

Семантический Грид для науки и инженерии

Исмагилов Е.А.

ДА-52М

Использование семантического грид для науки и инженерии

- ▶ Какие возможности семантической технологии представляют интерес для науки и инженерии:
 - ▶ отделение содержания от представления
 - ▶ машинная обработка
 - ▶ логический вывод
- ▶ Примеры задач, которые могут решаться с помощью семантической технологии:
 - ▶ возможность повторения уже проведенных экспериментов
 - ▶ поиск схожих по характеристикам ресурсов и сервисов
 - ▶ улучшенный поиск данных и информации
 - ▶ автоматизация выполнения экспериментов и вычислений

варианты использования семантического Грид

- ▶ интеграция данных
- ▶ обнаружение ресурсов
- ▶ поиск информации



Интеграция данных

▶ Подзадачи:

- ▶ виртуализация данных
- ▶ агрегирование данных
- ▶ стандартизация

▶ Рекомендации:

- ▶ стоит использовать в случае построения автономного ГРИД
- ▶ при разработке стоит обращать внимание на специфику данных
- ▶ учитывать усложнение системы и влияние на производительность
- ▶ в случае когда планируется использование множества форматов или постепенное расширение их количества, стоит использовать решение с использованием промежуточного преобразования

Обнаружение ресурсов

▶ Подзадачи:

- ▶ классификация вычислительных ресурсов и ресурсов данных, метрик выполнения, управления работами
- ▶ отслеживание, фильтрация, анализ и поиск сервисов и ресурсов
- ▶ разграничение объектов ресурсов с помощью онтологий
- ▶ составление информационного потока соответственно подходам оркестровки и хореографии
- ▶ интеграция схем, описание рабочей нагрузки
- ▶ распечатка данных и служебного ввода/вывода
- ▶ выбор методов решения проблем
- ▶ предоставление инфраструктуры для аутентификации, регистрации и управления доступом
- ▶ обеспечение семантической совместимости
- ▶ помощь в обеспечении самоуправляемости

Обнаружение Ресурсов

▶ Рекомендации:

- ▶ стоит использовать в случае построения автономного ГРИД
- ▶ для максимальной эффективности и гибкости семантические системы ОР должны соответствовать современным требованиям ГРИД: совместимость, масштабируемость, децентрализация и динамизм
- ▶ включить в систему поддержку сравнений как на основе ключевых слов, так и на основе семантического сравнения
- ▶ использовать обновление по требованию в случае использования централизованной системы регистрации

Поиск информации

▶ Подзадачи:

- ▶ эффективный поиск на крупных гетерогенных данных
- ▶ использование операций логического вывода

▶ Рекомендации:

- ▶ использовать в случае, когда требуется выполнять сложные, логически обоснованные запросы на гетерогенных данных
- ▶ учитывать усложнение системы и влияние на производительность
- ▶ обращать внимание на удобство пользовательского интерфейса

Недостатки Семантического Грид

- ▶ дублирование информации
- ▶ большая трудоёмкость создания онтологий
- ▶ более медленная работа из-за выполнения операций логического вывода
- ▶ отсутствие поддержки семантической технологии в большинстве популярных промежуточных слоях ГРИД
- ▶ отсутствие мотивации разработчиков
- ▶ невозможность гарантировать достоверность метаданных
- ▶ при добавлении семантической технологии в существующий проект нужно будет добавлять семантическое описание для существующих данных и ресурсов

Реализованные проекты

- ▶ MyGrid, Ontology Gateway, WorldWide Molecular Matrix – биоинформатика
- ▶ CASA, LEAD – метеорология
- ▶ CombeChem – вычислительная химия
- ▶ InteliGrid – платформа для семантического ГРИД

Спасибо за внимание