

УНИВЕРСАЛЬНЫЕ ИНСТРУМЕНТАЛЬНЫЕ ПРОГРАММНЫЕ КОНВЕЙЕРНО-ПАРАЛЛЕЛЬНЫЕ КОМПЛЕКСЫ: ПРЕДЪЯВЛЯЕМЫЕ ТРЕБОВАНИЯ, МОДЕЛИРОВАНИЕ СЧЕТА ПРИКЛАДНОЙ ЗАДАЧИ

Фраленко В.П., Талалаев А.А.

ФГБУН Институт программных систем им. А.К. Айламазяна Российской академии наук, Исследовательский центр мультипроцессорных систем (152021, Ярославская обл., Переславский р-н, с. Вельково, ул. Петра I, д. 4а), e-mail: alarmod@pereslavl.ru

В работе сформулированы основные требования, предъявляемые к современным инструментальным конвейерно-параллельным комплексам, предназначенным для решения широкого круга задач. В частности, поддержка распределенно-облачных сред, наличие удаленного кроссплатформенного графического интерфейса, верифицированность основных алгоритмов функционирования, отказоустойчивость и т.д. Новое решение позволит абстрагироваться от деталей низкоуровневой реализации алгоритмов за счет модульной внутренней организации и будет способно функционировать, в том числе и в облачных средах. Представлены предложения по моделированию с помощью E-сетей Петри счета прикладной задачи в таком комплексе. Управляющие позиции E-сети в этой модели определяют множество управляющих воздействий, влияющих на ход выполнения задачи в многопоточном окружении, формализуя множество доступных управляющих команд системы диспетчеризации нагрузки. Полученную модель предлагается применить при отработке интеллектуального планировщика, обеспечивающего динамическое распределение вычислительной нагрузки.

Ключевые слова: комплекс, конвейерно-параллельный, моделирование, сеть Петри, планировщик.

UNIVERSAL PIPELINE-PARALLEL SOFTWARE TOOLS: REQUIREMENTS AND APPLIED PROBLEM DECISION MODELING

Fralenko V.P., Talalaev A.A.

Ailamazyan Program Systems Institute of the Russian Academy of Sciences, Research Center for Multiprocessor Systems (152021, Yaroslavl region, Pereslavl area, Peter I st., 4a), e-mail: alarmod@pereslavl.ru

The paper formulates the basic requirements for up-to-date universal pipeline-parallel software tools, designed for a wide range of tasks. For example, distributed-cloud environments support, remote cross-platform graphical user interface, veracity of the main operation algorithms, fault tolerance, etc. The new solution will allow to abstract from low-level algorithms implementation details due to a modular internal organization and will be able to working even in cloud environments. The suggestions for based on Petri E-nets applied problem decision modeling in such complex are presented. E-net control positions in this model define a set of control actions that affect a task progress in multithreaded environment, formalizing many available control commands of load dispatching system. The resulting model is proposed to apply for intelligent scheduler working off that provides dynamic allocation of computational load.

Keywords: complex, pipeline-parallel, modeling, Petri net, scheduler.

Введение

Коллектив исследователей Лаборатории интеллектуального управления Исследовательского центра мультипроцессорных систем ИПС им. А.К. Айламазяна РАН в рамках работ по Государственному контракту разрабатывает научно-технический задел в области организации высокопроизводительных конвейерно-параллельных вычислений в облачных средах и программно-инструментальных средств на их основе, позволяющих снизить временные затраты на разработку прикладных решений и повысить эффективность использования ресурсов облачной среды при решении прикладных задач.

Для облегчения решения сложных научно-технических задач, требующих задействования значительных вычислительных ресурсов, практикуется использование параллельных прикладных программных комплексов. Низкоуровневая реализация такого комплекса требует больших временных затрат и привлечения высококвалифицированных кадров. Далее будут сформулированы предъявляемые к перспективному комплексу требования и исследован вопрос моделирования счета прикладной задачи.

Требования, предъявляемые к перспективному комплексу

На сегодняшний день не существует универсального программно-инструментального комплекса, который бы удовлетворял всем требованиям, возникающим при решении реальных прикладных задач в конвейерно-параллельном режиме:

- а) функционирование на разнородных аппаратно-программных платформах (в том числе распределенно-облачных средах) и гибкость адаптации к новым платформам;
- б) высокий уровень абстракции;
- в) функционирование в условиях наложенных ограничений на количество используемых аппаратных ресурсов;
- г) наличие механизмов автоматизации построения программ;
- д) наличие интеллектуального планировщика, обеспечивающего динамическое распределение вычислительной нагрузки;
- е) масштабируемость;
- ж) наличие удаленного кроссплатформенного графического интерфейса с возможностью конфигурирования для организации максимального эффекта обработки в режиме «реального времени»;
- з) верифицированность основных алгоритмов функционирования;
- и) отказоустойчивость (работоспособность при поломке отдельных вычислительных узлов, поддержка контрольных точек и т.д.);
- к) отсутствие привязки к конкретным типам данных и кругу решаемых задач, наличие гибких инструментов для доступа к обрабатываемым данным;
- л) возможность подключения внешних устройств ввода-вывода;
- м) возможность реализации специализированных алгоритмов анализа и обработки данных с наименьшими трудозатратами.

Необходимость создания нового программного комплекса, позволяющего абстрагироваться от деталей низкоуровневой реализации алгоритмов за счет модульной внутренней организации и функционирующего, в том числе и в облачных средах, не вызывает сомнений. Он найдет применение при решении актуальных задач обработки данных в различных областях науки. Некоторые решения в виде архитектуры

многоуровневого вычислительного ядра и простейшего планировщика уже были предложены в работе [4], перечисленные выше требования являются расширением выдвинутых тезисов. Далее предлагается обсудить один из возможных способов моделирования счета прикладной задачи в таком комплексе.

Моделирование счета прикладной задачи

Построение программных комплексов, предназначенных для обработки потоков данных широкого назначения, требует научно-обоснованного подхода к выработке проектных решений. Принятие верных решений невозможно без предварительного построения математических и имитационных моделей с целью получения первичных оценок качества. Моделирование тесно связано с формализованным описанием и изучением динамики проектируемых систем [5], что накладывает определенные ограничения на выбор инструментальных средств. Попробуем описать конвейерно-параллельный процесс решения прикладной задачи в таком комплексе с помощью модели на основе *E*-сети Петри [3].

Аппарат сетей Петри позволяет математически строго описывать процессы смены состояний в ходе решения задачи, обладает широкими выразительными возможностями, основанными на графическом динамическом отображении ситуаций. Функционирование сети Петри, как и конечного автомата, осуществляется путем смены состояний по определенным правилам [1,2]. Срабатывания переходов соответствуют реализациям определенных событий и приводят к изменению разметки позиций, т.е. к изменению состояния сети. Переход может сработать, если выполнены все условия реализации соответствующего события.

Процесс счета сложных прикладных задач имеет все свойства потоковых асинхронных программ. В потоковых асинхронных программах используется понятие распределенной памяти, состоящей из элементов специального типа — очередей (каналов). Каждая очередь приписана некоторой группе операторов, и указывается, какие операторы могут записывать в очередь и какие считывать из нее [2]. На представленном ниже рис. 1 показана *E*-сеть Петри, реализующая предлагаемый принцип управления решением задачи. В качестве фишек в данной сети следует рассматривать ресурсы (как аппаратные ресурсы, так и данные, подлежащие обработке), позиции сети описывают очереди ожидающих обработки ресурсов, а переходам соответствуют некоторые функции обработки/преобразования данных. При этом время срабатывания переходов является стохастическим параметром и заранее неизвестно.

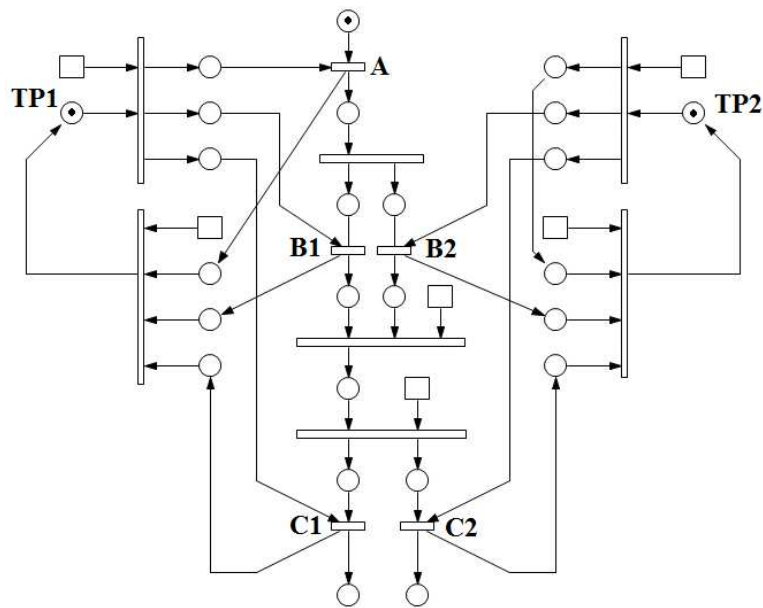


Рис. 1. *E*-сеть Петри имитационного моделирования

На схеме приведены блоки, соответствующие пулам потоков на вычислительных узлах (thread pools, TP1 и TP2) и определяющие доступные для решения прикладной задачи аппаратные ресурсы, представлены три класса модулей: класса А (непараллельная реализация поставщика ресурсов), класса В (модуль-обработчик данных, имеющий параллельную реализацию; представлен на схеме двумя взаимодействующими потоками В1 и В2) и класса С (непараллельная реализация потребителя ресурсов; представлен на схеме двумя независимыми потоками С1 и С2). Команды системы диспетчеризации устанавливают значения управляющих позиций *E*-сети, определяя последовательность срабатывания переходов (запуск программных модулей) и трассу передвижения меток (данных, подлежащих обработке). Специфичные для *E*-сетей типы переходов соответствуют дополнительным операциям над данными, выполняемым в ходе обработки задачи. *F*-переход, связывающий модули классов А и В, показывает, что данные от первого модуля поступят на вход всем копиям второго модуля. Комбинация переходов *X* и *Y*, связывающая модули классов В и С, показывает, что все данные, порождаемые параллельными копиями В1 и В2, представляют собой единый поток, который распределяется между копиями С1 и С2, в соответствии с некоторой стратегией распределения ресурсов (за что отвечает управляющая позиция *X*-перехода). Управляющие позиции *E*-сети в этой модели определяют множество управляющих воздействий, влияющих на ход выполнения задачи в многопоточном окружении, формализуя множество доступных управляющих команд системы диспетчеризации нагрузки.

Заключение

Сформулированные требования, предъявляемые к современным инструментальным программным конвейерно-параллельным комплексам, предназначенным для решения

широкого круга задач, существенно дополнены. В частности, в настоящее время необходимой становится поддержка распределенно-облачных сред, наличие удаленного кроссплатформенного графического интерфейса, верифицированность основных алгоритмов функционирования, отказоустойчивость и т.д.

Представление задачи в виде сети Петри позволяет получить наиболее полное описание процессов функционирования такого комплекса и создает основу для научно-обоснованного проектирования прикладных систем, обеспечивающего повышение эффективности использования параллельных вычислителей при решении сложных ресурсоемких задач. Полученную модель предлагается применить при отработке интеллектуального планировщика, обеспечивающего динамическое распределение вычислительной нагрузки.

Работа поддержана Министерством образования и науки Российской Федерации (Государственный контракт № 14.514.11.4109 по теме «Разработка алгоритмов и методов организации высокопроизводительных конвейерно-параллельных вычислений в облачных средах»).

Список литературы

1. Захаров Н.Г., Рогов В.Н. Синтез цифровых автоматов: Учебное пособие / Н.Г. Захаров, В.Н. Рогов. — Ульяновск: УлГТУ, 2003. — 135 с.
2. Котов В.Е. Сети Петри. — М.: Наука. Главная редакция физико-математической литературы, 1984. — 160 с.
3. Омаров О.М. Теория вычислительных процессов и структур. Учебное пособие. — Махачкала: РИО ДГТУ, 2005. — 268 с.
4. Талалаев А.А. Организация конвейерно-параллельных вычислений для обработки потоков данных // Информационные технологии и вычислительные системы. — 2011. - № 1. — С.8-13.
5. Питерсон Д. Теория сетей Петри и моделирование систем: Пер. с англ. — М.: Мир, 1984. — 264 с.

Рецензенты:

Славин О.А., д.т.н., заведующий лабораторией, ФГБУН Институт системного анализа Российской академии наук (ИСА РАН), г. Москва.

Знаменский С.В., д.ф.-м.н., заведующий лабораторией, ФГБУН Институт программных систем им. А.К. Айламазяна Российской академии наук, Исследовательский центр системного анализа, с. Вельково.