

## **CUBLIC — ПЛАТФОРМА ДЛЯ УДАЛЁННЫХ И РАСПРЕДЕЛЁННЫХ НАУЧНЫХ ВЫЧИСЛЕНИЙ.**

### **CUBLIC IS A PLATFORM FOR REMOTE AND DESTRIIBUTED SCINTIFIC COMPUTINGS.**

#### Авторы / Authors:

**Емельянов А.А., Климашов В.Ю., Савихин С.А., Терентьев А.Б. /**

**Emelyanov A.A, Klimashov V.Y., Savikhin S.A., Terentyev A.B.**

#### Аннотация / Annotation:

Настоящая статья описывает программную платформу CUBLIC, доступную для массового использования при решении широкого круга научных и прикладных задач. Комплекс программных средств для удалённого доступа к программной платформе обеспечивает быструю сборку заданий пользователем-непрограммистом, а среда исполнения — их эффективное выполнение на вычислительных архитектурах с массовым параллелизмом.

This article's about software platform CUBLIC, which is accessible to mass using in the solution of a broad circle of scientific and applied problems. The complex of software tools for remote access to software platform is ensure quick assemblage of the tasks by user-not programmer and their effective implementation on computing architectures with mass parallelism.

#### Ключевые слова / Key words:

вычислительный комплекс, grid, суперкомпьютерные технологии, распределённые вычисления;

computer complex, grid, supercomputer tehnologies, distributed computing.

Сфера суперкомпьютерных вычислений в последние годы приобретает всё большую и большую популярность. Повышенный спрос обуславливается многими причинами. Во-первых, без математического моделирования, имитационных экспериментов, просто без сложных расчётов в настоящее время не обходится практически ни одна как научная, так и инженерно-техническая отрасль. Современная наука всё чаще нуждается в сверхточных и сверхбыстрых вычислительных системах. Эта потребность остро ощущается в физике, химии, биологии, экономике, вычислительной математике, информатике, технике и многих других областях науки. Во-вторых, появляется всё больше и больше производителей аппаратных платформ как

зарубежных, так и отечественных, что положительно сказывается как на ассортименте предлагаемых решений, так и на их цене. Однако, стоимость вычислительных комплексов пусть даже средней производительности (до 20 Тфлопс) остаётся достаточно высокой. Совместно с аппаратными решениями также появляются и программные, направленные на упрощение разработки, отладки, контроля параллельных приложений, создания собственных вычислительных кластеров и т. д. В-третьих, со стороны Федеральных властей проводится активная политика продвижения суперкомпьютерных технологий, что открывает дополнительные возможности для их внедрения, использования и разработки. В то же время, обучение программистов, способных создавать качественные, эффективные вычислительные приложения, уже не кажется невыполнимой задачей — проводятся мастер-классы, школы, выпускаются учебные пособия. Однако, следует отметить, что задача всё равно остаётся чрезвычайно сложной и актуальной.

Существует ещё один вопрос, который, к сожалению, зачастую остаётся без внимания. Исторически сложилось, что основными потребителями суперкомпьютерных вычислений являются достаточно крупные организации, имеющие: 1) готовый отлаженный алгоритм вычислений, моделирования, расчётов; 2) свой штат профессиональных разработчиков (либо возможности пригласить специалистов), способных этот алгоритм эффективно реализовать. При этом отдельные учёные, небольшие научные и исследовательские организации зачастую не имеют достаточных возможностей воспользоваться существующими суперкомпьютерами. Связано это, в первую очередь, со спецификой научных исследований. Наш достаточно обширный опыт общения с учёными из разных сфер науки показывает, что проверка вновь придуманного алгоритма решения задачи, метода математического моделирования и т. д. проводится, как правило, в программах типа «MatLab», где можно достаточно быстро и просто описать разрабатываемый алгоритм, поменять некоторые параметры или полностью перестроить весь ход вычислений, если полученный результат по той или иной причине не удовлетворяет. То есть процессы разработки алгоритма и программы, его выполняющей, неразрывны и идут параллельно. Главный недостаток такого рода программ — невысокая (а в случае работы с большими объёмами данных, чрезвычайно низкая) скорость выполнения.

Представляемая программно-аппаратная платформа CUBLIC призвана максимально приблизить сложные научные вычисления к пользователям-прикладникам, не являющимся специалистами в программировании. Основная цель проекта — повышение доступности высокопроизводительных вычислений, снижение финансовых и временных затрат при решении научно-исследовательских и инженерно-технических задач в различных прикладных областях.

Такое видение вопроса сформировалось из понимания желаний учёных-исследователей, появившегося из-за достаточно плотного общения с ними. Современные исследования в областях физики, медицины, живых систем просто невозможны без обработки больших объёмов данных. Причём, в большинстве случаев обработка заключается не только в воспроизведении какого-либо алгоритма, а, в том числе, в создании некоторого нового метода, который требуется сразу же опробовать. Таким образом, учёным-исследователям нужен простой, удобный инструмент, позволяющий быстро реализовать требуемый алгоритм, настроить его соответствующим образом и быстро получить требуемый результат. При этом всё должно происходить без привлечения профессиональных программистов, а значит, разработка вычислительных приложений должна требовать только понимания используемого математического аппарата и наличия знаний в предметной области научного исследования.

Именно это и стало основным содержанием проекта CUBLIC — создание вычислительной платформы, позволяющей решать вычислительные задачи практически любой сложности без привлечения профессиональных программистов.

Алгоритм вычислительной задачи представляется в виде схемы с функциональными узлами, а именно: блоками, связями, указывающими направление передачи данных от одного блока к другому, и параметрами-настройками, специализирующими вычислительную функцию блока. Каждый блок выполняет определенную элементарную эффективно реализованную вычислительную функцию. Схема из нескольких блоков может быть преобразована в модуль — самостоятельный функционально независимый объект, который в основной схеме может применяться наравне с блоками на любом уровне вложенности.

Интуитивно понятный пользовательский интерфейс представляет собой браузерный редактор алгоритмов, использующий язык разметки HTML5 (поддерживается в браузерах Opera 10.6+, Google Chrome 6.0.4+, Firefox 3.6+, Internet Explorer 9+). Основным плюсом такого подхода является его независимость как от пользовательской аппаратной платформы, так и от используемой операционной системы. Собранные схемы-алгоритмы, файлы результатов хранятся на удалённом сервере, что позволяет работать с платформой в любой точке земного шара, где есть доступ к сети Интернет. Наряду с редактором был разработан личный кабинет для контроля над исполнением и завершением исполнения схем алгоритмов, а также над обменом данными между сервером проекта и компьютером пользователя.

На рис. 1 представлена схема алгоритма моделирования грозовой активности.

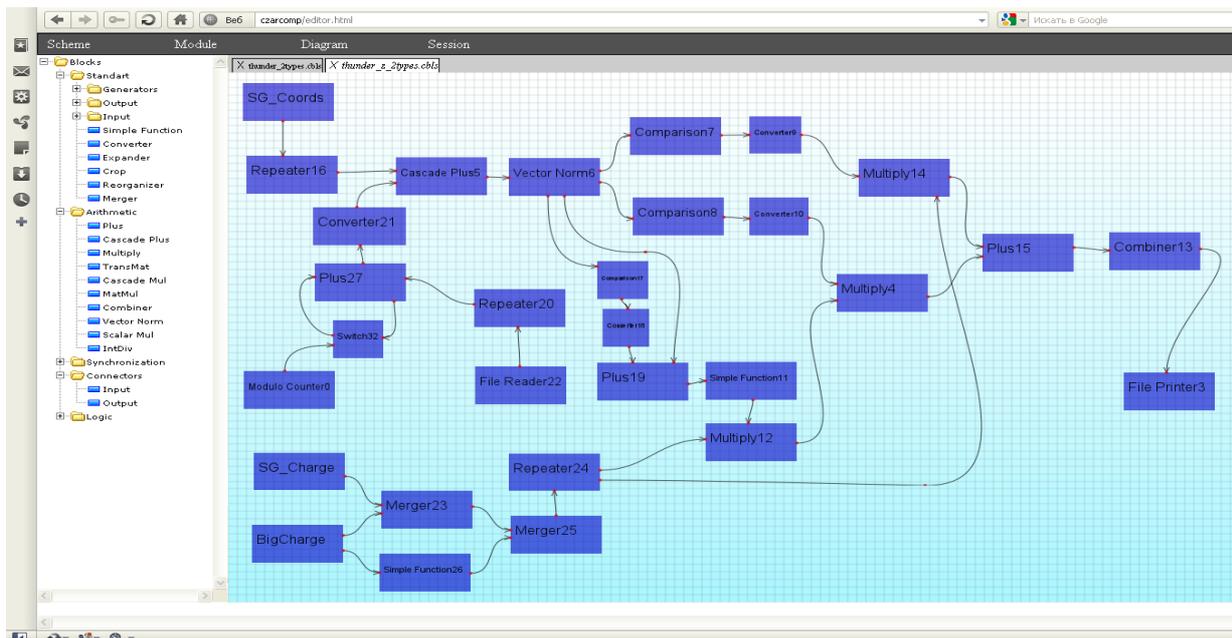


Рис.1 Схема алгоритма моделирования внутриоблачной грозовой активности.

В системе реализованы следующие виды элементарных блоков: ввода или генерации исходных данных, блоки синхронизации, векторно-матричных операций, элементарных функций, аналитических функций и блоки специализированных высокоуровневых алгоритмов. На базе вышеперечисленных блоков, реализуются модули, которые можно применять в решении сложных предметных задач.

Процесс низкоуровневого распараллеливания задачи (разделение общей задачи на независимые функциональные части, распределение их по вычислительным узлам, передача данных и синхронизация памяти) скрыт от пользователя и выполняется автоматически в рамках среды выполнения, в соответствии с конфигурацией вычислительного комплекса, доступностью аппаратных ресурсов, пользовательских привилегий. Таким образом, для решения прикладной задачи от специалиста-предметника требуется только владение основами математического моделирования.

При разработке проекта максимально использовались широко распространённые, открытые библиотеки, стандарты и протоколы: REST/JSON, MPI, X.509, OpenCL, boost и т. д. Кроме того, был разработан целый ряд собственных оригинальных решений по разбиению и распределению задач, распределению ресурсов, синхронизации процессов и данных.

Проект CUBLIC предполагает эффективное использование на аппаратных комплексах как в локальных, так и в распределённой кластерной или grid-конфигурациях, в том числе и на гибридной аппаратной платформе. Программная платформа включает в себя:

(1) комплекс инструментальных средств, обеспечивающих быструю сборку вычислительных приложений пользователем-непрограммистом, владеющим только основным математическим аппаратом в своей области;

(2) промежуточное программное обеспечение среды исполнения композитных пользовательских приложений, обеспечивающее эффективное выполнение вычислений на недорогой общедоступной аппаратной платформе, как в рамках традиционных кластеров, так и масштабируемых гетерогенных grid-сетях второго поколения;

(3) набор прототипов (шаблонов) прикладных функциональных модулей (компонентов) для использования в составе пользовательских композитных приложений, а также в качестве примеров. В настоящее время в качестве аппаратной платформы используются классические x86-процессоры и многоядерные GPU NVidia и AMD.

Проект CUBLIC уже успел пройти первичную апробацию конечными потребителями. Совместно со специалистами, занимающимися изучением биологических процессов головного мозга, с помощью CUBLIC были решены следующие задачи: моделирование внутреннего взаимодействия в нейрональных сетях, динамики роста нейрональных сетей, межнейронного взаимодействия с учётом периода рефрактерности, длины аксона, пластичности, а также реализованы алгоритмы исследования зависимости параметров нейрональной сети от структуры её внутренних связей. В рамках исследований в сфере атмосферного электричества собраны схемы моделирования внутриоблачной активности на первичных стадиях грозовых разрядов, расчёта фрактальной размерности грозовых облаков, симуляции роста проводящего кластера. В рамках совместных НИР проводятся реальные расчёты в областях физики живых систем, атмосферного анализа, акустики. В ходе совместных работ со специалистами-предметниками также разрабатываются высокоуровневые алгоритмы, которые в дальнейшем могут быть использованы при решении других задач.

Сайт проекта: <http://www.hopcomp.net>

Ознакомительная версия CUBLIC: <http://host1.hopcomp.net>.

Сведения об авторах / Information about authors:

**Емельянов Алексей Александрович / Emelyanov Aleksey Aleksandrovich**

ООО «НИЦ СВТ», главный конструктор, в 2010г. окончил НГТУ им. Алексева Р.Е., автор около 20 статей, область научных интересов – параллельные вычисления, математические модели, алгоритмы, e-mail: [leshiy@ivc.nnov.ru](mailto:leshiy@ivc.nnov.ru), тел. (831) 437-16-36

SCT LLC, main designer, graduated from Alekseev R.Y. NSTU in 2010, the author of near 20 articles, the fields of scientific researches're parallel computings, mathematical models, algorithmes, e-mail: [leshiy@ivc.nnov.ru](mailto:leshiy@ivc.nnov.ru), tel. (831) 437-16-36

**Климашов Виталий Юрьевич / Klimashov Vitaliy Yuryevich**

ООО «НИЦ СВТ», директор по маркетингу, в 2010г. окончил НГТУ им. Алексева Р.Е., автор около 20 статей, область научных интересов – высокопроизводительные вычисления, программирование, суперкомпьютерные технологии, e-mail: [klim@ivc.nnov.ru](mailto:klim@ivc.nnov.ru), тел. (831) 437-16-36

SCT LLC, marketing director, graduated from Alekseev R.Y. NSTU in 2010, the author of near 20 articles, the fields of scientific researches're high performance computings, programming, supercomputer technologies, e-mail: [klim@ivc.nnov.ru](mailto:klim@ivc.nnov.ru), tel. (831) 437-16-36

**Савихин Степан Александрович / Savikhin Stepan Aleksandrovich**

ООО «НИЦ СВТ», инженер-программист, в 2009г. окончил ННГУ им. Лобачевского Н.И., автор около 20 статей, область научных интересов – функциональный анализ, теория функций комплексных переменных, нелинейная динамика, вычислительная математика и методы оптимизации, e-mail: [savikhin@ivc.nnov.ru](mailto:savikhin@ivc.nnov.ru), тел. (831) 437-16-36

SCT LLC, engineer-programmer, graduated from Lobachevskiy N.I. NNSU in 2009, the author of near 20 articles, the fields of scientific researches're functional analysis, complex analysis, nonlinear dynamical systems, computational mathematics and mathematical optimization, e-mail: [savikhin@ivc.nnov.ru](mailto:savikhin@ivc.nnov.ru), tel. (831) 437-16-36

**Терентьев Александр Борисович / Terentyev Aleksandr Borisovich**

ООО «НИЦ СВТ», исполнительный директор, в 1985 г. окончил ГПИ им. Жданова А.А., в 1989 г. – ГГУ им. Лобачевского, автор около 40 статей, область научных интересов – параллельные вычисления, нелинейная динамика, e-mail: [ater@ivc.nnov.ru](mailto:ater@ivc.nnov.ru), тел. (831) 437-16-36

SCT LLC, executive director, graduated from Zhdanov A.A. GPI in 1985, Lobachevskiy N.I. GSU in 1989, the author of near 40 articles, the fields of scientific researches're parallel computings, nonlinear dynamical systems, e-mail: [ater@ivc.nnov.ru](mailto:ater@ivc.nnov.ru), tel. (831) 437-16-36