

# Технология параллельного программирования RiDE

М.О. Бахтерев, П.А. Васёв

Институт Математики и Механики УрО РАН

RiDE это технология программирования в параллельных распределенных средах на основе модели потока данных (dataflow, [1]). RiDE основана на анализе различных, в том числе и собственных, моделей потока данных [2]. Цель – упростить процесс создания параллельных программ, и сделать это не в ущерб эффективности исполнения вычислительных кодов.

Технология RiDE базируется на понятиях хранилища, задач и правил. Хранилище содержит в себе именованные данные, по отношению к которым доступны три операции – запись (создание), чтение и удаление (возможно в автоматическом режиме с распределенной сборкой мусора). Хранимые данные есть единицы информации с уникальными именами. Задачи выполняют программы, считывающие данные с определенными именами из хранилища, обрабатывают их и формируют новые данные, которые записываются в хранилище. Правила описывают взаимосвязи между задачами и содержимым хранилища, определяя тем самым поток данных параллельного вычисления. Более подробно: <http://www.ridehq.net>.

Описание вычислительных приложений в предложенных терминах представляется авторам более простой задачей, чем разработка и реализация параллельных схем работы в более традиционных терминах моделей MPI и OpenMP. Действительно, программисту необходимо, по сути, описать вычислительные процедуры, и зависимости между ними. И такое описание достаточно для автоматического формирования эффективного процесса параллельного исполнения программы в режиме совмещения счета и обменов (что является преимуществом dataflow).

Предлагаемая технология в перспективе позволит относительно просто реализовать проведение вычислительного эксперимента на гибридных архитектурах с динамическим изменением количества вычислительных узлов во время самого счета (что актуально при больших вероятностях сбоях на машинах экзафлопного класса), работу в существенно неоднородных коммуникационных средах, автоматическое создание контрольных точек, приостановку и продолжение вычисления прозрачным для программиста образом, использование распределенные хранилищ данных, а также обеспечивает ряд других преимуществ.

Авторы выражают надежду, что результатом развития системы RiDE станет повышение эффективности труда программистов, разрабатывающих приложения для современных неоднородных высокопроизводительных систем. На данный момент для технологии RiDE разработана методика [3], и силами компании LineAct ведется ее реализация [4].

## Литература

1. Dennis J. Data Flow Supercomputers // Computer, 1980, Vol.13, No.11, P.48-56.
2. Бахтерев М.О., Описание параллельных вычислений при помощи замыканий // Тезисы 10-го Международного семинара "Супервычисления и Математическое моделирование", РФЯЦ-ВНИИЭФ, Саров, с. 31-32, 2008.
3. М.О. Бахтерев, П.А. Васёв, А.Ю. Казанцев, И.А. Альбрехт, Методика распределенных вычислений RiDE // Параллельные вычислительные технологии (ПаВТ'2011): труды международной научной конференции (Москва, 28 марта – 1 апреля 2011 г.) [Электронный ресурс] – Челябинск: Издательский центр ЮУрГУ, 2011, с. 418–426.
4. M. Bakhterev, A. Kazantzev, P. Vasev, I. Albrecht, Dataflow-Based Distributed Computing System // Proceedings of the Euromicro PDP 2011 Work in Progress Session (Eds. E. Grosspietsch, K. Kloeckner) p.6-7, SEAA-Publications No. SEA-SR-29 Johannes Kepler University Linz (Austria), ISBN 978-3-902457-29-5.