

Макет отказоустойчивой информационной системы для облачного хранения наборов научных данных*

А. А. Сорокин¹, С. П. Королёв¹, С. И. Смагин¹, А. Н. Поляков²

¹Вычислительный центр ДВО РАН, Хабаровск, Россия

²НИЦ “Курчатовский институт”, Москва, Россия

e-mail: alsor@febras.net, serejk@febras.net,

smagin@ccfebras.ru, andrew@kiae.ru

На примере архивов сейсмологической информации рассмотрены проблемы хранения больших и сверхбольших наборов научных данных. Описан созданный экспериментальный образец (макет) облачной информационной системы для распределённого хранения инструментальных сейсмологических данных, развёрнутый на базе ресурсов географически удалённых центров данных Вычислительного центра ДВО РАН и НИЦ “Курчатовский институт”. Проведено исследование эффективности представленной системы, показавшее перспективность использования облачных вычислений для решения задач распределённого хранения и обработки больших наборов научной информации.

Ключевые слова: облачные вычисления, файл, структурированные наборы научных данных, архив, автоматизированная информационная система.

Введение

Увеличение качественных и количественных характеристик различных инструментальных средств наблюдения, а также повышение производительности вычислительных ресурсов приводят к росту объёмов экспериментальных данных и результатов расчётов, требующих их хранения и дальнейшего использования. В настоящее время петабайтные массивы данных уже получены в результате сканирования звёздного неба и моделирования истории и структуры Вселенной, дистанционного зондирования из космоса поверхности и геофизических исследований внутреннего строения Земли.

Существующие методы организации (реляционная модель) и хранения научных данных (в отрыве от обрабатывающих вычислительных систем) уже не способны эффективно поддерживать всё возрастающие потоки и разнообразие моделей представления данных, а также не всегда являются эффективными с точки зрения применения технологий распределённой и параллельной обработки данных, потребности в которых постоянно растут.

Меняются и подходы к проведению научных работ. Отмечается растущая тенденция к выполнению междисциплинарных исследований. В этом случае исследователям необходим доступ к большим наборам данных, хранящимся в разных географических точках [1–5]. Так, понимание и прогнозирование долгосрочных последствий экологических катастроф, например разлива нефти, требует детального анализа спутниковых

*Работа выполнена в рамках ГК № 07.514.11.4129 по ФЦП “Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2013 годы”.

изображений поверхности, химического состава, биологических и экологических характеристик океана, а также создания и использования сложных океанографических и атмосферных моделей. Для исследования глобальных проблем, таких как изменение климата или персонифицированная медицина, требуется создание ещё более сложных моделей, для чего необходимо получать, обрабатывать и анализировать массивы данных больших объёмов, в том числе в реальном времени.

Чтобы решать проблемы извлечения новых знаний из сверхбольших объёмов данных, исследователи всё чаще обращаются к организации распределённых и параллельных вычислений в облаке (data cloud) [6–8]. На практике облако создаётся на базе информационно-вычислительной инфраструктуры центров обработки данных, содержащей в достаточном количестве серверы и дисковые носители. Для надёжного хранения в прямом доступе наборов научных данных последние должны быть распределены и тиражированы на большом количестве таких ресурсов [9], для работы с которыми необходимы технологии и интерфейсы программирования приложений (API).

Объектом исследования настоящей работы является макет отказоустойчивой облачной информационной системы для распределённого хранения и работы с инструментальными данными сети сейсмологических наблюдений ДВО РАН. Данная система разработана и развёрнута на базе географически удалённых программно-аппаратных ресурсов Вычислительного центра ДВО РАН (ВЦ ДВО РАН) и НИЦ “Курчатовский институт” с использованием компоненты Swift (OpenStack Object Storage) программной платформы OpenStack, позволяющей создавать частные и публичные вычислительные облака и облачные хранилища данных [10, 11]. Представлено краткое описание разработанной информационной системы и результатов проверки её эффективности на реальных задачах хранения и обработки инструментальных данных сети сейсмологических наблюдений ДВО РАН.

1. Формулировка проблемы

В Дальневосточном отделении РАН создаётся сеть сейсмологических наблюдений (далее Сеть). Аппаратная часть её измерительного оборудования в настоящее время представлена семью современными наземными сейсмическими станциями, каждая из которых укомплектована трёхкомпонентным широкополосным сейсмометром RefTek 151-120, цифровым регистратором RefTek 130-01 и GPS-приемником [12].

Исходные измерения в процессе наблюдений записываются в файлы длиной 60 с в бинарном формате PASSCAL [13], а затем в отложенном режиме или в реальном времени посредством каналов связи с использованием протокола сетевого уровня RTP [14] передаются в Центр данных ВЦ ДВО РАН. Управление Сетью и работа с архивом инструментальных данных осуществляются с использованием автоматизированной информационной системы (АИС) “Сигнал-С” [15], разработанной на основе архитектуры “клиент-сервер”, где в качестве клиентов выступают цифровые регистраторы, установленные в пунктах наблюдений, а в качестве серверов — соответствующие программно-аппаратные компоненты АИС “Сигнал-С” (рис. 1).

В результате работы АИС за сутки из одного пункта наблюдений в ВЦ ДВО РАН передаётся в среднем 1440 файлов общим объёмом 50 МБ. Размер архива исходных данных, накопленного за весь период работы Сети, составляет более 162 ГБ с общим числом файлов около 5 млн единиц.

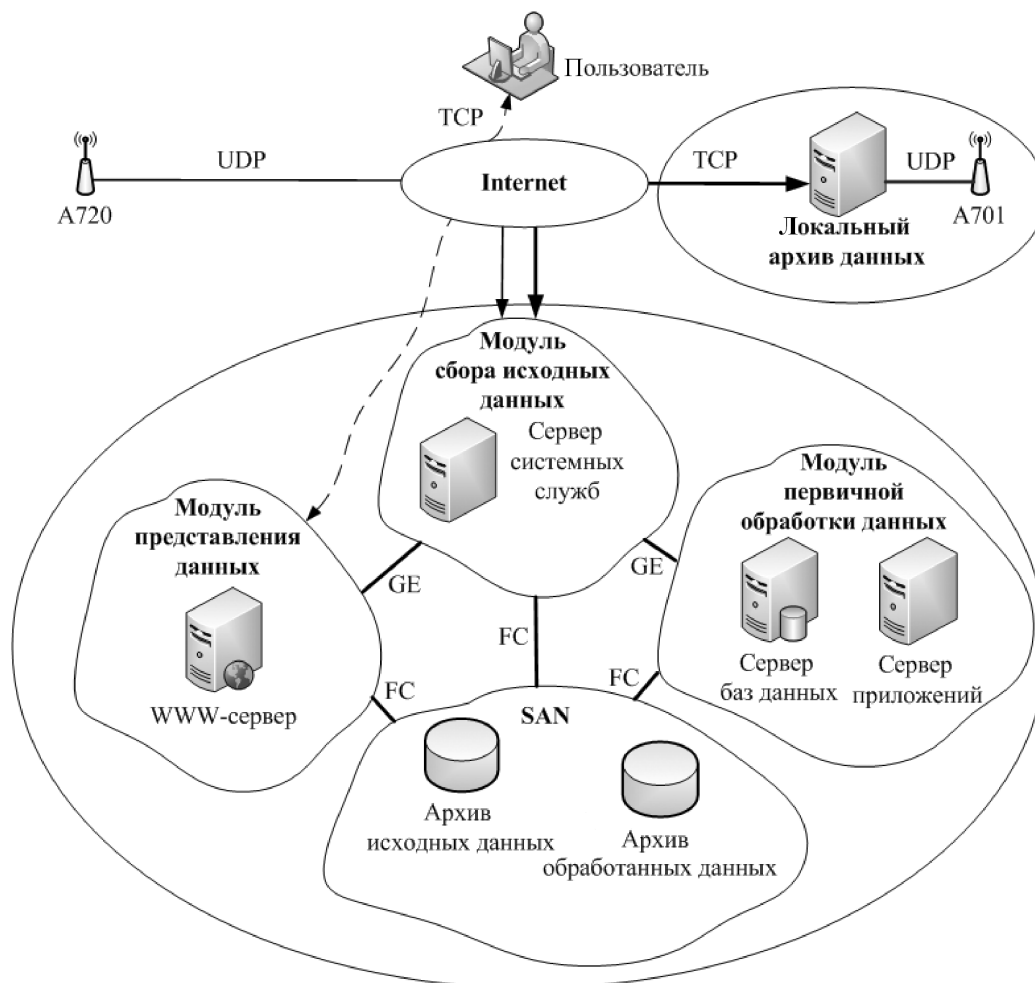


Рис. 1. Схема работы АИС "Сигнал-С"

С развитием Сети, её постепенным переходом на оперативный режим работы и соответственно с ростом количества наборов и объёмов передаваемых и обрабатываемых данных возникают задачи разработки отказоустойчивой системы сбора и хранения инструментальных сейсмологических данных, а также обеспечения гарантированного доступа к архивам наблюдений.

2. Методы и технологии решения проблемы

Эффективным решением вышеуказанных задач является применение облачных систем и сервисов, позволяющих предоставлять виртуальные ресурсы с требуемыми характеристиками и обеспечивать гибкий механизм их интеграции и взаимодействия с уже существующими прикладными информационными системами. В этом случае для выполнения низкоуровневых операций необходимы масштабируемые программно-аппаратные системы, обладающие высокой отказоустойчивостью, системой авторизации пользователей и обеспечивающие надёжный доступ к данным. В качестве такого инструмента был применён свободно распространяемый программный продукт OpenStack Swift (далее "Swift") [10, 11].

Swift — это система облачного хранения данных, которая позволяет создавать хранилища статических данных, обеспечивая при этом их избыточность и доступность на программном уровне. Целостность данных поддерживается на программном уровне за счёт регулярной репликации между узлами хранения. В результате можно формировать горизонтально масштабируемые облака больших объёмов данных без использования дорогих аппаратных RAID-решений с обеспечением высокого уровня доступности их ресурсов. Перечисленные функции системы и лицензионные условия ее использования выделяют Swift среди других платформ для организации распределённого хранения данных [8].

На базе инфраструктуры центров данных ВЦ ДВО РАН и НИЦ “Курчатовский институт” были развёрнуты две зоны Swift разной конфигурации с четырьмя вычислительными узлами в каждой. Общие требования к архитектуре — использование процессоров Intel с тактовой частотой не ниже 2.9 ГГц, оперативной памяти 4 ГБ и двух жёстких дисков для реализации системного программного обеспечения и непосредственного хранения данных. Связь между зонами осуществлялась посредством существующих в организациях Internet-соединений (рис. 2). В результате было организовано облачное хранилище в виде инфраструктуры как сервиса (IaaS) [16], которую необходимо интегрировать с АИС “Сигнал-С”.

Архитектура Swift состоит из следующих четырёх основных компонентов [17]:

- proxy server — обрабатывает внешние HTTP REST-запросы к облаку на выполнение определённой операции (авторизация, загрузка/выгрузка объекта, отображение информации о контейнере и т. д.), взаимодействует с вычислительными узлами, обеспечивая размещение объекта внутри облака;
- object server — предоставляет хранилище BLOB-объектов, реализует весь комплекс операций по хранению объекта внутри отдельного вычислительного узла и репликации с другими узлами в кластере;
- container server — хранит и обрабатывает список объектов внутри контейнеров, а также статистические данные по их количеству и объёму; указанная информация помещается в базу данных SQLite, реплицируемую между всеми узлами хранения;
- account server — осуществляет сопоставление контейнеров и зарегистрированных в системе пользователей.

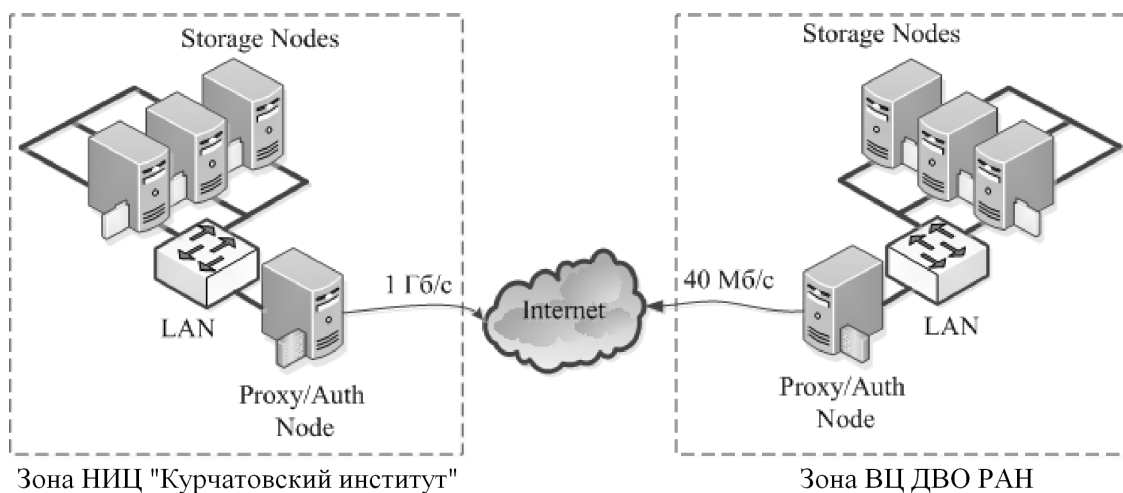


Рис. 2. Архитектура организованного облака данных

Основная единица хранения такого облака — объект, представляющий собой файл с информацией. Для систематизации группы объектов используется понятие контейнера — аналога директории внутри файловой системы. Для каждого загружаемого файла Swift производит сопоставление положения файла и его имени внутри SQLite базы данных. При этом формируются как минимум три копии объекта на разных вычислительных узлах, что обеспечивает их сохранность на случаи аппаратных сбоев оборудования. Любой объект может быть снабжен метаданными, хранящейся в расширенных атрибутах файла (*xattr*).

С применением языка программирования PHP и *Symfony framework*¹, позволяющего создавать приложения на основе шаблона Модель-Представление-Контроллер (*Model-View-Controller*) [18], была разработана модификация АИС “Сигнал-С”, получившая название “Сигнал-Облако”. Модификация состоит из пяти взаимосвязанных программных модулей, обеспечивающих работу с организованным облачным хранилищем через уже существующие пользовательские интерфейсы АИС “Сигнал-С”.

Модули, используя интерфейс программирования приложений (*API*) *HTTP RESTful* для *OpenStack Swift*, реализованный в виде библиотек для языков программирования PHP и Python, выполняют весь набор операций по созданию и удалению контейнеров, загрузке и выгрузке объектов, формируя в облаке идентичные по содержанию копии архивов инструментальных сейсмологических данных (рис. 3).

Получаемая от сети наблюдения измерительная информация помещается в облако в виде псевдоиерархической файловой системы. Это связано с тем, что в контейнере Swift нельзя реализовать стандартную структуру директорий обычной файловой системы и все операции осуществляются строго с бинарными объектами. Чтобы перенести в облако необходимую структуру хранения файлов и директорий, последние были реализованы как объекты с нулевым размером и типом данных “*application/directory*”, называемых *directory*-маркерами. Указанные маркеры создаются на этапе загрузки инструментальных данных в облако и позволяют в дальнейшем формировать из модулей АИС “Сигнал-С” запросы на выборку данных по требуемым параметрам (день, пункт наблюдений и т. п.) с последующей их обработкой и доступом к ним через пользовательские интерфейсы.

3. Проверка эффективности работы системы

Для проверки функционирования разработанной информационно-вычислительной системы был сформирован эталонный набор данных, состоящий из четырёхдневного архива инструментальных сейсмологических данных (с 4 по 7 июня 2012 г.) по трём станциям сети наблюдений ДВО РАН (Чегдомын, Ванино, Углегорск) с общим количеством файлов (объектов) 17 260 единиц и объёмом 600.2 Мбайт.

Для оценки надёжности, скорости доступа и масштабируемости программного комплекса созданы методики тестирования и проведено исследование производительности системы, включающее измерение времени выполнения операций по загрузке в облако и выгрузке из него эталонного набора данных.

Тестирование осуществлялось путём формирования запросов на получение данных через соответствующие формы пользовательского интерфейса автоматизированной ин-

¹<http://www.symfony-project.org/>

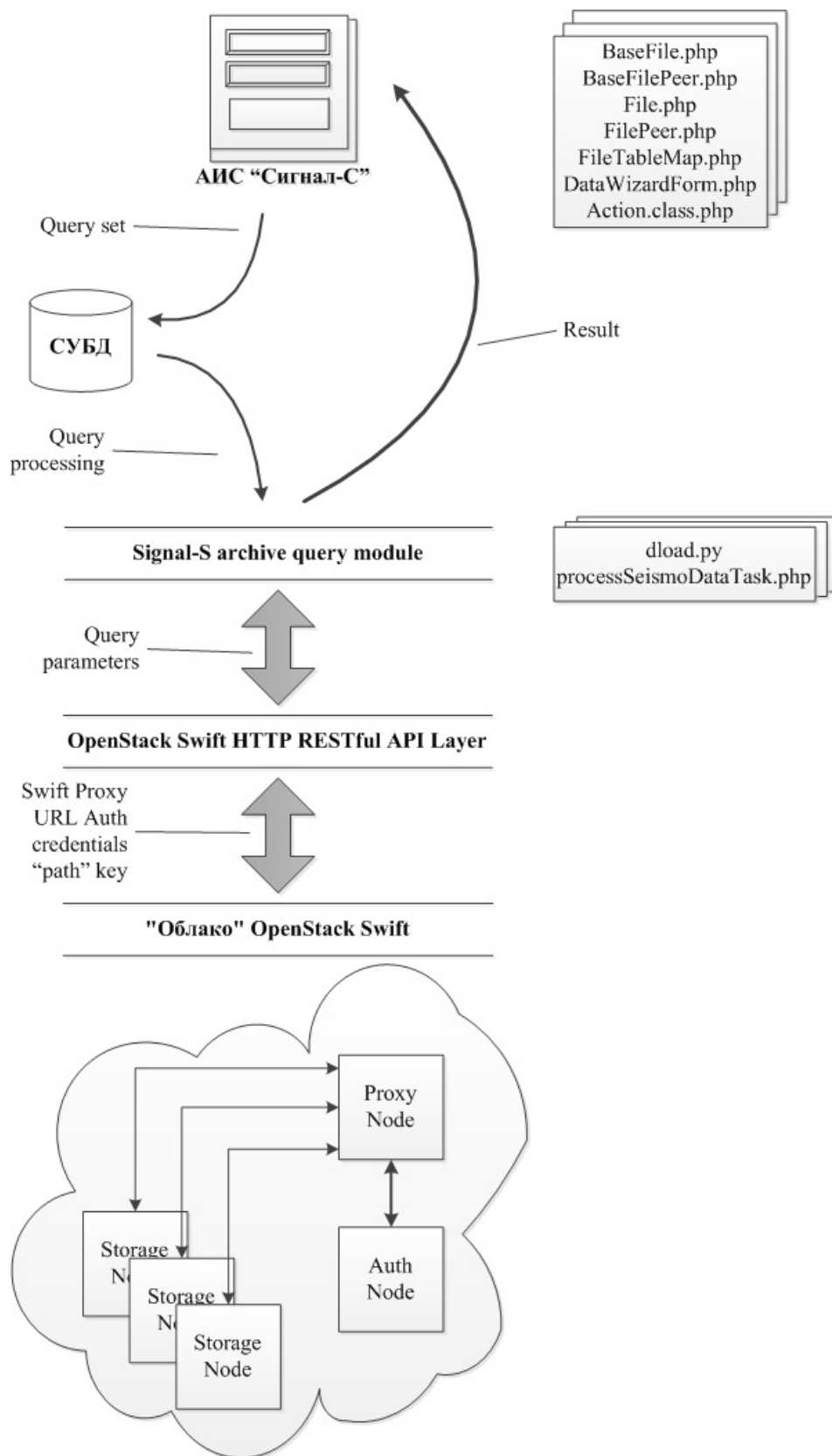


Рис. 3. Схема работы АИС "Сигнал-С" с разработанным программным расширением "Сигнал-Облако"

формационной системы “Сигнал-С” (рис. 4), компоненты которой размещены в центре данных ВЦ ДВО РАН.

При исследовании производительности системы было установлено время взаимодействия с зонами облачного хранилища:

| | Загрузка данных, с | Выгрузка данных, с |
|----------------------------------|--------------------|--------------------|
| Зона НИЦ “Курчатовский институт” | 8670 | 6527 |
| Зона ВЦ ДВО РАН | 7594 | 2613 |

Полученные данные позволяют сделать следующие выводы.

1. Время загрузки всего набора данных складывается из времени, затраченного на формирование списка файлов с данными по запрашиваемому периоду, и времени на накладные расходы, связанные с обработкой и передачей этой информации как внутри прикладной системы, так и в облаке. Его можно уменьшить за счёт использования более эффективных алгоритмов обработки данных и технологий взаимодействия компонентов единой системы.

2. Значительное отличие времени загрузки данных от времени их выгрузки из облака ($7594/2613 = 2.9$) связано с тем, что при выполнении первой операции происходят формирование/обновление большого числа объектов и заполнение SQLite-базы соответствующей им метаинформацией.

Сигнал-с
Автоматизированная Информационная Система

Главная | Новости | О системе | Программа | Ссылки | Помощь

Мастер запросов

Шаг 1 - Выбор даты

Начало временного отрезка
Окончание временного отрезка

Шаг 2 - Выбор станции

Станции
Все станции
Определите станцию (или выберите все станции) по которой будет отобрана информация согласно заданного временного интервала.

Шаг 3 - Выбор каналов

Каналы
ВНЗ/ВНН/ВНЕ
1-ый - ВНЗ, 2-ой - ВНН и 3-ий - ВНЕ

Шаг 4 - Выбор операций

сохранить файлы в корзину
 склеить файлы по заданным параметрам и поместить результат в корзину
 склеить файлы по заданным параметрам, сконвертировать и поместить результат в корзину

Добавить задание

Рис. 4. Экранная форма АИС “Сигнал-С”: формирование запроса на получение данных из архива, размещённого в облаке

3. Использование ресурсов удалённых вычислительных узлов увеличивает время на выполнение операций. В значительной степени это связано с естественными временными затратами на передачу данных между городами Хабаровск и Москва, где расположены узлы хранения. Для приложений, в которых этот показатель является важным, при конфигурации зон хранилища рекомендуется задавать максимальный приоритет использованию ближайших к пользователям узлов, обеспечивающих более высокую скорость работы с данными. Эти правила будут учтены проху-сервером при взаимодействии с узлами хранения. Работа облачного хранилища может базироваться на использовании как выделенных, так и Internet-каналов, учитывая, что первые обеспечивают гарантированную пропускную способность и, как правило, более короткий маршрут между узлами, а следовательно, и меньшее время обмена данными между ними.

4. Для определения отказоустойчивости комплекса в процессе выгрузки данных из облака были одновременно отключены несколько вычислительных узлов, причём выполняемая операция продолжилась и успешно завершилась. Эта характеристика является важной и необходимой при работе с данными, поступающими в информационную систему в режиме реального времени.

Заключение

Разработанная система позволяет создать отказоустойчивое облачное хранилище и сервисы доступа к инструментальным сейсмологическим данным на основе масштабируемой программной платформы OpenStack Swift. За счёт их интеграции с АИС «Сигнал-С» можно эффективно решать задачи сбора и хранения инструментальных сейсмологических данных, а также обеспечить гарантированный доступ пользователей к архивам наблюдений.

Проведённое тестирование макета информационной системы для распределённого хранения данных показало высокий уровень надёжности организованного хранилища и разработанного программного обеспечения. Параметры работы информационной системы обеспечивают эффективное решение поставленных задач. Это указывает на возможность применения созданного экспериментального образца системы и результатов проведённых исследований для сбора, хранения и обработки больших и сверхбольших наборов данных, в том числе в рамках комплексной программы фундаментальных научных исследований ДВО РАН «Современная геодинамика, активные геоструктуры и природные опасности Дальнего Востока России».

Список литературы

- [1] Сорокин А.А., Королёв С.П., Холобуда С.П., Малонушенко О.Г. Интеграция данных по катастрофическим явлениям Дальнего Востока России // Информационные технологии и высокопроизводительные вычисления. Хабаровск, 2011. С. 239–242.
- [2] Шокин Ю.И., Федотов А.М. Распределённые информационные системы // Вычисл. технологии. 1998. Т. 3, № 5. С. 79–93.
- [3] SAMET R., TURAL S. Web based real-time meteorological data analysis and mapping information system // Intern. J. of Education and Informat. Technologies. 2010. Vol. 4, iss. 4. P. 187–196.

- [4] ШОКИН Ю.И., ЖИЖИМОВ О.Л., ПЕСТУНОВ И.А. и др. Распределённая информационно-аналитическая система для поиска, обработки и анализа пространственных данных // Вычисл. технологии. 2007. Т. 12, Спец. выпуск 3. С. 108–115.
- [5] ГОРДОВ Е.П., ОКЛАДНИКОВ И.Г., ТИТОВ А.Г. Разработка Веб-ГИС системы для статистического анализа больших архивов пространственно распределённых данных при проведении комплексных геофизических исследований // Рос. научный электронный журн. 2009. № 1.
- [6] PALLICKARA S., PALLICKARA S., PIERCE M. Scientific data management in the cloud: A survey of technologies, approaches and challenges // Handbook of Cloud Computing. Springer Science+Business Media, 2010. P. 517–533.
- [7] ZHIZHIN M., MEDVEDEV D., MISHIN D. ET AL. Transparent data cube for spatiotemporal data mining and visualization // Grid and Cloud Database Management. Springer, 2011. P. 319–341.
- [8] ANTONOPOULUS N., GILLAM L. Cloud Computing: Principles, Systems and Applications. Springer, 2010. 396 p.
- [9] ЖИЖИН М.Н., ПОЛЯКОВ А.Н., ПОЙДА А.А., МЕДВЕДЕВ Д.П. Разработка распределённых алгоритмов и высокопроизводительной программной системы для облачного хранения, потоковой обработки и сбора в реальном времени сверхбольших наборов научных данных // Параллельные вычислительные технологии. Челябинск, 2012. С. 479–484.
- [10] REPPLE K. Deploying OpenStack. Creating Open Source Clouds. O'Reilly Media, 2011. 88 p.
- [11] PETCU D., MACARIU G., PANICA S., CRACIUN C. Portable cloud applications — from theory to practice // Future Generation Computer Systems. Elsevier, 2012.
- [12] ХАНЧУК А.И., КОНОВАЛОВ А.В., СОРОКИН А.А. и др. Инструментальное и информационно-технологическое обеспечение сейсмологических наблюдений на Дальнем Востоке России // Вестник ДВО РАН. 2011. № 3(157). С. 127–137.
- [13] 130 BROADBAND Seismic Recorder: Recording Format Specification. RefTek Manual. Refraction Technology, 2008.
- [14] RTPD Protocols: RTP and Client. RefTek Manual. Refraction Technology, 2008.
- [15] СОРОКИН А.А., КОРОЛЁВ С.П., МИХАЙЛОВ К.В., КОНОВАЛОВ А.В. Автоматизированная информационная система оценки состояния сети инструментальных сейсмологических наблюдений “Сигнал-С” // Информатика и системы управления. 2010. № 4(26). С. 161–167.
- [16] BHARDWAJ S., JAIN L., JAIN S. Cloud computing: A study of infrastructure as a service (IAAS) // Intern. J. of Engineer. and Informat. Technology. 2010. Vol. 2, No. 1. P. 60–63.
- [17] OPENSTACK Beginner's Guide: Ubuntu Precise Edition / A. Jha, D. Johnson, K. Murari, M. Raju, V. Cherian, Y. Girikumar. CSS Corp., 2012.
- [18] LEFF A., RAYFIELD J. Web-application development using the model/view/controller design pattern // 5th IEEE Intern. Conf. on Enterprise Distributed Object Comput. Seattle, 2001. P. 118.

*Поступила в редакцию 23 июля 2012 г.,
с доработки — 28 декабря 2012 г.*