ПРИНЦИПЫ РЕАЛИЗАЦИИ СИСТЕМЫ МОНИТОРИНГА И ДИНАМИЧЕСКОЙ БАЛАНСИРОВКИ НАГРУЗКИ ДЛЯ ВЫЧИСЛИТЕЛЬНЫХ ВЕБ-СЕРВИСОВ

Афанасьев А.П., Колпаков В.Г., Корх А.В., Курочкин И.И., Смирнов С.А., Рубцов А.О., Сухорослов О.В.
Институт системного анализа РАН

Одной из перспективных областей применения облачной модели Software as a Service (SaaS) является реализация вычислительных веб-сервисов (ВВС). Актуальность данного направления исследований обусловлена растущей потребностью в масштабируемых вычислительных сервисах, ориентированных на решение важных классов прикладных задач, как в науке, так и бизнесе.

Для эффективной балансировки нагрузки между узлами облачной вычислительной инфраструктуры (ОВС) ВВС требуется использование информации о текущей загрузке узлов, а также динамическое перераспределение (миграция) заданий в случае перегрузки узлов. Таким образом, необходимо создание интегрированной системы мониторинга и динамической балансировки нагрузки (СМДБН).

Для сбора информации о загрузке узлов ОВС предлагается использовать архитектуру с агентами. На каждом узле ОВС размещается специальный агент, осуществляющий сбор информации о локальном узле и передающий её в СМДБН через систему обмена сообщениями. Данный подход позволяет реализовать эффективный сбор произвольных метрик с большого количества узлов и их передачу в СМДБН и другие системы-потребители данной информации.

Функции приёма и хранения данных мониторинга выполняет сервис мониторинга, реализующий интерфейсы для доступа к этим данным. Полученные от агентов данные предлагается хранить в памяти, поскольку долговременное хранение данных не требуется для целей СМДБН и может быть реализовано в рамках системы мониторинга общего назначения. При этом, для учёта текущей загрузки узлов при реализации балансировки нагрузки, в рамках СМДБН достаточно хранить только последние значения метрик, полученных от каждого узла.

Потребителем информации, предоставляемой сервисом мониторинга, является модуль СМДБН, реализующий динамическую балансировку нагрузки между узлами ОВС. Данный модуль должен обеспечивать обработку событий следующих типов: поступление нового задания, завершение выполнения задания узлом, подключение нового узла, отключение узла, перегрузка узла. События первого типа поступают извне, от подсистемы ВВС, принимающей запросы клиентов сервиса. События других типов могут быть обнаружены путём анализа данных мониторинга. Данную функцию предлагается возложить на сервис мониторинга, обеспечив передачу обнаруженных событий в модуль балансировки нагрузки.

При поступлении нового задания модуль балансировки нагрузки должен попытаться подобрать узел для запуска задания. Для этого требуется информация о текущей загрузке узлов, которая доступна через программный интерфейс сервиса мониторинга. В случае, если узел подобран, то модуль балансировки должен инициировать запуск задания на данном узле. Для этого требуется обеспечить взаимодействие между модулем и подсистемой запуска заданий ВВС. В случае, если подобрать узел не удалось, то модуль балансировки должен сохранить задание в очереди неназначенных заданий.

При отключении узла модуль должен назначить выполнявшиеся узлом задания другим узлам или поместить задания в очередь. Для этого модуль должен хранить в памяти информацию о том, какие задания назначены каждому из узлов, а также всю информацию о каждом задании, необходимую для его перезапуска. При этом требуется информация о текущей загрузки узлов, которая доступна через программный интерфейс сервиса мониторинга.

При перегрузке узла модуль балансировки должен прекратить выполнение части заданий на узле, назначить остановленные задания другим узлам или поместить задания в очередь. Требования к хранимой и используемой информации здесь аналогичны ситуации с отключением узла.

При завершении задания узлом или подключении нового узла модуль балансировки может назначить данному узлу новое задание из имеющейся очереди. Также при завершении задания модуль может удалить из памяти всю информацию о данном задании.

Таким образом, модуль балансировки нагрузки должен иметь внутреннее состояние, включающее:
- информацию о всех незавершённых заданиях, необходимую для их перезапуска;
- очередь заданий, которые ожидают назначения узла;
- список незавершённых заданий, которые назначены каждому из узлов.

Данное состояние имеет небольшой размер и может храниться в оперативной памяти.

Работа выполняется в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2013 годы» при финансовой поддержке Минобрнауки, государственный контракт № 14.514.11.4041.