

УДК 681.324

ПАРАЛЛЕЛЬНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ НА БАЗЕ ВЫЧИСЛИТЕЛЬНОГО КЛАСТЕРА

© 2003 Г.П. Яровой¹ Ю.Н. Радаев² Ю.А. Родичев³ В.А. Салеев⁴

В статье обсуждается перспективный проект, связанный с созданием центра высокопроизводительных вычислений и организацией вычислительного кластера в Самарском государственном университете для решения фундаментальных и прикладных научных задач, требующих значительных вычислительных мощностей, обработки больших объемов информации, а также организации учебного процесса по современным методам параллельного программирования.

Введение

В Концепции развития Самарского государственного университета как центра науки, культуры и образования Поволжского региона [1] среди основных целей и задач отмечены следующие:

- создание условий для развития научных школ, ведущих разработку фундаментальных и прикладных научных проблем в рамках международных, российских и региональных программ;
- обеспечение ведущей роли Самарского государственного университета в реализации концепции информатизации вузов и других образовательных учреждений Самарской области.

Важным шагом в решении этих задач является становление в Самарском государственном университете новейших информационных технологий, которые бы не только дополнили уже имеющиеся и успешно развивающиеся на базе соответствующих подразделений университета, но и вывели бы университет в сфере информатизации на уровень, сопоставимый с

¹Яровой Геннадий Петрович (rector@ssu.samara.ru), кафедра радиопизики и компьютерного моделирования радиосистем Самарского государственного университета, 443011, г. Самара, ул. Акад. Павлова, 1.

²Радаев Юрий Николаевич (radayev@ssu.samara.ru), кафедра механики сплошных сред Самарского государственного университета.

³Родичев Юрий Андреевич (rodichev@ssu.samara.ru), информационно-вычислительный центр Самарского государственного университета.

⁴Салеев Владимир Анатольевич (saleev@ssu.samara.ru), кафедра общей и теоретической физики Самарского государственного университета.

уровнем ведущих образовательных центров России, и позволили выйти на лидирующие позиции среди вузов Самары.

Одним из приоритетных направлений развития науки, техники и технологии в Российской Федерации является направление: "Информационно-телекоммуникационные технологии и электроника", а в перечень критических технологий Российской Федерации включен пункт "Высокопроизводительные вычислительные системы" [2]. На наш взгляд, становление и развитие указанного выше направления в Самарском государственном университете является актуальной и соответствующей интеллектуальному и материально-техническому уровню университета задачей.

1. Вычислительные кластеры в науке и образовании

Последние годы во всем мире и в России происходит бурное внедрение вычислительных кластеров — локальных сетей с узлами из рабочих станций или персональных компьютеров, специально собранных для использования в качестве многопроцессорной вычислительной системы (суперкомпьютера). В качестве примера можно привести Московский государственный университет им. М.В. Ломоносова. По данным, представленным на информационном Интернет-портале [4], в Научно-исследовательском вычислительном центре МГУ функционирует 4 кластера с числом рабочих станций (Pentium II, Pentium III, 2xXeon, 2xAlpha) от 12 до 18 в каждом. Кроме этого, отдельные вычислительные кластеры созданы на физическом факультете (18 узлов, Pentium II) и факультете вычислительной математики и кибернетики (18 узлов, Pentium III) МГУ. В Самаре вычислительные кластеры имеются только в подразделениях Академии наук: в Самарском научном центре РАН (5 узлов, 2xAlpha, Pentium II) и Институте СОИ РАН (4 узла, Pentium II). Крупнейший вычислительный кластер в России (система МВС-1000М/384) находится в Межведомственном компьютерном центре и включает в себя 384 узла (2xAlpha). Мировой опыт создания вычислительных кластеров значительно шире: от скромных 20-30 узловых кластеров в учебных или научных лабораториях университетов до гигантских вычислительных систем, состоящих из 1000-2000 рабочих станций, созданных в рамках специальных проектов.

Для построения вычислительных кластеров обычно используются общедоступные компьютеры на базе процессоров Intel или AMD, стандартные сетевые технологии Ethernet или Fast Ethernet, свободно распространяемая операционная система (ОС) Linux [5] и коммуникационная библиотека MPI [6], реализующая связь между ветвями параллельного вычислительного процесса. Таким образом, сегодня вычислительные кластеры стали общедоступной и сравнительно дешевой альтернативой традиционным суперкомпьютерам. На многих классах задач и при достаточно большом числе узлов кластерные системы достигают производительности, сравнимой с суперкомпьютерной [7]. Еще раз подчеркнем, что главным успехом кластерных

вычислительных технологий является то, что параллельные вычисления, реальная альтернатива которым в области высокоскоростных вычислений в ближайшие годы не предвидится, стали доступны для многих образовательных и научно-исследовательских организаций.

Приходится далее констатировать, что парк вычислительной техники университета составляют в основном персональные компьютеры, а вычислительные кластеры отсутствуют. Тем не менее в Самарском государственном университете существует ряд научных направлений, для успешного развития которых повышение имеющегося в университете уровня вычислительных мощностей является совершенно необходимым условием.

В микроволновой электронике и теории волновых процессов (это направление развивается на кафедре радиофизики и компьютерного моделирования радиосистем) методика параллельных вычислений позволит решать граничные и эволюционные задачи для сложных волноводных структур с физическими средами, обладающими пространственно-временной зависимостью электродинамических параметров [8].

Научные исследования, ведущиеся на кафедре неорганической химии, требуют большого объема вычислений, в связи с чем при решении ряда задач (в частности, для квантово-химических расчетов твердого тела, моделирования структуры вещества методами молекулярной динамики и др.) могут быть использованы ресурсы университетского вычислительного кластера. Кроме того, квантово-химические программы для расчетов *ab initio* (Gaussian, HONDO) могут быть установлены на рабочих станциях кластера для их использования в учебном процессе.

В настоящее время на кафедре общей химии и хроматографии проводятся теоретические и экспериментальные исследования методом газовой хроматографии сорбционных и селективных свойств систем, содержащих термотропные жидкие кристаллы (ЖК) с различным типом мезофаз. Актуальной задачей, решаемой в рамках данного направления, является молекулярно-статистическое (МС) моделирование термодинамических характеристик индивидуальных и смешанных ЖК-систем (моно- и полимолекулярных слоев и фазовых пленок), включая теоретическое описание поведения немезогенов в таких системах (адсорбция, растворение), а также развитие теории анизотропных растворов [9]. Проведение расчетов даже в рамках простейших МС-моделей связано с необходимостью использования современной высокопроизводительной вычислительной техники, а детальное моделирование может быть реализовано только в рамках технологий параллельного программирования.

Научно-исследовательская работа, проводимая кафедрой механики сплошных сред, требует подключения технологий параллельного программирования для численного анализа нелинейных задач, с которыми приходится сталкиваться как в фундаментальных вопросах механики, так и в инженерной практике. Сложные прикладные расчеты нагружения деформируемых тел, обладающих сложными реологическими свойствами,

течений жидкости и газа с сопутствующими теплофизическими процессами традиционно выполняются на суперкомпьютерах или высокопроизводительных рабочих станциях, имеющих несколько гигабайт оперативной памяти и до нескольких десятков высокопроизводительных процессоров. Сейчас выполнить подобную вычислительную работу в рамках университета не представляется возможным из-за отсутствия высокопроизводительных рабочих станций и опыта программирования в многопроцессорных системах. В последнее время, в связи с быстрым ростом производительности широко распространенных компьютеров с 32-разрядной архитектурой Intel, появляется возможность проводить указанные расчеты также и на персональных компьютерах, использующих современные пакеты поддержки механики деформируемого твердого тела, подобные ANSYS, LS-DYNA⁵ и др. Наблюдающаяся в настоящее время тенденция к быстрому переоснащению компьютерами Pentium IV (2.4 ГГц) с оперативной памятью 1-2 Гб позволяет ставить и решать многие вычислительные задачи механики сплошных сред (в частности, задачи механики разрушения и накопления повреждений [10], математической теории пластичности, моделирования технологий нефтедобычи, а также прогнозирования изменений климата и поддержания экологического равновесия Самарского региона), организуя многопроцессорные сетевые кластеры на базе процессоров Intel.

На кафедре общей и теоретической физики в стадии развития находится проект по созданию научной группы для обработки экспериментальных данных с ускорителя протонов У-70 ГИЦ "Институт физики высоких энергий" г. Протвино [11]. Имеющийся научный задел и кадровый потенциал уже сейчас позволяют вести конкретную работу по обработке экспериментальных данных совместно с Отделом адронной физики ИФВЭ г. Протвино. Однако объем обрабатываемой информации и разумные требования на время ее обработки требуют вычислительной мощности, на 1-2 порядка превышающей возможности персонального компьютера Pentium IV. Отметим, что обработка больших объемов экспериментальных данных, в том числе в режиме реального времени, является одним из перспективных направлений использования вычислительного кластера. Другим перспективным направлением использования вычислительного кластера может быть решение нелинейных задач квантовой теории поля методом "решеток" [12].

Создание вычислительного кластера в университете позволит организовать полноценный учебный процесс и готовить специалистов в области

⁵LS-DYNA — многоцелевой конечно-элементный комплекс разработки Livermore Software Technology Corp. (LSTC) — предназначена для анализа высоконелинейных и быстротекущих процессов в задачах механики сплошных сред. Сфера приложений пакета LS-DYNA: нелинейная динамика, теплоперенос, теплообмен, термомеханика, разрушение и развитие трещин, контактные взаимодействия, акустика, связанный анализ (взаимодействие потоков жидкостей и газов с деформируемыми конструкциями, связанные термомеханические задачи). LS-DYNA содержит более 130 уравнений состояния деформируемых сред.

высокопроизводительных вычислений, спрос на которых прогнозируется в ближайшие годы.

2. Аппаратное обеспечение

Обсуждаемая ниже аппаратная конфигурация вычислительного кластера относится к его модельному сегменту и предполагает использование оборудования, уже выделенного для этой цели в Самарском государственном университете.

Представляется целесообразным создание кластера выделенных рабочих станций, т.е. сети, узлы которой выделены для работы в составе многопроцессорной вычислительной системы и не будут использоваться в других целях. Это предполагает компактное размещение компонент сети и оснащение программным обеспечением (ПО), ориентированным исключительно на управление и использование вычислительным кластером как целым. Пользователи при этом смогут осуществлять связь с кластером в режиме удаленного доступа со своих персональных компьютеров, используя их как рабочие терминалы кластера.

В качестве узлов вычислительного кластера будут использованы стандартные персональные ЭВМ на основе процессора Intel Pentium IV с тактовой частотой 2.4 ГГц, оперативной памятью 1 Гб и жесткими дисками по 80 Гб. Управляющей машиной является двухпроцессорная рабочая станция Intel Xeon 2x2.4 с оперативной памятью 1 Гб и двумя жесткими дисками по 80 Гб (число жестких дисков может быть увеличено до 8). Таким образом, вычислительный кластер будет однородным, что упростит и сделает более эффективными алгоритмы параллельного программирования. Кластер будет объединен с единой корпоративной сетью (ЕКС) университета и при необходимости может использовать ее ресурсы, а также иметь выход в глобальные сети.

Управление выполнением параллельных программ на узлах вычислительного кластера предполагает наличие сети управления, через которую управляющая машина взаимодействует с узлами сети. С другой стороны, при решении единой задачи на разных узлах кластера возникает необходимость обмена данными между узлами, который осуществляется через сеть коммуникаций. Физически сети управления и коммуникаций могут быть одной локальной сетью. В случае решения задач, связанных с обработкой больших объемов данных, для достижения максимальной интенсивности функции ввода/вывода данных могут быть переданы специальной сети ввода/вывода. В простом варианте исходные данные и результаты расчетов размещаются на едином файловом сервере, роль которого может выполнять более мощная управляющая машина. Более сложным, но иногда и более эффективным может быть использование для хранения данных локальных жестких дисков узлов кластера.

Принимая во внимание параметр цена/производительность, в качестве

сетевой технологии в кластере предполагается реализовать 1000-мегабитный вариант сети Ethernet — Gigabit Ethernet [13]. Для этого будет использован коммутатор (switch) D-Link с максимальной пропускной способностью до 1 Гб/с, узлы вычислительного кластера будут оснащены гигабитными сетевыми адаптерами. Принципиальная схема модели сегмента кластера показана на рис. 1.

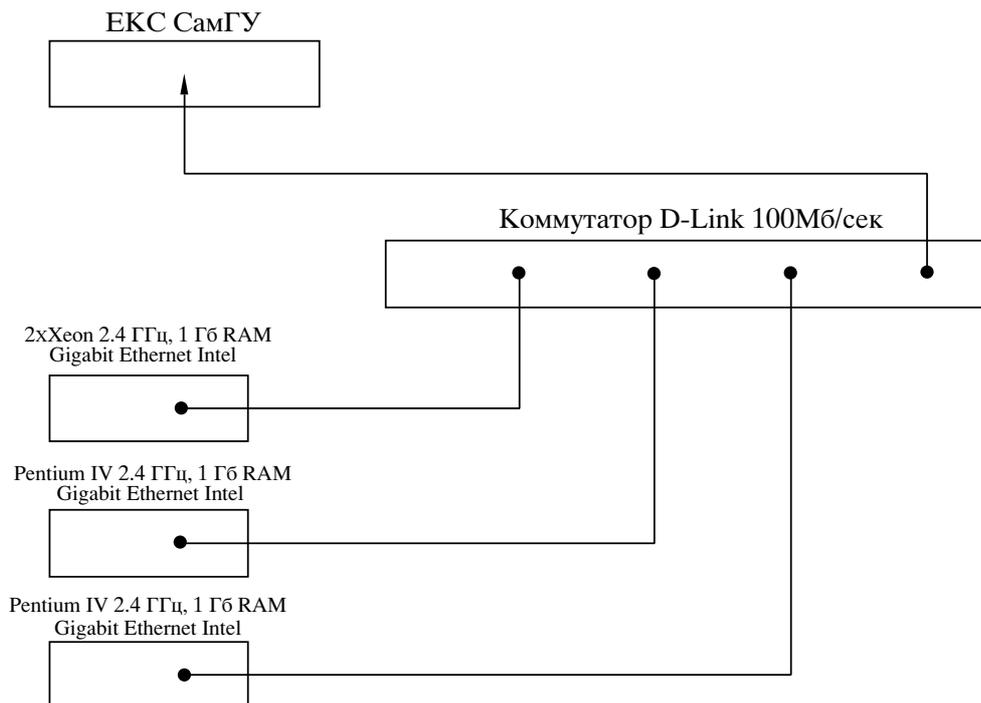


Рис. 1. Существующая модель сегмента кластера

3. Программное обеспечение

Набор необходимого программного обеспечения определяется задачами кластера: устойчивый многопользовательский и многозадачный режим работы, поддержка технологий параллельного программирования. Этим требованиям удовлетворяют операционные системы (ОС) семейства UNIX [14]. Учитывая мировой и российский опыт создания вычислительных кластеров, в качестве ОС университетского кластера предполагается ОС Linux, один из наиболее известных и свободно распространяемых диалектов ОС UNIX. В частности, уже созданный тестовый сегмент кластера, состоящий из управляющей машины (Intel Xeon) и трех узлов (Pentium IV), работает под управлением ОС Debian Linux [15], надежно зарекомендовавшей себя в работе с сетевыми приложениями.

В настоящее время наиболее распространенной технологией параллельного программирования в системах с распределенной памятью является стандарт MPI — Message Passing Interface [6], который поддерживает ин-

терфейс передачи сообщений при взаимодействии параллельных процессов внутри одной программы. В рамках MPI-программы каждый из параллельных процессов работает в своем адресном пространстве, при этом общие переменные или данные отсутствуют и посылка сообщений является основным способом взаимодействия между процессами. Библиотека MPI-интерфейса MPICH [16] стандарта MPI-1.1, реализуемая в ПО тестового сегмента кластера, имеет более 120 функций, которые поддерживают работу с расширениями языков последовательного программирования высокого уровня C, C++, Fortran 77(90) и работу загрузчика приложений.

В перспективе аппаратное и программное обеспечение вычислительного кластера позволяет использовать более современные подходы к разработке параллельных программ (например, DVM-модель, созданную в Институте прикладной математики им. М.В.Келдыша РАН [17]).

4. Реализация проекта

К настоящему времени удалось довести до рабочего состояния модель сегмента кластера. В работе по его созданию принимали участие сотрудники кафедры механики сплошных сред, кафедры общей и теоретической физики, кафедры радиофизики и компьютерного моделирования радиосистем. На рис. 1 показана принципиальная схема уже собранного сегмента вычислительного кластера. На управляющую машину (двухпроцессорная рабочая станция Intel Xeon) и рабочие станции (Pentium IV) установлена ОС Debian Linux, произведены конфигурирование системы и настройка компонент, установлены библиотеки интерфейса обработки параллельных программ MPI. Тестирование сегмента кластера показало его работоспособность как распределенной многопроцессорной системы.

Наиболее оптимальной, однако, может быть конфигурация кластера, показанная на рис. 2. Замена однопроцессорных Pentium IV на двухпроцессорные рабочие станции Intel Xeon, физическое разделение сети управления (Fast Ethernet, 100 Мбит) и более быстрой сети коммуникаций между узлами кластера (Gigabit Ethernet, 1000 Мбит) и возможность увеличения количества узлов кластера до 8 с числом процессоров 16 позволит, по крайней мере, на порядок увеличить быстродействие кластера при решении вычислительных задач по сравнению со стандартным персональным компьютером Pentium IV. Такая конфигурация позволит использовать кластер для учебных целей и решения вычислительных научных задач. Доступ к ресурсам кластера может быть осуществлен с любой рабочей станции ЕКС университета или сторонних организаций.

Сравнение параметров оптимальной конфигурации кластера (рис. 2) с параметрами существующих в России кластерных вычислительных систем [4] показывает, что университетский кластер будет находиться в одном ряду с аналогичными установками в ведущих вузах (МГУ, МФТИ, СПбГУ) и научно-исследовательских институтах Российской Федерации.

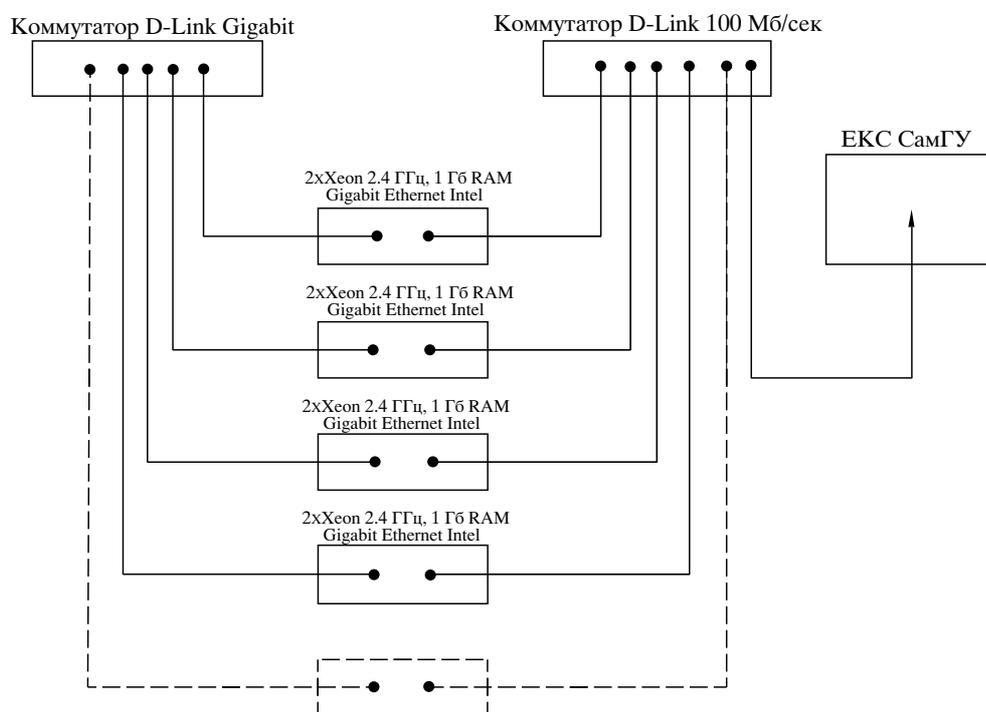


Рис. 2. Оптимальная модель общеуниверситетского кластера

В заключение заметим, что в мире резко возрос интерес к проблематике глобальных вычислительных сетей GRID [18], элементарными ячейками которых выступают отдельные кластерные системы, и широко распространяется понимание того, что внедрение таких сетей будет иметь громадное влияние на развитие человечества и обеспечит единство информационного поля ноосферы [19].

Литература

- [1] Яровой Г.П., Кабытов П.С., Гарькин В.П., Горелов Ю.Н., Симатова В.Н. Концепция развития Самарского государственного университета как центра науки, культуры и образования Поволжского региона (Новая редакция) // Вестник Самарского гос. ун-та. Гуманитарная серия. 2002. №1. С. 135-141.
- [2] Указы Президента РФ / Пр.-577, Пр.-578 от 30 марта 2002 г.
- [3] Латис А.О. Как построить и использовать суперкомпьютер. М: Бестселлер, 2003. 230 с.
- [4] Информационный Интернет-портал <http://parallel.ru>
- [5] Интернет сайт <http://www.linux.org>
- [6] Message Passing Interface Forum, Document for Standard MPI, 1993. Version 1.0 / <http://www.unix.mcs.anl.gov/mpi/>

- [7] Воеводин В.В., Воеводин Вл.В. Параллельные вычисления. СПб: БВХ-Санкт-Петербург, 2002. 609 с.
- [8] Неганов В.А., Раевский С.Б., Яровой Г.П. Линейная макроскопическая электродинамика. Т. 2. М.: Радио и связь, 2001; Зайцев В.В., Панин Д.Н., Яровой Г.П. Численный анализ отражений от слоя неоднородной плазмы // Физика волновых процессов и радиотехнические системы. 2000. Т.3. №1.
- [9] Онучак Л.А. и др. / Научный отчет по гранту РФФИ 01-03-32587. Самара, 2003.
- [10] Астафьев В.И., Радаев Ю.Н., Степанова Л.В. Нелинейная механика разрушения. Самара: Изд-во "Самарский университет", 2001. 632 с.
- [11] Салеев В.А., Васин Д.В. / Научный отчет по гранту "Интеграция", 30056/1535. Самара, 2002; Салеев В.А., Гюнтер Д.А. / Научный отчет по гранту "Интеграция", 33138/1456. Самара, 2003.
- [12] Макеенко Ю.Н. Метод Монте-Карло в калибровочных теориях на решетке // УФН. 1984. Т. 143. Вып. 2. 161 с.
- [13] Олифер В.Г., Олифер Н.А. Компьютерные сети. Принципы, технологии, протоколы. СПб: Питер, 2003.
- [14] Робачевский А. Операционная система UNIX. СПб: БВХ-Санкт-Петербург, 2000. 528 с.
- [15] Интернет сайт <http://debian.org>
- [16] Интернет сайт <http://www.mcs.anl.gov/mpi/mpich/>
- [17] DVM-система / <http://www.keldysh.ru/dvm/>
- [18] Foster I., Kesselman C. The GRID: Blueprint for a New Computing Infrastructure. San Francisco, CS, USA: Morgan Kaufmann Publishers, Inc., 1999.
- [19] Крюков В.А. Разработка параллельных программ для вычислительных кластеров и сетей // Информационные технологии и вычислительные системы, 2004 (принята к печати); <http://www.keldysh.ru/dvm/>.

Поступила в редакцию 18/XI/2003;
в окончательном варианте — 19/XII/2003.

PARALLEL INFORMATION TECHNOLOGY BASED ON A COMPUTING CLUSTER

© 2003 G.P. Yarovoy,⁶ Y.N. Radayev,⁷ Y.A. Rodichev,⁸ V.A. Saleev⁹

In the paper a prospective project on the institution of Center of High Performance Computing in Samara State University is presented. Cluster based computational technologies are discussed in an attempt to fulfill needs in high performance computing and storage resources caused by fundamental and research work carried out by a number of Departments of Samara State University. A model of computing cluster developed so far is described in further technical details. A concept of organizing parallel programming studies in Samara State University is considered.

Paper received 18/*XI*/2003;

Paper accepted 19/*XII*/2003.

⁶Yarovoy Gennadiy Petrovich (rector@ssu.samara.ru), Dept. of Radiophysics and Computer Modelling of Radiosystems, Samara State University, Samara, 443011, Russia.

⁷Radayev Yuri Nickolaevich (radayev@ssu.samara.ru), Dept. of Continuum Mechanics, Samara State University.

⁸Rodichev Yuri Andreevich (rodichev@ssu.samara.ru), Center of Computational Science and Technology, Samara State University.

⁹Saleev Vladimir Anatoljevich (saleev@ssu.samara.ru), Dept. of General and Theoretical Physics, Samara State University.