

*Автор:*

*Казанцева Татьяна Ивановна*

студентка

*Научный руководитель:*

*Чирышев Юрий Владимирович*

аспирант, ассистент

ФГАОУ ВО «Уральский федеральный университет

им. первого Президента России Б.Н. Ельцина»

г. Екатеринбург, Свердловская область

## **ОБЗОР АРХИТЕКТУРЫ РАСПРЕДЕЛЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ**

*Аннотация:* работа посвящена обзору и сравнению существующих решений для выполнения параллельных вычислений на базе многопроцессорных архитектур, выявлению их достоинств и недостатков применительно к разработке систем в области информационной безопасности.

*Ключевые слова:* параллельные вычисления, фреймворк, диспетчер, программа, Linux.

На сегодняшний день в сфере информационной безопасности существует большое количество задач, в том числе переборного типа [7], решение которых предполагает возможность применения технологии параллельных вычислений на основе специализированных суперкомпьютеров или кластерных архитектур, построенных на базе IBM-совместимых компьютеров. Большинство таких решений, демонстрирующих высокую производительность, имеют ограничения на коммерческое использование или не всегда могут быть применены за счет неприемлемо высокой стоимости. Вследствие этого, актуальной остается задача создания высокопроизводительной вычислительной системы на основе свободного программного обеспечения. Целью статьи является обзор существующих архитектур распределенных вычислений и средств разработки под данную архитектуру.

Как и любые системы, параллельные – могут классифицироваться по различным признакам. Основной характеристикой считается присутствие общей (англ. SMP–Symmetric Multi-Processing) или распределённой памяти (англ. MPP–Massively Parallel Processing). Гибридным соединением SMP и MPP является архитектура с неравномерной памятью (англ. NUMA– Non-Uniform Memory Access) [5].

Симметричная мультипроцессорная система -совокупность массива общей памяти, а также двух и более равноправных процессоров. Все прочие компоненты, такие как память, система ввода/вывода, операционная память – в единственном виде. Из равноправности процессоров вытекает и главный недостаток системы SMP – необходимость организации канала процессоры – память с очень высокой пропускной способностью.

Массивно-параллельная система – система, состоящая из однородных вычислительных узлов, включающих: один или несколько центральных процессоров, локальную память, сетевой адаптер и устройства ввода/вывода. К MPP добавляются специальные узлы ввода/вывода и управляющие узлы, это связано с увеличением скорости обработки данных и повышением надежности управления системой [6].

NUMA-система, включающая несколько отдельных процессоров, каждый из которых, кроме собственного кэша, обладает также локальной памятью [3].

Рассмотрим место кластерной архитектуры вычислительных систем в данной классификации. Кластер – это слабосвязанная система связи узлов. Такая система основана на базе шинной архитектуры или коммутатора. В роли узлов выступают обычные компьютеры с системой ввода/вывода, ОС, памятью и т. д. Основная классификация по цели применения кластеров включает следующие виды: кластеры с балансировкой нагрузки, отказоустойчивые и вычислительные кластеры.

Кластеры с балансировкой нагрузки служат для увеличения производительности обработки последовательностей большого числа вычислительных запросов, благодаря распределению запросов по узлам кластера.

Отказоустойчивые кластеры обеспечивают доступность системы, позволяя ей функционировать в случаях отказа каких-либо узлов кластера, за счет механизма подмены отказавшего узла другим автоматически.

Вычислительные кластеры позволяют значительно повысить скорость вычислений, благодаря большому числу узлов и возможности вести параллельную обработку задачи [5; 6].

Заметим, что в рассматриваемой классификации вычислительный кластер является MPP системой. Преимущества данной архитектуры в том, что она является легко масштабируемой, и экономически более эффективной, поскольку стоимость одной единицы FLOPS намного меньше, чем у высокопроизводительных систем, приведенных в работе [4]. Кроме этого, неоспоримым преимуществом является то, что в качестве такого узла можно использовать обычный персональный компьютер.

В силу того, что MPP архитектура не похожа на архитектуру обычного персонального компьютера, программирование под нее также отличается. Однако существуют библиотеки и фреймворки, которые позволяют упростить разработку программ под эту архитектуру. К таким решениям относятся:

1. MPI (Message Passing Interface) – спецификация для программного интерфейса передачи сообщений между процессами на кластере.

2. PVM (Parallel Virtual Machine) – пакет программ, позволяющий создавать и запускать много процессные программы на MPP архитектуре [5], является аналогом MPI.

3. Nadoop – фреймворк для разработки программ на кластерных системах, содержит большое количество утилит и библиотек [1].

4. Одним из типичных примеров реализации распределенных систем являются кластеры Beowulf [2].

Следует отметить, что кластер предназначен для выполнения приложений, учитывающих технологию параллельных вычислений. Программа, которая в последствии будет внедряться на кластер, должна уметь определять какие данные надо «кластеризовать» для распараллеливания задачи, а также определять их

минимальный структурный элемент. При этом кластерная система с минимальным набором функций должна включать следующие роли, каждая из которых решает свои задачи:

- администратор: организует доступ программистов к ресурсам кластера и корректирует приоритетность вычислительных задач;
- диспетчер: «кластеризует» задачи и данные, т.е. имеет возможность подсчёта узлов, а также сбора данных о производительности каждого узла;
- программист: способен постоянно отслеживать процесс выполнения задачи за счет циклического опроса состояния узлов;
- диспетчерский узел: выделяет каждому узлу их количество пропорционально вычислительной мощности этих узлов.

Таким образом, в результате проделанной работы произведен обзор и сравнение существующих архитектур и средств разработки распределенных систем, получены следующие результаты: данные решения развиваются достаточно быстро, но из-за усложненных аппаратных и программных составляющих, цена на эти системы может быть слишком высока и недоступная для многих пользователей. Возможным решением проблемы нехватки вычислительных ресурсов являются кластерные системы. Данные системы могут быть построены на основе более бюджетных технологий, с помощью существующих решений, на базе вычислительных систем общего назначения и с использованием программного обеспечения с открытым исходным кодом.

### ***Список литературы***

1. Hadoop. What Is Apache Hadoop [Electronic resource]. – Access mode: <http://hadoop.apache.org/> (access date: 15.04.2018).
2. Дьяченко Р.А. К вопросу организации кластеров на основе технологии Beowulf / Р.А. Дьяченко, Р.Х. Багдасарян, М.В. Руденко, И.С. Лоба // Научные чтения имени профессора Н.Е. Жуковского: Сборник научных статей. – 2015. – С. 189–191.

3. Леонов А.В. Сравнительный анализ параллельных вычислительных систем. Вычислительный кластер ФГУП НИИР // Труды научно-исследовательского института радио. – 2008. – №2. – С. 42–53.

4. Массово-параллельная архитектура [Электронный ресурс]. – Режим доступа: <http://www.parallel.ru/computers/classes.html#mpp> (дата обращения: 15.04.2018).

5. Симметричная мультипроцессорность [Электронный ресурс]. – Режим доступа: [http://citforum.ru/hardware/articles/art\\_5.shtml](http://citforum.ru/hardware/articles/art_5.shtml) (дата обращения: 15.04.2018).

6. Сбитнев Ю.И. Кластеры. – Екатеринбург, 2009 [Электронный ресурс]. – Режим доступа: <http://cluster.linux-ekb.info/clusters.pdf> (дата обращения: 15.04.2018).

7. Чернышев Ю.О Криптографические методы и генетические алгоритмы решения задач криптоанализа: монография / Ю.О. Чернышев, А.С. Сергеев, Е.О. Дубров, А.В. Крупенин, О.П. Третьяков. – Краснодар: ФВАС, 2013. – 138 с.