

МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ РАСПРЕДЕЛЕННОЙ МУЛЬТИАГЕНТНОЙ УНИВЕРСАЛЬНОЙ ДЕЦЕНТРАЛИЗОВАННОЙ МАСШТАБИРУЕМОЙ ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ

В.А. Сорокин, И.А. Ботыгин, А.С. Зензин, Д.С. Рохмистров, Е.П. Гордов

(г. Томск, Томский политехнический университет)

(г. Томск, Томский государственный университет)

vas1@tpu.ru, bia@tpu.ru, alezes1@rambler.ru, obozhe12@gmail.com, gordov@scert.ru

MATHEMATICAL SOFTWARE OF DISTRIBUTED MULTIAGENT DECENTRALIED UNIVERSAL SCALABLE COMPUTING SYSTEM

V.A. Sorokin, I.A. Botygin, A.S. Zenzin, D.S. Rohmistrov, E.P. Gordov

(Tomsk, Tomsk Polytechnic University)

(Tomsk, Tomsk State University)

Abstract. This paper reviews the concept of building an open, horizontally-scalable distributed information and computing environment, presented in the form of decentralized architecture resources (computing nodes, storage nodes, communication channels) with decentralized control of computations and data. The distributed infrastructure divided into two main components – the compute cluster and the storage cluster. Management of the cluster assigned to specialized frameworks – the infrastructure manager, network manager, computing manager, storage manager, authorization manager, web manager, upload manager. The experimental results of mathematical processing of meteorological parameters studies obtained with standard ultrasonic weather station on the developed information-computational environment are presented.

Keywords: distributed computing, grid-technologies, personal grid, virtual organization, distributed file system, big data

Введение. Специфика научных исследований, как процесса получения новых знаний, требует комплексного и междисциплинарного подхода при анализе больших объемов значимых и сопутствующих параметров изучаемых явлений [1-7]. Отсюда и первая проблема – консолидация инфокоммуникационной инфраструктуры (ИКТ-инфраструктуры) научных подразделений, объединивших свои усилия для решения поставленной научной задачи. И желательна такая консолидация на уровне виртуальной организации, чтобы можно было осуществлять эффективное мультизадачное управление проводимыми исследованиями.

В настоящей работе предложена концепция формирования открытой горизонтально масштабируемой распределенной информационно-вычислительной платформы, представляемой в виде децентрализованной архитектуры ресурсов (узлов вычислений, узлов хранения, каналов связи) с децентрализованным управлением вычислениями и данными.

Основные компоненты платформы. В функциональной структуре и технологической схеме взаимодействия основных компонентов информационно-вычислительной платформы выделены и спроектированы менеджер инфраструктуры, сетевой менеджер, менеджер вычислений, менеджер хранилища, менеджер авторизации, web-менеджер, менеджер планирования загрузки. Схема управления и алгоритмы функционирования распределенного хранилища данных базируются на концепции распределенной файловой системы.

В организационно-управленческом (эксплуатационном) аспекте платформа включает два кластера узлов – кластер узлов обработки (вычислительный кластер) и кластер узлов хранения (кластер хранения). В каждом из кластеров выделен главный узел – командный центр.

Командный центр вычислительного кластера – это вычислительный компонент платформы. Он представляет собой сервер, поддерживающий вычислительную работоспособность всей платформы. В его задачи входит сбор информации о доступных вычислительных

ресурсах и распределение нагрузки на множество узлов обработки, путем выполнения пользовательских инструкций. По сути, командный центр вычислительного кластера представляет функциональность, с помощью которой можно организовать параллельную обработку данных на вычислительных узлах и последующий сбор результатов для выдачи исследователю. Узел вычислительного кластера – это любой компьютер, входящий в состав системы и предоставляющий часть своих ресурсов для обработки. Узел получает и обрабатывает данные, руководствуясь инструкциями, полученными от исследователя. Инструкции, по сути, определяют сценарий работы инфраструктуры при выполнении конкретного приложения. Причем, данные из хранилища узел управления получает сам – в соответствии с правилами и схемами, заданными инструкциями исследователя.

Командный центр кластера хранения занимается обслуживанием хранилища данных – предоставляет интерфейс для размещения данных и доступа к ним. Узел кластера хранения только лишь хранит данные, санкционированные командным центром, и выдает их по авторизованному запросу узлам вычислений.

Достаточно важной проблемой является извлечение полученных данных из хранилища. Поскольку количество файлов в системе может быть очень велико, то необходим эффективный механизм поиска нужной информации. Для облегчения данной задачи с каждым файлом при занесении в систему формируется метаописание, содержащее информацию о содержимом файла. Таким образом, у каждого файла есть некоторый набор ключевых слов, связанных с содержимым файла, по которому можно осуществлять поиск необходимых данных. Все метаописания хранятся в файлах-каталогах, по которым командный центр осуществляет поиск. В каждом метаописании также содержится информация об узле, хранящем данный файл, так что можно легко найти нужный узел хранилища. Заметим, что для взаимодействия с файловым хранилищем была разработана специализированная система команд (API), с помощью которой любой исследователь может создать приложение – обработчик данных, которое будет получать необходимые файлы из распределенного файлового хранилища.

Все взаимодействие исследователей с платформой осуществляется через web-портал. Web-портал выполняет функцию графического интерфейса, который состоит из набора компонентов управления, предоставляющих исследователям простой и удобный доступ к сервисам и услугам платформы (инфокоммуникационным сервисам и услугам – сетевым службам, инфраструктуре предоставления услуг и управления сервисами вычислений, виртуальной инфраструктуре сетевого хранилища метеоданных – сервисам хранения метеоданных). Кроме того, через web-портал выполняются и функции системного административного управления платформой.

Инфраструктура средств обеспечения информационной защиты хранилища данных и доступа к результатам обработки базируется на использовании стандартных протоколов защищенной передачи данных, процедур идентификации исследователя и аутентификации, а также, при необходимости, на использовании защищенного канала передачи – модифицированной схемы виртуальной частной сети.

Заключение. Представлен комплекс программных средств для построения мультиагентной вычислительной платформы со слабо связанными совокупностями вычислителей (персональными компьютерами), предназначенной для информационной поддержки порталов при решении крупномасштабных научных задач, в частности, задач мониторинга, анализа и прогноза климатических процессов.

ЛИТЕРАТУРА

1. International Desktop Grid Federation. URL: <http://desktopgridfederation.org/> (дата обращения: 25.02.2016).

2. Афанасьев А.П., Евтушенко Ю.Г., Посыпкин М.А. Грид-системы из персональных компьютеров как резерв вычислительной мощности для решения оптимизационных и комбинаторных задач большой размерности. URL: http://2014.nscf.ru/TesisAll/0_NSCF_Plenar/13_156_PosypkinMA.pdf (дата обращения: 25.02.2016).
3. World Community Grid: сеть распределенных вычислений. URL: <http://habrahabr.ru/company/ibm/blog/156687/> (дата обращения: 25.02.2016).
4. World Community Grid. URL: <http://www.worldcommunitygrid.org/> (дата обращения: 25.02.2016).
5. Андреев А.Л., Посыпкин М.А., Заикин О.С. Гонку за терафлопсами пора заканчивать. URL: <http://trv-science.ru/2015/02/10/gonku-za-teraflopsami-pora-zakanchivat/> (дата обращения: 25.02.2016).
6. Афанасьев А.П., Бычков И.В., Заикин О.С., Манзюк М.О., Посыпкин М.А., Семенов А.А. Концепция многозадачной грид-системы с гибким распределением свободных вычислительных ресурсов суперкомпьютеров. URL: http://2014.nscf.ru/TesisAll/0_NSCF_Plenar/12_189_VichkovIV.pdf (дата обращения: 25.02.2016).
7. Мищенко П.В., Губарев В.В. Распределенная вычислительная система для подготовки специалистов в области высокопроизводительных вычислений. URL: http://2014.nscf.ru/TesisAll/5_Gridi_iz_rabochix_stanciy_i_kombinirovannie_gridi/04_185_MisheenkoPV.pdf (дата обращения: 25.02.2016).

ВЕКТОРИЗАЦИЯ РАСТРОВЫХ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ ПРЕОБРАЗОВАНИЯ ХАФА

Л.В. Степура., А.Ю. Дёмин

Научный руководитель: А.Ю. Дёмин, к.т.н., доцент кафедры ИПС ИК ТПУ

(г. Томск, Томский политехнический университет)

e-mail: blueberry7251@gmail.com

This article describes Hough algorithms for detecting lines and circles in pictures and gives basic definitions of such area. There is a detailed overview of the algorithm

Keywords: vectorization of image, Hough transform, Hough space, recognition of lines, recognition of circumferences.

Под векторизацией изображения понимается процесс преобразования из растрового вида информации в векторную модель представления. Векторное представление, в данном случае предполагается набор кривых – векторов и совокупность различных набора кривых, которые однозначно описывают изображение.

На преобразовании Хафа, которое было разработано в 1962 году, основываются множество методов распознавания объектов на растре. Важнейшими плюсами данного метода являются простота и наглядность. Распознавание сложных или простых геометрических объектов, таких как прямые, эллипсы, круги и прочее, на изображении – наиболее распространенная задача векторизации. Например, поиск прямолинейных сегментов изображений может использоваться в задаче навигации робота в незнакомой окружающей среде на основе видеоинформации от молекулярного источника. Также идентификация геометрических объектов может использоваться для решения задач распознавания рукописных символов или выделения границ металлических проводников на изображениях слоев цифровых интегральных микросхем и их технологических шаблонов[2].