

Добровольные распределенные вычисления в России: социологический анализ

В.Н. Якимец, И.И. Курочкин
Институт проблем передачи информации РАН
iakim@isa.ru, kurochkin@iitp.ru

Аннотация

В докладе изложены результаты первого в России выполненного в 2014 году социологического исследования сообщества краучеров страны, вовлеченных в добровольные распределенные вычисления. Описана организация опроса и приведены сведения о возможностях, предпочтениях и пожеланиях краучеров.

Ключевые слова: добровольные распределенные вычисления, краучеры, социологический опрос.

В рамках распределенных вычислений, представляющих собой способ решения трудоемких вычислительных задач с использованием компьютеров, объединенных в вычислительную систему, особый интерес представляют добровольные вычисления (volunteer computing). Это распределенные вычисления с использованием добровольно предоставленных вычислительных ресурсов. Известно, что современные вычислительные системы для добровольных вычислений строятся на базе грид-систем. Если не вдаваться в математические, технологические и информационные аспекты таких вычислений, а обратиться к термину «добровольные», то, если угодно, мы сталкиваемся с малоизученной стороной российского гражданского общества. Сообщество специалистов и различных заинтересованных сторон, вовлеченных в развитие и применение технологий распределенных информационно-вычислительных сред, а также в выполнение добровольных вычислений – это особый и своеобразный фрагмент гражданского общества.

Приведем примеры добровольных вычислений из разных областей знания. Так, в рамках естественных наук – это моделирование работы адиабатического квантового сверхпроводящего компьютера (AQUA@home), или проект в области астрономии, целью которого является определение формы и параметров вращения астероидов по данным фотометрических наблюдений (Asteroids@home), или проект по моделированию влияния выбросов углекислого газа на климат Земли (Climate Prediction); решение вычислительно сложных аэрокосмических задач (оптимизация траекторий спутников и зондов, моделирование экзосферы Луны (Constellation); моделирование поведения заряженных частиц в магнитном поле термоядерного реактора ITER

(EDGeS@Home); проект по обработке сигналов радиотелескопа для поиска радиосигналов внеземных цивилизаций (SETI@Home).

В области биологии и медицины можно назвать проект университета Вандербильда, основными целями которого являются выяснение структуры мембранных белков (мишени для лекарств), проектирование белков с новой структурой и функциями, понимание количественных отношений между химической структурой и биологической активностью (синтез лекарств) (Biochemical Library), проект, целью которого является исследование корреляций между последовательностями в трехмерной структуре генома (Correlizer); российский проект, связанный с вычислением структуры белка, скринингом баз данных биологически активных соединений (DrugDiscovery@Home); проект по поиску лекарства от рака (до апреля 2007 года реализовывался grid.org), а затем был передан в сеть World Community Grid.

Участников процесса добровольных распределенных вычислений называют кранчерами (crunchers).

Нас заинтересовал социологический срез российского сообщества кранчеров. Детального социологического исследования российских кранчеров найти не удалось.

В тоже время качественная реализация исследований и разработок в данной сфере определяется и тем, сколь эффективно работают коллективы ученых и специалистов (как правило, междисциплинарные), и тем, как участвуют в создании, внедрении и поддержке технологий распределенных информационно-вычислительных сред иные заинтересованные стороны и лица, представляющие интернет-сообщество, составляющее на сегодняшний день своеобразный срез гражданского общества.

В мире было выполнено два социологических опроса [1]:

- анкетирование, которое проводили создатели платформы BOINC (Калифорнийский университет) в 2006 году (более 30 тыс. респондентов, включая около 500 человек из России и Украины),
- опрос 2013 года, организованный IBM среди участников проекта World Community Grid (WCG), в котором было задействовано более 15 тыс. респондентов, в том числе около 300 русскоязычных.

В 2014 году на базе Центра распределенных вычислений Института проблем передачи информации РАН мы организовали и провели первое социологическое исследование мотиваций и предпочтений участников добровольных распределенных вычислений (далее ДРВ) в России [2, с.8; 3].

Рассмотрим некоторые результаты исследования.

Для опроса предпочтений российских кранчеров нами была разработана специальная анкета.

В ней содержалось 26 вопросов, разбитых по следующим разделам:

- опыт добровольных распределенных вычислений;
- проекты и интересы;
- национальные проекты;
- общие данные о кранчере;
- предпочтения и пожелания кранчеров в отношении проектов ДРВ;
- идентификация (по желанию).

Каждый из 15 вопросов анкеты был полужакрытым, чтобы помимо предложенных вариантов ответа у респондента была возможность вписать свой вариант. Опрос с использованием этой анкеты проводился в середине 2014 года.

Информационное обеспечение и привлечение волонтеров-кранчеров к участию в анкетировании было обеспечено администраторами сайта BOINC.RU. Им удалось задействовать практически все существующие русскоязычные интернет-ресурсы, посвященные ДРВ, и реализовать следующие действия:

- распространить объявление на сайте и озвучить тему и цель исследования на форуме BOINC.RU;
- сделать почтовую рассылку среди почти 800 зарегистрированных участников форума;
- разместить информацию в тематических форумах на сайтах: astronomy.ru, overclockers.ru, distributed.ru, distributed.org.ua, в группе социальной сети «ВКонтакте»;
- осуществить почтовую рассылку по списку адресов участников известных российских команд Russia Team и Crystal Dream (более 3000 адресов).

Этими действиями была обеспечена высокая репрезентативность результатов опроса.

В России активно участвуют в ДРВ около 4000 российских кранчеров, всего на анкету ответили почти 650 человек, что составляет более 16 %. В большинстве своем это мужчины (97%) в возрасте от 23 до 50 лет (87%) (рис.1) с высшим (80%) образованием (преимущественно (55%) техническим) (рис.2). Для сравнения отметим, что в опросе WCG участвовало 90% мужчин, а 68% респондентов были в возрасте от 25 до 54 лет.

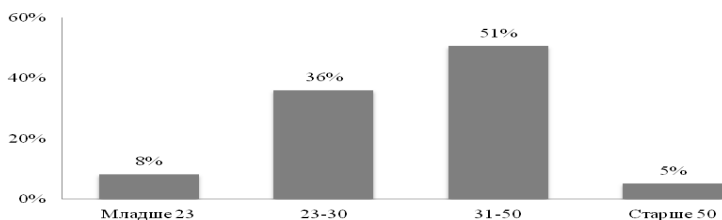


Рис. 1. Возраст респондентов (полных лет)

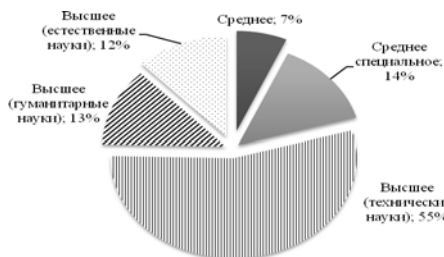


Рис. 2. Распределение респондентов по уровню образования

Подавляющее большинство респондентов (96%) – это кранчеры – непосредственные участники ДРВ, к основателям ДРВ-команд отнесли себя 3% опрошенных.

На рисунке 3 приведено распределение кранчеров по стажу занятий ДРВ. Половина из них занимается ДРВ более 3 лет, в том числе 17% имеют стаж более 7 лет.

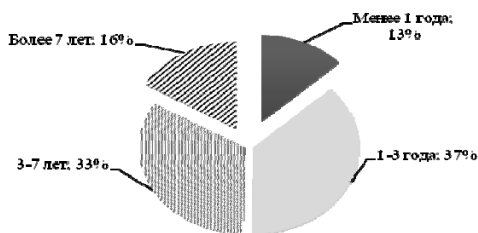


Рис. 3. Распределение респондентов по стажу занятий кранчерством

Интересно, что отвечая на вопрос о том, что их привлекает участвовать в ДРВ (разрешалось выбрать не один вариант ответа), почти 93% респондентов отметили свое стремление помочь научным исследованиям, 51% – свою причастность к научным открытиям, а 27% – просто «спортивный» интерес.

Выявлено, что мощности, которыми располагают кранчеры для ДРВ, распределились следующим образом:

- 1 или 2 компьютера стандартной конфигурации – 76%;
- 1 вычислительный сервер или много компьютеров – 17%;
- много серверов или компьютеров – 7%.

Интересны данные относительно той тематики проектов ДРВ, в которой участвуют кранчеры (рис.4). Здесь доминируют две области знаний: астрономия (57%) и биология/медицина (57%). Немалое количество респондентов (44%) участвуют в проектах из области математики и физики. Если учесть еще, что 17% участников опроса выбрали вариант ответа «любая тематика», то становится очевидным, что осуществляя ДРВ в избранной тематике, кранчеры интересуются работами в проектах других направлений.

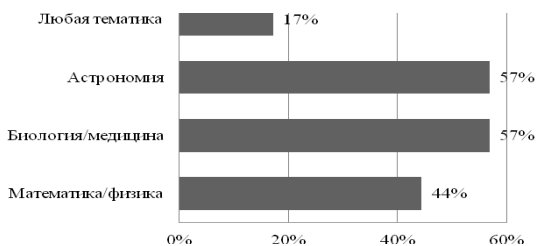


Рис.4. Распределение кранчеров по тематике проектов ДРВ

Заметим, что более трех четвертей опрошенных представлены на сайте boinc.ru, 40% – на сайтах проектов и почти четверть (23%) – на сайте команды

(рис.5). Такое предпочтение в отношении платформы BOINC респонденты объясняли тем, что это самая распространенная платформа (56%), 4% кранчеров сослались на то, что им посоветовали так поступить их коллеги, а 36% опрошенных ничего не знают о других платформах.

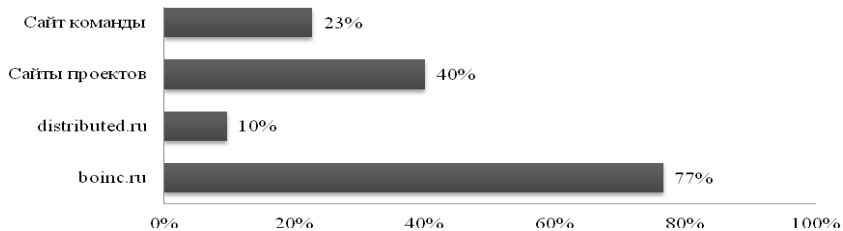


Рис. 5. Присутствие кранчеров на сайтах ДРВ

Данные социологического опроса позволяют сделать некий обобщающий вывод относительно российских кранчеров: это люди, с высоким уровнем образования и квалификации (в различных областях знаний); они осознанно подходят к выбору, как самого проекта ДРВ, так и к вопросам взаимодействия и взаимоотношений с его организаторами, а также предпочитают ощущать себя соучастниками исследований, быть в курсе проблем и успехов проекта, к которому присоединились; это партнеры для организаторов проектов.

Можно выстроить следующий усредненный портрет российского кранчера: это - мужчина (97%); в возрасте от 23 до 50 лет (87%), имеющий высшее (80%) техническое образование (55%), располагающий 1-2 компьютерами (76%), использующий платформу BOINC (96%) для оказания помощи науке (93%).

Задавался вопрос о том, как вообще у кранчеров возникает доверие к проектам ДРВ. Ответы кранчеров распределились так: если они имеют доступ к подробной информации о проекте (90%), или когда они могут познакомиться с публикациями проектных результатов (88%) или со ссылками на научные статьи (65%). (рис.6). И почти половина опрошенных (45%) испытывают доверие, если имеется обратная связь с разработчиками.

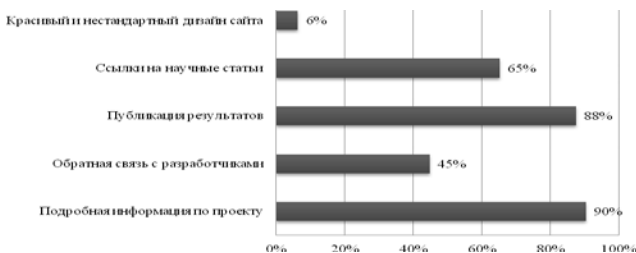


Рис. 6. Из чего складывается доверие кранчеров к проектам ДРВ

Были выявлены предпочтения российских кранчеров, на которые они опираются, участвуя в тех или иных проектах ДРВ. Так, более половины

опрошенных - 54 % кранчеров – готовы участвовать в реализации национальных проектов, если им будет предоставлена подробная информация о целях и задачах, в то же время 34% вообще готовы к сотрудничеству без предварительных условий (рис. 7).



Рис. 7. О готовности респондентов участвовать в национальных проектах

Из рисунка 8 видим, что для большинства кранчеров значимо, как определены цели и задачи ДРВ-проектов из разных отраслей. Чуть более трети из них готовы участвовать без каких-то условий.

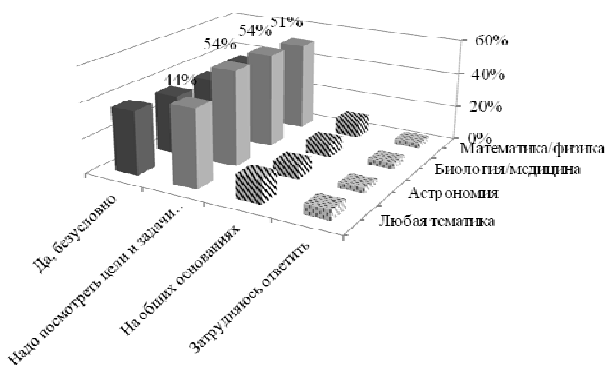


Рис. 8. О готовности участвовать в ДРВ-проектах из разных отраслей

На рисунке 9 сведена информация о том, какие результаты проектов ДРВ наиболее интересны кранчерам. Это фотографии, графики, гистограммы, ссылки на научные публикации (около трех четвертей опрошенных). Чуть более половины респондентов заинтересованы знать о ближайших планах разработчиков (56%) и о том, где проект упомянут в СМИ (55%). В меньшей степени кранчеры хотели бы «копаться» в табличных данных.

Интересы кранчеров к результатам ДРВ-проектов из разных отраслей не сильно отличались: визуализация (графики, фотографии и пр.), планы разработчиков и ссылки на публикации.

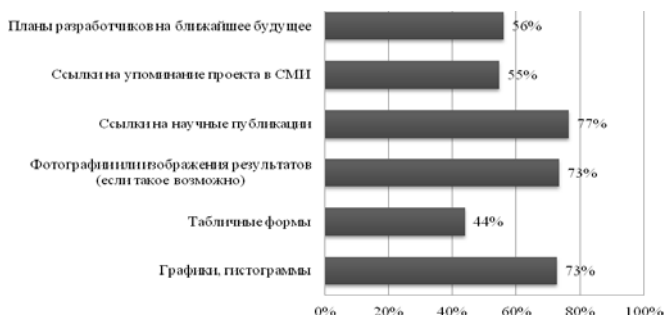


Рис. 9. Какие результаты проектов ДРВ интересны кранчерам

Приведем обобщенные пожелания кранчеров к новым проектам ДРВ (рис.10).

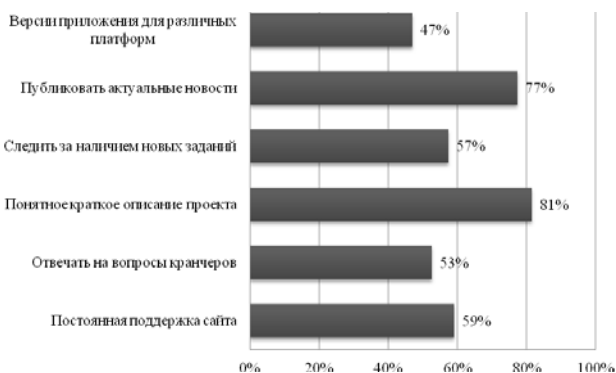


Рис.10. Доля кранчеров, отметивших пожелания к новым проектам ДРВ

Два важнейших из них – понятное краткое описание проекта (81%) и публикация актуальных новостей о нем (77%). Следующие три пожелания – постоянно поддерживать сайт (59%), следить за наличием новых заданий (57%), отвечать на вопросы кранчеров (53%). И последнее – создавать версии приложений для различных платформ (47%).

В заключении отметим, что выполненное исследование посвящено чрезвычайно важной, с точки зрения развития распределенных вычислений, теме - отношения интернет-сообщества к проблематике добровольных распределенных вычислений. Представленные результаты обработки, фактически, первого, проведенного в России, многофакторного социологического опроса, позволили определить предпочтения и пожелания кранчеров в отношении различных проектов ДРВ, что создает предпосылки для повышения эффективности работы проектных групп, а также очерчивает

перспективы повсеместного внедрения и поддержки распределенных информационно- вычислительных сред.

Литература

- [1] Андреев А.Л. Методы повышения популярности и привлечения участников в проектах добровольных распределенных вычислений на платформе VOINC. // Доклад на Национальном Суперкомпьютерном Форуме (НСКФ-2014), г. Переславль-Залесский, 25–27 ноября 2014 года. URL: http://2014.nscf.ru/TesisAll/5_Gridi_iz_rabochix_stanciy_i_kombinirovannye_gridi/05_211_AndreevAL.pdf (дата обращения: 22.04.2015).
- [2] Андреев А. Социология ГРИД. Троицкий вариант, № 166, 04 ноября 2014 года. С. 8
- [3] Курочкин И.И., Якимец В.Н. Исследование предпочтений участников распределенных вычислений. Центр распределенных вычислений ИППИ РАН. Презентация доклада на конференции ГРИД'14: Дубна, 2014. URL: <https://docs.google.com/file/d/0B22OtfxDnI5haFpYQ0t0ajdzSjQ/edit?pli=1> (дата обращения: 22.04.2015).

Volunteer distributed computing in Russia: sociological analysis

V. Yakimets, I. Kurochkin
Institute for Information Transmission Problems
of Russian Academy of Sciences

The paper outlined the results of the first made in Russia in 2014 sociological study of the crunchers community involved in volunteer distributed computing. The organization of the survey is described and information about the capabilities, preferences and wishes of the crunchers are given.

Keywords: volunteer distributed computing, crunchers, sociological survey