

## **ЧТО ТАКОЕ ГРИД?**

Три критерия

Ян Фостер

\* \* \*

## **WHAT IS THE GRID?**

A Three Point Checklist

Ian Foster<sup>1,2</sup>

<sup>1</sup> Mathematical and Computer Science Division, Argonne National Laboratory,  
Argonne, IL 60439

<sup>2</sup> Department of Computer Science, University of Chicago, Chicago, IL 60637

[foster@mcs.anl.gov](mailto:foster@mcs.anl.gov)

Недавняя вспышка коммерческого и научного интереса к грид-системам заставляет нас вернуться к этому вопросу. В самом деле, что такое Грид? В этой статье я предлагаю состоящий из трех пунктов список критериев для определения того, какую систему можно считать грид-системой. Также я рассматриваю решающую роль, которую должны играть стандарты в определении грид-системы.

В четком определении нуждаются самые разные системы - от никому не известных академических до в высшей степени популярных. Мы читаем о вычислительных сетях, сетях данных, сетях доступа, сетях знаний, биосетях, сенсорных сетях, кластерных сетях, университетских сетях, тера сетях и товарных сетях, и все они именуют себя грид-системами. Можно простить скептику, если он поинтересуется, является ли Грид чем-то большим, чем "концепция финансирования" или, поскольку в дело вовлекается промышленность, просто рыночным слоганом. В самом деле, если, установив планировщик в своей локальной сети, я создаю "грид-кластеры", тогда установка сетевой файловой системы в той же локальной сети разве не обеспечит мне "грид-память"? И разве моя рабочая станция, где соединены процессор, память, диск и сетевая карта, не представляет собой "ПК-грид"? Существуют ли вообще компьютерные системы, не являющиеся грид-системами?

В конечном счете, Грид нужно оценивать не по его архитектуре, а по научным и коммерческим результатам, полученным при решении прикладных задач. Тем не менее, если мы хотим, чтобы грид-компьютинг завоевал доверие, нужно ответить на поставленные выше вопросы и сфокусировать внимание на том, что же требуется для его развития и процветания. В этом и других отношениях наша ситуация похожа на ситуацию с интернетом в начале 1990-х годов. Тогда поставщики сетей заявляли, что частные сети, такие как SNA и DECNET, являются частью интернета, а другие утверждали, что каждая локальная сеть является формой интернета. Эта запутанная ситуация прояснилась только тогда, когда и для глобальных, и для локальных сетей был принят интернет-протокол (IP).

### **Ранние Определения**

В 1998 году Карл Кессельман и я попытались дать определение в книге "The Grid: Blueprint for a New Computing Infrastructure." Мы писали: "Компьютерная сеть (грид) - это аппаратно-программная инфраструктура, которая обеспечивает надежный, устойчивый, повсеместный и недорогой доступ к высокопроизводительным компьютерным ресурсам".

Конечно, мы не были первыми, кто заговорил о доступе по запросу к компьютерным ресурсам, данным и услугам. Например, в 1969 году Лен Клейнрок предсказывал будущее, правда, несколько преждевременно: "Вероятно, мы скоро увидим распространение "компьютерных коммунальных услуг", которые, подобно электричеству и телефону, придут в дома и офисы по всей стране".

В следующей статье, "The Anatomy of the Grid", написанной в соавторстве со Стивом Тьюке в 2000 году, мы уточнили это определение, учитывая социальные и политические аспекты, и заявили, что грид-компьютинг - это скоординированное разделение ресурсов и решение задач в динамически меняющихся виртуальных организациях со многими участниками. Ключевая концепция - это достижение договоренности о разделении ресурсов между поставщиками и потребителями и использование полученного пула ресурсов для различных целей. Мы отметили: "Разделение, о котором мы говорим, - это не просто обмен файлами, а еще и прямой доступ к компьютерам, программному обеспечению, данным и другим ресурсам, как этого требует ряд стратегий, принятых в промышленности, науке и инженерии для совместного решения задач и выделения ресурсов. Очевидно, что это разделение должно самым тщательным образом контролироваться поставщиками и потребителями ресурсов, должно быть ясно и четко определено, что именно находится в совместном использовании, кому разрешается пользоваться разделяемыми ресурсами и на каких условиях. Отдельные пользователи и/или институты, подчиняющиеся таким правилам, образуют то, что мы называем виртуальной организацией.

Мы также говорили о важности стандартных протоколов как о средстве, обеспечивающем взаимодействие и общность инфраструктуры. Приведенные выше определения можно представить в виде простого списка критериев, в соответствии с которыми грид - это система, которая:

- координирует использование ресурсов при отсутствии централизованного управления этими ресурсами (Грид интегрирует и координирует ресурсы и пользователей, которые находятся в разных местах, например, персональный компьютер пользователя и центральный компьютер, разные административные отделения одной компании или разные компании, и направляет участникам уведомления о гарантиях, страховке, платежах, членстве и т.д. Если это не так, мы имеем дело с локальной системой управления).
- использует стандартные, открытые, универсальные протоколы и интерфейсы (Грид строится на базе многоцелевых протоколов и интерфейсов, позволяющих решать такие фундаментальные задачи как аутентификация, авторизация, обнаружение ресурсов и доступ к ресурсам. Ниже я покажу, как важно, чтобы эти протоколы и интерфейсы были стандартными и открытыми. Если это не так, мы имеем дело со специализированной прикладной системой)
- должна нетривиальным образом обеспечивать высококачественное обслуживание (грид позволяет использовать входящие в его состав ресурсы таким образом, чтобы обеспечивалось высокое качество обслуживания, касающееся, например, таких параметров, как время отклика, пропускная способность, доступность и надежность, а совместное использование ресурсов различных типов, готовых соответствовать сложным запросам пользователей, делает выгоду от

использования комбинированной системы значительно выше, чем от суммы ее отдельных частей)

Конечно, этот список может вызвать справедливую полемику, в частности, могут возникнуть вопросы, что значит "централизованное управление", "стандартные, открытые, универсальные протоколы" и "качество обслуживания". Я скажу об этом ниже, но сначала позвольте мне применить мои критерии к нескольким кандидатам в "грид-системы".

Давайте сначала рассмотрим системы, которые согласно моим критериям нельзя квалифицировать как грид-системы. Системы управления кластером, такие как Sun Grid Engine, Load Sharing Facility или Portable Batch System, установленные на параллельной вычислительной машине или в локальной сети, гарантируют качество обслуживания и, таким образом, представляет собой мощный грид-ресурс. Однако, такую систему нельзя рассматривать как грид, поскольку она осуществляет централизованное управление хостами: она обладает полной информацией о состоянии системы и запросах пользователей, а также полностью контролирует отдельные компоненты. Нельзя рассматривать как грид и Web: ее открытые, универсальные протоколы поддерживают доступ к распределенным ресурсам, но не координируют использование этих ресурсов для получения интересующего нас качества обслуживания.

С другой стороны, к грид-системам первого поколения с достаточным основанием можно отнести такие разработки, как мультисайтовые планировщики, например, система MultiCluster, распределенные вычислительные системы, поддерживаемые Condor, Entropia и United Devices, использующие простаивающие рабочие станции; системы peer-to-peer (P2P), такие как Gnutella, которые поддерживают совместное использование файлов узлами-участниками; федеративный ресурс-брокер, обеспечивающий распределенный доступ к данным. Несмотря на то, что протоколы, используемые в этих системах, вероятно, слишком специализированы, чтобы отвечать второму критерию (и не являются, по большей части, ни открытыми, ни стандартными), каждая из этих систем действительно интегрирует распределенные ресурсы при отсутствии централизованного управления и предоставляет интересующее нас качество обслуживания, хоть и в ограниченных областях.

Соответствие систем трем критериям наиболее четко прослеживается применительно к крупномасштабным грид-разработкам в научном сообществе, таким как распределенные системы обработки данных, входящие в проекты "DataGrid" (GriPhyN, PPDG, EU DataGrid, iVDGL, DataTAG), NASA's Information Power Grid, распределенная система ASCI Supercomputer (DAS-2), связывающая кластеры пяти голландских университетов, DOE Science Grid и DISCOM Grid, связывающие системы в лабораториях министерства энергетики (DOE), и система TeraGrid, созданная для связи основных академических сайтов США. Каждая из этих систем интегрирует ресурсы многих институтов, каждая имеет собственную стратегию и механизмы, каждая использует открытые, универсальные (Globus Toolkit) протоколы для реализации и управления разделением ресурсов и поддерживает высокое качество обслуживания в

различных направлениях, включая безопасность, надежность и производительность.

Грид: Критерий необходимости межсетевых протоколов в моем списке критериев относится к понятию "грид", в то время как в названии этой статьи задается вопрос, что такое Грид. Это важное различие. Представляется, что в проекте Грид потребуется использование протоколов (интерфейсов и политики), которые являются не только открытыми и универсальными, но и стандартными. Именно стандарты позволят нам динамически устанавливать соглашения о разделении ресурсов с любой из заинтересованных сторон и тем самым создавать нечто большее, чем множество разрозненных, несовместимых, не взаимодействующих разбросанных систем. Стандарты так же важны, как и средства, делающие возможным создание универсальных сервисов и инструментальных средств.

С моей точки зрения, определение стандартных интергрид-протоколов - это единственная наиболее критичная проблема, стоящая сегодня перед грид-сообществом. К счастью, мы достигли большого прогресса. Что касается стандартов, то здесь мы имеем растущую эффективность деятельности Global Grid Forum. С практической стороны, шесть лет экспериментов и усовершенствований фактически привели к широкому использованию стандарта - открытого Globus Toolkit. И теперь основные усилия Global Grid Forum направлены на определение Открытой архитектуры грид-служб (OGSA - Open Grid Services Architecture), которая модернизирует и расширит протоколы Globus Toolkit, что позволит удовлетворить новые возникающие требования, сохраняя, в то же время, и работу с Web-службами. Компании IBM, Microsoft, Platform, Sun, Avaki, Entropia и United Devices выразили твердую поддержку OGSA. Надеюсь, в ближайшем будущем мы сможем утверждать: чтобы интегрироваться в среду Грид, нужно использовать интергрид-протоколы OGSA, точно так же, как для вхождения в интернет нужно (кроме всего прочего) использовать протокол IP. Как открытые, так и коммерческие продукты будут эффективно взаимодействовать в этом неоднородном, разнообразном грид-мире, создавая всеобъемлющую инфраструктуру, которая обеспечит успех различным грид-приложениям.

Оригинальный текст:

Ian Foster “WHAT IS THE GRID? A THREE POINT CHECKLIST”

<http://www-fp.mcs.anl.gov/~foster/Articles/WhatIsTheGrid.pdf>