

ГДЕ ВАШИ ОБЛАКА? ТАМ ПРОСТО ГРИД!

Ян Фостер

THERE'S GRID IN THEM THAR CLOUDS*

Ian Foster

<http://ianfoster.typepad.com/blog/2008/01/>

Вы, по-видимому, уже сталкивались в последнее время с потоком новостей, касающихся "cloud-компьютинга". [В Business Week имеется на эту тему большая статья](#) (с забавной и остроумной критикой). А Nick Carr написал об этом даже [книгу](#). Так о чём же это, что здесь нового, и какое это имеет значение для информационной технологии?

Основная идея, кажется, состоит в том, что в будущем мы не будем проводить вычисления на локальных компьютерах, мы будем вычислять на централизованных ресурсах, управляемых утилитами сторонних организаций. На что я отвечаю, Аллилуйя, считая, что имеется в виду конец запакованному программному обеспечению, а значит, и необходимости его распаковки и инсталляции.

Вряд ли стоит напоминать, что эта идея не нова. Еще в далеком 1960 году пионер программирования [Джон Маккарти](#) пророчески говорил, что "вычислительная деятельность может быть со временем организована как общественная услуга" – и рассматривал, как это могло бы быть сделано.

В середине 1990-ых годов был введен в обращение термин "[грид](#)" для описания технологий, которые бы позволяли пользователям посылать запросы к службе "вычислительного обеспечения" и почти сразу же начинать вычислять. [Вместе с другими специалистами](#) я пришел к заключению, что стандартизация протокола запроса поможет нам быстрее создать вычислительный грид, аналогичный по форме и удобству гриду электроснабжения. Впоследствии эти идеи были успешно и интересно развиты в различных исследованиях, что проявилось в создании, например, мощных федеративных систем ([TeraGrid](#), [Open Science Grid](#), [caBIG](#), [EGEE](#), [Earth System Grid](#),), которые удовлетворяют не только вычислительные потребности, но также и запросы на данные и программное обеспечение. Учреждения по стандартизации (например, [OGF](#), [OASIS](#)) определяют соответствующие стандарты. Очень часто термин «грид» использовался в промышленности как термин маркетинга для кластеров. Но ни одного жизнеспособного коммерческого грида компьютеринг-провайдерами предоставлено не было, по крайней мере, до последнего времени.

Так не является ли "cloud-компьютинг" просто новым названием грида? В информационной области, где технологии растут на порядок каждые пять лет, часто изобретая себя повторно, прямого и ясного ответа на такие вопросы не существует.

Да: цель одна и та же – снизить стоимость компьютеринга, увеличить надежность, увеличить гибкость, преобразуя компьютеры из объектов, которые мы покупаем, и которыми мы управляем сами, в нечто, управляемое третьей силой.

Но и нет: теперь все по-другому, чем 10 лет назад. Теперь нам просто необходимо анализировать очень большие массивы данных, что является мотивацией значительно увеличившегося спроса на компьютеринг. Осознав выгоды от перехода с мейнфреймов на коммерческие кластеры, мы обнаруживаем, что управление этими кластерами что-то уж слишком дорого. У нас есть доступная по цене виртуализация. И, что весьма существенно, тратятся многие миллиарды долларов такими, и подобными им, корпорациями, как Amazon, Google и Microsoft на создание по настоящему коммерческих гридов, содержащих сотни тысяч компьютеров. Перспектива получить по запросу доступ к 100000+ компьютерам в десятках центров данных, расположенных по всему миру, с помощью всего лишь кредитной карточки – не может не волновать. Полученный ресурс может быть приложен к решению проблем, требующих потенциально распределенных данных большого объема. Итак, мы работаем в новых, впечатляющих масштабах; работа в этих условиях может потребовать фундаментально других подходов к решению проблем. А также дает возможность справляться с совершенно новыми проблемами, а часто и применима только к ним.

И, тем не менее, да: проблемы, с которыми мы сталкиваемся в гриде и "облаках", в большинстве случаев одни и те же. И здесь, и там необходима возможность управлять большими ресурсными центрами; определять методы, с помощью которых пользователи обнаруживают, запрашивают и используют ресурсы, предоставляемые этими центрами; обеспечивать выполняемые на этих ресурсах вычисления, часто обладающие сложным параллелизмом. Детали различаются, но оба сообщества борются с решением большого числа одних и тех же вопросов.

К сожалению, по крайней мере, на настоящий момент, используемые для достижения этих целей методы, в теперешних коммерческих "облаках" не являются ни открытыми, ни общецелевыми; они, как правило, принадлежат частным компаниям и являются специализированными (например, анализ данных очень большого объема), ориентированными на выполнение конкретных внутренних задач разработавших их компаний. Мысль о возможной необходимости обеспечения интероперабельности для провайдеров как в гридах электроэнергии еще не стала очевидной. Грид-технологии и протоколы прямо ставят эти вопросы, и их нужно решать.

И что еще общего: мы видим один и тот же маркетинг. Первые "кластеры cloud-компьютинга" – удивительно напоминающие "грид-кластеры" недавнего прошлого – с ними уже можно познакомиться. Не видно ли на горизонте Oracle 11c?

Что нам обещает будущее? Рискну сделать несколько предсказаний, основываясь на моей вере в то, что экономика компьютеринга будет все более и более похожа на экономику электроснабжения. Ни электроснабжение, ни вычислительные гриды следующих поколений не будут похожи на вчерашние силовые электрические гриды. И те, и другие будут двигаться в направлении объединения микропродуктов и больших утилит, причем

постоянно будет увеличиваться количество маломасштабных производителей, сосуществующих с региональными производителями большого масштаба; нагрузка между ними будет распределяться динамически. Да, я знаю, что компьютеринг на самом деле мало похож на электричество, но я действительно верю, что мы, тем не менее, будем свидетелями параллельной эволюции, совершающейся под воздействием очень похожих сил.

При построении этого распределенного "облака" или "грида" ("гробла"?), мы будем вынуждены уметь предоставлять и конфигурировать по требованию интегральные "виртуальные системы", обеспечивая в точности те особенности, которые нужны конечному пользователю. Нам нужно будет определять протоколы, позволяющие пользователям и провайдерам служб обнаруживать запросы и уметь перенаправлять их другим провайдерам, осуществлять мониторинг и операции сохранения, разобраться с системой оплаты. Нам будут необходимы инструменты для менеджмента как нижележащих ресурсов, так и соответствующих распределенных вычислений. Нам будет необходимо централизованное масштабирование имеющихся на сегодняшний день cloud-утилит и интероперабельность нынешних грид-возможностей.

Некоторые из требуемых протоколов и инструментов будут предоставлены нам высокими профессионалами из Amazon и Google. Другие поступят от высоких профессионалов, работающих на гриде. Еще одни поступят от тех, кто глубоко интересуется гридом и "облаками". Очень интересно, до какой степени эти различные сообщества найдут общие точки соприкосновения или будут двигаться параллельными путями.

* Не совсем ясная культурологическая ссылка: фраза "Там много золота, где те холмы" была впервые произнесена, согласно некоторым источникам, старым старателем в фильме 1948 года "Сокровища Сьерра Мадре", с Humphrey Bogart в главной роли.