

## **Добровольные распределенные вычисления.**

*Как повысить привлекательность проекта. Некоторые аспекты взаимодействия с волонтерами.*

Александр Андреев, <http://boinc.ru>

Распределенные вычисления, с точки зрения их организаторов, являются всего лишь инструментом для повышения эффективности своей научной работы. При этом основные проблемы, с которыми им приходится сталкиваться, носят, прежде всего, чисто технический характер (подготовка расчетного приложения, инсталляция ПО, поддержание вычислительного парка в рабочем состоянии и т.п.). Все сказанное, в полной мере можно отнести к сегменту, так называемых, десктоп-грид систем корпоративного уровня, в том числе организованных с использованием программной платформы BOINC

Однако программная платформа BOINC позволяет достаточно просто организовать и проекты другого, более высокого (глобального) уровня, так называемые добровольные проекты распределенных вычислений, участвовать в которых могут все желающие из любого уголка нашей планеты. В таких проектах может быть задействовано несколько тысяч пользователей и десятки (и даже сотни) тысяч компьютеров. Причем время работы таких глобальных проектов практически ничем не ограничено. Проект может работать десятки лет, до тех пор, пока он поддерживается организаторами, пока он интересен добровольцам-волонтерам, предоставляющим свои компьютеры для расчетов, или пока не будет получен ожидаемый от проекта результат.

Таким образом, решение организовать открытый добровольный проект РВ добавляет его организаторам кроме технических еще и целый ворох проблем коммуникативного характера, заставляет организовывать взаимодействие с участниками проекта и учитывать некоторые (порой достаточно существенные) особенности волонтерского (как его еще называют – кранчерского) сообщества[1].

Следует понимать, что люди, которые будут участвовать в вашем проекте очень разные (по возрасту, образованию, интересам и пр.). Кроме того, в мире насчитывается более семи десятков проектов различной тематики, в которых участвуют сотни тысяч кранчеров, и ваш проект оказывается лишь одним и многих. Привлечь в свой проект как можно больше людей, а, следовательно, и

вычислительных ресурсов (ради чего всё и затевается) в таких условиях оказывается задачей непростой.

В этой статье делается попытка обобщить многолетний опыт участия в добровольных распределенных вычислениях (ДРВ) и рассказать об основных моментах, «стандартах» и принципах, которые необходимо учитывать и придерживаться при создании и развитии собственного открытого проекта добровольных РВ на платформе BOINC.

## **АУДИТОРИЯ.**

Конечно же, в первую очередь, организаторы должны понимать с кем им придется общаться и взаимодействовать в ходе работы своего проекта. Помочь в этом понимании может проведенное в мае 2014 года Центром распределенных вычислений ИППИ РАН совместно с сайтом BOINC.RU и Российским отделением IDGF социологическое исследование российских (в основном) кранчеров[2].

По результатам анкетирования можно констатировать, что потенциальными участниками вашего проекта будут в основном (97%) мужчины в возрасте от 23 до 50 лет с высшим (80%) образованием (преимущественно (55%) техническим). Большинство из них сможет подключить к проекту 1-2 стандартных компьютера (76%), хотя от 17% можно ожидать подключения и значительно большего количества.

При этом 93% из них участие в проектах объясняют желанием помочь научным исследованиям и ожидают от организаторов довольно подробной информации как о самом проекте (его целях и задачах), так и о ходе самих исследований.

На мой взгляд, основной вывод из этого должен быть таким:

- участники - это союзники для организаторов проекта;
- участники - это вполне сформировавшиеся люди, с достаточно высоким уровнем образования и квалификации (в различных областях);
- участники вполне осознанно подходят к выбору, как самого проекта, 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.

## **КОММУНИКАЦИЯ**

Для общения между участниками и организаторами существует несколько возможностей и самые основные из них это сайт и доска объявлений (форум) проекта. С них и надо начинать в первую очередь.

Кроме того, в качестве дополнительных площадок для привлечения внимания к своему проекту следует назвать специализированные ресурсы, посвященные тематике добровольных РВ (сайты статистики, сайты и форумы крупных команд, активные группы в социальных сетях).

## САЙТ

- **Оформление.**

Опыт показывает, что как таковое оформление сайта проекта не играет очень большой роли. Конечно, красиво и оригинально оформленный сайт, имеющий свою индивидуальность это приятно, оригинально и это можно только приветствовать. Но, на первоначальном этапе вполне можно ограничиться стандартным дизайном BOINC-сайта, установленным «по умолчанию», а работы по украшательству отложить на будущее, поскольку имеются гораздо более важные (и одновременно – более простые) составляющие, которыми стоит заниматься в первую очередь. Абсолютное большинство проектов именно так и делают.

Хотя, если имеется возможность придумать и изобразить какую-то эмблему или логотип своего проекта, то это будет вполне уместно.

- **Во-первых, язык сайта:**

Крайне желательно для российских проектов делать двуязычные (как минимум) варианты страниц: на родном языке организаторов и на английском. Благо функционал это позволяет.

- **Во-вторых, наполнение главной страницы:**

Главная страница сайта проекта является его лицом. На неё участники попадают в первую очередь. Поэтому в обязательном порядке на главной странице должны присутствовать краткое описание сути проекта, информация об организаторах и, даже на начальном этапе, какие-то новости.

### Примеры описания проектов:

1. [Проект SAT@home.](#)

*«SAT@home - исследовательский проект, использующий соединяемые через сеть Интернет компьютеры для решения трудных и практически важных задач (обращения дискретных функций, дискретной оптимизации, биоинформатики и*

*т.д.), которые могут быть эффективно сведены к задаче о выполнимости булевых формул. На данный момент в проекте анализируется шифр Vivium (ослабленная версия шифра Trivium).*

*Основатели SAT@home:*

*Институт динамики систем и теории управления Сибирского отделения РАН, лаборатория Дискретного анализа и прикладной логики - участник Международной федерации гридов из персональных компьютеров*

*Институт проблем передачи информации РАН, отдел Распределенных вычислений - участник Международной федерации гридов из персональных компьютеров»*

Здесь дается как краткое описание самого проекта, так и обобщенные сведения об организаторах. Более подробная информация о людях занятых в проекте SAT@home приводится на другой странице, названной «[Персоналии](#)», а о научной части проекта - на странице «[Наука](#)».

## 2. [Проект Einstein@home](#)

*«About Einstein@Home*

*Thank you for your interest in Einstein@Home!*

*Einstein@Home is a World Year of Physics 2005 and an International Year of Astronomy 2009 project supported by the American Physical Society (APS) and by a number of international organizations.*

*Einstein@Home uses your computer's idle time to search for weak astrophysical signals from spinning neutron stars (also called pulsars) using data from the LIGO gravitational-wave detectors, the Arecibo radio telescope, and the Fermi gamma-ray satellite. Einstein@Home volunteers have already discovered more than three dozens new neutron stars, and we hope to find many more in the future. Our long-term goal is to make the first direct detections of gravitational-wave emission from spinning neutron stars. Gravitational waves were predicted by Albert Einstein almost a century ago, but have never been directly detected. Such observations would open up a new window on the universe, and usher in a new era in astronomy.*

*To learn more about Einstein@Home, please explore the links under "Science information and progress reports" below, or read some of the popular articles linked from "Einstein@Home in the News" below.*

*If you want to participate, please follow the "Join Einstein@Home" instructions below. It takes just a minute or two to sign up, and little or no maintenance to keep Einstein@Home running. Einstein@Home is available for Windows, Linux and Macintosh OS X computers.*

**Bruce Allen**

*Director of Einstein@Home; Director, MPI for Gravitational Physics, Hannover; Professor of Physics, U. of Wisconsin – Milwaukee»*

В этом описании мы видим как расширенное описание сути проекта, так и объединенную с ним персонифицированную информацию об организаторах.

### 3. Проект [LHC@home \(LHS Sixtrack\)](#)

«*About SIXTRACK*

*SIXTRACK is a research project that uses Internet-connected computers to advance Accelerator Physics. Participate by downloading and running a free program on your computer.*

*SIXTRACK is based at [CERN](#). You can run it under the [LHC@home](#) application. Read more about the:*

*[SixTrack project](#)  
[SixTrack team](#) »*

Это описание можно, пожалуй, назвать образцом минимализма. Однако обратите внимание, что четыре ссылки, которыми «украшены» пять строчек описания, существенно расширяют круг предоставленной информации, как об истории самого проекта, так и об организаторах и самом расчетном приложении.

- **Новости**

Лента новостей, которая обычно находится в правой части главной страницы, является важным показателем жизнеспособности проекта. Пустая или давно не обновляемая информация в ленте новостей создает впечатление заброшенности как сайта, так и самого проекта.

Что же публиковать в новостях?

Да практически всё что угодно, относящееся к тематике проекта:

- информацию о начале и окончании очередного эксперимента;
- данные о достигнутых результатах;

- сведения о возникающих проблемах и их устранении;
- о возможной временной неработоспособности проекта и сроках возобновления работы;
- о проводимых в проекте соревнованиях;
- о публикации научных материалов, основанных на результатах расчетов;

Ну и многое другое, что организаторы могут посчитать для себя важным.

Некоторые новости могут получать свое расширенное описание и обсуждение в темах форума (Доски сообщений).

Так же как и описание для российского проекта целесообразно новости выкладывать в двуязычном варианте. Например, как это делается в проекте [SAT@home](#).

### **ФОРУМ (Доска сообщений)**

Форум – составная часть сайта проекта. И это основная площадка для общения как организаторов с участниками, так и участников между собой. Поэтому инсталлировать и организовывать работу форума нужно обязательно.

На этапе создания проекта можно «не заморачиваться» со структурой форума и оставить основные разделы: «**Science**», «**Number crunching**» и «**Cafe**». Нередко добавляют разделы «**Problems and Bug Reports**» и «**News**».

Главное для форума – регулярное посещение его представителями организаторов и ответы на вопросы участников.

### **ЧАСТО ЗАДАВАЕМЫЕ ВОПРОСЫ.**

Этот раздел также можно использовать для разъяснений нюансов как научной части проекта, так и каких-то технических его особенностей.

На начальном этапе можно самостоятельно составить несколько вопросов и ответы на них (может быть самого общего плана), а в дальнейшем пополнять раздел за счет вопросов, которые участники будут задавать в форуме.

Пример вопросов для первоначального заполнения этого раздела:

- Для чего понадобилось создание проекта ... ?
- Какие задачи решаются в проекте сейчас?
- Какие задачи уже решены в проекте?

- Предполагается ли закрытие проекта после завершения текущего эксперимента?

И несколько вопросов, которые будут заданы в обязательном порядке и к ответам на которые нужно готовиться с самого начала:

- Будут ли в проекте приложения для других операционных систем (Linux, Android, MacOS)?
- Будут ли в проекте приложения для ГПУ?

Тут желательно сразу продумать стратегию ответа. Если «ДА», то дать примерные сроки, а если «НЕТ», то указать причину (например, для приложений под GPU – что данный алгоритм не может быть эффективно реализован с использованием структуры графических ускорителей).

### А ЧТО ЖЕ ДАЛЬШЕ?

Но, допустим, все вышесказанное уже реализовано. Проект работает, народ присоединяется, задача решается ...

Но Вы чувствуете, что потенциал раскрыт далеко не полностью. Вы ощущаете в себе силы и желание для дальнейшего развития проекта и привлечения в него новых участников и компьютеров. О чем можно еще подумать?

1. Следует подумать о внесении вашего проекта в списки статистических сайтов. А для этого связаться с администраторами соответствующих сайтов. Это, в первую очередь, такие сайты как:
  - [BOINCstats.com](http://BOINCstats.com)
  - [Free-DC](http://Free-DC)

2. Хорошим способом раскрутки своего проекта является публикация новостей проекта непосредственно в BOINC-менеджере, во вкладке «Уведомления». Для этого используется файл - gui\_urls.xml. Есть определенные URL, которые появляются в BOINC-менеджере в левой полосе меню. Перед тегом <gui\_urls> нужно поставить код <rss\_feed>. Вот пример из проекта Asteroids@home:

```
<rss_feeds>
<rss_feed>
<url>http://asteroidsathome.net/boinc/notices.php?userid=<userid/>&auth=<rss_auth/></url>
```

```
<poll_interval>86400</poll_interval>  
</rss_feed>  
</rss_feeds>
```

После того, когда клиент свяжется с сервером, сервер получит новую информацию, что он может читать новые сообщения на RSS с периодом проверки, указанным в poll\_interval. На стороне сервера всё готово. Теперь, вы просто должны написать какие-то новости в форуме. Если вы пишете в разделе новостей, то кроме добавления электронной записи на пост, есть еще один флажок, уведомляющий что сообщение должно быть опубликовано в менеджере BOINC. Проверьте, что сообщение настроено как видимое для BOINC менеджера. Итак, если вы установили уведомления в менеджере BOINC, сконфигурировали XML для сообщений с кодом, которые позволяют ему быть показанными в менеджере BOINC, то сообщение появиться там.

3. Неплохо зарегистрироваться и поучаствовать в общении на основных международных и национальных командных сайтах. Кранчеры очень одобрительно относятся к фактам общения с ними организаторов проектов. А если это происходит не только на проектных форумах, а и на командных сайтах, то это большой «плюс» в раскрутке Вашего проекта.

Из российских сайтов можно рекомендовать следующие:

- Межкомандный российский сайт [BOINC.RU](http://BOINC.RU)
- [Форум команды TSC!Russia](#)
- [Сайт команды Russia](#)
- [Группа ВКонтакте «Программа BOINC. Распределённые вычисления.»](#)

## ЗАКЛЮЧЕНИЕ

Конечно, это далеко не все методы и способы повышения привлекательности нового проекта. Но даже если попытаться выполнить описанные рекомендации, эффект повышения привлекательности Вашего проекта, а значит, и роста его вычислительной мощности будет определенно заметен.

Кроме того, огромную пользу можно извлечь из общения с организаторами других проектов, а также активными участниками таких проектов. Их, конечно можно искать по сайтам и проектам, о которых упоминается в этой статье, но есть и более простой способ. В настоящее время большинство упомянутых лиц можно найти через



возможности [Российского отделения IDGF](#) (Международной федерации грид-систем из персональных компьютеров), участники которого пытаются поддерживать развитие в России распределенных вычислений, в первую очередь на платформе BOINC и объединять в своих рядах всех заинтересованных в этом деле лиц, участвует в различных научных конференциях, посвященным высокопроизводительным вычислениям и является одним из организаторов специализированной конференции [BOINC:FAST](#).

Дополнительные материалы:

1. [Добровольные распределенные вычисления в России глазами краучера. Современное состояние и перспективы развития. Андреев А.Л. \(Сайт BOINC.RU\) Доклад на Национальном Суперкомпьютерном Форуме 2013 г.](#)
2. [Исследование предпочтений участников распределенных вычислений. И.И. Курочкин, В.Н.Якимец \(Центр распределенных вычислений ИППИ РАН\). Презентация доклада на конференции GRID'2014 в Дубне](#)

Ссылки на ресурсы, упомянутые в статье:

**Проект SAT@home** - <http://sat.isa.ru/pdsat/index.php>

**Проект Einstein@home** - <http://einstein.phys.uwm.edu/>

**Проект LHC@home (LHS Sixtrack)** - <http://lhathomeclassic.cern.ch/sixtrack/>

**Сайт статистики BOINCstats.com** - <http://boincstats.com/>

**Сайт статистики Free-DC** - <http://stats.free-dc.org/>

**Сайт BOINC.RU** - <http://boinc.ru>

**Форум команды TSC!Russia** - <http://forums.overclockers.ru/viewforum.php?f=21>

**Сайт команды Russia** - <http://distributed-russia.ru/news.php>

**Группа ВКонтакте «Программа BOINC. Распределённые вычисления.»** - <http://vk.com/boinc>

**Сайт Российского отделения IDGF** - <http://desktopgrid.ru/>

**Сайт конференции BOINC:FAST** - <http://boincfast.ru/>