

## **РАЗРАБОТКА СИСТЕМЫ УПРАВЛЕНИЯ ВЫЧИСЛИТЕЛЬНЫМ КЛАСТЕРОМ**

**Д.М. Бречка**

к.т.н., доцент кафедры кибернетики, e-mail: dbrechkawork@yandex.ru

**К.В. Пугин**

студент, e-mail: ria.freelander@gmail.com

**С.В. Стопкин**

студент, e-mail: stopkin.sergey@gmail.com

Факультет компьютерных наук, Омский государственный университет

### **Аннотация.**

**Ключевые слова:** суперкомпьютер, кластер, программа, Java, сервер, разработка, приложение.

### **Введение**

Вычислительный кластер МВС-1000/128 поступил в распоряжение факультета компьютерных наук ФГБОУ ВПО «ОмГУ им. Ф.М. Достоевского» в июле 2011 года. Кластер состоит из 64 вычислительных модулей, связанных между собой сетью на основе 100baseT Ethernet. Каждый вычислительный модуль содержит в себе два процессора DEC Alpha 21 164 с частотой 667 МГц. Кластер применяется в основном для проведения занятий со студентами факультета компьютерных наук, а также для проведения научно-исследовательской работы студентов.

Для организации более эффективной работы кластера было принято решение о проведении его модернизации. Одной из задач модернизации кластера является установка современного программного обеспечения на вычислительные модули. Решение данной задачи описывается в [1].

Другой задачей модернизации является разработка системы, позволяющей осуществлять удалённый доступ к вычислительному кластеру, а также выполнять компиляцию и запуск программ, написанных с использованием технологии MPI. Данная статья посвящена описанию этой системы.

Для достижения поставленной цели требуется решить ряд задач:

1. Определение программно-аппаратной конфигурации кластера.
2. Выбор средств разработки.
3. Разработка структуры системы управления кластером.
4. Реализация и тестирование системы управления кластером.

Далее в статье описывается решение каждой из указанных задач.

## 1. Описание программно-аппаратной конфигурации МВС-1000/128

### 1.1. Описание аппаратной конфигурации

Вычислительный кластер МВС-1000/128 состоит из 64 вычислительных модулей, связанных между собой сетью на основе 100BaseT Ethernet. Модули содержатся в трёх стойках. На рисунке 1 приведена схема вычислительного кластера.

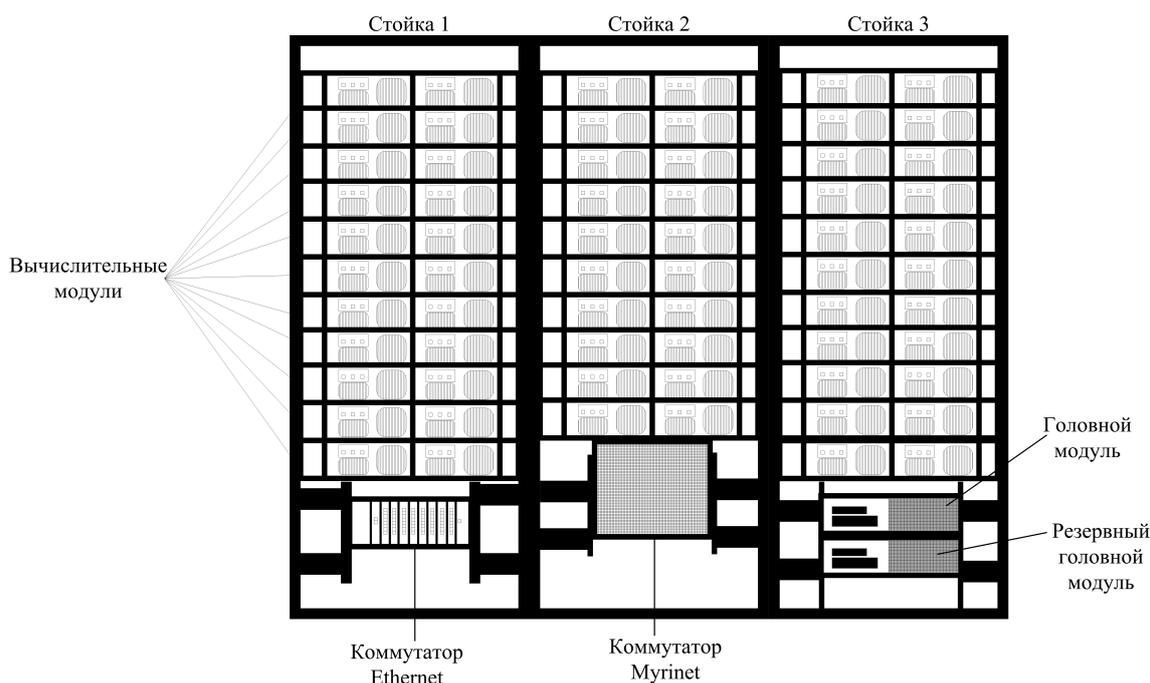


Рис. 1. Схема вычислительного кластера МВС-1000/128

Каждый вычислительный модуль содержит в себе два процессора DEC Alpha 21164 с частотой 667 МГц. Общий объем оперативной памяти модуля составляет 2048 Мбайт. Подробное описание аппаратного обеспечения модуля приведено в таблице 1.

### 1.2. Описание программной конфигурации

Изначально на каждом вычислительном модуле была установлена операционная система Red Hat Linux Alpha 7.2. Для поддержки выполнения параллельных программ на вычислительных модулях установлены: библиотека передачи сообщений MPICH, компиляторы Compaq Fortran и Compaq C/C++, а также библиотека Compaq Extended Math Library (CXML).

В 2011 году сотрудниками факультета компьютерных наук была проведена модернизация программного обеспечения вычислительного кластера [1]. В результате модернизации на все вычислительные модули стойки 3 (рис. 1) была

Таблица 1. Аппаратное обеспечение вычислительного модуля

Тип микропроцессора	DEC Alpha 21164
Разрядность микропроцессора	64
Набор инструкций	RISC
Количество ядер микропроцессора	1
Частота микропроцессора	667 МГц
Напряжение питания микропроцессора	2 В
Объем кэша данных	64 Кбайт
Объем кэша команд	64 Кбайт
Количество микропроцессоров на плате	2
Количество оперативной памяти	2048 Мбайт
Тип оперативной памяти	DIMM SDRAM PC100+ ECC
Количество памяти, доступной микропроцессору	2048 Мбайт
Тип материнской платы	Api UP2000
Наличие коммуникационных портов	LPT, 2 x COM, SCSI, IDE, 2 x PS/2
Тип сетевого адаптера	PCI - 100BaseT
Тип графического адаптера	PCI VGA 4 Мбайт
Тип жесткого диска	IDE ATA133
Объем жесткого диска	20 Гбайт
Максимальная потребляемая мощность	600 Вт

установлена операционная система Debian Linux 5.0 (версия ядра 2.6.22). Для поддержки выполнения параллельных программ на модулях стойки 3 установлены:

- пакет ssh для защищенного соединения с удаленным компьютером;
- монитор процессов Linux htop;
- коллекция компиляторов gcc;
- библиотека передачи сообщений OpenMPI.

Программное обеспечение вычислительных модулей стоек 1 и 2 оставлено без изменений.

### 1.3. Описание конфигурации сети

Физической топологией сети, связывающей вычислительные модули, является звезда. Центральный узел звезды – коммутатор Hewlett-Packard ProCurve 8000M, все вычислительные модули кластера подключены к данному коммутатору. Также к коммутатору подключены головной модуль, резервный головной модуль и рабочий компьютер администратора МВС-1000/128. Все подключённые узлы принадлежат сети с IP-адресом 10.0.0.0 и маской 255.0.0.0. МВС-1000/128 имеет возможность выхода во внешнюю сеть факультета компьютерных наук и Интернет, в качестве шлюза используется головной модуль.

## 2. Определение требований к системе управления кластером и выбор средств разработки

Перед началом разработки системы управления кластером были определены требования к этой системе. Требования таковы:

- 1) система должна иметь web-интерфейс;
- 2) система должна позволять удалённо запускать параллельные программы на модулях вычислительного кластера;
- 3) система должна представлять отчёты о результатах работы параллельных программ;
- 4) система должна предоставлять возможность исправления ошибок в коде параллельной программы;
- 5) система должна предоставлять следующие возможности администрирования:
  - ввод консольных команд Linux для управления кластером;
  - просмотр состояния модулей кластера;
  - выключение и перезагрузка модулей кластера;
- 6) система должна поддерживать многопользовательскую работу;
- 7) система должна отделять функции администрирования от функций обычных пользователей.

Для реализации системы управления кластером был выбран язык программирования Java [2]. Данный язык обладает следующими преимуществами, важными для разработки системы управления:

- 1) кроссплатформенность — позволяет запускать Java-программы на различных программных и аппаратных платформах, в том числе на Alpha;
- 2) поддержка web-технологий — Java обладает набором инструментов для разработки web-интерфейсов и создания клиент-серверных приложений;
- 3) объектно-ориентированное программирование — упрощает разработку структуры приложения.

При разработке системы управления кластером использовалась седьмая версия языка Java. Для организации работы серверной части системы управления кластером было решено использовать контейнер сервлетов Apache Tomcat [3].

### 3. Архитектура системы управления кластером

Функциональные компоненты системы управления кластером сгруппированы в три пакета: web, database и system, функции пакетов соответствуют их названиям.

На рисунке 2 представлена диаграмма пакета web. Этот пакет служит для организации web-интерфейса к системе.

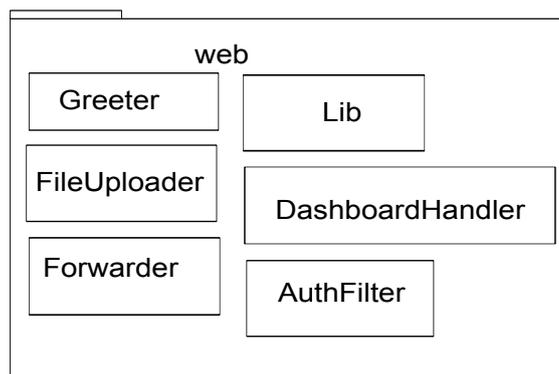


Рис. 2. Пакет web

Пакет web содержит 6 классов, данные классы отвечают за выполнение следующих функций.

1. Класс AuthFilter — реализует процедуру авторизации. Проверяет куки браузера на наличие имени пользователя и пароля в виде хэша, а затем сверяется с конфигурационным файлом на наличие такого пользователя. В случае нахождения пропускает пользователя дальше.
2. Класс Lib — служит для построения хэша пароля и получения нужной куки из пула. Класс содержит две утилитарные функции Answer — построение готовой к встраиванию таблицы после выполнения программы, и Status — построение готовой к встраиванию таблицы рабочих машин для MPI.
3. Классы Forwarder и Greeter — вспомогательные классы авторизации. Greeter перенаправляет незарегистрированных пользователей на страницу для ввода пароля, а Forwarder — зарегистрированных на основную страницу.
4. Класс DashboardHandler — обрабатывает все запросы с web-интерфейса.
5. Класс FileUploader — обрабатывает файлы для загрузки на сервер.

На рисунке 3 представлен пакет database. Данный пакет содержит служебные классы для обработки строк и структур данных.

Классы пакета database выполняют следующие функции.

1. Класс DataStruct — создаёт структуру из трёх объектов, имеющих доступ к системе.
2. Класс Storage — вспомогательный класс, создаёт список из объектов DataStruct для удобства обработки.

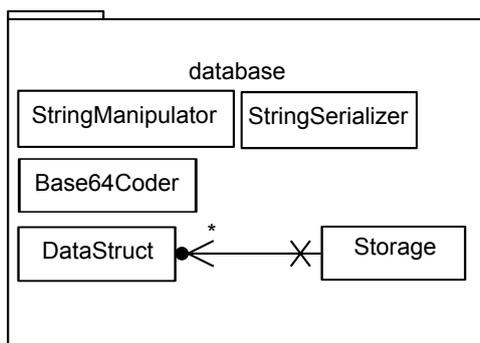


Рис. 3. Пакет database

3. Классы Base64Coder и StringSerializer — служат для работы со строками в кодировке Base64.
4. Класс StringManipulator — реализует поиск по строке.

Пакет system (рис. 4) реализует компоненты, взаимодействующие с операционной системой.

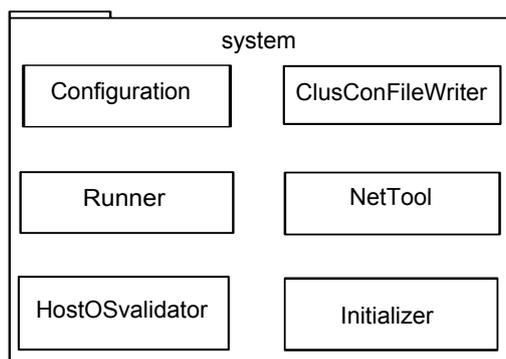


Рис. 4. Пакет system

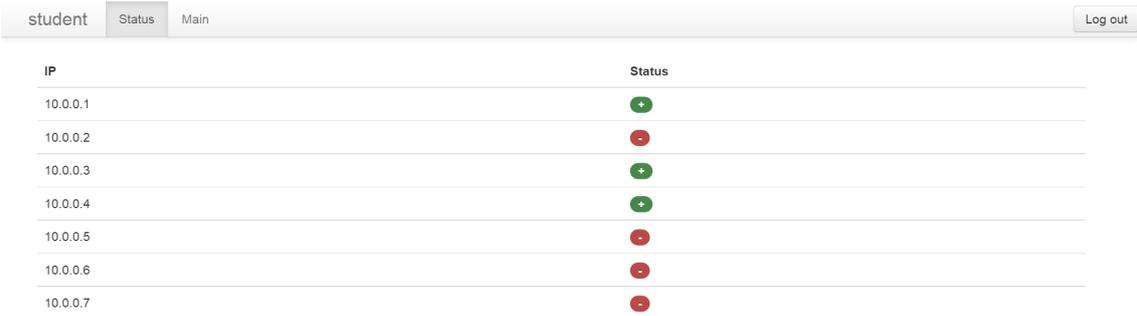
Классы пакета system реализуют следующий функционал.

1. Класс Initializer — инициализирует сервер путём взаимодействия с Tomcat.
2. Класс Configuration — предоставляет интерфейс для работы с конфигурационным файлом. Конфигурационный файл хранится в формате JSON [4].
3. Класс HostOSValidator — проверяет операционную систему на совместимость с сервером.
4. Класс NetTool — составляет список рабочих IP-адресов при помощи системной команды ping.
5. Класс ClusConFileWriter — содержит все функции для работы с файлами, необходимые системе, например, файл IP-адресов, файл конфигурации и т.д.
6. Класс Runner — реализует функции взаимодействия системы управления кластером с MPI.

#### 4. Интерфейс системы управления вычислительным кластером

Система управления развёрнута на головном модуле кластера. Для осуществления доступа к системе необходимо в адресной строке браузера ввести URL-адрес, либо IP-адрес головного модуля с указанием порта и стартового каталога.

Для начала работы с системой управления необходимо пройти процедуру авторизации. Система поддерживает два вида ролей пользователей: обычный пользователь и администратор. Для работы в режиме обычного пользователя необходимо ввести логин и пароль. Интерфейс обычного пользователя состоит из двух окон. Первое окно «Status» отображает состояние модулей вычислительного кластера и их IP-адреса (рис. 5). Модуль может находиться в двух состояниях: «готов к использованию» (идентифицируется знаком «+» на зелёном фоне) и «не готов» (идентифицируется знаком «-» на красном фоне). Готовые к использованию модули доступны для запуска на них параллельных программ.



IP	Status
10.0.0.1	+
10.0.0.2	-
10.0.0.3	+
10.0.0.4	+
10.0.0.5	-
10.0.0.6	-
10.0.0.7	-

Рис. 5. Окно «Status» системы управления кластером

Второе окно — «Main» — предоставляет для пользователя основную функционал системы управления (рис. 6). Каждый пользователь вычислительного кластера имеет собственную директорию на головном модуле. Содержимое этой директории отображается в окне «Main». Система управления позволяет совершать следующие действия в директории пользователя:

- создание нового каталога;
- загрузка файлов на кластер;
- просмотр файла;
- редактирование файла;
- переименование (перемещение) файлов и каталога;
- копирование файлов и каталогов;
- удаление файлов и каталогов;
- обновление содержимого каталога.

Для запуска исполняемого файла необходимо дважды щёлкнуть левой кнопкой мыши на пиктограмме этого файла. После этого откроется окно, в котором

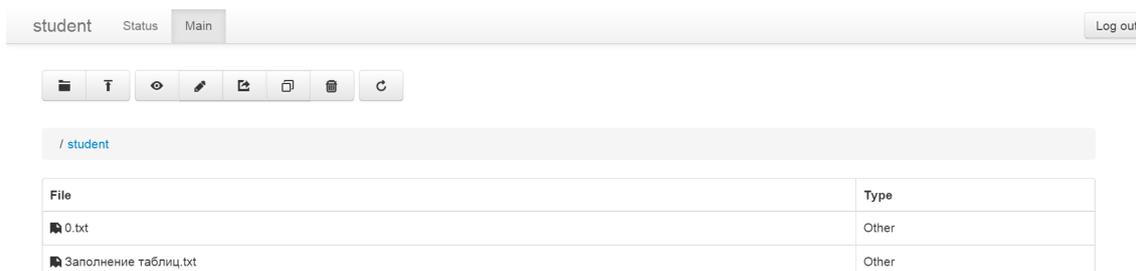


Рис. 6. Окно «Main» системы управления кластером

необходимо выбрать модули, используемые для работы программы. Консольный вывод, возвращаемый программой, отображается в окно браузера.

Для административных ролей пользователя система управления предоставляет дополнительное окно — «CLI» (рис. 7). В этом окне доступны следующие средства администрирования:

1. User Management — управление пользователями. Позволяет создавать и удалять пользователей, изменять имена пользователей и пароли, а также изменять роли пользователей.
2. Main Settings — основные настройки. Позволяет задавать диапазон IP-адресов модулей, которые будут контролироваться системой ClusCon. Также позволяет задавать базовую директорию для средств MPI (сетевая директория).
3. Status Management — управление состояниями. Позволяет выключать и перезагружать модули кластера, а также производить перепоключение сетевой файловой системы.
4. Console — консоль. Позволяет вводить консольные команды Linux для управления кластером.

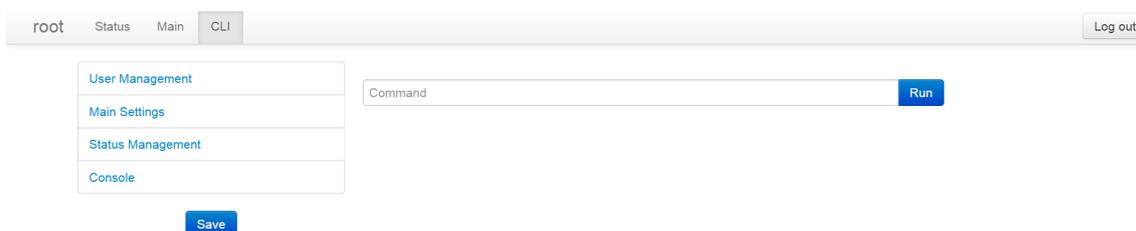


Рис. 7. Окно «CLI» системы управления кластером

## Заключение

В ходе работы для вычислительного кластера МВС-1000/128 была разработана система управления, позволяющая осуществлять удалённый доступ к кластеру, удалённо запускать параллельные программы на кластере и проводить

удалённое администрирование кластера. Система управления была реализована на языке Java с использованием контейнера сервлетов Apache Tomcat. Тестирование системы, проведённое сотрудниками факультета, показало соответствие функционала системы предъявленным требованиям.

Разработанная система управления существенно упрощает доступ к вычислительным ресурсам кластера, что позволяет сделать кластер доступным для более широкого круга пользователей. Таким образом, можно сделать заключение о высокой практической значимости проделанной работы.

## ЛИТЕРАТУРА

1. Стопкин С.В., Пугин К.В., Бречка Д.М. Установка программного обеспечения на вычислительный модуль с архитектурой процессора Alpha // Математические структуры и моделирование. 2013. Т. 27, № 1. С. 75–81.
2. Нотон П., Шилдт Г. Полный справочник по Java. М. : «Диалектика», 1997.
3. Apache Tomcat 7 Documentation URL: <http://tomcat.apache.org/tomcat-7.0-doc/index.html> (дата обращения: 10.01.2014).
4. Introducing JSON: <http://json.org/> (дата обращения: 10.01.2014).

## DEVELOPMENT OF THE COMPUTER CLUSTER CONTROL SYSTEM

**D.M. Brechka**

Ph.D.(Eng.), Ass. Prof., e-mail: [dbrechkawork@yandex.ru](mailto:dbrechkawork@yandex.ru)

**K.V. Pugin**

student, e-mail: [ria.freelander@gmail.com](mailto:ria.freelander@gmail.com)

**S.V. Stopkin**

student, e-mail: [stopkin.sergey@gmail.com](mailto:stopkin.sergey@gmail.com)

Omsk State University n.a. F.M. Dostoevskiy

**Abstract.** The article describes the control system for computing cluster MVS-1000/128, developed by the staff of the Faculty of Computer Science OmsSU n.a. F.M. Dostoevsky. The main aim is to develop a computer cluster control system, according to the specified requirements. The paper describes the architecture of the system, its interface and the basics of working with the system. The main result is the development of a working computer cluster control system.

**Keywords:** supercomputer, cluster, program, Java, server, development, application.