

Гончаренко В.С. — рецензент Гемба О.В.

УНК “Институт прикладного системного анализа” НТУУ “КПИ”, Киев, Украина

Использование онтологий для повышения точности подбора ресурсов в Грид

Актуальной на сегодняшний день задачей для Грид-систем является усовершенствование методов координации и распределения ресурсов. Учитывая специфику Грид-систем, ресурсы могут очень отличаться друг от друга как аппаратной, так и программной конфигурацией, а также политиками доступа и использования пользователями. Сбор, отслеживание, а главное, наиболее эффективный подбор требуемых ресурсов для задачи, становятся непростыми задачами.

Информация о каждой задаче, выполняемой в Грид-системе, содержит определенные требования к процессорам, памяти и программному окружению в виде атрибутов. На данный момент проверка соответствия определенного ресурса поставленным требованиям выполняется по лексикографическому совпадению атрибутов.

Одним из способов повышения точности подбора ресурсов является использование семантических сопоставлений на основе онтологий [1]. Для реализации данного подхода предлагается специализированная база знаний и алгоритм работы системы подбора ресурсов Грид (“матчмейкер”).

Специализированная база знаний включает в себя:

- базовую онтологию ресурсов вычислительных систем и их характеристик
- знания о программном и аппаратном обеспечении
- набор правил для выполнения подбора ресурса, заданный на языке TRIPLE
- знания о доступных в данный момент ресурсах, которые обновляются динамически, используя информацию с GIS (Grid Information System)

Рассмотрим работу матчмейкера на следующем примере. Допустим, Грид-система включает два ресурса:

- ОС SunOS 5.8, 64 процессора с частотой 1900 Mhz;
- ОС Linux 6.2, 640 процессоров с частотой 1666 Mhz;

Если на вход данной системы поступит запрос ресурса с атрибутом OpSys = “UNIX” и требованием совместимости с библиотекой MPI, классический алгоритм подбора ресурсов вернет пустой результат. Использование специализированной базы знаний и логического вывода позволяет получить информацию о том, что и Linux и SunOs являются подмножествами Unix. Более того, база знаний позволяет проверить совместимость библиотеки MPI с обеими ОС. Также может быть проверено соблюдение требований к памяти и процессору. Так как оба ресурса подходят для выполнения задачи, выбирается ресурс с наиболее высокой частотой процессоров.

Очевидно, что средства логического вывода позволяют выполнять более точный подбор ресурсов под требования задачи, что повышает эффективность использования доступных ресурсов Грид-системы. Но данный метод имеет ряд ограничений в использовании, такие как необходимость наполнения и поддержания в актуальном состоянии правил базы знаний, увеличение времени подбора ресурсов, и, соответственно, непроизводительных расходов ресурсов системы. Первую проблему можно решить с помощью соответствующего программного обеспечения, выполняющего визуализацию и редактирование содержимого базы знаний. Что касается снижения накладных расходов при выборе ресурсов, вместо SPARQL для логического вывода можно использовать систему Algernon [2]. Повышение производительности алгоритмов сопоставления ресурсов является темой дальнейших исследований.

Литература. 1. Andreas Harth, Yu He, Hongsuda Tangmunarunkit, Stefan Decker, Carl Kesselman. “A Semantic Matchmaker Service on the Grid”. 2. Balachandar R. Amarnath, Thamarai Selvi Somasundaram, Mahendran Ellappan, Rajkumar Buyya “Ontology-based Grid resource management” // Softw. Pract. Exper. 2009;39:1419–1438.