

# ПРОГРАММИРОВАНИЕ СТРУКТУРНО СЛОЖНЫХ РАСПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ И ПРОЦЕССОВ УПРАВЛЕНИЯ В НЕДЕТЕРМИНИРОВАННОЙ КОМПЬЮТЕРНОЙ СРЕДЕ<sup>1</sup>

Ю.С. Затуливетер

Глобальная Сеть охватывает всё большее число сфер жизнедеятельности. Благодаря опережающему развитию массовых компьютерных и сетевых технологий число решаемых в ней задач быстро растёт. Они всё полнее отражают суть обязательных проблем жизнеобеспечения – сбалансированное функционирование и устойчивое развитие социосистемы. Однако системные возможности Сети, весьма слабо организованной в части управления распределёнными вычислениями, со всё большей очевидностью отстают от её количественного расширения. Всё заметнее и жёстче в последствиях системный дисбаланс между стихийным ростом потоков и объёмов глобально распределённой информации и явно отстающими от требований времени возможностями их своевременной, с необходимой степенью полноты, переработки [5].

Многоплановая интрига отношений между задачами и компьютерными технологиями их решения – одна из главных сюжетных линий компьютерного прогресса. Сегодня, в условиях глобального информационного пространства, – это уже и кибернетические факторы зарождения и роста новейших тенденций прогресса социального [3].

## ЗАДАЧИ "БОЛЬШИЕ" И "МАЛЫЕ"

Компьютерные задачи и аппаратно- программные средства их решения неразрывно связаны. Прогресс в постановке и применении задач стимулирует развитие средств их решения. Верно и обратное: успехи технологий в развитии аппаратных или программных средств расширяют границы понимания и возможности решения новых задач.

Все компьютерные задачи можно условно расположить между двумя полюсами – "большими" задачами и "малыми".

Уровень требований больших задач к средствам решения находится либо на пределе текущих возможностей, либо превышает их. Такие задачи, как правило, уникальны. Они требуют особых усилий и дорогостоящих решений. Наряду с достижением поставленных целей, они стимулируют развитие имеющихся технологий производства компьютерных компонентов или появление новых.

Для решения больших задач нужен инструмент. Когда задач становится относительно много и появляются временные рамки достижения результатов, необходим универсальный инструмент их исчерпывающего решения. Первым таким инструментом 60 лет назад стал универсально программируемый компьютер, воплотивший модель Дж. фон Неймана – классическую компьютерную аксиоматику универсального автоматического счёта. Она получила идеальное воплощение в микропроцессорной архитектуре, стала основой компьютерной революции и продолжает оставаться единым логическим стандартом и системным ядром, соединяющим воедино компьютеры и программы в условиях массового производства.

Развитие больших задач и инструментария их решения сначала шло в применении к трудоемким в вычислительном отношении алгоритмам решения особо важных научно-технических задач. От компьютеров к высокопроизводительным суперкомпьютерам, к локальным сетям. Компьютеры обеспечивали повышение вычислительной производительности за счет увеличения рабочих частот и внутреннего распараллеливания архитектур. Сети приносили внешнюю компоненту пространственного параллелизма, обеспечивая информационную привязку вычислительной среды к распределённой в пространстве окружающей среде. Это – и физическая (материальная), и человеческая (информационная) среды.

Малые задачи стали играть заметную роль и оказывать стимулирующее влияние на развитие технологий по мере освоения массовых применений, ориентированных на обслуживание потребностей миллионов пользователей, многочисленных малых и средних организаций/ предприятий. Развитие малых задач стимулирует технологии к снижению себестоимости компьютерных и программных средств, к повышению их доступности, улучшению эксплуатационных характеристик. Примеры: персональные компьютеры (ПК), ориентированные на массового пользователя ОС, локальные сети.

Решение в ранних сетях локальных задач обозначили стратегические направления на системную интеграцию распределенных машинных (вычислительных) и человеческих (интеллектуальных) ресурсов. Далее – от локальных сетей к глобальным, к единому информационному пространству. Технологии ответили новыми поколениями компьютерного/ сетевого оборудования и программных средств, обеспечившими освоение новых сфер массового применения. Примеры: Web-технологии, сети мобильной связи и коммуникаций, встраиваемая электроника и др.

<sup>1</sup> Работа поддержана Российским фондом фундаментальных исследований (РФФИ), проект 08-07-00334.

Разнообразие траекторий развития компьютерного прогресса определяется во взаимодействиях уникальных "больших" задач с массовостью и неограниченным разнообразием "малых". В ходе компьютерного прогресса наиболее удачные средства решения больших задач становятся достоянием малых. Миллионы пользователей ПК – общедоступного инструмента решения малых задач – с лагом запаздывания в 10-15 лет становятся обладателями ранее уникальных суперкомпьютерных возможностей. На освобождающие "вакансии" рекордных задач приходят новые постановки. Поначалу критерием "большой" задачи была вычислительная сложность, определяющаяся объемом арифметических действий. По мере освоения массовых, в том числе сетевых, приложений, вовлекающих информационную активность большого числа людей, наряду с вычислительной всё большую роль играет структурная сложность. Людям свойственно воспринимать и отражать мир через сложные, динамически развивающиеся информационные структуры.

С установлением глобального информационного пространства масштабы понятия "большие задачи" стали совершенно иными. По количественным, качественным и социальным показателям. Количество и разнообразие "малых задач" в Сети растут лавинообразно. Они решаются посредством разнородных средств, несовместимых ни по базовым моделям, ни по формам представления компьютерной информации (программ и данных). Это крайне затрудняет интеграцию существующих решений. В отсутствие единой и универсальной модели распределённых вычислений сложность системных проблем интеграции в неоднородном информационно- вычислительном пространстве с увеличением размеров и числа задач растет комбинаторно. "Проклятие" размерности служит непреодолимым барьером на путях интеграции любых совокупностей малых задач в масштабах всего глобального информационного пространства.

Таким образом, стихийный количественный рост всей совокупности разнородных постановок и решений малых задач создаёт на общесистемном уровне принципиально новую "большую" задачу. Она состоит в необходимости создания единого системного ядра, способного стать универсальной и общедоступной логической основой для глобально распределённых вычислений в условиях недетерминированной компьютерной среды.

#### НОВЫЕ КЛАССЫ ЗАДАЧ

Глобализация информационного пространства кардинально меняет условия функционирования и развития социосистемы и её частей. При этом возникают исторически беспрецедентные задачи управления и самоорганизации социально значимых процессов [3].

По ходу вовлечения в пространство Сети различных сфер жизнедеятельности сопутствующая глобально распределённая компьютерная информация де-факто, с растущей мерой представительности, отражает текущее состояние социосистемы в целом и её частей. Открывающийся в компьютерной среде полный (теоретически) доступ к текущему состоянию динамических систем открывает возможности построения и исследования математических моделей с гарантированным качеством управления. Воплощение этих моделей в глобальной компьютерной среде, с её совокупным (практически неограниченным) функциональным и вычислительным потенциалом открывает принципиально новые перспективы повышения качества управления процессами функционирования и развития социосистемы. Совокупный вычислительный потенциал всех связанных сетями компьютеров открывает перспективы для своевременной и полной переработки циркулирующих в Сети потоков информации и выработки (с учетом текущего состояния социосистемы) управляющих воздействий, в целях повышения эффективности функционирования социосистемы и обеспечение её устойчивого развития [4].

Сверхбольшие потоки и объёмы информации, циркулирующей в Сети, уже существенно превышают пропускную способность сложившихся в прежние времена социальных механизмов управления и самоорганизации. Избыточные, своевременно не перерабатываемые потоки и объёмы, ведут к ускоренному воспроизводству информационного шума, снижающего наблюдаемость и устойчивость социосистемы и её частей, действенность привычных процессов управления. Здесь, в глобальном информационном пространстве, скрываются пока ещё мало изученные причины опережающего роста рисков утраты управляемости экономическими, техногенными и другими социально значимыми процессами [5].

Для выправления ситуации требуются новые подходы к управлению социосистемой в условиях единого информационного пространства, основанные на решении задач компьютерной глобализации парадигмы управления [4]. Она направлена на системно целостную реализацию во всей совокупности сетевых вычислительных ресурсов действий полного цикла управления: съём, первичная обработка и сбор входной информации, доставка к местам накопления, переработка в соответствии с моделями управления (с учетом внутреннего состояния), выработка управляющих воздействий и их доставка к местам исполнения, переход в ожидание реакции управляемой системы.

В результате компьютерная среда в целом должна обрести единые общесистемные свойства и универсальные механизмы, которые соединят распределённые вычислительные ресурсы любой совокупности связанных компьютеров в целостный, универсально программируемый агрегат. Он должен надёжно функционировать в условиях существенной недетерминированности глобальной компьютерной среды и предо-

ставлять регулярные, универсально программируемые и массово доступные средства для решения непрерывного спектра взаимосвязанных задач управления социально значимыми процессами.

#### ОСОБЕННОСТИ НОВЫХ ЗАДАЧ

Глобальная компьютерная среда охватывает всё больше распределённых объектов, функционирующих как составные части социосистемы и нуждающихся в эффективном управлении. Вовлекая в процессы управления и самоорганизации многие миллионы людей, она меняет условия функционирования систем управления, открывает дополнительные возможности для их массового воплощения.

Для систем с массовыми каналами интерактивного взаимодействия характерна высокая структурная и динамическая сложность данных и программ. Такие системы должны оперировать не только со скалярами или регулярными структурами типа вектор или матрица, но в большей степени со структурированными данными общего вида, представляемые произвольно изменяющимися графами. Системы такого класса будем называть графодинамическими. Суть графодинамических систем в том, что их текущие значения входов, выходов и внутреннего состояния представляются посредством динамических структур общего вида. В ходе функционирования систем преобразования графов могут приводить к изменению состава вершин, связей между вершинами и значений с ними связанных. Этот класс динамических систем был впервые введен в рассмотрение в [1]. Для таких систем требуются новые принципы и методы формализации объектов, постановки и решения задач управления ими.

Глобальная Сеть – главный стимул создания и широкого распространения графодинамических систем [4], поскольку охватывает всё больше применений, связанных с глобально распределёнными объектами, функционирующими в условиях непрерывного и массового производства и использования сложно организованной информации с быстро меняющимися структурами. По разнообразию и объёмам применений потенциальный спрос на такие системы практически не ограничен.

Отсутствие регулярных методов постановки и формализации графодинамических задач, а также средств их решения посредством глобально распределённых структурно-сложных вычислений делает это новое направление исследований весьма актуальным.

#### ТРЕБОВАНИЯ К НОВЫМ ЗАДАЧАМ И СРЕДСТВАМ ИХ РЕШЕНИЯ

Глобальная Сеть приносит качественно новые приоритеты и требования к масштабам постановки компьютерных задач, методам, способам и средствам их решения. Накопление и развитие глобально распределённой информации создаёт единый для решения всех задач контекст. На смену прежнему принципу презумпции локальности решаемых задач, который предполагает изначальную независимость или слабую связность задач и решений, приходит принцип презумпции их сильной связности/зависимости [4]. Реализованная в Сети связность каждой порции информации (независимо от её расположения) предполагает потенциальное наличие функциональной зависимости между ними. Все задачи и решения (в том числе собственно программы, процессы, значения данных) в Сети составляют глобальный, динамически меняющийся информационный контекст. Потенциально каждая задача может изменить этот контекст и, одновременно, сама зависит от него. Это значит, что постановки и решения таких задач должны производиться и корректироваться с учётом непрерывно меняющегося глобального информационного контекста в соответствии с принципом "всё зависит от всего".

Глобальная Сеть обладает полной связностью по данным. Пересылка информации системно "прозрачна" (например, в протоколе ТСП/IP). Но этого не достаточно для глобализации свойства универсальности, замкнутого во внутренних ресурсах компьютеров. Для распространения свойства универсальности с внутрикомпьютерных ресурсов на любую совокупность компьютеров необходима глобальная связность не только по данным, но и по управлению вычислительными действиями в распределённых процессах (с предельной глубиной управления – вплоть до элементарных действий).

Чрезвычайная разнородность форм представления и способов обработки глобально распределённой информации остаётся одной из главных причин её недоступности для разработки и полномасштабного применения общих методов математического и компьютерного решения проблем в едином, математически однородном формализме.

Для формально строгого отображения текущего состояния решаемых в компьютерной среде задач глобально распределённая компьютерная информация должна обрести единые и универсальность формы своего представления, обладающие следующими свойствами:

- математическая однородность и структурная целостность в представлении данных и программ;
- функциональная полнота в исполнении программ;
- инвариантность относительно машинной среды (машино-независимость);
- свободная масштабируемость данных, программ, процессов и систем обработки/ управления в условиях глобальной компьютерной среды.

## ФУНДАМЕНТАЛЬНЫЕ АСПЕКТЫ ПРОБЛЕМ ГЛОБАЛИЗАЦИИ УНИВЕРСАЛЬНЫХ ВЫЧИСЛЕНИЙ

Совокупный функциональный и вычислительный потенциал миллиардов универсально программируемых компьютерных устройств, связанных сетями, – стационарных, мобильных, встроенных – практически неисчерпаем. Однако общедоступное использование программистами глобально распределённого компьютерного интеллекта остаётся невозможным. В отличие от каждого универсального компьютера, из которых составлены глобальные сети, вся совокупность или любое подмножество компьютеров не образуют детерминированный, универсально программируемый, системно целостный и технологически замкнутый вычислительный агрегат. Причины носят общесистемный характер.

В работе выделены и рассмотрены два фундаментальных аспекта проблемы:

- отсутствие единой замкнутой модели структурно целостного и функционально полного программирования и исполнения распределённых вычислений, действие которой на уровне универсального логического базиса можно распространить на любые подмножества связанных сетями универсальных компьютеров;
- существенная и, в общем случае, неустранимая недетерминированность совокупных вычислительных ресурсов глобальных сетей.

В основе индустрии компьютеров и программ лежит универсальный логический стандарт вычислений, представленный классической компьютерной аксиоматикой (модель Дж. фон Неймана). Однако её главное системообразующее свойство (универсальной программируемости вычислений) изначально замкнуто во внутренних ресурсах компьютеров [6,4]. В работе представлена коррекция модели Дж. фон Неймана, которая позволяет распространить свойство универсальной программируемости на структурно сложные распределённые вычисления в любой совокупности сетевых вычислительных ресурсов [6,4].

В новой аксиоматике математически однородного поля компьютерной информации коррекция осуществлена посредством математически замкнутой модели исчисления древовидных структур, которое реализовано в языке и системе программирования ПАРСЕК [8,9]. В рамках этой модели открываются перспективы общедоступного решения (программирования) задач в сколь угодно большой совокупности связанных компьютеров.

Логическое расширение вычислительного пространства на неограниченное число связанных компьютеров требует рассмотрения задачи организации надёжных вычислений в изначально ненадёжной компьютерной среде [7]. В работе формулируются математические задачи существенного повышения надёжности распределённых вычислений и процессов управления в недетерминированной компьютерной среде посредством внесения в исполняемые процессы относительно малой вычислительной избыточности.

### К ЕДИНОЙ МОДЕЛИ ГЛОБАЛЬНО РАСПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ

Классическая компьютерная аксиоматика в модели Дж. фон Неймана постулирует простые правила, которые обеспечивают универсальность машинных вычислений во внутренних ресурсах изолированных компьютеров. Действие этой модели не распространяется универсальным образом на сетевые ресурсы, что позволяет утверждать, что системообразующий потенциал классической аксиоматики исчерпан.

Общепризнанной универсальной модели, которая бы обобщила классическую аксиоматику и распространила свойство универсальной программируемости с внутренних ресурсов на любую совокупность связанных сетями компьютеров, пока нет [6,4,5]. Нет и конструктивного понимания какой должна быть эта модель. Этим объясняется отсутствие *универсальных* и *простых в использовании* инструментов решения "больших" задач обработки глобально распределённой информации в компьютерной среде. Поэтому совокупные вычислительные ресурсы глобальных сетей не составляют системно целостный универсально программируемый вычислительный агрегат. Глобальная Сеть в целом (как система) продолжает оставаться "глупее" каждого из компьютеров, из которых она составлена.

В отсутствие общей модели (единой логической основы универсально программируемых распределённых вычислений) для построения проблемно-ориентированных систем обработки глобально распределённой информации применяются технологии, обеспечивающие построение частных решений. Они нацелены на большие задачи принципиально ограниченные корпоративными целями. Как правило, это проблемно-ориентированные системы обработки глобально-распределённой информации, циркулирующей в замкнутых сообществах. Профессионально-производственные сообщества используют Grid-технологии. В других случаях широко используются пиринговые файлообменные системы.

Grid-технологии [2] лояльны классической аксиоматике и строятся как многоуровневые надстройки над стандартными сетевыми протоколами и операционными системами. Они обеспечивают управление глобально распределёнными вычислительными ресурсами с целью комплексирования для решения строго очерченных корпоративных задач. В результате получают громоздкие многослойные программные комплексы с большим числом степеней свободы, управление которыми требует сложного администрирования на системных и сетевых уровнях.

Grid-инструментарий создается по сильно различающимся моделям глобально распределенных вычислений. Развивается стихийно как хаотическая, неконтролируемая совокупность разнородных, трудновосместимых технологий, поэтому не обладает свойством системной целостности. Отсюда чрезмерная трудоемкость и себестоимость. Grid-технологии применяются, как правило, в рамках больших корпоративных проектов или особо важных применений, для которых можно аккумулировать требуемые финансовые и интеллектуальные ресурсы.

Масштабы корпоративных систем заведомо ограничены сверху из-за комбинаторной сложности проблем преодоления принципиально неустранимого информационного шума, который непрерывно производится в виде избыточного разнообразия форм представления и способов обработки компьютерной информации [4,6].

В [4,6] показано, что причины непрерывного воспроизводства информационного шума скрыты в постулатах универсальных вычислений модели Дж. фон Неймана. Для устранения причин информационного шума автором предложена коррекция классической компьютерной аксиоматики, осуществлённая посредством формализма исчисления древовидных структур [4,6]. Обобщая модель Дж. фон Неймана, данное исчисление поднимает степень абстракции правил универсальных вычислений на уровень математически замкнутого формализма, логически несвязанного с машинными вычислительными ресурсами. Правила универсальных вычислений определяются на древовидных структурах, представляющих как данные, так и программы. Машино-независимость правил позволяет распространить свойства универсальной программируемости с внутренних ресурсов компьютеров на любую совокупность компьютеров связанных сетями.

Универсальный язык и система программирования ПАРСЕК [8,9] дают простейшую (процедурную) реализацию исчисления древовидных структур. Программирование осуществляется посредством математически полного набора операций с древовидными и только древовидными структурами, которые представляют как данные так и программы. В ходе решения задач, программист формирует необходимые для решения задач деревья и функции их преобразования, придает им смысловую интерпретацию.

Язык и система программирования ПАРСЕК используются для программирования структурно-сложных распределенных вычислений, исполняемых в глобальной компьютерной среде. Кроме того, система ПАРСЕК предоставляет компьютерный формализм для реализации идей графодинамики [1] и в своем развитии может быть использована для программирования графодинамических систем с управлением, реализуемых в глобальной компьютерной среде.

#### НАДЕЖНЫЕ ВЫЧИСЛЕНИЯ В НЕНАДЁЖНОЙ КОМПЬЮТЕРНОЙ СРЕДЕ

Ненадёжность компьютерной среды объясняется принципиально неустранимой недетерминированностью вычислительных ресурсов глобальных сетей. Её причины можно разделить на два класса – внутренние и внешние.

Внутренние связаны с технической ненадёжностью вычислительного и сетевого оборудования. Внешние связаны с ситуационными факторами, которые состоят:

- в принципиальной недоступности отдаленных вычислительных ресурсов для применения дополнительных аппаратных или программных средств повышения их надежности;
- в непрерывном и случайным образом меняющемся составе и конфигурации связей работоспособных компьютеров;
- в стихийных или преднамеренных деструктивных воздействиях (физического или информационного характера) на вовлечённые в распределенные вычисления сетевые ресурсы.

Проблемы организации универсально программируемых распределенных вычислений с требуемым уровнем надёжности исполнения в недетерминированной компьютерной среде до настоящего времени мало исследованы. Для их решения предлагается подход, основанный на новой компьютерной аксиоматике в формализме исчисления древовидных структур [6,4]. Этот подход дополняется программными методами повышения надёжности распределенных вычислений и процессов управления до требуемого уровня в условиях непрерывно меняющегося состояния работоспособности ресурсов глобальной сети. Повышение надёжности обеспечивается за счёт введения на прикладном программном уровне вычислительной избыточности, которая осуществляется посредством параллельного исполнения на разных компьютерах идентичных вычислительных фрагментов с последующим сравнением в каждом из них промежуточных результатов. При этом не требуется каких-либо специальных или дополнительных средств повышения надёжности собственно сетевых вычислительных ресурсов.

Предлагаемый подход включает теоретическое обоснование решений, которое осуществляется через построение и исследование вероятностных и ситуационных моделей, учитывающих внутренние и внешние факторы ненадёжности, соответственно.

Предварительные оценки для внутренних факторов показывают, что надёжность распределённых вычислений растёт экспоненциально от увеличения степени вычислительной избыточности. Так, при надёж-

ности компьютеров равной 0.999 введение порядка 100-кратной вычислительной избыточности даёт повышение надёжности вычислительного процесса до 200 девяток после запятой [7].

Ситуационные модели направлены на повышение надёжности за счет динамической реконфигурации распределённых вычислений, направленной на минимизацию рисков полной утраты возможностей исполнения вычислительных или управляющих процессов.

#### ЛИТЕРАТУРА:

1. М.А. Айзерман, Л.А. Гусев, И.М. Смирнова, С.В. Петров. Динамический подход к анализу структур, описываемых графами (Основы графодинамики) // Автоматика и Телемеханика, 1977, №7 (С.135-151), №9 (С.123-136).
2. Д. Корягин, В. Коваленко. Организация grid: есть ли альтернативы? // Открытые системы. 2004. №12. С.34-41. <http://www.osp.ru/os/2004/12/184888/>.
3. Ю.С. Затуливетер. Информационная природа социальных перемен. М.: СИНТЕГ. Серия "Информация и Социум", 2001. 132 с.
4. Ю.С. Затуливетер. Проблемы глобализации парадигмы управления в математически однородном поле компьютерной информации // Проблемы управления. 2005. №1. Ч.І. С.1-12. №2. Ч.ІІ. С.13-23. (<http://zvt.hotbox.ru>).
5. Ю.С. Затуливетер. Проблемы управляемого перехода к новейшим инновационным укладам // Труды VII международной конференции "Идентификация систем и задачи управления", SICPRO`08, Москва 28-31 января 2008 года. Институт проблем управления им. В.А.Трапезникова РАН, г.Москва. С. 1988 - 2039.
6. Ю. Затуливетер. На пути к глобальному программированию // Открытые системы. 2003. № 3. (<http://www.osp.ru/os/2003/03/182704/>).
7. Ю.С. Затуливетер, В.В. Карибский, Н.В. Лубков. Проблема организации надежных процессов управления с применением ненадежных вычислительных сред // Тезисы докладов. Четвертая международная конференция "Проблемы управления в чрезвычайных ситуациях" г. Москва, 11 января, 1997. М.: Институт проблем управления РАН, 1997. С.160-163. ([http://zvt.hotbox.ru/bezop\\_96.htm](http://zvt.hotbox.ru/bezop_96.htm)).
8. Ю.С. Затуливетер, Т.Г. Халатян. ПАРСЕК - язык компьютерного исчисления древовидных структур с открытой интерпретацией. Стеновый вариант системы программирования. М.: Институт проблем управления РАН, 1997. 71с.
9. Ю.С. Затуливетер, А.В. Топорищев. Язык Парсек: программирование глобально распределенных вычислений в модели исчисления древовидных структур // Проблемы управления. 2005. №4. С. 12-20.