

На правах рукописи

Васильев Алексей Владимирович



**Разработка методов и программных средств
реализации компонентной вычислительной среды для
параллельного и распределённого моделирования
режимов систем газоснабжения**

05.13.11 – Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Москва – 2013

Работа выполнена в Российском государственном университете нефти и газа имени И. М. Губкина на кафедре Автоматизированные системы управления.

Научный руководитель: доктор технических наук, доцент
Сергей Александрович
Сарданашвили

Официальные оппоненты: доктор технических наук,
главный научный сотрудник
ОАО «Газпром Промгаз»
Евгений Романович
Ставровский

кандидат технических наук,
главный инженер
ООО «Информтрансгаз»
Владимир Семёнович
Панкратов

Ведущая организация: Институт систем энергетики имени
Л.А. Мелентьева Сибирского отделения
Российской академии наук

Защита диссертации состоится «26» марта 2013 г. в 15 часов 00 минут в аудитории 308 на заседании диссертационного совета Д 212.200.14 при Российском государственном университете нефти и газа им. И.М. Губкина. Ленинский проспект, 65, Москва, ГСП-1, 119991, Россия.

С диссертацией можно ознакомиться в библиотеке РГУ нефти и газа имени И.М. Губкина.

Автореферат разослан «21» февраля 2013 г.

Отзывы и замечания по автореферату в двух экземплярах, заверенные печатью, просьба высылать по вышеуказанному адресу на имя ученого секретаря диссертационного совета.

Ученый секретарь
диссертационного совета Д 212.200.14,
доктор технических наук, профессор

А.В. Егоров

Общая характеристика работы

Актуальность темы исследования. Единая система газоснабжения (ЕСГ) России управляется распределённой иерархической Автоматизированной Системой Диспетчерского Управления (АСДУ).

В работе диспетчерских служб (ДС) всех уровней активно применяются многочисленные программно-вычислительные комплексы (ПВК) систем поддержки принятия диспетчерских решений (СППДР), а также компьютерные диспетчерские тренажеры, основанные на моделировании режимов систем газоснабжения (СГ).

В настоящее время реализуется проект модернизации АСДУ ЕСГ России. Для информационного обеспечения всех подсистем АСДУ создаются централизованные источники актуализированных данных и распределённые средства доступа к ним – единое информационное пространство (ЕИП). Для вычислительного обеспечения потребностей АСДУ создан ряд мощных центров обработки данных (ЦОД).

Новые требования предъявляются и к ПВК СППДР:

- быстрое решение наиболее вычислительно-сложных задач моделирования СГ на основе параллельных и распределённых вычислений и с использованием всех доступных вычислительных ресурсов: от персональных компьютеров ДС до ЦОД;
- работа в многопользовательском, распределённом режиме, соответствующем иерархии АСДУ;
- интеграция с единым информационным пространством;
- обеспечение согласованности результатов работы различных ПВК, организация их взаимодействия на основе стандартных протоколов информационного обмена.

Реализацию перечисленных требований кардинально затрудняет ряд исторически сложившихся факторов:

- использование в каждом ПВК собственного расчётного модуля, взаимодействие с которым осуществляется по закрытому интерфейсу, затрудняет согласование результатов работы различных ПВК, их совместное использование, интеграцию с ЕИП;
- ориентация расчётных модулей на работу в рамках одного компьютера и последовательную организацию вычислительных процессов, что не соответствует современным требованиям;
- монолитность архитектуры расчётных модулей – длительная разработка на основе процедурного подхода привела к высокой степени взаимопроникновения между их элементами, что крайне затрудняет осуществление модернизации.

Один из путей преодоления проблем, связанных с перечисленными факторами – разработка унифицированного обеспечения иерархической, распределённой СППДР программно-вычислительными ресурсами для решения задач ДС всех уровней на базе единой вычислительной среды – единого вычислительного пространства (ЕВП) АСДУ ЕСГ.

В настоящей диссертации спроектирована и реализована такая вычислительная среда.

Степень разработанности темы исследования. В работе отмечен и проанализирован вклад в областях науки, связанных с темой диссертационного исследования, следующих отечественных и зарубежных учёных:

–Альтшуль А.Д., Берман Р.Я., Бобровский С.А., Вольский Э.Л., Галиуллин З.Т., Григорьев Л.И., Константинова И.М., Леонов Д.Г., Меренков А.Н., Митичкин С.К.,

Новицкий Н.Н., Сарданашвили С.А., Селезнев В.Е., Ставровский Е.Р., Сухарев М.Г., Чарный И.А., Хасилев В.Я., Швечков В.А., Юфин В.А., Яковлев Е.И. и др. в области математических методов, вычислительных алгоритмов моделирования и оптимизации режимов транспорта газа и их программной реализации в ПВК СППДР;

–Воеводин В.В., Таненбаум Э.С., Флинн М., Хоар Т., Хьюит К., Шохам Й., и др. в области теоретических основ и программных технологий программирования параллельных, распределённых, высокопроизводительных вычислений, создания распределённых систем;

–Буч Г., Гамма Э., Лисков Б., МакКоннел С., Мартин Р., Страуструп Б., Фаулер М., и др. в области объектно-ориентированного программирования, принципов, шаблонов и методологий проектирования программного обеспечения.

Цель и задачи диссертационной работы. Целью диссертационной работы является разработка моделей, методов, алгоритмов и программная реализация мультизадачной вычислительной среды комплекса моделирования режимов систем газоснабжения на основе применения технологий параллельных и распределённых вычислений.

Для достижения поставленной цели были решены следующие задачи:

1. Анализ текущего состояния и направлений развития АСДУ ЕСГ России, единого информационного пространства системы диспетчерского управления ЕСГ, архитектур построения ПВК моделирования режимов СГ, технологий программирования параллельных и распределённых вычислений.

2. Проектирование компонентной среды организации параллельных и распределённых вычислительных процессов решения задач моделирования режимов СГ и алгоритмов создания вычислительных сервисов в этой среде.

3. Разработка архитектурных решений и алгоритмов для вычислительных сервисов решения задач численного моделирования режимов СГ, адаптации модели режимов СГ к фактическим режимам.

4. Программная реализация мультизадачной вычислительной среды и вычислительных сервисов ее поддержки.

5. Проверка корректности и измерение производительности разработанного программного обеспечения в однопользовательском, сетевом, распределённом режимах.

6. Разработка методологии перехода ПВК СППДР к использованию вычислительных сервисов, созданных на базе мультизадачной вычислительной среды, осуществление интеграции с одним из ПВК СППДР в тестовом режиме.

При этом задачи, поставленные в данной диссертации, в области программных средств обеспечения систем поддержки принятия диспетчерских решений АСДУ Единой системы газоснабжения России решаются впервые.

Научная новизна. Основные научные результаты работы состоят в следующем.

1. Разработаны архитектурные решения, методы, алгоритмы и программная реализация компонентной вычислительной среды параллельного и распределённого моделирования режимов СГ, обеспечивающей эффективное использование вычислительных ресурсов в однопользовательском, сетевом, распределённом режимах, механизмы интеграции с ПВК СППДР на основе открытого предоставления вычислительных сервисов.

2. Разработаны специализированные сервисы, управляющие вычислительными процессами и ресурсами, вычислительные сервисы моделирования режимов СГ, поз-

воляющие многократно ускорить решение расчётных задач, реализуя параллельные вычисления общего назначения на графических картах.

3. Разработан вычислительный сервис решения многоуровневых итерационных задач, организующий параллельное моделирование множества однородных и разнородных СГ в распределённой среде; показана его эффективность при использовании «облачных» вычислительных ресурсов *Amazon Web Services Elastic Compute Cloud*.

Практическая ценность работы заключается в создании базовой многозадачной вычислительной среды для решения наиболее сложных расчётных задач моделирования режимов распределённых СГ большой размерности, позволяющей повысить эффективность организации вычислительных процессов решения задач ДУ на основе создания единого вычислительного пространства АСДУ.

Разработанная компонентная вычислительная среда параллельного и распределённого моделирования режимов СГ впервые позволяет реализовать решение задач моделирования связанных технологических режимов смежных разнородных технологических комплексов: системы добычи, магистрального транспорта, подземного хранения и распределения газа.

Методология и методы исследования. В работе применены следующие основные методы:

- математического и численного моделирования и решения расчётных задач диспетчерского управления для систем газоснабжения;
- объектно-ориентированного проектирования, программирования гетерогенных параллельных вычислений, создания распределённых систем.

Положения, выносимые на защиту:

–Разработаны архитектурные решения в области построения мультизадачной вычислительной среды, основанной на реализации сервис-ориентированного подхода средствами компонентного программного обеспечения промежуточного уровня и явном управлении вычислительными ресурсами.

–Осуществлена алгоритмическая и программная реализация ядра вычислительной среды – сервисов управления типизированными комплектами вычислительных ресурсов и сервиса диспетчеризации вычислительных ресурсов.

–Разработаны архитектурные решения, осуществлена алгоритмическая и программная реализация вычислительного сервиса моделирования режимов систем газоснабжения (СГ), организующего параллельное моделирование объектов СГ, в том числе с использованием вычислений общего назначения на графических картах.

–Разработаны архитектурные решения, осуществлена алгоритмическая и программная реализация вычислительного сервиса решения многоуровневых вычислительных итерационных задач ДУ СГ в распределённой среде.

Степень достоверности и апробация результатов. Основные результаты диссертационной работы докладывались на следующих конференциях и семинарах:

1. 9-я Всероссийская конференция молодых учёных, специалистов и студентов по проблемам газовой промышленности «Новые технологии в газовой промышленности» (Москва, РГУ Нефти и газа, 2011 г.).
2. 9-я Всероссийская научно-техническая конференция «Актуальные проблемы развития нефтегазового комплекса России» (Москва, РГУ Нефти и газа, 2012 г.).

3. 66-я Международная молодёжная научная конференция «Нефть и газ 2012» (Москва, РГУ Нефти и газа, 2012 г.).

4. Всероссийский научный семинар с международным участием «Математические модели и методы анализа и оптимального синтеза развивающихся трубопроводных и гидравлических систем» (Вышний Волочек, УПЦ Залучье, 2012 г.).

5. 5-я Международная конференция «Компьютерные технологии поддержки принятия решений в диспетчерском управлении газотранспортными и газодобывающими системами Диском 2012» (Москва, ВНИИГаз, 2012 г.).

Публикации. Материалы диссертации опубликованы в 8 печатных работах [1-8] из них 3 статьи в рецензируемых журналах [1-3] и 5 тезисов докладов.

Личный вклад автора. Содержание диссертации и основные положения, выносимые на защиту, отражают персональный вклад автора в опубликованные работы. Подготовка к публикации полученных результатов проводилась совместно с соавторами, причем вклад диссертанта был определяющим. Все представленные в диссертации результаты получены лично автором.

Структура и объем работы. Диссертация состоит из введения, трех глав, выводов, литературы из 160 наименований и 1 приложения. Работа изложена на 123 страницах основного текста и 19 страницах приложения. Текст работы содержит 38 рисунков и 11 таблиц.

Содержание работы

Во введении обоснована актуальность диссертационной работы, сформулирована цель и аргументирована научная новизна исследований, показана практическая значимость полученных результатов, представлены выносимые на защиту научные положения.

В первой главе *«Анализ проблем в области разработки и развития компьютерных программно-вычислительных комплексов моделирования систем газоснабжения»* рассмотрено современное состояние функциональной подсистемы компьютерного моделирования и поддержки принятия диспетчерских решений в составе АСДУ ЕСГ России.

Единая система газоснабжения (ЕСГ) России представляет собой сложный производственно-технологический комплекс, объединяющий объекты добычи, переработки, транспорта и подземного хранения газа в единый иерархически управляемый процесс.

Иерархическое управление ЕСГ осуществляется автоматизированной системой диспетчерского управления (АСДУ), важнейшей проблемой модернизации которой является создание единого информационного и вычислительного пространств.

ЕВП может быть создано на базе единой вычислительной среды, обеспечивающей полную поддержку технологий параллельных и распределённых вычислений, механизмов информационного взаимодействия и интеграции, с последующей разработкой необходимых вычислительных сервисов на её базе.

До последнего времени актуальным оставалось решение следующих задач:

– разработка архитектуры, алгоритмической и программной реализации единой вычислительной среды, обеспечивающей основу для разработки унифицированного обеспечения распределённой иерархической СППДР АСДУ вычислительными сервисами на базе технологий параллельных и распределённых вычислений;

- разработка вычислительных сервисов решения конкретных расчётных задач моделирования СГ на базе единой вычислительной среды;
- разработка методологии интеграции ПВК СППДР с новыми вычислительными сервисами, реализация интеграционных проектов, тестирование корректности и производительности реализованных решений.

Следующие главы диссертации посвящены решению указанных проблем.

Во второй главе *«Разработка компонентной среды организации параллельного и распределённого вычислительного процесса решения расчётных задач СППДР»* предложена архитектура вычислительной среды, определены методы и средства её построения, предложен общий алгоритм разработки вычислительных сервисов на её основе, спроектированы вычислительные процессы решения ряда конкретных задач моделирования режимов СГ.

На основании проведённого в первой главе анализа предметной области сформулированы основные требования к разработанной в диссертации компонентной, мультизадачной вычислительной среде:

- модульность, гибкость, расширяемость и развиваемость;
- организация в распределённой среде вычислительных процессов произвольной конфигурации, в том числе соответствующей распределённой иерархии АСДУ ЕСГ;
- минимизация накладных расходов удалённого взаимодействия в распределённой среде;
- высокая производительность решения вычислительно-сложных задач моделирования СГ за счёт полной поддержки технологий параллельного и распределённого программирования, в том числе вычислений общего назначения на графических картах;
- унифицированное вычислительное обеспечение СППДР АСДУ за счёт предоставления вычислительных сервисов на основе стандартных протоколов;
- наличие развитых механизмов информационного взаимодействия для интеграции с ЕИП АСДУ.

Данные требования явились основой для последующей разработки проектных, алгоритмических и программных решений реализации ЕВП.

Архитектура построения вычислительной среды

Для реализации перечисленных выше требований при разработке архитектуры вычислительной среды были приняты три ключевых принципа.

1. Сервис-ориентированный подход на основе использования компонентного программного обеспечения промежуточного уровня (middleware).
 2. Обеспечение явного управления вычислительными ресурсами.
 3. Разработка вычислительных сервисов для среды по унифицированному алгоритму, позволяющему максимально использовать её преимущества.
- Перечисленные принципы последовательно раскрыты в работе.

Сервис-ориентированный подход средствами компонентного middleware

Сервис-ориентированный подход заключается в разработке программного обеспечения на основе использования слабо связанных компонентов, оснащённых стандартизированными интерфейсами для взаимодействия по стандартизированным протоколам в распределённой среде – сервисов.

Для реализации сервис-ориентированного подхода применяются различные технологии, наиболее распространённая из них – использование SOAP (Simple Object Access Protocol) Web-сервисов. Однако, на основе анализа различных технологий, принято решение использовать в диссертации компонентное обеспечение промежуточного уровня (middleware), характеризующееся более высокой эффективностью.

Сервис-ориентированный подход, реализованный на базе middleware, обеспечивает:

- высокую степень модульности системы – она строится из набора слабо связанных компонентов с чётко определёнными интерфейсами, что создаёт предпосылки для её гибкости;
- независимость от физического расположения компонент; инвариантность к платформе; поддержку различных языков программирования;
- простоту программирования – middleware предоставляет поддержку исполнения синхронных и асинхронных удалённых вызовов, развитую объектную модель, дополнительные инструментальные средства построения распределённых компонентных систем;
- высокую производительность – middleware, в отличие от Web-сервисов, использует эффективные бинарные протоколы, многократно снижающие накладные расходы взаимодействия в распределённой среде.

Сервисы в распределённой системе выполняются компьютерами, связанными сетями передачи данных, – узлами, каждый из которых:

- содержит набор вычислительных ресурсов: центральных процессоров (ЦП); GPGPU-видеокарт (General Purpose Graphics Processing Unit); других устройств, способных осуществлять вычисления;
- работает под управлением той или иной операционной системы, выполняет ряд приложений, написанных с использованием тех или иных языков программирования и технологий параллельного программирования;
- содержит middleware runtime – среду времени выполнения компонентного программного обеспечения промежуточного уровня, выполняющую функции адаптера между прикладными сервисами и низкоуровневыми механизмами удалённого взаимодействия.

Каждый сервис в системе строится из двух частей: интерфейса и исполнителя. Интерфейс сервиса описывает множество предоставляемых сервисом методов на специализированном языке описания интерфейсов. Исполнитель сервиса реализует методы, объявленные в интерфейсе, функционирует в рамках вычислительных ресурсов узла распределённой системы, принимая и обрабатывая удалённые вызовы методов от клиентов.

Сервисы способны как предоставлять некоторую функциональность для клиентов, так и использовать другие сервисы, то есть одновременно находиться в роли и клиента, и сервера.

Для использования методов, предоставляемых сервисом, клиент генерирует объект-посредник – локальный для клиента объект, реализующий интерфейс сервиса – и использует его. Посредник устанавливает связь с исполнителем сервиса и переадресует локальные вызовы удалённому серверу прозрачно для клиента.

Схема взаимодействия сервисов в распределённой среде представлена на рисунке 1.



Рисунок 1 – Схема взаимодействия сервисов в распределённой среде

Традиционные подходы к управлению вычислительными ресурсами

В работе проанализированы и выделены следующие основные подходы к управлению вычислительными ресурсами в распределённых системах.

Прозрачное управление вычислительными ресурсами балансировщиком нагрузки. Балансировщик принимает удалённые вызовы методов сервисов от клиентов и переадресует их узлам распределённой системы, обеспечивая равномерное распределение нагрузки. Выделяется два типа сервисов.

1. Без сохранения состояния – вся необходимая информация для исполнения сервиса передаётся при каждом вызове. В этом случае распределение нагрузки организуется наиболее просто – сервер приложений переадресует вызов клиента наименее загруженному узлу.

2. С сохранением состояния – клиент выполняет последовательность операций, меняющих состояние удалённого объекта и зависящих друг от друга. В этом случае распределение нагрузки осуществляется более сложно. Применяются различные стратегии хранения состояний объектов в оперативной памяти, на жёстком диске, в базе данных.

Управление ресурсами вычислительного кластера штатной системой очереди. Для совместной работы множества пользователей с вычислительными кластерами используется следующая система: для запуска расчёта необходимо оформить вычислительное задание, состоящее из запускаемой программы и спецификации необходимых для её выполнения вычислительных ресурсов, и поставить его в очередь, используя штатную систему управления заданиями. Такая система планирует очередность

выполнения поставленных в очередь заданий, чтобы максимизировать общую загрузку кластера.

Явное управление вычислительными ресурсами

Для решения вычислительно-сложных многоуровневых итерационных задач моделирования СГ характерны требования, выходящие за рамки традиционных подходов.

Монопольный доступ клиента к вычислительному ресурсу. В каждый конкретный момент времени с вычислительным ресурсом может работать только один клиент, что целесообразно по ряду причин:

– большинство расчётных задач моделирования СГ является вычислительно сложными и могут полностью использовать потенциал ресурса – в такой ситуации разделение вычислительного ресурса между несколькими расчётными процессами на основе многозадачности приводит к снижению производительности;

– для некоторых вычислительных устройств, в частности CUDA-видеоадаптеров, использование в режиме многозадачности невозможно из-за недостаточно развитых механизмов прерываний, защиты памяти – в таком случае организация монопольного доступа необходима.

Выделение вычислительным сервисам, использующим технологии параллельных и распределённых вычислений, соответствующих вычислительных ресурсов. Для функционирования высокопроизводительных вычислительных сервисов требуются специализированные вычислительные ресурсы. Например, для сервиса, использующего вычисления общего назначения на графических картах, требуется соответствующий видеоадаптер; для сервиса, использующего технологию программирования MPI (Message Passing Interface), требуется несколько узлов с настроенной средой выполнения MPI.

Прямой доступ клиента к серверу. В течение монопольной работы клиента с вычислительным ресурсом целесообразно организовать между ними прямой, без посредников, доступ.

Поддержка операций с сохранением состояния удалённого вычислительного ресурса без накладных расходов организуется автоматически за счёт монопольной работы клиента с вычислительным ресурсом. В том числе реализуется сохранение состояния памяти сопроцессоров удалённого вычислительного ресурса.

Управление конфигурацией вычислительного процесса в распределённой среде с использованием наиболее подходящих типов вычислительных ресурсов: от персональных компьютеров до ЦОД.

Перечисленные требования реализованы следующим образом.

Все сервисы в системе подразделены на две категории: служебные и вычислительные.

– Вычислительные сервисы решают расчётные задачи.

– Служебные сервисы организуют взаимодействие клиентов, сервисов и вычислительных ресурсов в распределённой среде.

Введено понятие *комплекта вычислительных ресурсов* как некоторой совокупности вычислительных ресурсов: от подмножества вычислительных ресурсов одного узла до объединения нескольких узлов.

Каждому комплекту вычислительных ресурсов присваивается *тип комплекта вычислительных ресурсов*. На основе типа сопоставляются вычислительные сервисы и подходящие для их выполнения комплекты вычислительных ресурсов.

В рамках каждого комплекта вычислительных ресурсов запускается *служебный сервис управления типизированным комплектом вычислительных ресурсов*, предоставляющий методы для запуска вычислительных сервисов соответствующего типа в рамках управляемого комплекта.

Для явного управления вычислительными ресурсами на одном из узлов вычислительной среды запускается *служебный сервис диспетчеризации ресурсов*. Его интерфейс предоставляет клиенту возможность установления прямого монопольного соединения с сервисом управления комплектом вычислительных ресурсов заданного типа – для этой операции введён термин «захват вычислительного ресурса».

Схема функционирования явного управления типизированными комплектами вычислительных ресурсов приведена на рисунке 2.



Рисунок 2 – Схема явного управления комплектами вычислительных ресурсов

Первоначально выполняется этап формирования вычислительной среды. На узлах распределённой системы формируются типизированные комплекты вычислительных ресурсов и запускаются соответствующие служебные сервисы управления комплектами ресурсов. На одном из узлов запускается сервис диспетчеризации комплектов ресурсов. Сервисы управления комплектами вычислительных ресурсов регистрируются в сервисе диспетчеризации, передавая ему свои объекты-посредники. Полученные посредники собираются сервисом диспетчеризации в типизированные пулы (для каждого типа объектов создаётся свой пул). Тем самым формируется множество комплектов вычислительных ресурсов, управляемых сервисом диспетчеризации.

Для осуществления своих вычислительных потребностей клиент инициирует следующую последовательность действий.

1. Клиент определяет, какой тип комплекта вычислительных ресурсов поддерживает выполнение необходимой ему операции и запрашивает у сервиса диспетчеризации объект-посредник сервиса управления комплектом вычислительных ресурсов необходимого типа.

2. Сервис диспетчеризации обеспечивает установление прямого монопольного соединения между клиентом и сервисом управления комплектом ресурсов запрошенного типа, передавая клиенту объект-посредник, изъятый из соответствующего пула.

3. Клиент использует методы сервиса управления комплектом вычислительных ресурсов для запуска необходимых вычислительных сервисов на захваченном комплекте вычислительных ресурсов.

4. Клиент выполняет удалённые вызовы методов запущенных вычислительных сервисов для выполнения необходимой работы.

5. Клиент освобождает захваченный вычислительный ресурс, и сервис диспетчеризации возвращает посредник его сервиса управления в соответствующий пул.

Контроль над организацией вычислительного процесса в распределённой среде достигается за счёт выбора клиентом сервиса диспетчеризации, у которого он запрашивает необходимый комплект вычислительных ресурсов – тем самым клиент определяет пул, из которого ему будет выделен комплект вычислительных ресурсов.

В работе рассмотрены комбинации взаимного расположения клиентов, сервиса диспетчеризации, множества управляемых им вычислительных ресурсов, формирующие следующие конфигурации и режимы работы вычислительных процессов: однопользовательский режим; клиент-серверный режим; режим с использованием вычислительных ресурсов клиентов в локальной сети; распределённый режим, в частности, соответствующий распределённой иерархии АСДУ; комбинированный режим.

Технологии реализации архитектурных решений

С целью выбора наиболее подходящих технологий реализации предложенной архитектуры в работе проведён сравнительный обзор современных технологий гетерогенного параллельного программирования, распределённого программирования, создания распределённых систем. Осуществлён выбор, представленный в таблице 1.

Таблица 1 – Технологии реализации вычислительной среды

Категория	Технология	Назначение
Программирование систем с общей памятью	OpenMP, Intel TBB, библиотеки численных методов	Программирование сложных, разветвлённых алгоритмов
Программирование высокопараллельных вычислений	CUDA, библиотеки численных методов	Программирование простых алгоритмов с высокой степенью параллелизма
Программирование вычислительных кластеров	MPI	Программирование наиболее вычислительно сложных задач
Компонентное middleware	ZeroC ICE	Использование в качестве базы построения распределённой системы

Выбранные технологии в большей степени, нежели их конкуренты, отвечают следующим требованиям: кросс-платформенность; высокоуровневость модели программирования; производительность. Они предпочтительны при разработке вычислительных сервисов для предложенной в работе среды, однако модульность системы позволяет при необходимости использовать и любые другие технологии программирования.

Проектирование, алгоритмическая и программная реализация разработанных вычислительных сервисов

В работе проведён анализ структуры задач моделирования режимов СГ и выделен базовый комплекс расчётных задач, включающий в себя следующие основные процедуры:

- расчёт свойств газа и газового потока;
- процедуры моделирования режимов отдельных объектов СГ: трубопроводы, крановые системы, аппараты воздушного охлаждения, газоперекачивающие агрегаты (ГПА);

- процедуры моделирования подсистем СГ: трубопроводные системы, компрессорные цеха (КЦ), магистральные газопроводы, газотранспортные системы.

На основе процедур базового комплекса расчётных задач решаются задачи более высокого уровня – в частности, задачи оптимизации и прогнозирования режимов СГ, мониторинга фактического состояния СГ, ситуационного анализа, идентификации эмпирических параметров модели.

Многие высокоуровневые расчётные задачи, основанные на использовании базового комплекса, представляют собой многоуровневые иерархические вычислительные процессы, в которых моделирование режима отдельной СГ является основным вычислительным элементом.

Эффективное решение таких задач требует организации одновременного моделирования нескольких режимов СГ в распределённой среде с использованием параллельных вычислений для моделирования каждого из режимов.

В данной диссертационной работе спроектированы и реализованы распараллеленные и распределённые вычислительные сервисы на примере решения следующих задач:

- моделирование режимов СГ;
- идентификация эмпирических параметров модели режимов СГ.

Общая схема моделирования режимов систем газоснабжения

Для решения задачи моделирования режимов СГ топология системы задаётся условно-ориентированным «расчётным графом» (дуги имеют направление потоков газа, но оно может изменяться в процессе расчётов)

$$G = (V, E), |V| = m, |E| = n. \quad (1)$$

V – множество узлов графа, E – множество дуг.

Дугам графа соответствуют моделируемые технологические объекты, участвующие в перекачке газа (трубопроводы, крановые системы, ГПА, КЦ и так далее), – узлы графа соединяют входы/выходы дуг – соответствующих моделируемых объектов.

В краевых узлах графа задаются параметры газового потока, необходимые для расчёта режима.

Для каждого моделируемого объекта на основании математического представления физических законов моделируемого процесса формируется расчётная модель, которую можно условно представить в виде функционала:

$$\{T_{\text{ВЫХ}}, q\} = f(P_{\text{ВХ}}, P_{\text{ВЫХ}}, T_{\text{ВХ}}, \text{Gas}, \text{Environment}, \text{Object}), \quad (2)$$

где q – расход газа по дуге, $P_{\text{ВХ}}, P_{\text{ВЫХ}}$ – давления на входе и выходе дуги, $T_{\text{ВХ}}, T_{\text{ВЫХ}}$ – температуры газа на входе и выходе дуги, Gas – компонентный состав газового потока, Environment – параметры внешней среды, Object – параметры объекта: нормативно-справочные (паспортные) данные; параметры фактического состояния; режимно-технологические ограничения.

Для сокращения выкладок используется упрощённый вид функционала (2):

$$q(P_{\text{ВХ}}, P_{\text{ВЫХ}}, \bar{A}) \quad (3)$$

где \bar{A} – множество прочих параметров объекта.

В общем виде функционал (3) представляет собой итерационную вычислительную процедуру.

Решение задачи состоит в определении поля давлений и температур во всех узлах графа, обеспечивающего баланс потоков по дугам во всех узлах графа, и соблюдение всех режимно-технологических ограничений расчётных объектов.

Эта задача решается итерационным (l – номер текущей итерации) *методом узловых давлений*, состоящим из последующих этапов.

1. Задание начального поля давлений и температур газа $\{P_k^0, T_k^0\}$ в узлах графа.
2. Расчёт расхода по каждой на основе функционала (3).
3. Проверка наличия существенных дисбалансов в узлах графа.

$$|\delta q_k^{l+1}| < \varepsilon_q, \quad (4)$$

где k – номер узла, ε_q – точность балансирования потоков. Если дисбалансов нет – алгоритм завершён, иначе – переход к следующему шагу.

4. Формирование системы уравнений баланса потоков газа в каждом i -м узле графа в соответствии с 1-м законом Кирхгофа:

$$\delta q_k^{l+1} = \sum_{i=1}^{N_k} q_{i,k}^{l+1} (P_{i,\text{ВХ}}^{l+1}, P_{i,\text{ВЫХ}}^{l+1}) - Q_k = 0, \quad (5)$$

где Q_k – заданные притоки или отборы газа в k -м узле, N_k – количество инцидентных k -му узлу дуг.

Функционал (3) на $l+1$ -й итерации представляется линейной составляющей разложения в ряд Тейлора:

$$q_i^{l+1}(P_{i,\text{ВХ}}^l, P_{i,\text{ВЫХ}}^l) \simeq q_i^l + \frac{\partial q_i^l}{\partial P_{i,\text{ВХ}}^{l+1}} \Delta P_{i,\text{ВХ}}^{l+1} + \frac{\partial q_i^l}{\partial P_{i,\text{ВЫХ}}^{l+1}} \Delta P_{i,\text{ВЫХ}}^{l+1} \quad (6)$$

и подставляется в соответствующие уравнения (5). В результате формируется система линейных алгебраических уравнений относительно расчётных компонент ΔP^{l+1} .

5. Решение системы линейных уравнений (5), расчёт нового поля давлений газа в узлах графа:

$$P_i^{l+1} = P_i^l + g \cdot \Delta P_i^{l+1}, \quad (7)$$

где g – регуляризационный параметр, одинаковый для всех выражений (7), начальное значение g равно единице.

6. Расчёт по каждой дуге расхода и температуры потока газа при новом поле давлений по функционалу (3).

Если для каких-то объектов происходит нарушение режимно-технологических ограничений, осуществляется коррекция значения (7) посредством уменьшения параметра g .

7. Вычисление расчётного суммарного дисбаланса потоков газа в узлах графа

$$F^{l+1} = \sum_{k=1}^m \delta q_k^{k+1} \quad (8)$$

Определение параметра g для выполнения условия $F^{l+1} < F^l$.

Переход к проверке условия завершения алгоритма (шаг 3).

Пример сервиса моделирования режимов систем газоснабжения

Вычислительный сервис моделирования режимов систем газоснабжения реализует метод узловых давлений с распараллеливанием этапов расчёта расходов по дугам графа на шагах (2), (6) и решения системы уравнений на шаге (5).

Расчёт расходов потоков газа по дугам графа по давлениям в краевых узлах дуг на итерационных шагах осуществляется параллельно, поскольку расчет режима каждой дуги является независимой процедурой. Для эффективного решения этой задачи разработано специализированное решение. Его схема представлена на рисунке 3.

Вычислительный сервис моделирования режимов систем газоснабжения реализует метод узловых давлений с распараллеливанием этапов расчёта расходов по дугам графа на шагах (2), (6) и решения системы уравнений на шаге (5).

Расчёт расходов потоков газа по дугам графа по давлениям в краевых узлах дуг на итерационных шагах осуществляется параллельно, поскольку расчет режима каждой дуги является независимой процедурой. Для эффективного решения этой задачи разработано специализированное решение. Его схема представлена на рисунке 3.

Все дуги расчётного графа разбиваются на группы по типам объектов – формируются векторы однотипных расчётных объектов.

Для каждого вектора расчётных объектов разработан свой менеджер параллельного моделирования, осуществляющий эффективное размещение в памяти и параллельное моделирование управляемых объектов. Для координации совместной работы всех менеджеров параллельного расчёта разработан сервис суперменеджер.

Параллельное решение системы уравнений выполняется разработанным вычислительным сервисом, основанным на оптимизированной библиотеке численных методов.

Для вычислительного сервиса решения задачи моделирования режимов СГ разработаны все необходимые сервисы: сервис загрузки параметров объектов расчётной схемы из хранилища; сервисы менеджеров параллельного моделирования режимов дуг всех типов, составляющих расчётную схему; сервис суперменеджер, координирующий параллельную работу менеджеров расчёта режимов векторов дуг схемы; сервис параллельного решения систем уравнений.

Схема взаимодействия созданных сервисов представлена на рисунке 4.

Для работы вычислительных сервисов, используемых сервисом моделирования режимов СГ, требуются следующие вычислительные ресурсы:

– Многоядерный центральный процессор для организации вычислительного процесса моделирования; параллельного моделирования расчётных объектов, обла-

дающих сложными вычислительными моделями, соответствующими менеджерами параллельного моделирования; работы сервиса суперменеджера;

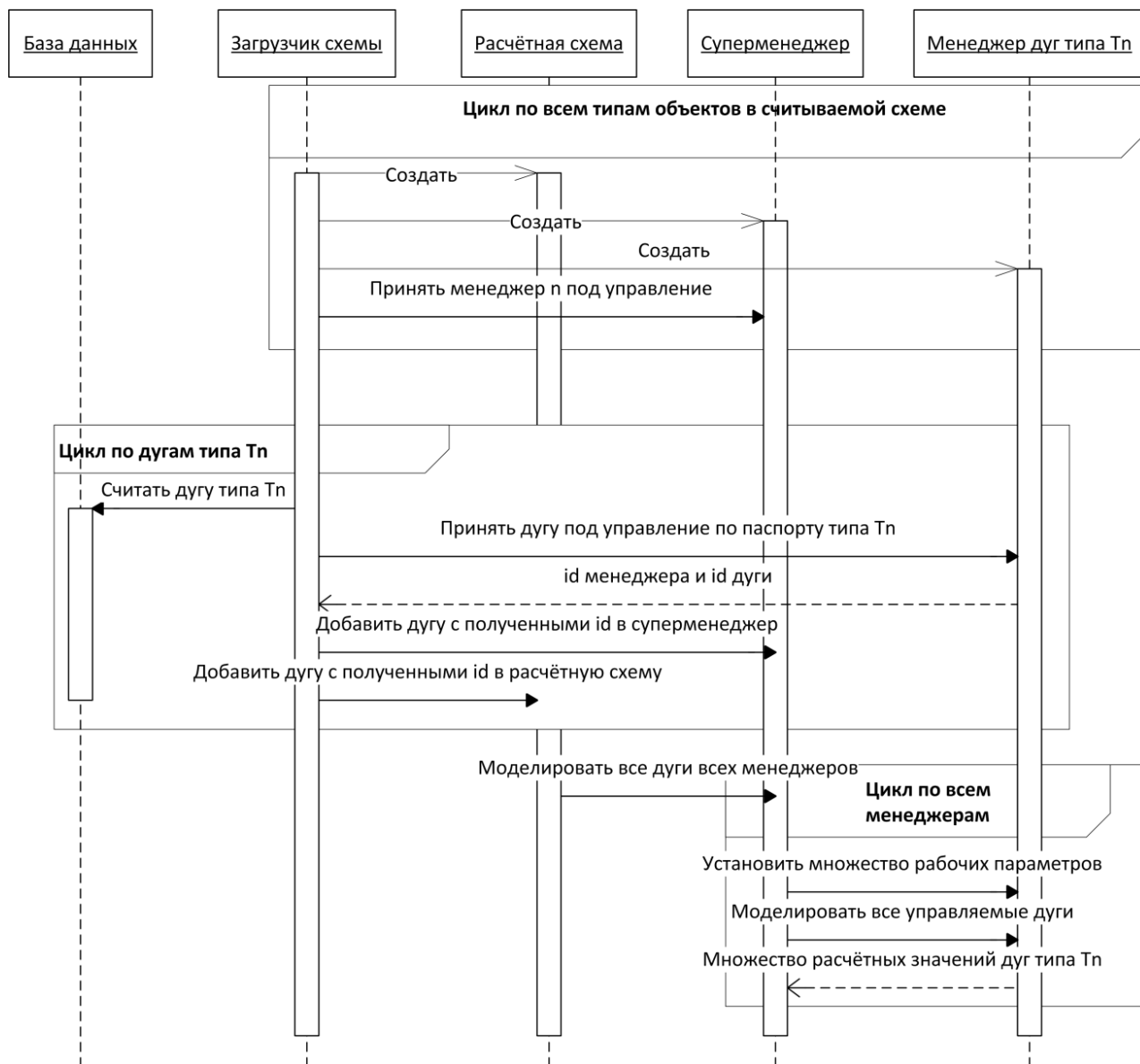


Рисунок 3 – Параллельное моделирование объектов расчётной схемы

–GPGPU-видеокарта CUDA для параллельного моделирования режимов многочисленных объектов, обладающих простыми вычислительными моделями, таких как трубопроводы; решения систем линейных алгебраических уравнений с использованием оптимизированной библиотеки численных методов.

Таким образом, для эффективного выполнения вычислительного сервиса моделирования режимов СГ необходим комплект вычислительных ресурсов, содержащий центральный процессор и CUDA-видеокарту. Тип такого комплекта вычислительных ресурсов назван «ЦП + CUDA».

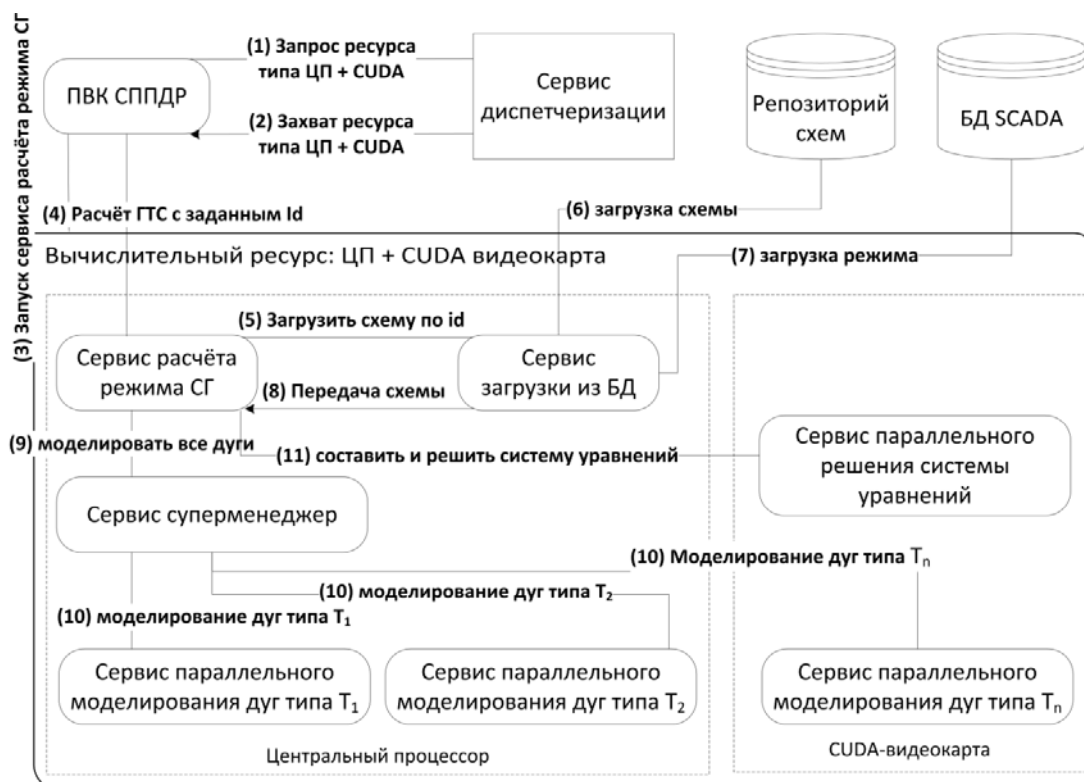


Рисунок 4 – Сервис расчёта режима системы газоснабжения

Пример сервиса идентификации эмпирических параметров модели режимов систем газоснабжения

Идентификация параметров модели режима СГ производится для повышения адекватности расчётных результатов модели, то есть их приближения к измеряемым значениям параметров газового потока в контрольных точках системы. Решение этой задачи требует большого объёма вычислений и имеет потенциал для многократного ускорения с использованием параллельных и распределённых вычислений в разработанной автором вычислительной среде.

Исходных данных задачи те же, что и для расчёта режима СГ, а также дополнительная, избыточная информация – множество значений замеров параметров газового потока в некоторых контрольных узлах СГ. В качестве замеров часто используются:

- P_i^*, T_i^* – замеры давления и температуры газа на входах и выходах КЦ, ГПА;
- Q_i^* – расходы газового потока, измеряемые на газоизмерительных станциях (ГИС).

Состав эмпирических параметров модели СГ определяется типами технологических объектов системы. Для трубопроводных систем такими параметрами являются:

- $K_{то}, K_{эф}$ – коэффициенты теплообмена и гидравлической эффективности трубопроводов, а для компрессорных цехов:

- K_{η}, K_N – интегральные по КЦ поправочные коэффициенты на газодинамические характеристики центробежных нагнетателей (ЦБН) однотипных ГПА, а также на индивидуальные коэффициенты технического состояния ЦБН ГПА.

Для математического описания рассматриваемой задачи введём следующие обозначения: $\vec{a} \in \mathbb{R}^{n_a}$, вектор идентифицируемых параметров расчётной модели, $\vec{p}^* \in \mathbb{R}^{n_p}$ и $\vec{p} \in \mathbb{R}^{n_p}$ – векторы измеряемых и расчётных параметров режима соответственно.

Множество расчётных параметров \vec{p} модели процесса является функцией идентифицируемых параметров \vec{a} :

$$\vec{p} = f(\vec{a}) \quad (9)$$

Задача идентификации эмпирических параметров модели в общем виде сводится к минимизации некоторого критерия рассогласования \vec{p}^* и \vec{p} , в качестве которого можно использовать различные критерии математической статистики, в частности сумму нормированных квадратов разностей разноразмерных параметров

$$F = \min_{\vec{a}} \sum_{r=1, \dots, n_p} \left(\frac{p_r^* - p_r(\vec{a})}{p_r^*} \right)^2 \quad (10)$$

Для решения задачи (10) применяются различные численные методы – в частности, *метод чувствительности*, основанный на итерационном поиске компонент искомого вектора эмпирических параметров:

$$\vec{a}^{l+1} = \vec{a}^l + g \cdot \Delta \vec{a}^{l+1}, \quad (11)$$

где g – регуляризационный параметр.

Условием минимума (10) является система уравнений

$$\frac{\partial F^{l+1}}{\partial \Delta a_k^{l+1}} = 0, \text{ для } k = 1, \dots, n_a \quad (12)$$

Функционал (9) на $l+1$ -ой итерации представляется в виде линейной составляющей разложения в ряд Тейлора:

$$p_r(\vec{a}^{l+1}) = p_r(\vec{a}^l) + \sum_{i=1}^{n_a} \left(\frac{\partial p_r(a^l)}{\partial a_i^l} \Delta a_i^{l+1} \right) \quad (13)$$

Подставляя (13) в (12) и выполняя ряд преобразований, получаем систему линейных уравнений относительно Δa_i^{l+1} :

$$\sum_{i=1}^{n_a} \sum_{r=1}^{n_p} \left(\frac{1}{(p_r^*)^2} \frac{\partial p_r^l}{\partial a_i^l} \frac{\partial p_r^l}{\partial a_k^l} \right) \Delta a_i^{l+1} = \sum_{r=1}^{n_p} \frac{1}{(p_r^*)^2} (p_r^* - p_r^l) \left(\frac{\partial p_r^l}{\partial a_k^l} \right), k = 1 \dots n_a \quad (14)$$

Процедура метода чувствительности строится на основе итерационного формирования и решения системы линейных уравнений (14), вычисления поправок Δa_i^{l+1} к идентифицируемым параметрам модели режима СГ.

Для численной аппроксимации производных в (14) требуется вычисление $n_a + 1$ режимов СГ при следующих значениях вектора эмпирических параметров модели:

$$\vec{a} = (a_1, a_2, \dots, a_k + \delta, \dots, a_{n_a}); \vec{a}_k^+ = (a_1, a_2, \dots, a_k + \delta, \dots, a_{n_a}), k = 1 \dots n_a \quad (15)$$

Размерность задачи многократно возрастает, когда имеется временной ряд замеров, и функция цели задачи (10) содержит сумму рассогласований контрольных параметров по всем временным слоям $t = [t_0, t_j]$.

На каждой внешней итерации минимизации функции цели (10) при соответствующих состояниях вектора идентифицируемых параметров \vec{a} расчеты режимов СГ могут выполняться независимо, поэтому имеется возможность их распараллеливания. Расчёт даже одного режима СГ с помощью разработанного в данной диссертации вычислительного сервиса полностью использует вычислительные ресурсы отдельного узла среды моделирования. Поэтому необходимо организовать одновременный расчёт множества режимов СГ на нескольких узлах распределённой системы – многоуровневый итерационный вычислительный процесс в распределённой среде.

Схема работы вычислительного сервиса, на примере задачи идентификации эмпирических параметров модели режима, в распределённой вычислительной среде представлена на рисунке 5.

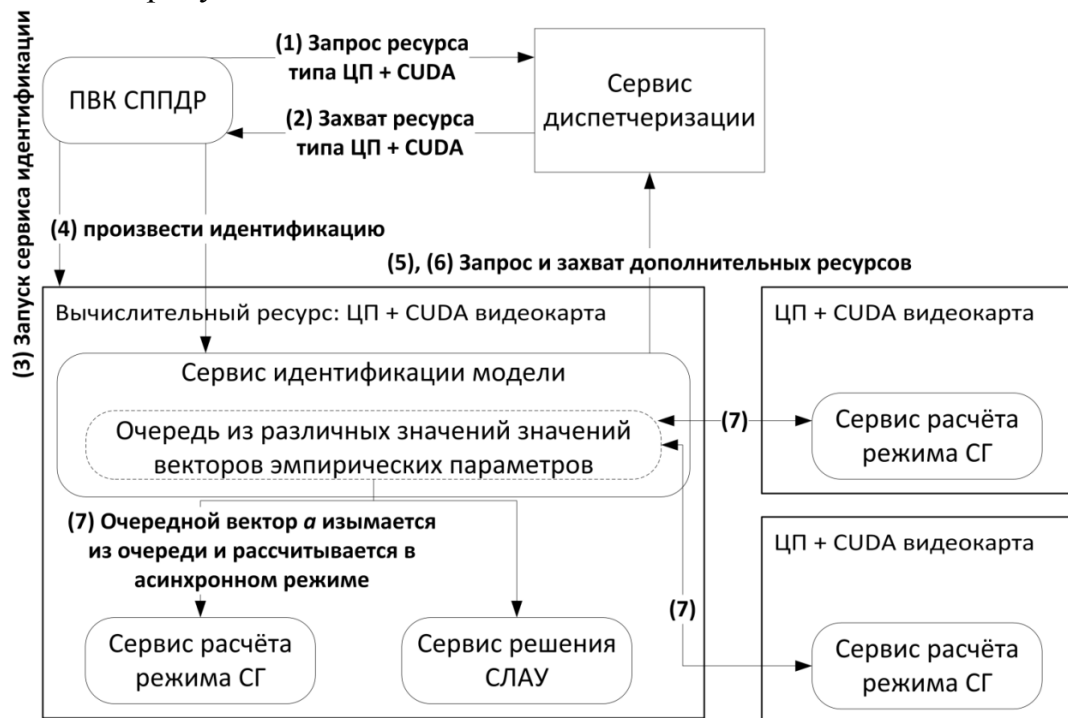


Рисунок 5 – Схема работы вычислительного сервиса идентификации эмпирических параметров модели СГ в распределённой среде

На этапах (1) – (6) осуществляется формирование вычислительной среды: запуск сервиса идентификации, захват им дополнительных вычислительных ресурсов типа «ЦП + CUDA» для моделирования режимов СГ при различных значениях вектора эмпирических параметров \vec{a} .

На этапе (7) реализуется итерационный многоуровневый процесс идентификации параметров модели СГ в распределённой среде.

Сервис идентификации запускает на каждом из захваченных ресурсов сервис расчёта режима СГ и вызывает метод загрузки адаптируемой СГ на каждом из них.

На каждой итерации метода чувствительности сервис идентификации формирует очередь из всех значений вектора идентифицируемых параметров модели (15), которые необходимо рассчитать для вычисления частных производных $\frac{\partial p_r(a^l)}{\partial a_i^l}$ в (14).

Сервис идентификации выполняет асинхронный запуск расчёта режима СГ на каждом из доступных сервисов моделирования режима СГ. Как только один из вычислительных ресурсов заканчивает свой расчёт СГ, сервис идентификации извлекает из очереди очередную компоненту вектора \vec{a} и снова загружает его работой.

После исчерпания очереди сервис идентификации формирует систему уравнений (12), решает её с помощью сервиса решения систем уравнений, формирует новое состояние вектора идентифицируемых параметров модели \vec{a} , проверяет условие завершения алгоритма и переходит к следующей итерации либо завершает работу.

Решающее значение в эффективности описанного алгоритма имеет использование операций с сохранением состояния. На каждом из комплектов вычислительных ресурсов моделируемая СГ загружается только один раз. Для каждого расчёта режима передаётся только новое состояние вектора эмпирических параметров \vec{a} . Без возможности сохранения состояния удалённого вычислительного ресурса было бы необходимо для каждого расчёта режима загружать СГ целиком, что привело бы к чрезмерным накладным расходам.

В третьей главе «Программная реализация сервисов моделирования систем газоснабжения на основе разработанной вычислительной среды» описываются разработанные управляющие и вычислительные сервисы решения конкретных задач моделирования СГ, приводится анализ проведённых вычислительных экспериментов, рассматриваются механизмы интеграции разработанных вычислительных сервисов с ПВК СППДР.

На основе предложенной в работе вычислительной среды разработаны как управляющие, так и вычислительные сервисы:

- сервис моделирования режимов СГ, сервис идентификации эмпирических параметров модели СГ;
- семейство взаимозаменяемых вычислительных сервисов для решения систем уравнений: на базе *Intel MKL* для плотных матриц, *Intel MKL* для разреженных матриц, библиотеки *CuSp* для разреженных матриц на *CUDA*;
- обобщённые менеджеры параллельного моделирования произвольных расчётных объектов на центральном процессоре с использованием *OpenMP*, *Intel TBB*; на *CUDA*-устройствах с использованием библиотеки *thrust*;
- оптимизированный менеджер параллельного моделирования трубопроводов на *CUDA* с использованием эффективного для работы *GPU* расположения данных в памяти.

Выполнено тестирование корректности и производительности разработанных вычислительных сервисов с использованием созданного автором механизма тестирования. В качестве базового тестового примера использовалась схема трубопроводной системы Саратов – Горький, которая включает в себя 155 моделируемых объектов, 48 краевых узлов (поставщики/потребители), 18 идентифицируемых параметров модели

режима $K_{то}$, $K_{эф}$. На основе базовой схемы построены тестовые примеры размерности, превосходящей исходную до двухсот раз.

Тесты проведены для всевозможных комбинаций вычислительных сервисов, размеров тестовых примеров, взаимного расположения клиента и сервера. Вычислительные сервисы выполнялись на следующих современных системах:

- ноутбук с ЦП Intel Core i5 (4 ядра) и CUDA-устройством GeForce GTX 460M;
- компьютер с ЦП intel Core i7 (8 ядер) и двумя GeForce GTX 460;
- виртуальная машина в облаке Amazon Web Services (AWS) Elastic Compute Cloud с двумя ЦП Intel Xeon X5570 (4 ядра) и двумя CUDA-устройствами Tesla M2050;
- кластер из четырёх виртуальных машин в облаке.

Выполненные расчеты показали высокую эффективность разработанных в диссертации архитектурных решений, алгоритмов, методов и их программной реализации.

Основные результаты и выводы

1. Спроектирована и реализована компонентная мультизадачная вычислительная среда, позволяющая осуществить унифицированное вычислительное обеспечение СППДР АСДУ на основе параллельных и распределённых вычислений, стандартных протоколов информационного взаимодействия, что является предпосылкой для создания единого вычислительного пространства АСДУ ЕСГ.

2. Впервые разработана архитектура и программно реализован базовый вычислительный сервис итерационного моделирования режимов систем газоснабжения, на основе распараллеленного расчета режимов объектов СГ, с использованием вычислений общего назначения на графических картах.

3. Разработана архитектура и программно реализован вычислительный сервис, организующий параллельное моделирование множества режимов СГ в распределённой среде, на примере многоуровневой итерационной процедуры решения задачи идентификации эмпирических параметров модели режимов СГ.

4. Разработанные сервисы параллельного и распределенного моделирования впервые предоставили возможность решать задачи совместного моделирования единых технологических режимов не только систем большой размерности (тысячи и десятки тысяч расчетных элементов), примерами которых являются региональные газотранспортные системы, но и систем газоснабжения, объединяющих такие разнообразные технологические комплексы, как системы добычи, транспорта, подземного хранения и распределения газа.

5. Выполнено тестирование разработанных вычислительных сервисов в однопользовательском, сетевом, распределённом режимах, на примере реальной системы газоснабжения, показавшее:

- корректность результатов разработанных вычислительных сервисов;
- многократное ускорение этапа моделирования объектов СГ сервисом моделирования режимов СГ за счёт использования вычислений общего назначения на графических картах с использованием технологии CUDA;
- масштабируемость многоуровневых итерационных процессов, основанных на многократном моделировании режимов СГ в распределённой среде – как в локальной сети, так и при использовании удалённых «облачных» вычислительных ресурсов AWS Elastic Compute Cloud, несмотря на заметные накладные расходы удалённого взаимодействия (порядка 10% от общего времени расчёта).

6. Разработаны унифицированные компоненты методологии архитектурной и программной модернизации ПВК моделирования СГ для их последующей интеграции с разработанными вычислительными сервисами, а также в единое информационное и вычислительное пространство АСДУ.

7. В тестовом режиме выполнена интеграция разработанных вычислительных сервисов с ПВК «Веста», внедрённым в ПДС ООО «Газпром трансгаз Санкт-Петербург, Саратов, Уфа», ООО «Газпром добыча Ноябрьск».

Публикации по теме диссертации

1. Васильев А.В., Сарданашвили С.А. Программная архитектура для использования параллельных вычислений при численном балансировании потоков газа на расчётном графе газотранспортной системы // Автоматизация, телемеханизация и связь в нефтяной промышленности. 2012. №1. С. 38-42.
2. Васильев А.В., Леонов Д.Г., Сарданашвили С.А., Швечков В.А. Архитектурные решения интеграции расчетных комплексов моделирования режимов систем газоснабжения в единое информационное пространство автоматизированной системы диспетчерского управления ОАО «ГАЗПРОМ» // Газовая промышленность. 2012. – Т. 683, № 12. С. 62-66.
3. Васильев А.В. Проектирование и реализация компонентной среды параллельного и распределённого моделирования систем газоснабжения // Автоматизация, телемеханизация и связь в нефтяной промышленности. 2013. – №2. С. 32-38.
4. Васильев А.В., Сарданашвили С.А. Программная архитектура для эффективного использования гетерогенных параллельных вычислений при численном моделировании газотранспортных систем // Тезисы докладов девятой Всероссийской конференции Новые технологии в газовой промышленности. – 2011. С. 14.
5. Васильев А.В., Сарданашвили С.А. Программная архитектура расчётно-вычислительного комплекса поддержки диспетчерского управления режимами газотранспортных систем // IX Всероссийская научно-техническая конференция Актуальные проблемы развития нефтегазового комплекса России, тезисы докладов, часть 2. – 2012. С. 92.
6. Васильев А.В. Эффективное численное моделирование транспорта газа многопользовательским сервером с использованием пулов типизированных ресурсов // Сборник тезисов 66-й Международной молодежной научной конференции Нефть и газ 2012. – 2012. С. 8.
7. Васильев А.В., Швечков В.А., Митичкин С.К. Технологии параллельных вычислений в ПВК «Веста» моделирования режимов систем газоснабжения // V Международная научно-техническая конференция ДИСКОМ-2012, тезисы докладов. – 2012. С. 53.
8. Васильев А.В., Леонов Д.Г., Митичкин С.К., Швечков В.А. Архитектурные решения создания единого вычислительного пространства распределённой многоуровневой СППДР, основанные на эволюционном развитии клиент-серверной среды ПВК «Веста» // V Международная научно-техническая конференция ДИСКОМ-2012, тезисы докладов. – 2012. С. 57.

Подписано в печать 18.02.2013
Объём – 1 п.л.

Формат 60x90/16
Тираж 100 экз.

Заказ № 571

119991, Москва, Ленинский просп., 65
Отдел оперативной полиграфии
РГУ нефти и газа имени И.М. Губкина