

Щёлковский муниципальный район Московской области

исследование на тему
«Проекты в облаках»

выполнил ученик 8 класса
МАОУ Свердловской СОШ №2
Черников Николай

руководитель
учитель информатики и ИКТ
Васина Алла Витальевна

2017 год

ОГЛАВЛЕНИЕ

Введение	4
1. Облачные технологии как современные услуги.....	5
2. Добровольные распределенные (облачные) вычисления.....	8
3. Мое участие в научных проектах через распределенные вычисления.....	11
Выводы.....	12
Литература	13
Приложение. Раздаточный мотивационный материал.....	14

Объект исследования: удаленные облачные сервисы и их возможности.

Предмет исследования: использование распределенных облачных вычислений для решения сложных научных задач и проектов.

Актуальность исследования: облачная революция, происходящая сейчас, требует от нас, школьников, не только иметь представление о новейших технологиях, но и использовать их для того чтобы сделать мир лучше. Облачные технологии предоставляют особые инструменты и сервисы. Современный школьник-исследователь, не знакомый с облачными возможностями, на наш взгляд, не способен конкурировать в меняющемся мире.

Цель исследования: описать некоторые облачные технологии, открывающие перед современным школьником двери в мир современных научных исследований.

Задачи исследования:

- определить понятие облачных технологий и сервисов;
- выделить модели и виды облачных услуг и ресурсов;
- рассмотреть распределенные вычисления как возможность помощи ученым в научных проектах;
- практически участвовать в нескольких проектах и сделать выводы о результатах.

Гипотеза исследования: современные облачные технологии позволяют каждому человеку вносить свой вклад в научные исследования. Для этого не нужно обладать специальными знаниями, а только однажды настроить свой компьютер.

Введение

Облачные сервисы современный и быстрорастущий сегмент информационных технологий, обеспечивающий пользователю удаленный доступ к аппаратным мощностям или программному обеспечению, расположенному в сети. Облака дают возможность пользователю использовать нужные инструменты независимо от его местонахождения. Необходимо лишь доступ к сети Интернет. Кроме того, работающий в облаке человек не хранит программное обеспечение и файлы на своем компьютере и не зависит от его мощности своей машины.

Сейчас мы не просто стоим на пороге облачной революции, мы уже стремительно вступаем в нее. Области применения облачных технологий расширяются с каждым днем. Развитие компьютерных мощностей и линий связи привело к созданию масштабных вычислительных комплексов и центров хранения данных. Современные телекоммуникации стали доступны каждому, услуги в области ИТ предоставляются удаленно.

Мне кажется, что учась в школе важно не просто успевать за современными тенденциями в мире технологий, но и целенаправленно использовать их для улучшения мира.

Облачные технологии как современные услуги

В чем же состоит суть облачных технологий? Организации, предоставляющие облачные сервисы вкладывают средства в аппаратные ресурсы и программное обеспечение, предоставляя обычным пользователям дистанционный доступ к ним на платной или бесплатной основе. Если потребности пользователей меняются, он может увеличивать или уменьшить необходимый для работы объем аппаратных и программных ресурсов. Конечно, доступ к этому ресурсу открыт везде, где есть достаточная скорость сети. Кроме того, такие сервисы доступны и с мобильных устройств, что очень важно современному специалисту.

Концепция облачных вычислений предполагает оказание следующих типов услуг своим пользователям:

- *все как услуга (Everything as a Service)*. При таком виде сервиса пользователю будет предоставлено всё - от программно аппаратной части и до управления бизнес процессами, включая взаимодействие между пользователями, от пользователя требуется только наличие доступа в сеть Интернет. На мой взгляд, данный вид сервиса это более общее понятие по отношению к нижеприведенным услугам, являющимися более частными случаями.

- *инфраструктура как услуга (Infrastructure as a service)*. Пользователю предоставляется компьютерная инфраструктура, обычно виртуальные платформы (компьютеры) связанные в сеть. Которые он самостоятельно настраивает под собственные цели.

- *платформа как услуга (Platform as a service)*. Пользователю предоставляется компьютерная платформа, с установленной операционной системой и приложениями.

- *программное обеспечение как услуга (Software as a service)*. Данный вид услуги обычно позиционируется как «программное обеспечение по требованию», это программное обеспечение развернутое на

удаленных серверах и пользователь может получать к нему доступ посредством Интернета, причем все вопросы обновления и лицензий на данное программное обеспечение регулируется поставщиком данной услуги. Оплата в данном случае производится за фактическое использование программного обеспечения. Так компания Adobe развернула сервис Криэйтив Клауд (Creative Cloud) на основе которого любой пользователь за плату может воспользоваться продуктами компании, такими как PhotoShop, Adobe Illustator и другими.

- *аппаратное обеспечение как услуга (Hardware as a Service)*. В данном случае пользователю услуги предоставляется оборудование, на правах аренды которое он может использовать для собственных целей. Данный вариант позволяет экономить на обслуживании оборудования, на основе которого можно развернуть свою собственную инфраструктуру с использованием наиболее подходящего программного обеспечения.

- *рабочее место как услуга (Workplace as a Service)*. В данном случае компания использует облачные вычисления для организации рабочих мест своих сотрудников, настроив и установив все необходимое программное обеспечение, необходимое для работы персонала.

- *данные как услуга (Data as a Service)*. Основная идея данного вида услуги заключается в том, что пользователю предоставляется дисковое пространство, которое он может использовать для хранения больших объемов информации. Самый широко используемый сервис такого вида это Гугл-диск.

- *безопасность как сервис (Security as a Service)*. Данный вид услуги предоставляет возможность пользователям быстро развертывать, продукты позволяющие обеспечить безопасное использование веб-технологий, безопасность электронной переписки, а также безопасность локальной системы, что позволяет пользователям данного сервиса экономить на развертывании и поддержании своей собственной системы безопасности.

При этом существуют следующие виды облаков:

- *публичные облака*, доступ к которым осуществляется любым пользователем Интернет;
- *частные облака*, доступ к которым организован только определенному кругу лиц;
- *общественные облака* доступны потребителям с одинаковыми потребностями;
- *гибридные облака* вбирают в себя свойства трех предыдущих типов.

На сегодня самыми крупными проектами на рынке облачных технологий являются:

- облачная система Windows Azure от компании Microsoft (операционная система, база данных и программный комплекс);
- Amazon Web Service от компании Amazon (предоставление виртуальных машин, ресурсов хранения и баз данных);
- сервисы Web 2.0.

Добровольные распределенные (облачные) вычисления

Все облачные сервисы дают возмездную или безвозмездную возможность использовать чужие компьютерные мощности или программы для собственной работы. Тем интереснее было узнать, что существует и широко используется и обратная схема - так называемые распределенные вычисления. Распределённые вычисления это способ решения трудоёмких вычислительных задач с использованием нескольких компьютеров, чаще всего объединённых в параллельную вычислительную систему. Масштабная задача делится на кусочки, и рассылается на клиентские компьютеры, где необходимые для решения этой задачи вычисления проводятся только во время простаивания компьютера.

Меня заинтересовало использование распределенных облачных вычислений для крупных научных проектов. Моделирование в научных исследованиях применяется в различных науках. Особенно моделирование развилось в двадцатом веке с появлением компьютеров. Необходимость использования метода моделирования определяется тем, что многие объекты непосредственно исследовать или вовсе невозможно, или же это исследование требует много времени и средств. Часто бывает неудобным и невозможным рассмотреть реального объекта, процесса или явления, ведь они бывают иногда многогранны и сложны. Тогда лучшим способом их изучения и становится построение модели, отображающей лишь какую-то грань реальности, потому более простой. И многовековой опыт развития науки доказал на практике плодотворность такого подхода. Современные научные модели сложны и требуют огромных аппаратных ресурсов. И здесь на выручку исследователям приходят распределенные облачные вычисления, в которых участвуют миллионы пользователей персональных компьютеров по всему земному шару.

Нужно отметить, что большинство проектов основано на добровольном и безвозмездном предоставлении вычислительных ресурсов. Участник

проекта соглашается на использовании своих простаивающих мощностей без какой либо платы.

Концепция таких вычислений была создана в 60-е годы двадцатого века, реализованы впервые в 90-е. Первым проектом стал SETI@home компьютерная реализация известного проекта SETI по поиску внеземных цивилизаций. Как работает эта программа? Мощнейший радиотелескоп в Аресибо записывает космический шум в различных диапазонах, затем сервер, куда поступает информация, делит все это на блоки, и рассылает по компьютерам пользователей, расположенных по всему миру.

Сейчас существует более двухсот научных проектов различной тематики и направленности, привлекающих всех желающих внести свой вклад в развитие современной науки, просто предоставив простаивающие компьютерные мощности. Проекты реализуются на базе крупных научных центров, таких как Университет Беркли, Стенфордский университет. Полный список проектов размещен на сайте *distributedcomputing.info*.

Для участия в работе проекта необходимо установить клиентскую программу, которая и будет проводить все необходимые вычисления, работая в фоновой режиме (Рисунок 1).

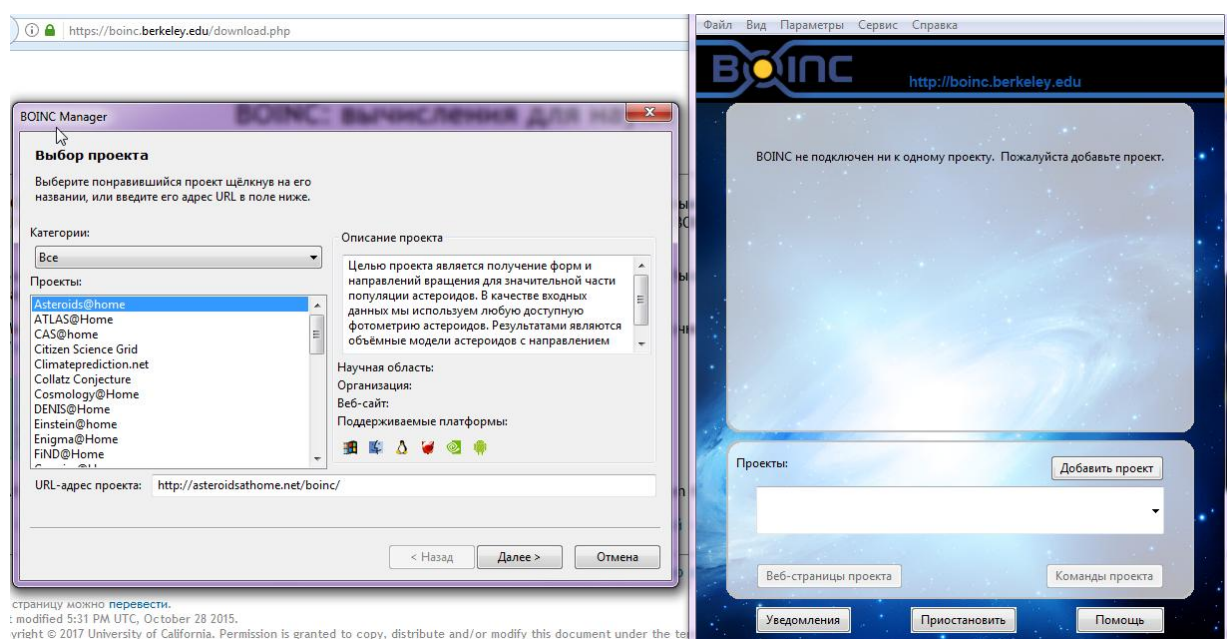


Рис. 1. Клиент BOINC

Знание предмета научного проекта, конечно, не требуется! Моим выбором стала открытая программная платформа BOINC (Berkeley Open Infrastructure for Network Computing), предоставляющая неиспользуемое время вашего компьютера для научных проектов, таких как: SETI@home, Climateprediction.net, Rosetta@home, World Community Grid и многих других.

Моё участие в научных проектах через распределенные вычисления

После инсталляции клиента BOINC и регистрации на сайте *worldcommunitygrid.org* мой персональный компьютер приступил к расчетам. Я смог добавлять любые научные проекты, входящие в сферу интересов Университета Беркли. Сегодня мой компьютер занят составлением карты маркеров рака, решает задачу борьбы со смертоносными вирусами Эбола и ВИЧ.

Удаленный сервер разделяет задачу на небольшие кусочки и рассылает их всем участникам. Моему компьютеру достается такой кусочек, и он приступает к моделированию. Если какой-либо машине в сети удастся найти искомую учеными конфигурацию - она отправляет ее на основной сервер. Закончив с отправкой, машина берется за новые расчеты. Процесс продолжается до тех пор, пока не будут найдены все варианты. Причем ученые получают с основного сервера только успешные результаты. Их мало, единицы. И поэтому каждую из них можно досконально изучить. А значит, найти лекарство от болезни.

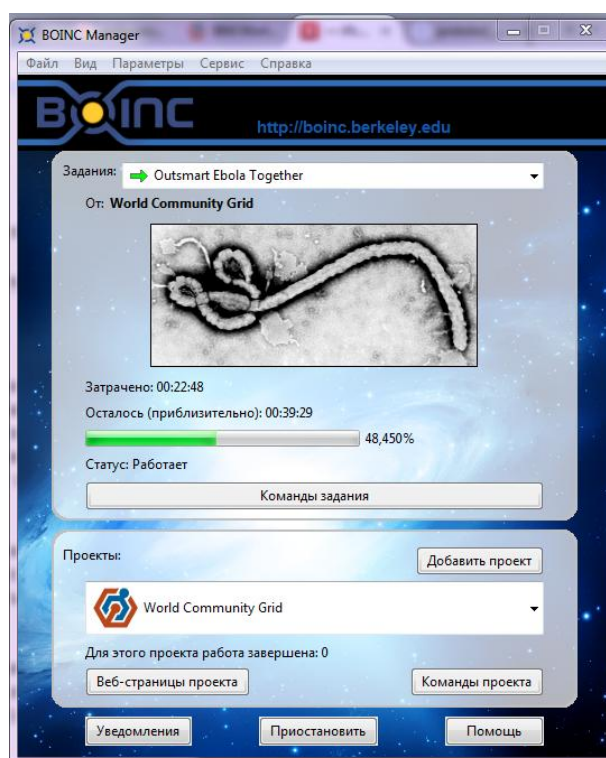


Рис. 2. Работа над проектом «Перехитрим Эболу вместе»

Замечаю ли я, что мой компьютер стал работать хуже? Нет. Понимаю ли я что помогаю бороться с болезнями – да. Некоторые люди, рассуждающие в форумных беседах о том, что придется платить за излишне потребленное компьютером электричество, вызывают во мне недоумение и непонимание. На мой взгляд, участие в таких проектах это - то малое, что каждый из нас может сделать для развития современной науки. На сегодняшний день в мире зарегистрировано почти три миллиона участников распределенных вычислений, обычных людей, таких как я или мои одноклассники, мои родители, их друзья. Россия в списке стран участниц занимает 10-е место по количеству активных пользователей и динамично увеличивает свою долю.

Такие проекты распределённых вычислений в интернете, как SETI@Home и Folding@Home обладают не меньшей вычислительной мощностью, чем самые современные суперкомпьютеры. Интегральная производительность проектов на платформе BOINC по данным на 17 декабря 2016 года составляет 28,7 петафлопс. Для сравнения, пиковая производительность самого мощного суперкомпьютера Sunway TaihuLight — 93 петафлопс.

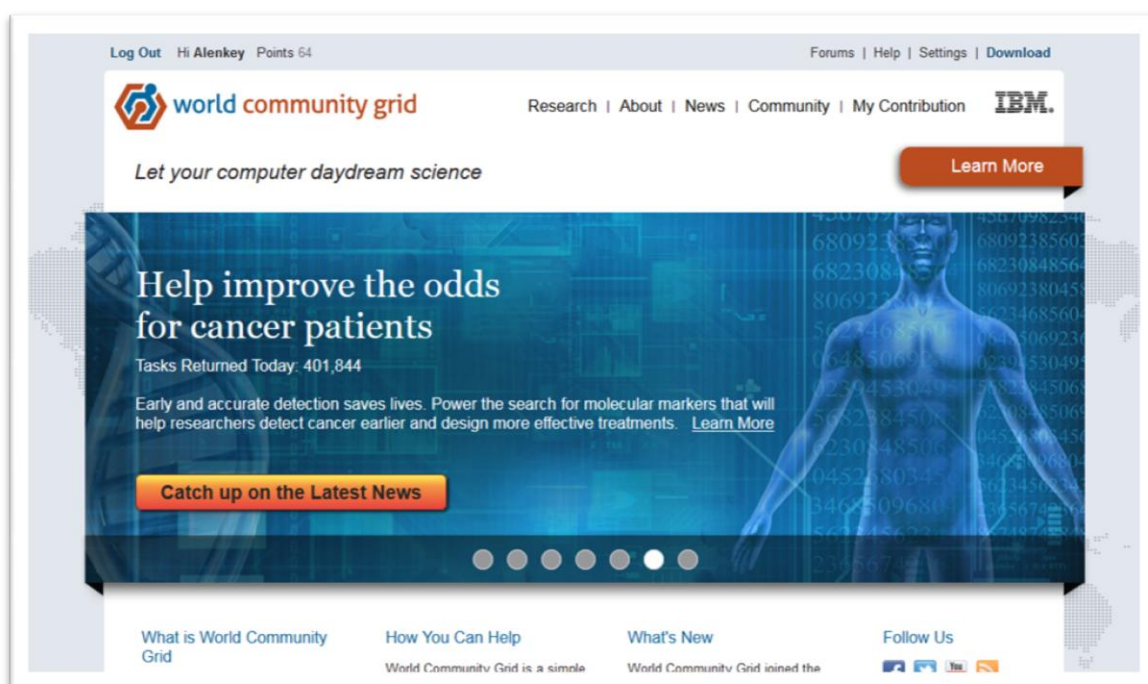


Рис. 3. Сайт проектов Всемирного общества распределенных вычислений

Как я отметил выше участие в вычислениях безвозмездное. Поэтому, важно заинтересовать потенциальных участников, привлечь их к работе над проектом, объяснить необходимость участия как можно большего количества машин. В рамках работы над своим исследованием я подготовил раздаточный материал (Приложение), в который поместил описание возможностей распределенных вычислений для науки, адреса основных ресурсов для начала работы. Мой научный руководитель учитель информатики Васина Алла Витальевна помогает в продвижении этой идеи в моей школе.

В следующем году я хочу продолжить работу над продвижением участия в распределенных вычислениях в школьной среде и в социальных сетях. В частности я узнал, что реализация подобных вычислений возможна на Android, то есть можно использовать наши смартфоны. Хотелось бы собрать статистику участия в научных проектах ребят нашей школы.

Выводы

В результате проведенного исследования **мы выполнили, поставленные перед собой задачи:** определить понятие облачных технологий и сервисов;

- выделили модели и виды облачных услуг и ресурсов;
- рассмотрели распределенные вычисления как возможность помощи ученым в научных проектах;
- на сегодняшний день практически участвуем в нескольких проектах.

В заключении нужно сказать, что современный школьник обязан иметь представление об облачных технологиях и уметь ими пользоваться. Одной из уникальных возможностей таких технологий является простота, с которой каждый из нас может помочь прогрессивным научным исследованиям по всему миру. Достаточно просто включить свой компьютер.

Литература

1. Что такое облачные (рассеянные) технологии. [Электронный ресурс] / Блог для «Чайников» - Режим доступа: <http://albas.ru/cloud-computing/tchto-takoe-oblatchne-rasseyane-tehnologii.html>;
2. Что такое облачные вычисления (CloudComputing)? [Электронный ресурс] /LiveBuiness – Режим доступа: http://www.livebusiness.ru/tags/CLOUD_COMPUTING;
3. Российские распределенные вычисления на платформе BOINC [Электронный ресурс] / - Режим доступа http://www.boinc.ru/#distributed_computing;
- 4.Сравнение биомедицинских проектов распределенных вычислений [Электронный ресурс] – Режим доступа: http://www.boinc.ru/Doc/Bio_Diffs.htm;
5. С миру по нитке: Суперкомпьютер. [Электронный ресурс] / Популярная Механика – Режим доступа: <http://www.popmech.ru/technologies/9137-s-miru-po-nitke-superkompyuter>.

Приложения. Раздаточный мотивационный материал

Знаете ли вы, что компьютер действительно не работает на все 100%? Думаем, что большая часть из вас об этом знает, а если не знает, то догадывалась. Дело в том, что большинство времени ресурсы компьютера используются менее, нежели на 5%. Оказывается, что к примеру, во время чтения текста с экрана, или в перерывах между нажатием клавиш при печати ресурсы ПК практически не используются (загрузка ресурсов в такие моменты составляет от 5% и ниже). Возможно, вас заинтересует тот факт, что время простоя вашего железного друга можно использовать с пользой?

В общем, один из способов загрузить свой ПК полезной работой (ничего не теряя при этом) является возможность стать частью важнейших исследований, проводящихся научно-исследовательскими институтами со всего мира. Эти исследования настолько серьезны, и нуждаются в таком количестве вычислений, что одного-двух-тысячи даже очень мощных компьютеров не хватит.

Именно поэтому еще в конце прошлого века был придуман способ (концепция – 60-е годы, воплощение – 90-е), как можно обойти ограничение в лимите вычислительной мощности компьютерной сети отдельно взятого института. Выходом из ситуации является организация системы проектов распределенных вычислений. Распределенные (или облачные) вычисления, это способ обработки данных, в которой масштабная задача делится на кусочки, и рассылается на клиентские компьютеры, где необходимые для решения этой задачи вычисления проводятся только во время простаивания компьютера.

Один из наиболее популярных "распределенных" проектов, SETI@home, который предлагает любому человеку сделать свой вклад в важнейшую задачу поиска внеземных цивилизаций. Как работает эта программа? Мощнейший радиотелескоп в Аресибо записывает космический шум в различных диапазонах, затем сервер, куда поступает информация,

делит все это на блоки, и рассылает по компьютерам пользователей, расположенных по всему миру. Для участия в работе проекта необходимо просто установить клиентскую программу, которая и будет проводить все необходимые вычисления. Стоит отметить, что с вашей стороны не требуется ни знание астрономии, ни понимание принципов работы программы. Все, что нужно - подключение к Сети и работающая в фоновом режиме программа проекта, которая, как уже говорилось выше, будет работать только тогда, когда простаивает компьютер.

Сайт с проектами, к вычислениям которых можно присоединиться сегодня: distributedcomputing.info.

ПО для организации добровольных вычислений:

- Apache Hadoop;
- BOINC — открытая инфраструктура для распределённых вычислений Университета Беркли (Berkeley Open Infrastructure for Network Computing), распространяемая под лицензией LGPL;
- Condor (англ.);
- Globus Toolkit (англ.) — набор программ, значительно упрощающий создание и управление распределёнными вычислениями.

Самые популярные проекты:

- Biochemical Library - выяснение структуры мембранных белков (мишени для лекарств), проектирование белков с новой структурой и функциями (фолдинг), понимание количественных отношений между химической структурой и биологической активностью (синтез лекарств).
- Correlizer - проект, целью которого является исследование корреляций между последовательностями в трехмерной структуре генома.
- DrugDiscovery@Home - Российский проект, фолдинг белка (вычисление его структуры), скрининг баз данных биологически активных соединений.
- DNA@Home - поиск последовательностей в молекулах ДНК, отвечающих за различные гены, с использованием вероятностных

алгоритмов; исследование механизма транскрипции генов с использованием молекул полимеразы; исследование ДНК бактерий, вызывающих туберкулез и бубонную чуму.

- evo@home - проект распределённых вычислений, целью которого является применение генетических алгоритмов для фолдинга белков.

- FightMalaria@Home - моделирование докинга протеинов малярии.

- Folding@Home - проект по расчёту третичной структуры белков.

- Neurona@Home - моделирование поведения больших клеточных автоматов, составленных из нейронов.

- NRG - проект в области молекулярного распознавания, вычислительной биологии и докинга.

- Rosetta@home - фолдинг белка с самой низкой энергией.

- Volpex@UH - имитация поведения белков в клеточной среде с целью разработки лекарственных препаратов.

- Wildlife@Home - анализ видеоданных записи жизни в дикой природе.

- World Community Grid - исследования, связанные с генетикой, а также с различными тяжёлыми заболеваниями человека. Решение научно-технических проблем, связанных с экологией, здоровьем, питанием и другими проблемами человечества.

- AQUA@home — моделирование работы адиабатического квантового сверхпроводящего компьютера;

- Asteroids@home — проект в области астрономии, целью которого является определение формы и параметров вращения астероидов по данным фотометрических наблюдений;

- Climate Prediction — проект по моделированию влияния выбросов углекислого газа на климат Земли;

- Constellation — проект для решения вычислительно сложных аэрокосмических задач (оптимизация траекторий спутников и зондов, моделирование экзосферы Луны);

- Cosmology@home — поиск модели, лучшим образом описывающей нашу Вселенную, а также определение диапазона моделей, согласующихся с современными астрономическими и физическими данными;

- EDGeS@Home — проект в области физики плазмы, моделирование поведения заряженных частиц в магнитном поле термоядерного реактора ITER;

- Einstein@Home — проект по проверке гипотезы Эйнштейна о гравитационных волнах с помощью анализа гравитационных полей пульсаров или нейтронных звёзд;

- Enabling Grids for E-science — проект, направленный на построение грид-инфраструктуры, которая сможет использоваться в многочисленных научных исследованиях в Европе;

- LHC@home — расчёты магнитной подсистемы ускорителя заряженных частиц — Большого адронного коллайдера (англ. *LHC*);

- MilkyWay@home — создание высокоточной трехмерной модели звёздных потоков, обращающихся вокруг центра Галактики, на основе данных цифрового обзора SDSS;

- Muon1 DPAD — расчёты при проектировании мюонного ускорителя — Neutrino Factory/Muon Collider;

- SETI@Home — проект по обработке сигналов радиотелескопа, для поиска радиосигналов внеземных цивилизаций;

- Solar@Home — проект, целью которого является создание более эффективных солнечных батарей;

- Spinhenge@home — проект в области нанотехнологий, исследование магнитных молекул;

- theSkyNet POGS — проект, целью которого является построение спектрального атласа ближайшей части Вселенной в области длин волн от ближнего инфракрасного излучения до ультрафиолета по данным GALEX, Pan-STARRS1 и WISE.