

УДК 004.822:004.89:519.63

ИНФОРМАЦИОННО-АНАЛИТИЧЕСКАЯ ПОДДЕРЖКА РЕШЕНИЯ ВЫЧИСЛИТЕЛЬНО СЛОЖНЫХ ЗАДАЧ НА СУПЕРКОМПЬЮТЕРАХ

Загорулько Юрий Алексеевич

к.т.н., зав. лабораторией, e-mail: zagor@iis.nsk.su,

Загорулько Галина Борисовна

н.с., e-mail: gal@iis.nsk.su,

Шестаков Владимир Константинович

м.н.с., e-mail: shestakov@iis.nsk.su,

Институт систем информатики им. А.П. Ершова СО РАН,
630090 г. Новосибирск, проспект Академика Лаврентьева, 6.

Сапетина Анна Федоровна

м.н.с., e-mail: sapetina@sscc.ru,

Институт вычислительной математики и математической геофизики СО РАН,
630090 г. Новосибирск, проспект Академика Лаврентьева, 6.

Аннотация. В статье представлен подход к разработке информационно-аналитического интернет-ресурса для поддержки решения вычислительно сложных задач математической физики на суперкомпьютерах. Основу этого ресурса составляет многоуровневая онтология, которая строится на основе отологий научного знания и научной деятельности. Для эффективной информационно-аналитической поддержки пользователей очень важно иметь подробное систематизированное описание параллельных алгоритмов и архитектур, достаточно полное описание доступных программных компонентов, реализующих параллельные алгоритмы и фрагменты параллельного кода, имеющих в наличии параллельных архитектур и используемых в них устройств, а также публикаций и информационных ресурсов, описывающих данную проблемную область.

Ключевые слова: решение вычислительно сложных задач на суперкомпьютерах, параллельные алгоритмы, параллельные архитектуры, онтология, информационно-аналитический интернет-ресурс.

Цитирование: Загорулько Ю.А., Загорулько Г.Б., Шестаков В.К., Сапетина А.Ф. Информационно-аналитическая поддержка решения вычислительно сложных задач на суперкомпьютерах // Информационные и математические технологии в науке и управлении. 2020. № 3 (19). С. 86- 95. DOI:10.38028/ESI.2020.19.3.009.

Введение. Сейчас в России и за рубежом активно эксплуатируются высокопроизводительные суперкомпьютеры, на которых можно решать задачи большой вычислительной сложности. В частности, суперкомпьютеры широко применяются при математическом моделировании различных физических явлений и процессов. Однако современные суперкомпьютеры редко используются рядовыми исследователями, так как им довольно трудно решать на них свои задачи по той причине, что они не обладают знаниями об архитектуре суперкомпьютера и особенностях реализации на ней алгоритмов решения

этих задач. В связи с этим актуальной задачей становится создание средств интеллектуальной поддержки решения вычислительно сложных задач на суперкомпьютерах [12].

Такая интеллектуальная поддержка должна основываться на использовании знаний о рассматриваемой проблемной области, включая методы и алгоритмы решения задач на суперкомпьютерах, а также опыта решения задач на суперкомпьютерах, представленного в виде методик, программных компонентов (фрагментов кода) и экспертных правил.

Первым шагом к интеллектуальной поддержке могло бы стать обеспечение информационно-аналитической поддержки пользователя, состоящее в предоставлении пользователю удобного содержательного доступа к информации обо всех доступных методах и алгоритмах решения задач на суперкомпьютере, о возможностях и ограничениях каждого из них, о характеристиках их реализаций. Средством такой поддержки может быть информационно-аналитический интернет-ресурс, базирующийся на многоуровневой онтологии проблемной области.

Данная работа посвящена разработке такого интернет-ресурса, обеспечивающего информационно-аналитическую поддержку решения вычислительно сложных задач математической физики на суперкомпьютерах.

1. Обзор информационных ресурсов, поддерживающих решение вычислительно сложных задач на суперкомпьютерах. На данный момент в России существуют и развиваются несколько информационных ресурсов, поддерживающих решение задач на суперкомпьютерах. Наиболее значимыми из них являются AlgoWiki [3] и parallel.ru [6].

Информационный ресурс AlgoWiki [2] позиционируется как открытая энциклопедия в сети Интернет по свойствам алгоритмов и особенностям их реализации на различных программно-аппаратных платформах, начиная от мобильных до эксафлопсных суперкомпьютерных систем. Заявленная цель ресурса AlgoWiki – дать исчерпывающее описание каждого алгоритма, которое поможет оценить его потенциал применительно к конкретной вычислительной платформе. Для этого в AlgoWiki для каждого алгоритма приводится его описание, а также описание его параметров, необходимых как для последовательной, так и параллельной численной реализации, с указанием наиболее затратных по времени частей. Кроме того, AlgoWiki предоставляет ссылки на уже готовые пакетные решения.

Web-ресурс parallel.ru [6] нацелен на информирование пользователей о событиях, происходящих в суперкомпьютерном сообществе (новые технологии, программные продукты и средства разработки параллельного кода, конференции, новые суперкомпьютеры). Кроме этого, данный ресурс оказывает информационно-консультационные услуги в области высокопроизводительных вычислений, такие, как обучение технологиям параллельного программирования, проектирование и настройка кластерных вычислительных систем, предоставление вычислительных ресурсов для проведения реальных вычислений различной сложности и интенсивности и т.п.

Из зарубежных ресурсов подобного типа можно назвать HPCwire [4], который позиционируется его создателями как новостной и информационный ресурс № 1, рассказывающий о самых быстрых компьютерах в мире и людях, которые ими управляют. Этот ресурс ориентирован на специалистов в области науки, техники и бизнеса, заинтересованных в высокопроизводительных и ресурсоемких вычислениях. В HPCwire

освещается много различных тем: от последних новостей и новейших технологий в мире высокопроизводительных вычислений до новых тенденций, экспертного анализа и эксклюзивных функций.

Ряд зарубежных ресурсов также предоставляет пользователям доступ к высокопроизводительным кластерам и обучает работе на них. Все это делается, как правило, на платной основе. Так, например, через сайт группы высокопроизводительных исследовательских вычислений (HPRC group) из Техасского университета [5] можно получить доступ к трем кластерам с общей пиковой производительностью в 947 TF и высокопроизводительным хранилищем в 13,5 PB. Сайт HPRC group также предоставляет консультации, техническую документацию и обучение пользователей этих ресурсов.

Упомянутые выше российские и зарубежные ресурсы систематизируют информацию, относящуюся к суперкомпьютерам и высокопроизводительным вычислениям, не прибегая к онтологиям, что значительно снижает их возможности, как в плане представления знаний и данных, так и удобства доступа к ним. Представление такой информации в виде онтологии позволяет обеспечить не только удобный доступ к ней, но и эффективную поддержку пользователя в выборе оптимальных алгоритмов и параллельных архитектур при решении его прикладных задач за счет возможностей логического вывода на онтологии.

Наиболее близкий к предлагаемому в статье подход описан в работе [8]. Авторы рассматривают использование онтологий и механизмов вывода на онтологии для оказания помощи пользователям в решении вычислительно сложных задач на разнородных вычислительных архитектурах, в частности, на кластерах, снабженных сопроцессорами NVIDIA GPGPUs и Xeon-Phi в дополнение к традиционному процессору Intel. Здесь механизмы вывода на онтологии помогают находить лучшие из возможных решений посредством комбинирования оборудования, программного обеспечения и стратегий планирования. В этой работе демонстрируется применение предложенного подхода к решению задач биоинформатики.

2. Онтология проблемной области. Онтология проблемной области «Решение вычислительно сложных задач математической физики», используемая в создаваемом информационно-аналитическом интернет-ресурсе (ИАИР), построена на основе базовых онтологий научного знания и научной деятельности. Эти онтологии были ранее разработаны в нашем коллективе в рамках технологии построения интеллектуальных научных интернет-ресурсов [2].

Онтология решения вычислительно сложных задач математической физики является многоуровневой. Рассмотрим верхний уровень этой онтологии (рис. 1). Решаемая пользователем *Задача* связана с *Объектами исследования*. Основными объектами в рассматриваемой области являются *Физические объекты* и *Физические явления*, которые изучаются в определенных *Разделах науки* и описываются *Фундаментальными законами природы*, в свою очередь выводимыми на основе *Экспериментов и наблюдений*. *Объекты исследования* рассматриваются в виде приближенной *Физической модели*, описываемой *Математической моделью*, которая формализует *Фундаментальные законы природы*, задается *Системой уравнений* и верифицируется в ходе *Экспериментов и наблюдений*.

Система уравнений разрешается *Численным методом*, который, в свою очередь, реализуется тем или иным *Параллельным алгоритмом*, определяемым, в том числе, *Структурой данных*. *Параллельный алгоритм*, реализующий *Численный метод*,

оптимизируется под имеющуюся в распоряжении пользователя параллельную *Архитектуру* и представляется с помощью определенной *Технологии параллельного программирования*. Итоговое представление *Параллельного алгоритма* кодируется программным *Кодом*, состоящим из набора *Программных компонентов* и исполняемым на параллельной *Архитектуре*.

Описанная онтология, помимо перечисленных выше основных понятий рассматриваемой проблемной области и их взаимосвязей, включает классы онтологии научной деятельности, служащие для представления дополнительной информации о разработанных *Численных методах* и *Параллельных алгоритмах*. В частности, в ней представлена информация о том, в рамках какой деятельности и для какой предметной области они были разработаны и где используются. Кроме того, онтология позволяет описывать информацию о выполняемых исследованиях и об информационных ресурсах, создаваемых и используемых в этих исследованиях, об ученых, сообществах, организациях, вовлеченных в процесс таких исследований, о публикациях, посвященных данным исследованиям. Для этих целей служат классы *Область использования*, *Раздел науки*, *Деятельность*, *Публикация*, *Событие*, *Персона*, *Организация*, *Информационный ресурс*, *Географическое место*, а также отношения «являетсяРезультатомДеятельности», «относитсяКРазделуНауки», «используетсяВДеятельности», «являетсяАвторомПубликации», «описываетсяВПубликации», «представленНаРесурсе» и др.

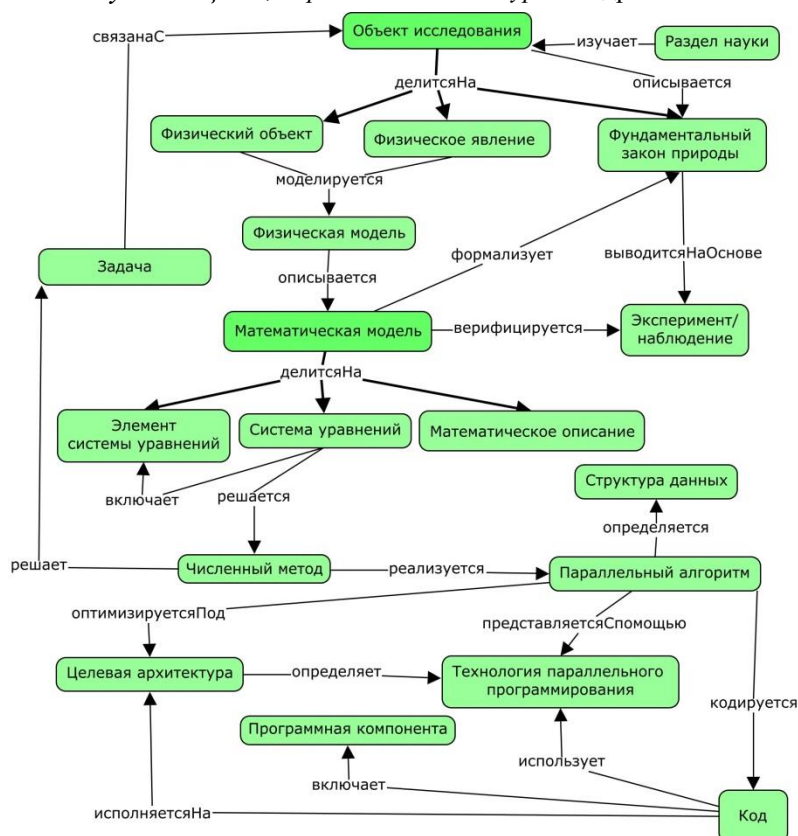


Рис.1. Верхний уровень онтологии «Решение вычислительно сложных задач математической физики».

Онтология решения вычислительно сложных задач математической физики описывается на языке OWL [7]. Она включает в себя описание классов и их свойств – атрибутов классов (Data Properties) и бинарных отношений между объектами классов (Object Properties), а также экземпляры классов (индивидов), составляющих наполнение базы знаний

конкретными задачами, методами, алгоритмами, программными компонентами и элементами параллельных архитектур.

Для систематизации и облегчения ввода экземпляров понятий онтологии была разработана система паттернов онтологического проектирования [11]. Такие паттерны представляют собой документально зафиксированные описания проверенных на практике решений типовых проблем онтологического моделирования. Они создаются для того, чтобы упорядочить и облегчить процесс построения онтологий и помочь разработчикам избежать типичных ошибок онтологического моделирования [9]. На Рис. 2 показан структурно-содержательный паттерн, описывающий один из основных классов рассматриваемой онтологии – *Численный метод*.

С содержательной точки зрения такой паттерн представляет собой семантическую окрестность центрального понятия, которым в данном случае является класс *Численный метод*. Для этого понятия определены свойства – атрибуты и отношения. Атрибуты представляются как Data Properties для тех свойств, значения которых имеют стандартный тип данных (*Название, Описание, Устойчивость к возмущениям, Абсолютная точность, Порядок точности, Вычислительная сложность*), или как Object Properties для свойств со значениями из перечислимого типа данных (*Вид решения, Тип сетки, Дискретизация расчетной области, Подход к представлению решения, реализующий метод Пакет*). Отношения задают связи объектов рассматриваемого класса с объектами других классов и представляются как Object Properties.

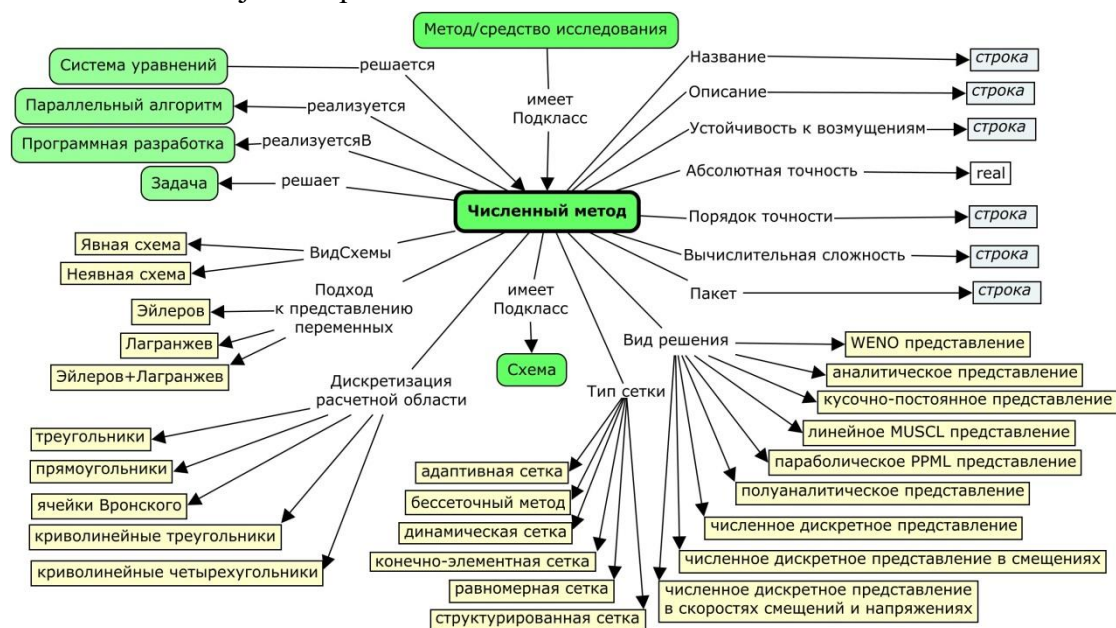


Рис. 2. Структурно-содержательный паттерн для задания экземпляров понятия «Численный метод»

Для построения онтологий и задания аксиом использовался редактор онтологий Protégé 5.2 [10]. Логический вывод в OWL-онтологиях осуществляется на основе аксиом, заданных в онтологии, средствами одной из машин вывода (Pellet, FaCT++, HermiT).

3. Реализация информационно-аналитического интернет-ресурса. На основе описанной выше онтологии был разработан информационно-аналитический интернет-ресурс для поддержки решения вычислительно сложных задач математической физики на суперкомпьютере. На рис. 3 показана страница данного ИАИР. В левой части этой страницы

представлены понятия онтологии, организованные в иерархии по отношению «общее-частное». При выборе понятия в центральной части страницы отображается список объектов выбранного класса, а также классов-потомков. Выбор какого-либо объекта из списка позволит отобразить описание свойств этого объекта (атрибутов и связей с другими объектами) в табличном или графическом виде. На рис. 3 представлено описание *Схемы MUSCL*. Из описания можно узнать, что данная схема имеет 1-ый *Порядок точности по времени* и 2-ой по пространству. MUSCL – это *Явная схема*, которая использует сетки разного типа, а также *Лагранжев* и *Эйлеров Подходы к представлению переменных*. Объекты, с которыми связан рассматриваемый объект, представлены на его странице гиперссылками. Переходя по этим ссылкам, можно выполнять навигацию по контенту интернет-ресурса и получать описания интересных объектов.

Для создания ресурса была использована технология разработки интеллектуальных научных интернет-ресурсов [2], которая предоставляет оболочку интернет-ресурса, набор базовых онтологий и методику построения онтологии с их использованием. Методика предполагает разработку системы понятий онтологии (Т-Box) средствами редактора Protégé.



Главная Онтология О ресурсе 🔍 Выход (admin)

ПОДДЕРЖКА РЕШЕНИЯ ВЫЧИСЛИТЕЛЬНО СЛОЖНЫХ ЗАДАЧ НА СУПЕРКОМПЬЮТЕРАХ

Свойства объекта

Название	MUSCL - Монотонная восходящая схема для законов сохранения
Описание	Явная монотонная противопотоковая схема для законов сохранения, обеспечивает высокоточные численные решения даже в тех случаях, когда решения имеют толчки, разрывы или большие градиенты
Порядок точности	1-ый по времени, 2-ой по пространству
Тип сетки	адаптивная сетка, динамическая сетка, конечно-элементная сетка, равномерная сетка, структурированная сетка
Подход к представлению переменных	Лагранжев подход, Эйлеров подход
Вид схемы	Явная схема

Связи объекта

реализует

Метод / средство исследования

[Метод Годунова](#)

[Метод конечных объемов](#)

реализуется в Программной разработке

Программная разработка

[Heracles](#)

[ART](#)

[AREPO](#)

© ИВМИГ, ИСИ 2019–2020 Ресурс разработан при финансовой поддержке РФФИ (проект №19-07-00085)

Рис. 3. ИАИР для поддержки решения вычислительно сложных задач математической физики на суперкомпьютерах

Наполнение ресурса, ввод конкретных объектов и их свойств (А-Box) может быть выполнено как в редакторе Protégé, так и в специализированном редакторе данных [1], входящем в состав оболочки интернет-ресурса. Этот редактор поддерживает работу с онтологическими паттернами содержания. При задании нового объекта некоторого класса редактор отображает семантическую окрестность типового объекта данного класса (его

свойства и связи с другими объектами), позволяя пользователю задавать значения свойств создаваемого объекта или выбирать эти значения/связанные объекты из списка возможных значений/объектов. На рис. 4 представлена страница редактора, которая отображает структурно-содержательный паттерн, показанный на рис. 2, а также атрибуты и отношения, унаследованные классом *Численный метод* от родительского класса *Метод/средство исследования*. В основном, это свойства, описанные в онтологии научной деятельности.

Значения атрибутов стандартных типов вводятся пользователем с клавиатуры. Доменные значения атрибутов выбираются из выпадающего списка. Объекты, с которыми связан создаваемый или редактируемый объект, также выбираются из выпадающего списка.

Редактирование экземпляра
Свойства

Название	ru	MUSCL - Монотонная восходящая схема для законов сохранения	✕
Описание	ru	Явная монотонная противопотоковая схема для законов сохранения	✕
Дата возникновения			
Устойчивость к возмущениям			
Порядок точности	ru	1-ый по времени	✕
	ru	2-ой по пространству	✕
Абсолютная точность			
Вычислительная сложность			
Тип сетки		адаптивная сетка	✕
		динамическая сетка	✕
		конечно-элементная сетка	✕
		равномерная сетка	✕
		структурированная сетка	✕
Подход к представлению переменных			✕
		структурированная сетка	✕
Вид решения		адаптивная сетка	
Вид схемы		бессеточный метод	✕
Дискретизация расчетной области		динамическая сетка	
		конечно-элементная сетка	
		равномерная сетка	

СВЯЗИ

- используется в Деятельности
- решает Задачу
- представлен на Ресурсе
- использует
- использует Метод
- реализует

Метод / средство исследования: MUSCL - Монотонная восходящая...

Метод / средство исследования: Метод Годунова

Метод / средство исследования: MUSCL - Монотонная восходящая...

Метод / средство исследования: Метод конечных объемов

Рис. 4. Редактор данных. Паттерн для задания экземпляров понятия «Численный метод»

Заключение. В статье представлен подход к разработке информационно-аналитического интернет-ресурса, предназначенного для поддержки решения вычислительно сложных задач математической физики на суперкомпьютерах. Данный ресурс обеспечивает пользователя подробным систематизированным описанием параллельных алгоритмов и архитектур и предоставляет ему содержательный доступ к детальным описаниям имеющихся в наличии параллельных архитектур и используемых в них устройств, доступных программных компонентов, реализующих параллельные алгоритмы и фрагменты

параллельного кода, а также к публикациям и информационным ресурсам, описывающим данную проблемную область.

Использование онтологии в качестве базиса такого ресурса позволяет обеспечить не только удобный доступ ко всей перечисленной выше информации, но и эффективную поддержку пользователя в выборе оптимальных алгоритмов и параллельных архитектур при решении его прикладных задач за счет возможностей логического вывода на онтологии.

Работа выполнена при частичной финансовой поддержке РФФИ (гранты № 19-07-00085 и № 19-07-00762).

СПИСОК ЛИТЕРАТУРЫ

1. Ахмадеева И.Р., Серый А.С., Шестаков В.К. Некоторые особенности реализации платформы для построения информационно-аналитических интернет-ресурсов // Информационные и математические технологии в науке и управлении. 2017. № 3 (7). С. 168-175.
2. Загорулько Ю.А., Загорулько Г.Б., Боровикова О.И. Технология создания тематических интеллектуальных научных интернет-ресурсов, базирующаяся на онтологии // Программная инженерия. 2016. № 2(7). С. 51-60.
3. Сайт AlgoWiki. Открытая энциклопедия свойств алгоритмов. Режим доступа: <https://algowiki-project.org/ru/> (дата обращения: 15.07.2020)
4. Сайт HPCwire. Режим доступа: <https://www.hpcwire.com/> (дата обращения: 15.07.2020)
5. Сайт HPRC. Режим доступа: <https://hprc.tamu.edu/> (дата обращения: 15.07.2020)
6. Сайт Parallel.ru. Режим доступа: <https://parallel.ru/> (дата обращения: 15.07.2020)
7. Antoniou G., Harmelen F. Web Ontology Language: OWL // Handbook on Ontologies. Berlin: Springer Verlag. 2004. Pp. 67-92.
8. Compton M., Barnaghi P., Bermudez L., Garcia-Castro R., Corcho O., Cox S., Graybeal J., Hauswirth M., Henson C., Herzog A., et al. The SSN Ontology of the W3C Semantic Sensor Network Incubator Group. Web Semant. Sci. Serv. Agents World Wide Web. 2012. Vol. 17. Pp. 25-32.
9. Gangemi A., Presutti V. Ontology Design Patterns // Handbook on Ontologies. Berlin: Springer. 2009. Pp. 221-243.
10. Protégé. Режим доступа: <https://protege.stanford.edu> (дата обращения: 05.01. 2019).
11. Zagorulko, Y., Borovikova, O., Zagorulko, G., 2018. Development of Ontologies of Scientific Subject Domains Using Ontology Design Patterns // Communications in Computer and Information Science. 2018. Vol. 822. Pp. 141-156.
12. Zagorulko G., Zagorulko Y., Glinskiy B., Sapetina A. Ontological Approach to Providing Intelligent Support for Solving Compute-Intensive Problems on Supercomputers // Kuznetsov S., Panov A. (eds) Artificial Intelligence. RCAI 2019. Communications in Computer and Information Science. 2019. Vol. 1093. Pp. 363-375.

UDK 004.822:004.89:519.63

**INFORMATION AND ANALYTICAL SUPPORT OF SOLVING
COMPUTE- INTENSIVE PROBLEMS ON SUPERCOMPUTERS**

Yury A. Zagorulko

Dr., Head of Laboratory "Artificial Intelligence", e-mail: zagor@iis.nsk.su,

Galina B. Zagorulko

Researcher, e-mail: gal@iis.nsk.su,

Vladimir K. Shestakov

Junior Researcher, e-mail: shestakov@iis.nsk.su,

A.P. Ershov Institute of Informatics Systems

Siberian Branch of the Russian Academy of Sciences

6, Acad. Lavrentjev pr., 630090, Novosibirsk, Russia.

Anna F. Sapetina

Junior Researcher, e-mail: sapetina@sscc.ru,

Institute of Computational Mathematics and Mathematical Geophysics

Siberian Branch of the Russian Academy of Sciences,

6, Acad. Lavrentjev pr., 630090, Novosibirsk, Russia.

Annotation. The paper presents an approach to the development of an information-analytical Internet resource to support of solving compute-intensive problems of mathematical physics on supercomputers. The basis of this resource is a multilevel ontology, which is built using ontologies of scientific knowledge and scientific activity. For effective information and analytical support of users, it is very important to provide a detailed systematized description of parallel algorithms and architectures, a fairly complete description of available software components that implement parallel algorithms and fragments of parallel code, available parallel architectures and devices used in them, as well as publications and information resources describing this problem area.

Keywords: solving compute-intensive problems on supercomputers, parallel algorithms, parallel architectures, ontology, information-analytical Internet resource

References

1. Ahmadeeva I.R., Seryj A.S., Shestakov V.K. Nekotorye osobennosti realizacii platformy dlya postroeniya informacionno-analiticheskikh internet-resursov [Some features of implementation of the platform for building information-analytical Internet resources] // Informacionnye i ma-tematicheskie tekhnologii v nauke i upravlenii = Information and mathematical technologies in science and management. 2017. № 3 (7). Pp. 168-175. (in Russian).
2. Zagorulko Yu.A., Zagorulko G.B., Borovikova O.I. Tekhnologija sozdaniya tematicheskikh intellektual'nykh nauchnykh internet-resursov, bazirujushhajasja na ontologii [Technology for building subject-based intelligent scientific internet resources based on ontology] // Programmaja inzhenerija = Software Engineering. 2016. №. 2. Pp. 51-60. (in Russian)
3. AlgoWiki website. Open Encyclopedia of Algorithm Properties. Available at:

<https://algowiki-project.org/ru/>, accessed 15.07.2020 (in Russian)

4. HPCwire. website. Available at: <https://www.hpcwire.com/>, accessed 15.07.2020
5. HPRC website. Available at: <https://hprc.tamu.edu/>, accessed 15.07.2020
6. Parallel.ru. website. Available at: <https://parallel.ru>, accessed 15.07.2020 (in Russian)
7. Antoniou, G., Harmelen, F.: Web Ontology Language: OWL. // Handbook on Ontologies. Berlin: Springer Verlag. 2004. Pp. 67-92.
8. Compton M., Barnaghi P., Bermudez L., Garcia-Castro R., Corcho O., Cox S., Graybeal J., Hauswirth M., Henson C., Herzog A., et al. The SSN Ontology of the W3C Semantic Sensor Network Incubator Group. Web Semant. Sci. Serv. Agents World Wide Web. 2012. Vol. 17. Pp. 25-32.
9. Gangemi A., Presutti V. Ontology Design Patterns // Handbook on Ontologies. Berlin: Springer. 2009. Pp. 221-243.
10. Protégé. <https://protege.stanford.edu>, accessed 05.01. 2019.
11. Zagorulko Y., Borovikova O., Zagorulko G., Development of Ontologies of Scientific Subject Domains Using Ontology Design Patterns. // Communications in Computer and Information Science. 2018. Vol. 822. Pp. 141-156.
12. Zagorulko G., Zagorulko Y., Glinskiy B., Sapetina A. Ontological Approach to Providing Intel-ligent Support for Solving Compute-Intensive Problems on Supercomputers // Kuznetsov S., Panov A. (eds) Artificial Intelligence. RCAI 2019. Communications in Computer and Information Science. 2019. Vol. 1093. Pp. 363-375.