15].

Один из подходов для обработки больших данных ориентирован на кластерную систему хранения данных (Isilon). Российская компания Sky Technics Group (STG) ввела в эксплуатацию Isilon для одной из крупнейших нефтяных компаний России. Главными достоинствами подхода являются не только производительность и масштабируемость, но и простота управления, связанная с автоматизацией большинства процессов [16].

В компании "**Сургутнефтегаз**" данные из ERP (Enterprise Resource Planning) в реальном времени передаются в СУБД HANA (High-Performance Analytic Appliance). HANA — высокопроизводительная NewSQL платформа для хранения и обработки данных, которая обеспечивает работу со сложными аналитическими запросами. Этот подход стал реальной инновацией, благодаря которой удалось на порядок сократить основные параметры обработки информации и попутно упростить архитектуру информационных систем. В компании функционируют 29 приложений SAP, и появление резидентных технологий позволило обслуживать их с помощью единой СУБД. Нужно отметить и то, что одним из крупнейших внедрений SAP HANA в мире является инсталляция серверов SAP HANA для компании Shell. Серверы SAP HANA сокращают цикл бизнес-процессов Shell от нескольких дней до нескольких часов [17].

Норвежская специальная служба по информационным технологиям Ziebel объявила, что она завершила сбор данных по нефтегазовой отрасли по заказу компании, желающей остаться неизвестной. В течение восьми месяцев были собраны 1708 Тбайт данных о скважинах в Северном море. Данные были получены с помощью Z-System Ziebel, которая использует волоконно-оптические технологии композитного стержня для доступа и визуализации стволов скважин в режиме реального времени. Z-System также может получить информацию, относящуюся к оптимизации потока скважин, контролю рисков целостности, моделированию пласта и интенсификации добычи нефти с минимальным производством. Z-System обеспечивает хранение и представление этих данных в эффективной и доступной форме. Данные хранятся в пяти серверах Ziebel, каждый из которых включает в себя 36 жестких дисков и имеет емкость 120 Тбайт [18].

Компания **ExxonMobil** считается инициатором многих технологических достижений по сбору и анализу сейсмических данных для моделирования сложной геологической среды, выявления вероятности сбоя во время бурения при одновременном снижении воздействия на окружающую среду [19]. В данный момент в компании ExxonMobil исследования в основном включают в себя следующие направления [20]:

- разработка методов статистического вывода данных с использованием сложных моделей для выявления недостающих данных;
- разработка алгоритмов распознавания образов для многомерных наборов данных;
- построение моделей для временных рядов и пространственно-временных данных;
- разработка высокопроизводительных кодов в режиме реального времени.

PetroChina по величине производства нефти и газа и в плане доходов является одной из передовых компаний в мире. PetroChina стала использовать EMC Isilon для анализа больших данных, полученных из северо-западного филиала своего научно-исследовательского института нефтегазовой разведки и разработки (PetroChina NWGI).

Система EMC Isilon была разработана корпорацией EMC (Electro-Motive Diesel). EMC — один из мировых лидеров в области информационных технологий. PetroChina NWGI консолидировала ряд месторождений нефти и газа, используя модель Isilon X-Series, предложенную EMC и снабжаемую электроэнергией от операционной системы OneFS для сбора и распределения приложений на один общий ресурс хранения, что упрощает управление большими данными для своих критически важных операций. С помощью системы Isilon PetroChina NWGI может масштабировать производительность и емкость линейно или независимо друг от друга по мере изменения потребностей бизнеса, обеспечивая максимальную гибкость в решении массовости и пропускной способности потребностей инфраструктуры ИТ для больших данных.

PetroChina NWGI выполняет ряд геолого-геофизических исследований и разработок по всему Китаю, сосредоточенных на выявлении, добыче, производстве и распределению ресурсов нефти и газа по всему миру. До Isilon традиционные системы хранения

PetroChina NWGI не могли идти в ногу с экспоненциальным ростом данных организаций и интенсивных требований производительности, снижая производительность рабочих процессов и наращивание эксплуатационных расходов.

Используя Isilon, PetroChina NWGI может предоставить своим ученым высокий уровень быстрого одновременного доступа к сейсмическим данным и приложениям, улучшая взаимодействие и операционную эффективность. С 2015 г. PetroChina начала сотрудничать с компанией Alipay в области мобильных платежей, больших объемов данных и других областях. Сотрудники компании признались, что сотрудничество с Alibaba положило начало облачным вычислениям для всестороннего повышения добычи нефти и оказывает значительное влияние на форму обслуживания нефтяных станций [21].

Согласно отчету компании Industrial Internet Insights за 2015 г., большинство компаний нефтегазовой промышленности (65 %) используют большие аналитические данные только для выявления эксплуатационных проблем. Тем не менее, одна треть (29 %) из 250 опрошенных руководителей компании призналась, что для исследования, прогнозного анализа и оптимизации своего бизнеса использует большие данные [22]. Из отчетов компаний [22, 23] ясно, что применение технологий больших данных в нефтегазовой промышленности находится все еще в экспериментальной стадии. Лишь немногие компании используют большие данные для принятия решений. Важность анализа больших данных в нефтегазовой отрасли по-прежнему определяется коммерческими задачами — использования имеющихся данных для повышения эффективности и экономичности и по возможности исключения потенциальных и фактических потерь при операционной деятельности.

Заключение

Разведка и добыча нефти являются сложным и критически важным заданием для поставщиков энергии. Чтобы эффективно и точно находить новые ресурсы и оптимизировать производство из существующих резервуаров, приходится каждый год увеличивать расходы предприятий. Исследования показали, что анализ больших данных является ключевой задачей для увеличения производства. С учетом высокого уровня производительности, который необходимо обеспечивать в нефтегазовой отрасли, первоочередными задачами являются также снижение затрат и безопасность рабочих и окружающей среды.

Исследования показали, что как крупные корпорации, так и независимые компании стремятся эффективно использовать большие данные, чтобы снизить расходы в цепочке поставок, а также оптимизировать стратегии закачивания скважин и уменьшения простоев. В распоряжении компаний оказываются огромные объемы данных из самых различных сфер, и для решения проблем, связанных с большими данными, существуют разные подходы. В ближайшем будущем лидерами будут только те компании, которые смогут эффективно использовать новые возможности анализа больших данных.

Подводя итог, нужно отметить, что значительного снижения затрат можно добиться в первую очередь за счет эффективного использования больших данных. И, наконец, помимо огромных прибылей для нефтяных компаний и выгоды для потребителей есть немалые шансы на то, что большие данные в базах данных нефтяных компаний когда-нибудь пригодятся и в других областях, на первый взгляд совершенно не связанных с нефтегазовой добычей.

Список литературы

1. **Suthaharan S.** Big Data classification: problems and challenges in network intrusion prediction with machine learning // ACM SIGMETRICS Performance Evaluation Review. 2014. Vol. 41. № 4. P. 70—73.
2. **Farris A.** How big data is changing the oil & gas industry // Analytics Informs. 2012, P. 20—47. URL: <http://viewer.zmags.com/publication/90ffcc6b#/90ffcc6b/21>
3. **Chevron** Drills Down Big Data Assets, 2012. URL: http://www.datanami.com/2012/05/10/chevron_drills_down_big_data_assets/
4. **Парошина И.** Большая нефть, "большие данные", большие деньги // Компьютерра. 2013. URL: <http://www.computerra.ru/88216/oil-bigdata/>
5. **Holdaway K.** Harness oil and gas Big Data with analytics: optimize exploration and production with data-driven models. John Wiley Sons Inc. 2014. 384 p.
6. **Кузнецов В. М., Шехтман Г. А., Череповский А. В.** Методика наблюдений в многоволновой сейсморазведке // Технологии сейсморазведки. 2013. № 2. С. 37—59. URL: http://ts.sbras.ru/ru/articles/2013-02_037.pdf
7. **Oracle** enterprise architecture white paper. Improving oil & gas performance with big data. URL: <http://www.oracle.com/us/technologies/big-data/big-data-oil-gas-2515144.pdf>
8. **Global** Onshore Oil and Gas Market 2015-2019 with BP, Chevron, ExxonMobil, Royal Dutch Shell & Total Dominating. URL: <http://www.prnewswire.co.uk/>
9. **Bernard M.** Big Data In Big Oil: How Shell Uses Analytics To Drive Business Success // Forbes Tech. Forbes, Inc. 26 May 2015. URL: <http://www.forbes.com/sites/bernardmarr/2015/05/26/big-data-in-big-oil-how-shell-uses-analytics-to-drive-business-success>.
10. **Baaziz A., Quoniam L.** How to use Big Data technologies to optimize operations in upstream petroleum industry // International Journal of Innovation, 2013, vol. 1, no. 1, pp. 19—25.
11. **Big Oil's** Big Data Push Changing the Future of Energy, 2013. URL: <http://blogs.wsj.com/cio/2013/01/02/big-oils-big-data-push-changing-the-future-of-energy/>
12. **Chevron** and Schlumberger sign data analytics contract, 2015. URL: <http://www.oilreviewmiddleeast.com/information-technology/chevron-and-schlumberger-sign-data-analytics-contract>
13. **Salas A.** Decision Space Geosciences (DSG) Software — End-to-End Solution // 14th International Congress of the Brazilian Geophysical Society and EXPOGEF, Rio de Janeiro. August 3—6, 2015. P. 3.
14. **Walters H.** From the digital oilfield to big data analytics, 2015. URL: <http://www.adjacentoilandgas.co.uk/it-cyber-security/from-the-digital-oilfield-to-big-data-analytics/22100/>
15. **Khakwani M. Sh.** Enterprise Data Solution Leveraging Data Warehousing for Big Data Veracity at Saudi Aramco // Proceedings of the 6th International Conference on Advances in Databases, Knowledge and Data Applications. Chamonix, France. April 20—24, 2014. P. 74—79.
16. **Смирнов Н.** Big Data для нефтяников // Computerworld Россия. 2012. №18. URL: <http://www.osp.ru/cw/2012/18/13016696/>
17. **Мельник О.** Больше данных — больше денег? // CRN. 2013. № 3 (77). URL: <http://www.i-teco.ru/press/publications/579/>

18. **Venables M.** Big oil, big data // Oil & Gas Technology. 2015. URL: <http://www.oilandgastechology.net/upstream-news/big-oil-big-data>

19. **Cohen K.** Big data in the oil patch, ExxonMobil's Perspectives Blog, 2013. URL: <http://www.exxonmobilperspectives.com/2013/07/31/big-data-in-the-oil-patch/>

20. **Data Analytics and Optimization.** URL: <http://www.kdnuggets.com/jobs/13/09-25-exxonmobil-data-analytics-optimization.html>

21. **PetroChina Extracts Maximum Value from Big Data Using EMC Isilon,** 2011. URL: <http://luxembourg.emc.com/about/news/press/2011/20110607-01.htm>

22. **Industrial Internet Insights Report for 2015.** URL: https://www.accenture.com/us-en/_acnmedia/Accenture/next-gen/reassembling-industry/pdf/Accenture-Industrial-Internet-Changing-Competitive-Landscape-Industries.pdf

23. **URL:** http://influencemap.org/site/data/000/089/Influence-Map_Oil_Sector_October_2015.pdf

R. M. Aliguliyev, Head of Department, r.aliguliyev@gmail.com,

I. Ja. Alakperova, Head of Sector, airada.09@gmail.com

Institute of Information Technology of Azerbaijan National Academy of Sciences, Baku

Big DATA Problem in Oil and Gas Industry: Current State and Prospects

Topic big data in oil and gas industry is multifaceted. The biggest oil and gas companies of the world have long had to deal with big data for decision-making. To increase production and resist the competitors, oil and gas companies implement effective big data analytics. The article reveals the problems associated with big data, the classification of the data sources in the oil and gas industry. Also have been explored approaches used by biggest oil and gas companies to solve the problems associated with the storage and processing of big data.

It was confirmed that the volume and variety of data play a key role in increasing business opportunities giant oil and gas companies. Proposals for effective analysis of big data to predict the market prices in the investigation, drilling and materilano logistics and other tasks. Howled found that, taking into account the high level of performance, which should be provided in the oil and gas industry, as a priority are reducing costs and security of workers and the environment. The future prospects of using complete, accurate and fast processing of large data collected in the oil and gas companies have been identified that have become part of the architecture of information.

Keywords: oil & gas companies, big data, information technology, data analysis, analytics, complex analytical queries

References

1. **Suthaharan S.** Big Data classification: problems and challenges in network intrusion prediction with machine learning, *ACM SIGMETRICS Performance Evaluation Review*, 2014, vol. 41, no. 4, pp. 70–73.

2. **Farris A.** How big data is changing the oil & gas industry, *Analytics Informa*, 2012, pp. 20–47. <http://viewer.zmags.com/publication/90ffcc6b#90ffcc6b/21>

3. **Chevron Drills Down Big Data Assets,** 2012. URL: http://www.datanami.com/2012/05/10/chevron_drills_down_big_data_assets/

4. **Paroshina I.** Bolshaya neft, "bolshie dannie", bolshie denqi, *Kompyuterra*, 2013. URL: <http://www.computerra.ru/88216/oil-big-data/> (in Russian).

5. **Holdaway K.** *Harness oil and gas Big Data with analytics: optimize exploration and production with data-driven models.* John Wiley Sons Inc. 2014. 384 p.

6. **Kuznetsov B. M., Shextman Q. A., Cherepovckiy A. B.** Metodika nablyudeniya b mnogoovolnovikh seysmorazvedke, *Tekhnologii seysmorazvedki*, 2013, no. 2, pp. 37–59. URL: http://ts.sbras.ru/ru/articles/2013-02_037.pdf (in Russian).

7. **Oracle** enterprise architecture white paper. *Improving oil & gas performance with big data.* URL: <http://www.oracle.com/us/technologies/big-data/big-data-oil-gas-2515144.pdf>

8. **Global Onshore Oil and Gas Market 2015-2019 with BP, Chevron, ExxonMobil, Royal Dutch Shell & Total Dominating,** URL: <http://www.prnewswire.co.uk/>

9. **Bernard M.** Big Data In Big Oil: How Shell Uses Analytics To Drive Business Success. *Forbes Tech., Forbes, Inc. 26 May 2015.* URL: <http://www.forbes.com/sites/bernardmarr/2015/05/26/big-data-in-big-oil-how-shell-uses-analytics-to-drive-business-success>.

10. **Baaziz A., Quoniam L.** How to use Big Data technologies to optimize operations in upstream petroleum industry, *International Journal of Innovation*, 2013, vol. 1, no. 1, pp. 19–25.

11. **Big Oil's Big Data Push Changing the Future of Energy,** 2013. URL: <http://blogs.wsj.com/cio/2013/01/02/big-oils-big-data-push-changing-the-future-of-energy/>

12. **Chevron and Schlumberger sign data analytics contract,** 2015. URL: <http://www.oilreviewmiddleeast.com/information-technology/chevron-and-schlumberger-sign-data-analytics-contract>

13. **Salas A.** Decision Space Geosciences (DSG) Software — End-to-End Solution, *14th International Congress of the Brazilian Geophysical Society and EXPOGEF*, Rio de Janeiro, August 3–6, 2015, p. 3.

14. **Walters H.** *From the digital oilfield to big data analytics,* 2015. URL: <http://www.adjacentoilandgas.co.uk/it-cyber-security/from-the-digital-oilfield-to-big-data-analytics/22100/>

15. **Khakwani M. Sh.** Enterprise Data Solution Leveraging Data Warehousing for Big Data Veracity at Saudi Aramco, *Proceedings of the 6th International Conference on Advances in Databases, Knowledge and Data Applications.* Chamonix, France. April 20–24, 2014. pp.74–79.

16. **Smirnov N.** Big Data dlya neftnyanikov, *Computerworld Russian.* 2012, no. 18. URL: <http://www.osp.ru/cw/2012/18/13016696/> (in Russian).

17. **Melnik O.** Bolshie dannie — bolshe deneq? CRN, 2013, no. 3 (77). URL: <http://www.i-teco.ru/press/publications/579/> (in Russian).

18. **Venables M.** Big oil, big data, *Oil & Gas Technology.* 2015. URL: <http://www.oilandgastechology.net/upstream-news/big-oil-big-data>

19. **Cohen K.** Big data in the oil patch, *ExxonMobil's Perspectives Blog*, 2013. URL: <http://www.exxonmobilperspectives.com/2013/07/31/big-data-in-the-oil-patch/>

20. **Data Analytics and Optimization.** URL: <http://www.kdnuggets.com/jobs/13/09-25-exxonmobil-data-analytics-optimization.html>

21. **PetroChina Extracts Maximum Value from Big Data Using EMC Isilon,** 2011. URL: <http://luxembourg.emc.com/about/news/press/2011/20110607-01.htm>

22. **Industrial Internet Insights Report for 2015.** URL: https://www.accenture.com/us-en/_acnmedia/Accenture/next-gen/reassembling-industry/pdf/Accenture-Industrial-Internet-Changing-Competitive-Landscape-Industries.pdf

23. **URL:** http://influencemap.org/site/data/000/089/Influence-Map_Oil_Sector_Octobe