

Параметры ресурсного запроса в распределенной системе

Т.С. Дьячук

Запорожский национальный технический университет, г. Запорожье, Украина

Введение

Для эффективного использования ресурсов в Grid-среде необходимо учитывать характеристики заданий, поступающих в систему. Эти характеристики отражаются в параметрах ресурсного запроса.

При планировании ресурсов [1] важной задачей является определение приоритетов пользовательских запросов на вычислительные ресурсы. Приоритет определяет порядок выполнения задания.

Подходы для определения приоритета

– В системе OurGrid применяется схема предпочтений, при которой каждый узел ведет учет выполненных работ для другого узла. И в зависимости от этого может использовать чужие ресурсы для своих задач. Таким образом, приоритет запроса зависит от приоритета узла, с которого этот запрос поступил [2].

– Для определения приоритета запроса может быть использован экономический подход к Grid – вычислениям, который является системой учета потребления ресурсов [3].

Запросы на ресурсы бывают локальные и глобальные. Необходимо распределять имеющиеся ресурсы между локальными и глобальными заданиями, учитывая их приоритеты. Локальные пользователи имеют преимущество на использование локальных ресурсов. Приоритет локальным запросам назначает администратор. Приоритет глобального запроса определяется, исходя из его параметров. Возможна ситуация, при которой глобальный запрос на ресурсы будет иметь больший приоритет, чем локальный.

С одной стороны ресурсы не должны простаивать, но приносить наибольшую прибыль, а с другой задания пользователей должны быть решены как можно быстрее с наименьшими затратами. Цена ресурсов должна меняться во времени в зависимости от спроса и внутренних потребностей. Путем повышения цены возможно прекращение поступления глобальных запросов. А пользователь должен иметь возможность повысить приоритет своего запроса с помощью поднятия предложенной платы за ресурсы, и таким образом конкурировать с локальными запросами. Все эти факторы формируют цену и соответственно приоритет ресурсного запроса.

Факторы, формирующие цену использования ресурсов:

- объем и характеристики ресурса;
- отношение реальной производительности ресурса к эталонным показателям;
- часы пик;
- спрос со стороны локальных запросов;
- приоритет локального задания, претендующего на данный ресурс.

Параметры ресурсного запроса

s – число параллельных ветвей задания;

t – ориентировочное время выполнения ветви задания на эталонном ненагруженном ресурсе, при запуске задания происходит пересчет времени в соответствии с параметрами исполнительных ресурсов;

T_{max} – максимальное время ожидания выполнения задания;

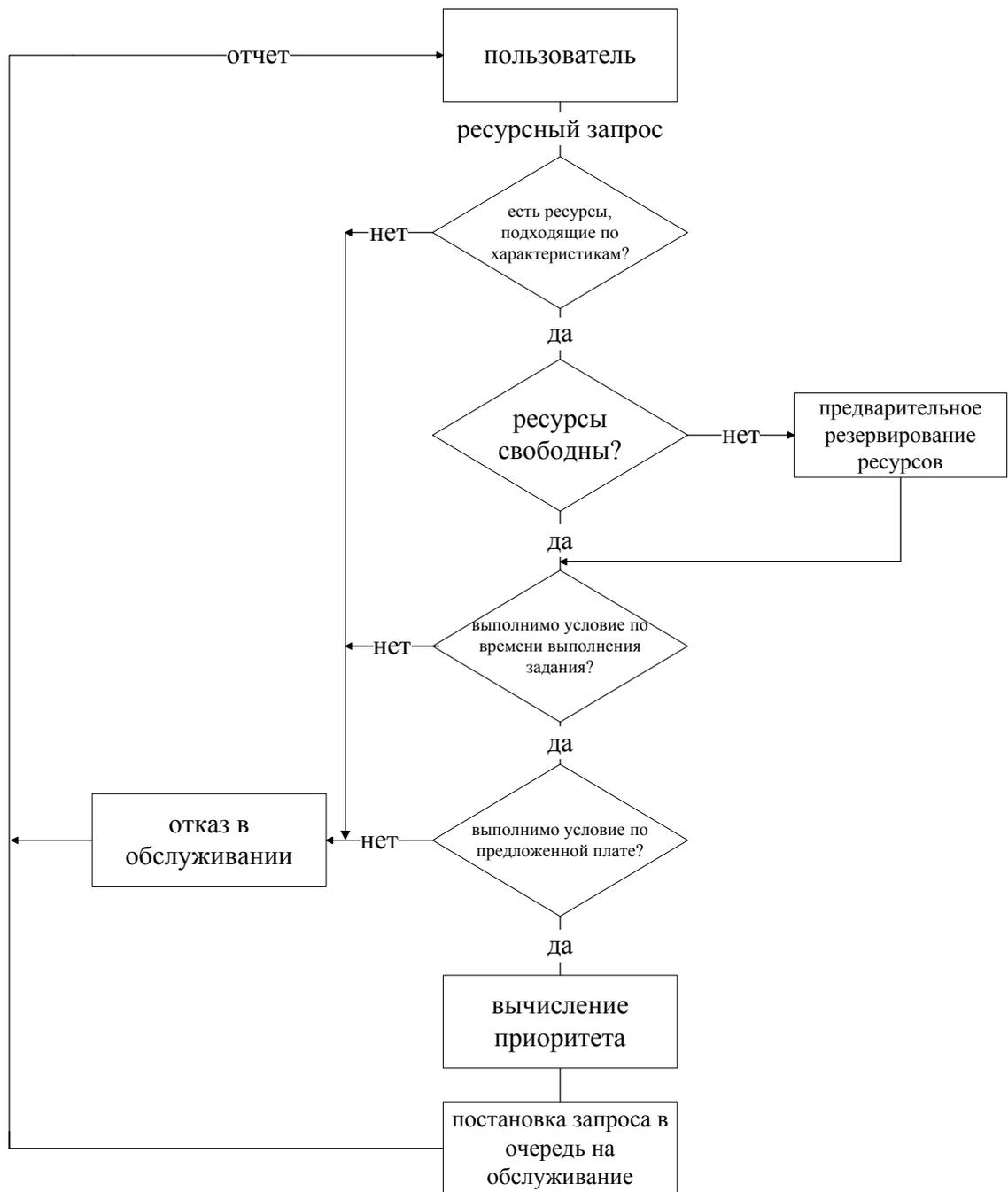


Рисунок 1. Алгоритм обслуживания ресурсного запроса.

$(p1, p2, \dots)$ – вектор, содержащий необходимые характеристики ресурса для выполнения ветви задания;

Pay – предложенная плата за использование ресурсов.

Таким образом, ресурсный запрос Z имеет вид:

$$Z(s, t, Tmax, (p1, p2, \dots), Pay) \quad (1)$$

Введем следующие обозначения:

g_i – стоимость i -го ресурса в единицу времени;

u_i – нормирующий коэффициент для определения реального времени выполнения ветви задания, который учитывает реальную производительность i -го ресурса и текущую нагрузку.

После подбора подходящих по характеристикам ресурсов проверяется выполнение условий по времени выполнения задания и предложенной плате.

Прогнозируемое время счета задания:

$$T_{count} = \max_i (t \cdot u_i) \quad (2)$$

T_{wait} – время ожидания выполнения задания.

Время выполнения задания:

$$T_{max} \geq T_{count} + T_{wait} \quad (3)$$

Стоимость выполнения задания:

$$Cost = \sum_{i=1}^s g_i \cdot (t \cdot u_i) \quad (4)$$

Приоритет запроса:

$$Pr = \frac{Pay}{Cost} = \frac{Pay}{\sum_{i=1}^s g_i \cdot (t \cdot u_i)} \quad (5)$$

Алгоритм обслуживания ресурсного запроса представлен на рисунке 1.

Выводы

В работе проведен анализ и предложен метод расчета приоритета ресурсного запроса. Сортировка запросов согласно приоритетам необходима для решения задачи планирования ресурсов.

Литература

1. Дьячук Т.С. Планирование ресурсов в распределенной системе //Высокопроизводительные параллельные вычисления на кластерных системах (НПС–2008). Труды 8-ой Международной конференции. Казань, ноябрь 17-19, 2008 – Казань: изд. КГТУ.–2008.–с. 28-30.
2. Cirne W., Brasileiro F., Andrade N., Costa L.B., Andrade A., Novaes R., Mowbray M. Labs of the World, Unite!!! – <http://copin.ufcg.edu.br/twiki-public/pub/OG/OurPublications/LabsoftheWorldUnitev19.pdf>
3. Buyya R., Giddy J.,Abramson D. A case for economy Grid architecture for service-oriented Grid computing// Proceedings of the Heterogeneous Computing Workshop, April 2001 – <http://gridbus.org/papers/ecogrid.pdf>