

## **Грид: приложение семантического WEB'a**

Кэрол Гобле,  
факультет компьютерных наук Манчестерского университета

Дэвид де Руж,  
школа электроники и компьютерных наук при Саутгемптонском университете

## **The Grid: An Application of the Semantic Web**

Carole Goble Department of Computer Science  
University of Manchester  
Oxford Road  
Manchester, M13 9PL, UK

[carole@cs.man.ac.uk](mailto:carole@cs.man.ac.uk)

David De Roure  
Dept of Electronics & Computer Science  
University of Southampton  
Southampton,  
SO17 1BJ, UK

[dder@ecs.soton.ac.uk](mailto:dder@ecs.soton.ac.uk)

<http://www.semanticgrid.org/documents/sigmod/ami9.pdf>

## Содержание

<b>Резюме.....</b>	<b>2</b>
<b>Что такое Грид? .....</b>	<b>3</b>
Грид-службы .....	4
<b>Что такое семантический Web?.....</b>	<b>5</b>
Грид и семантический Web.....	6
<b>Семантический Web для инфраструктуры грида .....</b>	<b>6</b>
Службы семантического грида.....	6
Интеграция информации.....	6
<b>Семантический Web для грид-приложений.....</b>	<b>8</b>
Источники, качество, доверие и доказательство .....	8
<b>Исследовательские проблемы и возможности.....</b>	<b>9</b>
<b>Признательности .....</b>	<b>12</b>
<b>Ссылки.....</b>	<b>12</b>

## Резюме

Грид представляет собой развивающуюся платформу, которая поддерживает создание по требованию "виртуальных организаций" для координированного разделения ресурсов и решения задач на глобальном уровне. Современные запросы научных приложений к исполняющим платформам весьма значительны, а масштаб и сложность научных данных является для баз данных трудной проблемой. Грид начинает использовать разработанные для Web-служб технологии, понимая их потенциальные возможности, и предпринимает попытки извлечь пользу из технологий семантического Web'a; в свою очередь грид и его высококвалифицированные пользователи создают такие приложения, которые благотворно влияют на развитие семантического Web'a.

## Что такое Грид?

*Грид это "гибкое, надежное, скоординированное разделение ресурсов между динамическими группами отдельных лиц, институций и ресурсов, которые мы называем виртуальными организациями." Фостер и др., 2001[1].*

Научные и инженерные разработки в настоящее время производятся при совместной деятельности ученых с привлечением гетерогенных компьютерных ресурсов, информационных систем и инструментов, разрозненных в географическом и организационном отношении. "Грид" это новая платформа обеспечения координированного распределения ресурсов и решения задач большого масштаба для приложений, обрабатывающих огромный объем данных в режиме интенсивного исполнения [1]. Само название возникло по аналогии с гридом электроснабжения: компьютерные ресурсы и ресурсы данных предполагается получать по Интернету бесшовно, прозрачно и динамично по мере надобности, так же, как мы пользуемся электричеством. Таким образом, главной мотивацией грида является упрощение рутинного взаимодействия ресурсов в интересах выполнения крупномасштабных научных и инженерных задач.

В самом начале грид был нацелен на разделение вычислительных мощностей и ресурсов, ориентируясь на актуальные запросы науки и инженерии. "Метакомпьютерные" проекты начала 90 годов занимались построением виртуальных компьютеров на базе сетевых компьютерных систем. В качестве основных потребителей были (и по сей день продолжают быть) приложения фундаментальной науки. Например, транснациональные эксперименты, такие, как поиск бозонов Хиггса (Higgs boson)<sup>1</sup> на большом адронном коллайдере. На этом устройстве генерируются петабайты данных в течение нескольких секунд, и комплексный анализ может занимать месяцы вычислительной обработки [2].

Увеличение вычислительной мощности за счет объединения большого числа географически удаленных систем поднимают вопросы масштабируемости и гетерогенности. С масштабируемостью связан ряд трудных проблем: неизбежность отказа компонент, необходимость автоматизации, необходимость обращать внимание на локальность ресурсов в связи с латентностью сети и увеличивающееся число организационных границ, что повышает значимость вопросов аутентификации и доверия.

Приложения большого масштаба могут появляться в результате композиции других приложений, что увеличивает сложность систем. Задачи гетерогенности решаются промежуточным программным обеспечением, таким, как Globus Toolkit [3], обеспечивающим однородность с помощью стандартного набора интерфейсов для нижележащих ресурсов.

Раннее промежуточное программное обеспечение грида использовало линейку протоколов, таких, как LDAP, для служб каталогов и запросов к хранилищам файлов [4], GridFTP для надежной передачи данных большого объема и SSL для обеспечения надежности. Функциональность более высокого уровня, такая, как толерантная репликация масштабируемых данных [5], пользуется этими протоколами. Было также обращено внимание на преобладающее использование грида в отношении данных, а не

---

<sup>1</sup> По стандартной, ныне существующей модели физики самая мельчайшая частица всего - Хиггс бозон, или "божественная частица". Существует пока лишь в теории, которую хотят доказать с помощью строящегося сверхпроводникового суперускорителя частиц.

объема вычислений; например, брокер запросов памяти предоставляет приложениям унифицированный доступ к распределенному хранилищу файлов [6]. Впрочем, деятельность исследователей и разработчиков грида обычно была сосредоточена на приложениях, где данные хранились в файлах; слабо поддерживались транзакции, доступ к реляционным базам данных или распределенная обработка запросов [7].

Теперь в грид-сообществе активно разрабатываются фундаментальные механизмы для интерактивного взаимодействия ресурсов любого типа, включая документы, базы данных, инструменты, архивы и людей. Поддержка взаимодействия данных сосредоточена на корректном и скоординированном доступе к базам данных из грид-приложений [8]. Частично это связано с применением гридов научными дисциплинами, отличными от физики элементарных частиц, (такими, как биология, наука о Земле, химия, астрономия), для которых большее значение имеет не объем данных, а необходимость их интеграции (смотри Buttler et al).

### Грид-службы

Промежуточное программное обеспечение гридов должно обеспечить динамическое и прозрачное построение новых возможностей из распределенных служб. Инженерия новых грид-приложений предполагает повторное использование существующих компонент и информационных ресурсов так, чтобы ассемблирование и координация этих компонент были достаточно гибкими. Отчасти по этой причине в гридах преобладает тенденция опоры не на набор протоколов, а на ориентированный на службы подход: с использованием Открытой Архитектуры Грид Служб (OGSA, Open Grid Services Architecture) [9]. Он объединяет Web-службы с требованиями и методами грида.

Из требований грида вытекает существенное развитие грид-службами свойств Web-служб. Конфигурации грид-служб должны быть:

- *динамичными и подвижными.* Совокупность служб (баз данных, сенсоров, вычислительных серверов), участвующих в сложном анализе, может подключаться и отключаться по мере прекращения их использования;
- *соответствующими надобностям ad hoc.* У объединения служб нет централизованного места пребывания, централизованного управления и постоянных доверительных отношений;
- *большими.* В любое время в оркестровке могут принимать участие сотни служб;
- *долгосрочными.* Моделирование может длиться неделями.

Из этих требований вытекает жесткая необходимость обеспечить восстанавливаемость после сбоев, надежность, безопасность и производительность [9]. В то время как Web-службы по умолчанию доступны и не имеют состояний, грид-службы по умолчанию преходящи и обладают состояниями.

Грид-службы в широком смысле образуют четырехярусную структуру:

1. остов (надежность, транспортировка данных, сертификация, удаленный доступ, сетевой мониторинг, право собственности и цифровые водяные знаки, аутентификация);
2. база (распределение ресурсов, обработка метаданных, определение версий);

3. высокий уровень (рабочая нагрузка, управление базами данных, персонализация);
4. приложение (упорядочение последовательности генов, база данных Swiss-Prot<sup>2</sup>, алгоритм обнаружения генов).

Каждый ярус существенно использует метаданные. Гибкое ассемблирование грид-служб требует наличия информации о функциональности, доступности и интерфейсах самых различных служб. Обнаружение служб и брокерское обслуживание используют описания на уровне метаданных [10]. Композиция служб управляется и поддерживается описаниями метаданных [11] В реализации полного видения грид-служб метаданные являются ключевым аспектом.

### Что такое семантический Web?

*Семантический Web – это "...расширение традиционного Web'a в направлении существенно лучшего определения смысла информации, позволяющего компьютерам и людям эффективнее выполнять совместную работу. Мы хотим, чтобы данные на Web'e были определены и связаны ссылками так, чтобы их можно было легче находить, интегрировать, автоматизировать и повторно использовать в различных приложениях... чтобы данные были разделяемыми и могли обрабатываться как автоматизированными средствами, так и людьми". World Wide Web Consortium [12].*

Конечная амбициозная цель состоит в создании такой среды, где программные агенты могут динамически обнаруживать и опрашивать ресурсы, а затем взаимодействовать с ними. Агенты должны уметь справляться с возникающими виртуальными проблемами интеллектуализированной среды, обнаруживать новые факты и выполнять изощренные задания, получаемые от людей [13]. Простые метаданные и простые запросы как-то нам помогают в решении задач интеграции информации, но эта помощь далеко недостаточна [14]. В том или другом случае, простом или сложном, автоматизированная обработка содержимого Web'a требует явной, связанной с ресурсом машиннообрабатываемой семантики для описания того, что относится к "о чем" и что относится к "для чего".

Семантический web можно представлять в виде трехярусной структуры:

1. остов, состоящий из уникальной глобальной *идентификации* ресурса; *метаданных* для подтверждения сведений о ресурсах и обслуживания запросов на такие сведения, и простого языка описания метаданных и знаний. Метаданные и знания содержатся в *онтологиях*, обеспечивающих взаимопонимание и общепринятый словарь для метаданных и правил *вывода* новых метаданных и знаний;
2. базовые службы для выполнения стандартных умозаключений и целенаправленного просмотра метаданных и онтологий, объяснения выполненных умозаключений. К этому ярусу относится также управление доверительностью, агенты, поисковые машины, серверы онтологий;
3. прикладные службы; например, служба агента бюро путешествий.

<sup>2</sup> Swiss-Prot – это биологическая база данных протеиновых последовательностей. Создана в 1986 году. (прим. перевод.)

Прилагается много усилий к разработке языка и технологий для остова и базовых служб; здесь стоит отметить RDF(S)[15], DAML+OIL[16] и OWL[17].

### **Грид и семантический Web**

До недавнего времени сообщества грида и семантического Web'a работали отдельно, несмотря на сближающиеся характеристики их видения соответствующих задач. Обоим сообществам необходимы вычислительная доступность и разделяемость метаданных для поддержки автоматического обнаружения информации, интеграции и агрегации. Оба оперируют в глобальной, распределенной и изменяющейся среде.

Базовые службы семантического Web'a могут стать базовыми службами грида. Остов семантического Web'a является тем средством, с помощью которого в гриде могут быть представлены метаданные: как в отношении инфраструктуры грида, управляя аппаратурой остова грида и его службами базового и высокого уровня, так и в отношении грид-приложений, представляя знания и операционное ноухау прикладной области.

### **Семантический Web для инфраструктуры грида**

#### **Службы семантического грида**

Описание службы существенно для автоматического обнаружения и поиска, выбора, сопоставления, композиции и взаимодействия, мониторинга вызова и исполнения. Все эти операции зависят от метаданных служб. Классификация служб на основе предоставляемой ими функциональности признана в различных сообществах как эффективный способ нахождения подходящих служб, например UDDI. Анализ описания служб играет определенную роль при классификации и сопоставлении служб. В системе Condor [11] механизм сопоставления используется при выборе вычислительных ресурсов. В архитектуре, где службы крайне подвижны, а конфигурация служб постоянно реорганизуется, очень важно знать, может ли одна служба безопасно быть заменена другой.

На момент написания этой статьи положение дел в описании грид-служб с помощью семантики сводится к использованию имен, приписываемых элементам Тип порта и Тип службы в документе WSDL, ссылающемся на документ спецификации [9]. Сближение семантического Web'a и Web-служб уже привлекло к себе внимание (смотрите в этой статье *Bussler et al*). DAML+OIL было исследовано в [1] и [10]. Здесь онтология служб является расширением онтологий DAML-S [18]. Классификация служб более выразительна, чем простые иерархии UDDI, службы ищут и сравнивают методом отнесения к категориям, используя описания служб. Впрочем, грид-службы динамически создают и ликвидируют экземпляры служб; им присущ характер мягкой регистрации и они могут создавать долгосрочные конфигурации служб. Как технологии семантического Web'a могут описывать и обнаруживать грид-службы, является проблемой, еще ждущей адекватного решения.

#### **Интеграция информации**

Для решения сложных, задаваемых учеными вопросов требуется изощренный анализ и синтез сведений из различных, независимо разрабатываемых и гетерогенных ресурсов. В биологии, например, сотни репозитариев данных активных служб имеют

различные форматы, интерфейсы, структуры, содержательные объемы и т.д. (смотрите в этой статье *Bussler et al*). Web и грид данных гарантируют определенный уровень интероперабельности при поиске данных и доступе к ним. Очередной уровень интероперабельности состоит не только в обеспечении доступа к данным, но и в понимании того, что означают эти данные, с тем, чтобы их можно было объединять в сущностно корректном аспекте; необходимо обеспечить автоматизированную поддержку этого процесса интеграции [19].

Обычно ученые связывают ресурсы двумя способами:

1. *оркестровка рабочей нагрузки*: согласование процессов, т.е. координация рабочей нагрузки и построение цепочки служб с использованием систематического плана, это демонстрация экспериментов в компьютерном моделировании; оно позволяет нам представлять экспериментальный процесс e-ученого явно;
2. *интеграция баз данных*: динамически распределенная обработка запросов, т.е. создание интегрированных баз данных при посредстве виртуальной федерализации или промежуточного хранилища данных [20].

Информационное посредничество не сводится к традиционным научным базам данных. Вычислительные ресурсы обнаруживаются, включаются и отключаются динамически и прозрачно для пользователя. Проблема посредничества между различными брокерскими моделями для вычислительных грид-ресурсов, такими, как Unicore и Globus, очень напоминает посредничество между двумя схемами баз данных.

Семантический Web и технологии баз данных открывают очень большие возможности. Общая информационная модель агрегации результатов, взятых из различных ресурсов или инструментов, могла бы воспользоваться доменными онтологиями RDF для семантического посредничества между схемами баз данных [19]; ввод и вывод в приложениях и рабочие элементы рабочей нагрузки могли бы взять DAML+OIL/RDF(S).

Доменные онтологии и правила можно привлечь к управлению параметрами аппаратуры и алгоритмов, а также и конфигурирования, допускающего достаточную свободу выбора. С привлечением аппарата логического вывода может быть значительно повышена эффективность планов выполнения рабочей нагрузки и различных способов комбинирования служб, что также обеспечивает семантическую корректность работающих композиций [21].

Итак, семантические Web-службы могут быть использованы для:

- классификации вычислительных ресурсов и ресурсов данных, метрик выполнения, управления работами;
- интеграции схем, описания рабочей нагрузки;
- распечатки данных и служебного ввода/вывода;
- выбора методов решения проблем и интеллектуальных порталов;
- инфраструктуры для аутентификации, регистрации и управления доступом.

Рассматривая вопрос с другой стороны, мы можем представить себе, что базовые службы и прикладные службы семантического Web'a реализованы как грид-службы.

## **Семантический Web для грид-приложений**

Конечной целью грида является поддержка обнаружения знаний. Семантический Web часто представляют как глобальную базу знаний. Рассмотрим следующий сценарий: некий ученый задает вопрос: "какие протеины суперсемейства АТР мы находим у мышей?" и может получить один из следующих ответов: (a) у нее есть разрешение на доступ к базе данных Swiss-Prot по коду доступа к протеинам; (b) InterPro является базой образцов, но необходимо разрешение и предварительная оплата; (c) проект Attwood занимается нуклеотидами при связывании протеинов (протеиновое суперсемейство АТРase это вид протеинов с нуклеотидным связыванием); (d) две недели назад Смит опубликовал новую статью о чем-то подобном в Nature Genetics; (e) сотрудник вашей лаборатории Джон уже задавал этот вопрос на прошлой неделе.

Ученому могут сообщить об оборудовании или настройке параметров алгоритма, помочь при выборе и планировании подходящих экспериментов и ресурсов, учитывая цели и общепринятую лучшую практику. Ему могут гарантировать, что будут делаться только такие заключения, которые полностью соответствуют используемой технике. Все это является приложениями семантического Web'a или для семантического Web'a; они включают индивидуализированных агентов или службы, семантические порталы служб, рекомендационные системы и большое разнообразие других служб знаний [22].

Научное сообщество встретило Web с большим удовольствием. В результате появилось большое количество обычных публикаций, содержащих информацию без сопровождающего аппарата. У многих ресурсов очень простой поисковый интерфейс без API или языка запросов и только визуальный интерфейс "укажи и кликни". Научное знание часто является обычной публикацией с сопровождающими текстовыми "аннотациями". Традиция чтения и интерпретации текстов с трудом поддается переходу на автоматизированную обработку и при огромном объеме доступных данных становится нежизнеспособной. Семантический Web как раз и нацелен на то, чтобы сделать вычислительно недоступное доступным и автоматизировать обнаружение информации.

## **Источники, качество, доверие и доказательство**

Как результаты, так и способы их получения одинаково ценны. Откуда мы получаем данные, кто их создал, когда, почему и как они были получены, так же важно для пользователя и провайдера службы, как и сами данные [23]. Этими вопросами занимаются такие приложения, как Доказательства, Доверие и Цифровые подписи на семантическом Web'e (смотри Maximilien и Finin). В молекулярной биологии данные по мере прохождения через бесчисленные базы повторно копируются, корректируются и трансформируются. За публикуемыми данными активно следят как автоматы, так и люди. Сложные компиляции программ получают результаты из баз данных. Возможными приложениями семантического Web'a и технологиями баз данных будут приложения по сопровождению аннотаций комментариями и ссылками на источники, по объяснению, почему выбранные параметры использовались при работе алгоритма и каковы эти параметры.

Аналитические высказывания относятся также и к качеству. Научное знание контекстуально и субъективно. Контексты меняются, а мнения противоречивы. Новая информация может как поддерживать, так и противоречить господствующей ортодоксии, приводя к ревизии верований. Анализ этих аналитических высказываний может привести к появлению нового знания, но этот анализ должен быть для всех доступен, иначе ученые не будут им пользоваться.

Итак, мы можем использовать семантический Web для:

- аннотирования результатов, рабочей нагрузки, записей баз данных и выбранных при анализе параметры с: персональными примечаниями, данными о создании, путями прохождения информации, объяснениями или претензиями;
- подсоединения *in silico* и "на коленке" экспериментальных компонент, таких, как: литература, замечания, базы данных, промежуточные результаты, наброски, изображения, рабочая нагрузка, проводящий эксперимент сотрудник, производственная лаборатория, конечный отчет;
- описания людей, лабораторий, литературы, средств и научных знаний.

Для надежного хранения научные знания дублируются и архивируются. Очень важно уметь представлять себе срез состояния умов в данный момент в данном месте для осознания научной точки зрения, господствующей в это время. Возникает вопрос: каков информационный мусор при сборе "семантического грида", и как мы в дальнейшем сможем найти этот срез?

Грид-службы приходят и уходят, вот почему служба оповещения о событиях является базовой грид-службой. По мере развития коллекций данных и аналитических приложений проследить последствия изменений становится все труднее. При модификации баз данных ученые вынуждены заново формулировать свои запросы, новые знания изменяют фактическую область анализа. Ошибки или потерявшая достоверность информация продолжают распространяться и с трудом поддаются элиминации. Онтологии и правила также меняются. Новые верования заставляют менять онтологии, но прежние выводы и заключения, перестающие быть справедливыми, при этом не исчезают автоматически (и как вообще мы следим за влиянием этих изменений?). Они должны продолжать сосуществовать и быть доступными. Мониторинг событий и сущностей может быть описан с помощью онтологий; с помощью триггеров<sup>3</sup> баз данных можно реализовать механизм оповещения.

## Исследовательские проблемы и возможности

Развитие Web'a стимулировалось физикой элементарных частиц. Данное сообщество физиков было хорошо организованным микрокосмом всего общества. У него была определенная и четко артикулированная необходимость в распространении информации, и у него была группа понимающих людей, готовых к сотрудничеству с помощью нового средства и желающих делать это. Сегодня развлечения с гридом чем-то напоминают нам Web недавно минувших лет. Кое-как, но грид развертывается, большей

<sup>3</sup> Триггеры базы данных могут использоваться для дополнения системы управления базой данных. Например, триггер базы данных может быть создан для того, чтобы разрешить модификацию таблицы только в обычные часы работы.

частью за счет энтузиазма членов научного сообщества (и здесь снова Физика высоких энергий), появляются стандарты и кое-какая коммерческая поддержка. То же самое можно было бы сказать о текущем положении дел в развертывании семантического Web'a, хотя и не ясно имеется ли такой же драйв, который имел место для Web'a и грида.

Тем временем сам Web привлек к себе всесторонний интерес и продолжает интенсивно развертываться; например: сдвиг в сторону коммуникации компьютер-человек (HTML) и компьютер-компьютер (XML), возникновение парадигмы web-служб. Требования одной из побудительных сфер, e-коммерции, совпадает с вектором требований e-науки. Но типичное грид-приложение не является типичным Web-приложением. Грид-приложение может включать большое количество процессов, взаимодействующих скоординированным образом, тогда как типичная Web-транзакция пока все еще обслуживается небольшим количеством хостов (например, сервер, кэш, браузер). Более того, грид-процессы постоянно появляются и исчезают. Обеспечение желаемого поведения распределенной системы большого масштаба требует решения таких технических проблем, с которыми самому Web'у не приходилось сталкиваться; впрочем, Web-службы заставляют нас двигаться в аналогичном направлении.

Грид полагается на службы, умеющие работать с метаданными. Семантический Web требует умеющего работать с метаданными Web'a. Точно так же, как в последние несколько лет, компоненты информационных систем подвигли на поддержку HTML и XML, также необходимо теперь их поддерживать для создания и сопровождения метаданных. Препятствий предостаточно. Ручное создание метаданных проблематично. Различные обстоятельства и ошибки не всегда позволят людям выдавать аккуратные метаданные; всегда будут альтернативные и равнокорректные описания. Грид-приложения должны и могут автоматизировать работу с *высококачественными* метаданными. Первостепенное значение имеет решение проблемы создания, управления и объединения онтологий.

Симбиоз грида и семантического Web'a достаточно очевиден, перспективы их развития связаны. От семантического Web'a грид заимствует остов и службы для работы со своей семантикой. Сам семантический Web стимулируется мощным развитием грид-приложений и имитирует инфраструктуру грида. Базовые службы семантического Web'a - серверы онтологий, генераторы метаданных, согласование онтологий и т.д. – могут быть реализованы как Web-службы (Рис 1)

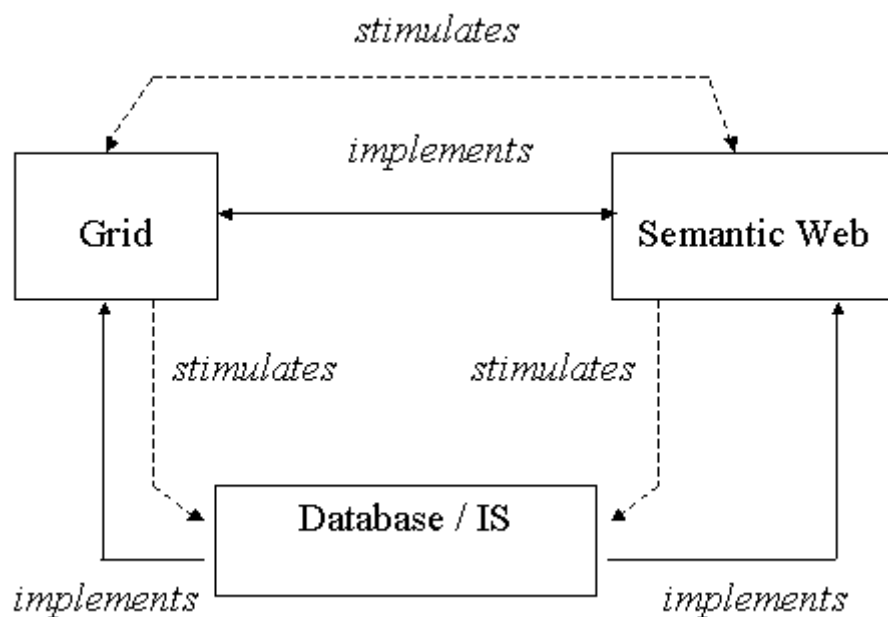


Рис. 1: Грид, Семантический web и DB/IS

Технологии баз данных и информационных систем являются существенными элементами при построении такой глобальной и распределенной инфраструктуры, которая необходима как для грида, так и для семантического Web'a. Для этого требуется, чтобы грид-сообщество, занимающееся компьютерингом и приложениями, обратило должное внимание на семантический Web. Это относится и к вертикальным проектам, где технологии семантического Web'a могут быть применены в области приложений, и к разработке промежуточного программного обеспечения, которое может быть встроено в инфраструктуру семантического Web'a. Внедрение новых технологий стоит денег, а выгода не всегда проявляется сразу. Сообщество семантического Web'a должно позаботиться о поддержке начального импульса, не забывая, что многие разработчики традиционного грида придерживаются системно ориентированной точки зрения и смотрят на технологии знаний как на нечто второстепенное. Одним из барьеров является различное понимание того, что можно сделать сразу и к чему можно отнести с позиции "поживем-увидим".

Почему специалисты по семантическому Web'у и базам данных должны интересоваться гридом?

1. Это хороший пример типа приложения, присущего семантическому Web'у. Суть грида в порождении новых больших возможностей при *широкомасштабной интеграции ресурсов*. Предоставляемая гридом высочайшая автоматизация с необходимостью требует "универсально доступной платформы, позволяющей как автоматизированным средствам, так и людям разделять и обрабатывать данные".
2. Это реальное приложение, в нем особое внимание обращено на развертывание и производительность, оно масштабно, обладает сформированным сообществом пользователей. Такие приложения

необходимы при становлении семантического Web'a и нуждаются в технологиях баз данных.

3. Потенциал грида неотделим от технологий семантического Web'a. Уже занимаясь самым нижним уровнем грида, его разработчики признают, что им приходится иметь дело с возникающими "информационными островками" и это требует интероперабельных решений.
4. Гриду будет интересно, что происходит на семантическом Web'e, где будут решаться задачи, требующие глубокого исследования и имеющие специфическое значение и для грида. Маловероятно, чтобы решения этих задач принадлежали исключительно области грид-компьютинга - в некоторых приложениях семантического Web'a они будут очевидны.
5. Это самодостаточный мир четко организованного сообщества, работающего с общими инструментами и стандартами.
6. Некоторые стороны деятельности семантического Web'a могут стать приложениями грид-компьютинга, это относится, например, к поиску, извлечению информации, трансляции, выбору информации мультимедиа.

От партнерства семантического Web'a и грида можно ожидать захватывающего зрелища. В продвижении вперед у каждого свои трудности, но успех одного окажется успехом и другого. Обоим нужны DB/IS, оба будут заниматься этими исследованиями.

## Признательности

Данная работа спонсировалась EPSRC и DTI по программе е-науки UK. Автор хочет поблагодарить всех своих коллег по этому проекту. Особенно благодарю Nigel Shdbolt, Nick Jennings и Sean Bechhofer.

## Ссылки

- [1] I. Foster, C. Kesselman, S. Tuecke *The Anatomy of the Grid: Enabling Scalable Virtual Organizations*, International Journal of Supercomputer Applications, 15(3), 2001.
- [2] S. Bethke, M. Calvetti, H.F. Hoffmann, D. Jacobs, M. Kasemann, D. Linglin *Report of the Steering Group of the LHC Computing Review*, CERN/LHCC/2001-004, February 2001.
- [3] I. Foster and C. Kesselman, *Globus: A Metacomputing Infrastructure Toolkit*, Int. Journal of Supercomputer Applications, 11(2): 115-128, 1997.
- [4] K. Czajkowski, S. Fitzgerald, I. Foster, C. Kesselman. *Grid Information Services for Distributed Resource Sharing*. Proceedings of the Tenth IEEE International Symposium on High-Performance Distributed Computing (HPDC-10), IEEE Press, August 2001.
- [5] A. Chervenak, E. Deelman, I. Foster et al *Giggle: A Framework for Constructing Scalable Replica Location Services*, SC2002, November 11-16, 2002, Baltimore, Maryland.

- [6] A. Rajasekar, M. Wan and R. Moore, *MySRB & SRB - Components of a Data Grid*, The 11th International Symposium on High Performance Distributed Computing (HPDC-11) Edinburgh, Scotland, July 24-26, 2002.
- [7] P. Watson, *Databases and The Grid*, UK e-Science Programme Technical Report Number UKeS-2002-01.
- [8] N. Paton, M. Atkinson, V. Dialani, D. Pearson, T. Storey, P. Watson, *Database Access and Integration Services on the Grid*. UK e-Science Programme Technical Report Number UKeS-2002-03.
- [9] I. Foster, C. Kesselman, J. Nick and S. Tuecke, *The Physiology of the Grid: Open Grid Services Architecture for Distributed Systems Integration*, GGF4, Feb. 2002. See <http://www.globus.org/alliance/publications/papers/ogsa.pdf>
- [10] C. Wroe, R. Stevens, C. Goble, A. Roberts, M. Greenwood, *A suite of DAML+OIL Ontologies to Describe Bioinformatics Web Services and Data*, International Journal of Cooperative Information Systems *in press*.
- [11] R. Raman, M. Livny, and M. Solomon. *Matchmaking: An extensible framework for distributed resource management*. Cluster Computing: The Journal of Networks, Software Tools and Applications, 2:129-138, 1999.
- [12] W3C Semantic Web Activity Statement, <http://www.w3.org/2001/sw/Activity>
- [13] J. Hendler, *Agents and the Semantic Web*, IEEE Intelligent Systems Journal, March/April 2001 (Vol. 16, No. 2), pp. 30-37.
- [14] B. McBride, "Four Steps Towards the Widespread Adoption of a Semantic Web", in Proceedings of the First International Semantic Web Conference (ISWC 2002), Sardinia, Italy, June 9-12, 2002. LNCS 2342, pp 419-422.
- [15] I. Horrocks, *DAML+OIL: a reason-able web ontology language*, in Proceedings of EDBT 2002, March 2002.
- [16] *Resource Description Framework* <http://www.w3.org/RDF>
- [17] OWL Web Ontology Language 1.0 Reference <http://www.w3.org/2001/sw/WebOnt/>
- [18] A. Ankolekar, M Burstein, J Hobbs, Lassila, D Martin, S McIlraith, S Narayanan, M Paolucci, T Payne, K Sycara, H Zang *DAML-S: Semantic Markup for Web Services* in Proceedings of the International Semantic Web Working Symposium (SWWS), July 30-August 1, 2001.
- [19] C.A. Goble, *Supporting Web-based Biology with Ontologies*, in Proceedings of the Third IEEE ITAB00 Arlington, VA (November 2000), pp. 384-390.

[20] I. Foster, J. Voeckler, M. Wilde, and Y. Zhao. *Chimera: A Virtual Data System for Representing, Querying and Automating Data Derivation*. Proceedings of the 14th Conference on Scientific and Statistical Database Management, Edinburgh, Scotland, July 2002.

[21] J. Cardoso and A. Sheth, *Semantic e-Workflow Composition*, Technical Report, LSDIS Lab, Computer Science, University of Georgia, July 2002.

[22] D. De Roure, N. Jennings, N. Shadbolt. *Research Agenda for the Semantic Grid: A Future e-Science Infrastructure*, UK e-Science Programme Technical Report Number UKeS-2002-02.

[23] P. Buneman, S. Khanna, K. Tajima, W-C Tan. *Archiving Scientific Data*, SIGMOD Conference 2002.