

ГЕОИНФОРМАТИКА

УДК 004.21, 004.9 + 551 + 622

РАЗРАБОТКА СЕРВИСА ОБЛАЧНЫХ ВЫЧИСЛЕНИЙ И ОБРАБОТКИ ДАННЫХ О СЕЙСМОСОБЫТИЯХ В ГЕОМЕХАНИКО-ГЕОДИНАМИЧЕСКИ АКТИВНЫХ УГЛЕДОБЫВАЮЩИХ РАЙОНАХ КУЗБАССА

В. П. Потапов¹, В. Н. Опарин², О. Л. Гиниятуллина¹, И. Е. Харлампенков¹

¹Институт вычислительных технологий СО РАН, Кемеровский филиал, E-mail: kembict@gmail.com,
ул. Рукавишниковая, 21, 650025, г. Кемерово, Россия

²Институт горного дела им. Н. А. Чинакала СО РАН, E-mail: oparin@misd.nsc.ru,
Красный проспект, 54, 630091, г. Новосибирск, Россия

Дано новое решение проблемы более эффективного использования информационного сервиса для облачных вычислений в задачах контроля геомеханико-геодинамической безопасности в тектонически активном регионе с повышенной сейсмичностью. Апробация выполненной разработки осуществлена на примере анализа сейсмичности территории Кемеровской области за период 01.01.2006–31.12.2009 гг. с применением метода построения траекторий миграции приведенного центра сейсмо-энерговыведения, а также для выделенной зоны в пределах области площадью 150 км². Показано, что реализация разработанного сервиса с использованием облачного хостинга позволяет в несколько раз сократить время работы тестируемого алгоритма обработки сейсмической информации.

Геомеханико-геодинамическая безопасность, угледобывающие районы Кузбасса, геоинформационные технологии, web-сервисы, облачные хостинги, сейсмичность, быстрдействие

Увеличение сейсмической активности в ранее асейсмических районах, вызванное существенными изменениями в состоянии земных недр в результате добычи полезных ископаемых, приводит к необходимости оценки общей геомеханико-геодинамической обстановки угледобывающих районов, особенно в таком крупнейшем в России, как Кузбасс [1–5]. Обычно при решении данной задачи сейсмологами осуществляется регистрация и анализ сейсмограмм из очаговых зон землетрясений и горных ударов и построение соответствующих карт сейсмического районирования. Эта информация объединяется в сейсмологические бюллетени или каталоги и рассматривается как справочник, содержащий сведения о произошедших землетрясениях.

В этом случае сейсмические события рассматриваются с точки зрения классических методов обработки данных как временные ряды. В то же время с позиции геоинформатики сейсмологический бюллетень представляет собой источник пространственных данных: координаты динамических событий отвечают за их пространственную привязку, а дата и время возникно-

вения, глубины очагов и магнитуды источников являются атрибутивными компонентами. Подобный переход к пространственным данным позволяет использовать современные геоинформационные системы и соответствующие методы вычислений для их анализа. При таком подходе важной задачей становится поиск закономерностей во взаимном расположении групп сейсмических событий в различных регионах [5, 6].

В Кемеровском филиале Института вычислительных технологий СО РАН (КФ ИВТ СО РАН) разработана геоинформационная система обработки сейсмических событий горнопромышленных регионов [1], осуществляющая расчет траекторий миграции (по В. Н. Опарину) центров сейсмоэнерговыведения для выбранной территории на основе методики, разработанной в Институте горного дела СО РАН [2, 5]. В настоящее время для широкого применения данная реализация сервиса обладает рядом ограничений: низкая “отказоустойчивость”, сложности с масштабированием, а также высокое потребление системных ресурсов и недостаточная производительность при проведении большого объема расчетов. С учетом роста количества обращений к этому сервису возникла необходимость в разработке нового программного решения для данной задачи [1, 2], которое позволило бы не только устранить выявленные “узкие” места, но и обеспечить расширяемость сервиса. Современный этап развития горной геоинформатики и основные подходы к созданию геоинформационных систем позволяют дать эффективное решение данной проблемы в рамках “облачного” подхода для реализации нового поколения распределенных геоинформационно-вычислительных сервисов [7, 8].

ОБЛАЧНЫЕ ВЫЧИСЛЕНИЯ В РЕАЛИЗАЦИИ РАСПРЕДЕЛЕННЫХ ГЕОИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНЫХ СЕРВИСОВ

В настоящее время под термином “облачные вычисления” обычно понимают возможность получения необходимых вычислительных мощностей по запросу из сети. Ярким примером могут служить поисковые системы, интерфейс которых очень прост, но они предоставляют пользователям большие вычислительные ресурсы для поиска нужной информации. Современные крупные центры обработки информации не только хранят и обрабатывают определенные по качеству данные, но и дают возможность для создания собственных виртуальных Data-центров. Существует много определений “облачных вычислений”. Зачастую они расходятся в своем значении и акцентах [6–12]. В рамках развиваемого в КФ ИВТ СО РАН научного геоинформационного направления под термином “облачные вычисления” понимается комплекс требований по предоставлению услуг (сервисов) через Интернет [7]. Это позволяет строить облака как в локальных, так и глобальных сетях, ориентируясь на их конкретную специфику и предоставляемые типы услуг.

Для реализации облачных платформ обычно используется сценарий, связанный с размещением различных web-проектов и интернет-сервисов в виде “хостинга”, для которого характерны требования к эластичности облачной платформы, необходимости размещения больших объемов данных, балансировке нагрузки, использованию технологии Content Delivery Network (CDN). В этом случае IaaS используется только для “особых проектов”, которые не могут по разным причинам быть реализованы на платформе PaaS, например из-за большого унаследованного кода. SaaS также иногда используется в “хостинге”, а поскольку кастомизации в SaaS практически нет, то нет и уникальности интернет-сервисов. SaaS используется только для быстрого размещения простых web-решений (сервис типа wordpress.com).

Для сценария “хостинг” наиболее перспективным представляется использование PaaS, где выполняется кастомизированная разработка интернет-сервисов. Реализации PaaS включают средства для создания, тестирования, развертывания, размещения и управления приложениями,

а также для поддержки жизненного цикла их разработки. Web-интерфейсы инструментов создания приложений основаны на таких технологиях, как HTML, JavaScript и др. Провайдеры PaaS часто включают услуги для управления параллельной обработкой, масштабируемостью, отказоустойчивостью и безопасностью. Важной особенностью является интеграция с web-службами и базами данных. Поддержка протокола обмена структурированными сообщениями в распределенной вычислительной среде (Simple Object Access Protocol, SOAP) и других интерфейсов позволяют приложениям PaaS создавать комбинации web-сервисов, которые называют mashup. Способность формировать и распространять код между специализированными, предопределенными или распределенными командами увеличивает производительность системы в целом. Однако, несмотря на выбор в качестве платформы реализации облачного сервиса обработки геодинамических данных хостинга PaaS, возникают определенные вопросы в выборе инструментов реализации приложения.

В настоящее время в сети Internet существует достаточно много хостингов для разработки приложений [13–15]. Для задач разработки специализированных приложений обработки геомеханико-геодинамических данных выделены хостинги Jelastic и OpenShift, которые обладают следующими возможностями: автоматическое масштабирование, поддержка любых JVM приложений, в том числе написанных на Java 6 и 7, система контроля версий, поддержка базы данных PostgreSQL с модулем Postgis, контроль и анализ приложений [14]. В дальнейшем наиболее удобным с позиции разработки данного приложения признан хостинг Jelastic.

Важный вопрос при создании системы — выбор инструмента отображения пространственных данных. Одним из ключевых компонентов информационной системы обработки пространственных данных является инструмент визуализации геоданных. В данной подсистеме можно выделить клиентскую часть, которая непосредственно отвечает за визуализацию и формирование основ для построения пользовательского интерфейса. В качестве клиентского компонента можно использовать Google Maps, Oracle Maps, OpenLayers. Карты Google — это свободный доступ к набору приложений, построенных на основе бесплатного картографического сервиса и технологии, предоставляемых компанией Google. В рамках данного решения используется формат kml, но отсутствует поддержка спецификаций OGC, в частности WMS и WFS [16]. Oracle Maps является дополнением к Oracle MapViewer и служит для отображения в браузере генерируемых им карт, позволяя пользователям управлять отображаемой информацией, наносить на электронные слои свои пометки, рисовать различные фигуры, но библиотека не соответствует стандартам WMS и WFS [17].

Библиотека OpenLayers позволяет создать web-интерфейс для отображения картографических материалов, представленных в различных форматах и расположенных на различных серверах. Поддерживается добавление слоев, предоставляемых WMS и WFS серверами. Возможно использование данных картографических сервисов Google. Такое решение является программным обеспечением с открытым исходным кодом [18]. Исходя из перечисленных возможностей и с учетом специфики данных разрабатываемой системы, использовалась библиотека OpenLayers.

РЕАЛИЗАЦИЯ ОБЛАЧНОГО СЕРВИСА

При проектировании сервиса обработки геомеханико-геодинамических данных выбрана традиционная для web-приложений трехуровневая архитектура (клиент – сервер приложений – сервер базы данных) с добавлением картографического сервера GeoServer, обеспечивающего публикацию пространственных данных (рис. 1).

В качестве хранилища данных используется СУБД PostgreSQL, дополненная расширением PostGIS для управления пространственными данными. Такое решение обеспечивает приемлемый уровень надежности и быстродействия системы.

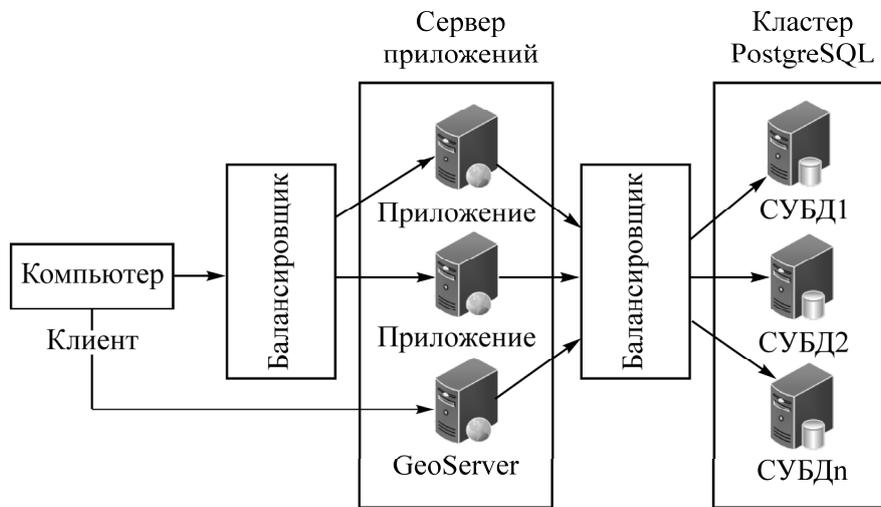


Рис. 1. Архитектура web-сервиса обработки геодинамических данных

В роли сервера приложений выступает Apache Tomcat 7, сопряженный с технологией создания web-приложений Java. В данном компоненте сосредоточен код приложения, отвечающий за обработку пользовательских запросов, формирование интерфейса, генерацию запросов к картографическому серверу. Вычислительный модуль представляет собой независимый компонент, функционирующий в рамках сервера приложений. В данном решении имеется программный код основных методов обработки геоданных, используемых в прикладной информационной системе. Для взаимодействия с внешними подсистемами применяется SOAP — протокол обмена структурированными сообщениями в распределенной вычислительной среде. Такой подход позволяет легко наращивать функционал системы и предоставляет возможность превратить вычислительный модуль в независимый web-сервис.

Клиент построен на основе комбинации HTML и JavaScript библиотеки extJs. В качестве средства отображения электронных карт используется библиотека GeoExt, которая включает в себя OpenLayers. Для получения информации об отдельных элементах векторных слоев используется технология AJAX, а для взаимодействия с вычислительной подсистемой применяется JavaScript реализация протокола SOAP.

При проектировании сервиса особое внимание уделялось структуре базы данных для хранения сейсмобюллетеней. База данных состоит из трех независимых таблиц — по одной для каждого источника данных. В структуру таблиц введен атрибут geom, который необходим для хранения пространственных данных в СУБД PostgreSQL с применением расширения PostGIS. Данное поле позволяет осуществлять привязку сведений о произошедшем динамическом событии к определенной точке на местности с помощью географических координат “широта–долгота” в проекции WGS84.

Программное решение сервиса разработано на языке программирования Java. Для вычислений с использованием пространственных данных применялась библиотека Net Topology Suit, доступ к основным функциям которой реализован через web-интерфейс сервиса. Интерфейс пользователя включает в себя электронную карту, панель инструментов и панель управления.

Панель инструментов представляет собой меню для изменения режимов отображения карты (масштабирование, перемещение, включение и выключение слоя космоснимков). Панель управления содержит функциональные блоки: отображения списка доступных на текущий момент информационных слоев и манипулирования ими. Каждому слою соответствует элемент управления, позволяющий включать или отключать его визуализацию на карте. Пространственные объекты, нанесенные на карту, являются интерактивными, атрибутивная информация предоставляется по запросу пользователя.

Для выполнения расчетов запускается вычислительный модуль на панели инструментов, который выводит окно ввода параметров обработки. При этом необходимо указать координаты центра области расчетов, ее размеры (ширина, длина), выбрать источник и временной интервал.

Запуск вычислительного модуля приводит к вызову серии удаленных функций, выполняющих расчет центров миграции сейсмоэнерговыведения и фрактальных размерностей [1, 2]. Результаты их работы передаются на уровень web-интерфейса, где происходит отрисовка траектории перемещения приведенных центров сейсмоэнерговыведения [1], формируется итоговая таблица результатов вычислений.

СРАВНЕНИЕ ВРЕМЕНИ ВЫПОЛНЕНИЯ СЦЕНАРНЫХ РАСЧЕТОВ С ИСПОЛЬЗОВАНИЕМ ОБЛАЧНОГО СЕРВИСА

Разработанный сервис использован для анализа совокупности сейсмических событий с позиций [1, 2] по некоторым регионам Сибири с целью сравнения быстродействия программы в “облаке” и “вне облака”. Выбраны соответствующие данные, предоставленные Территориальным центром мониторинга, лабораторного контроля прогнозирования чрезвычайных ситуаций природного и техногенного характера (ТЦМП, г. Кемерово, www.spasko.ru) и сайтом Международного сейсмологического центра (ISC, www.isc.ac.uk). Для проведения исследований использовались локальный и облачный хостинги, характеристики которых приведены ниже (отметим, что ограничение требуемых информационных ресурсов в облачном хостинге присутствует только в бесплатной версии):

	Локальный хостинг	Облачный хостинг
Процессор, МГц	2400	400–3000
Оперативная память, МБ	4000	256–1.88

При проведении сравнения на первом этапе для заданной территории выполнялась выборка всех сейсмических событий из базы данных за указанный временной интервал. Осуществлялось преобразование исходных форматов данных источников в локальный формат сервиса: выделялись только поля, связанные непосредственно с расчетом, и строилась соответствующая динамическая таблица. Выполнялся расчет траектории миграции центров сейсмоэнерговыведения, который передавался в модуль вычисления фрактальных размерностей. Результаты выводились на экран пользователю, при этом контроль скорости вычисления проводился с помощью библиотеки профилирования приложений для локального хостинга и средствами консоли администрирования облачного хостинга. Далее приведены результаты изменения скорости тестовых расчетов для территории Кемеровской области. Расчеты траекторий миграций центров сейсмоэнерговыведения для других областей Сибири (Алтайский край, Республика Хакасия, Томская и Иркутская области) показали аналогичные результаты, что говорит о стабильной работе облачного сервиса.

Примеры результатов расчета тестовых данных в облачном сервисе. В первом тесте “окно расчетов” покрывает всю площадь Кемеровской области и охватывает временной период с 01.01.06 по 31.12.09 г. Время выполнения расчетов: на локальном хостинге 7 мин 28 с, на облачном — 4 мин 44 с. При увеличении нагрузки облачный хостинг увеличил мощность процессора

до 2382 МГц (рис. 2а). В втором тесте “окно расчетов” покрывает территориальную зону только в 150 км² от центра области и охватывает временной период с 01.01.06 по 31.12.09 г. Время выполнения расчетов: на локальном хостинге 47 с, на облачном — 15 с. При увеличении нагрузки облачный хостинг увеличил мощность процессора до 524 МГц (рис. 2б).

Заметим, что в упомянутых тестах использовался бесплатный вариант облачного хостинга, который ограничивает ресурсы до 15 клаудлетов (1 клаудлет равен 128 МБ оперативной памяти и 200 МГц процессора). При использовании платного варианта облачного хостинга полученные результаты дают весьма значимую финансовую выгоду.

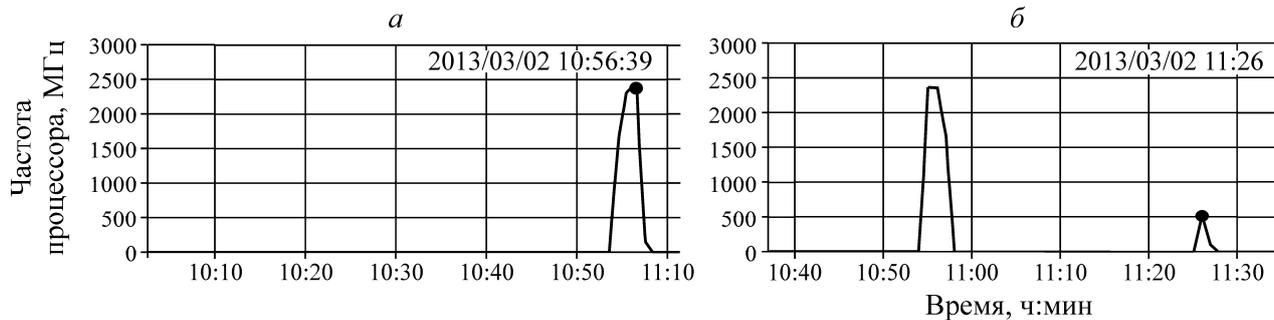


Рис. 2. Увеличение мощности процессора во времени: а — в первом тесте; б — во втором

ВЫВОДЫ

Развитие современных облачных геоинформационных технологий привело к возникновению принципиально нового класса сетевой продукции — web-сервисов на основе облачных хостингов. Это позволяет значительно увеличить производительность региональной геоинформационной системы, а также отказаться многим горнодобывающим компаниям от необходимости создания и поддержки собственных вычислительных мощностей при формировании и развитии локальных мониторинговых систем геомеханико-геодинамической безопасности промышленных предприятий. Получены следующие основные результаты:

- установлено, что реализация геоинформационных вычислений на базе облачных технологий наиболее предпочтительна при использовании PaaS-платформ, предоставляющих широкие возможности комбинации web-сервисов, параллельной обработки, масштабирования, отказоустойчивости;

- для задач разработки специализированных приложений обработки геомеханико-геодинамических данных с учетом специфики форматов пространственных данных и вычислительных алгоритмов наиболее удобным для размещения облачного сервиса признан хостинг Jelastic, который поддерживает JVM-приложения, средства интеграции сторонних модулей, спецификации OGC и соответствие стандартам WMS и WFS;

- выбор трехуровневой архитектуры для реализации облачного сервиса обработки геомеханико-геодинамических данных позволил встроить картографическую подсистему как для этапа выбора исходных параметров по координатам, так и для вывода результатов расчетов. Сочетание предложенной архитектуры и облачной платформы дает возможность оперативно реагировать на увеличение числа пользовательских запросов, добавляя новые вычислительные ресурсы только на критическом слое, связанном с выполнением расчетов, обеспечивая тем самым масштабируемость системы в целом;

- проведенные тестовые расчеты на созданном облачном сервисе позволили оценить гибкость данного подхода к построению систем геоинформационных вычислений, его результативность и увеличение производительности в 2 раза по сравнению с локальными реализациями.

СПИСОК ЛИТЕРАТУРЫ

1. **Опарин В. Н., Потапов В. П., Попов С. Е., Замаев Р. Ю., Харлампенков И. Е.** Разработка распределенных ГИС-средств мониторинга миграций сейсмических проявлений // ФТПРПИ. — 2010. — № 6.
2. **Опарин В. Н., Потапов В. П., Гиниятуллина О. Л., Харлампенков И. Е.** Фрактальный анализ траекторий миграции геодинамических событий в Кузбассе // ФТПРПИ. — 2012. — № 3.
3. **Опарин В. Н., Еманов А. Ф., Востриков В. И., Цибизов Л. В.** О кинематических особенностях развития сейсмоземиссионных процессов при отработке угольных месторождений Кузбасса // ФТПРПИ. — 2013. — № 4.
4. **Еманов А. Ф., Еманов А. А., Фатеев А. В., Лескова Е. В., Шевкунова Е. В., Подкорытова В. Г.** Техногенная сейсмичность разрезов Кузбасса (Бачатское землетрясение 18 июня 2013 г.) // ФТПРПИ. — 2014. — № 2.
5. **Деструкция земной коры и процессы самоорганизации в областях сильного техногенного воздействия** / [В. Н. Опарин и др.]; отв. ред. Н.Н. Мельников. — Новосибирск: Изд-во СО РАН, 2012.
6. **Потапов В. П., Опарин В. Н., Логов А. С., Замаев Р. Ю., Попов С. Е.** Геоинформационная система регионального контроля геомеханико-геодинамических ситуаций на основе энтропийного анализа сейсмических событий (на примере Кузбасса) // ФТПРПИ. — 2013. — № 3.
7. **Бычков И. В., Опарин В. Н., Потапов В. П.** Облачные технологии в решении задач горной геоинформатики // ФТПРПИ. — 2014. — № 1.
8. **OnLine Maps with API and Webservice.** Ed prof. Michael Peterson, Springer, New York, Heidelberg, 2012.
9. **Above the Clouds.** A Berkeley View of Cloud Computing. USA, California, 2009.
10. **Massino Cafaro, Giovanni Alliso,** eds. Grid, Clouds and Virtualization, London, New York: Springer, 2011.
11. **Сороко В.** Облачные вычисления и сервисы: классификация, основные функции, преимущества и недостатки / International Conference of developers and users of free/open source software. August 2014, http://lvee.org/en/reports/LVEE_2011_03.
12. **Романченко В.** Облачные вычисления на каждый день // 3Dnews. Daily Digital Digest [эл. журнал]. http://www.3dnews.ru/editorial/cloud_computing.
13. **Сводная таблица** Google App Engine vs. Microsoft Windows Azure vs. Salesforce.com (Force.com), <http://cloud-computing.findthebest.com/compare/15-17-19-123/Google-App-Engine-vs-Microsoft-Windows-Azure-vs-Salesforce-com-Force-com>.
14. **Jelastic.** <https://jelastic.com>.
15. **OpenShift.** <https://www.openshift.com>.
16. **API Google Карт.** <https://developers.google.com/maps/?hl=ru>.
17. **Oracle Fusion Middleware MapViewer.** <http://www.oracle.com/technetwork/middleware/mapviewer/overview/index-090261.html>.
18. **OpenLayers: Free Maps for the Web.** <http://openlayers.org>.

Поступила в редакцию 1/II 2015