ПАРАДИГМА ПЛАНИРОВАНИЯ ОДНОРОДНЫХ ЗАДАЧ СЛУЧАЙНОГО ПОИСКА

П.Е. Голосов, М.В. Козлов, Ю.Е. Малашенко, И.А. Назарова, А.Ф. Ронжин

Администратору вычислительной машины необходимо организовать выполнение набора вычислительных работ, про которые он знает, что для их выполнения необходимо применить одну и ту же программу при разных начальных данных, разбитых на неделимые фрагменты. Находящаяся в его распоряжении вычислительная машина является многократно делимым ресурсом, каждая часть которого может выполнять каждую работу целиком или по отдельным множествам фрагментов независимо от других работ и частей.

Время выполнения каждой работы неизвестно до ее завершения. Его среднее ожидаемое значение при монопольном выделении работе всего ресурса линейно связано с числом неделимых фрагментов из которых она состоит. Это число в дальнейшем называется объемом работы. Время выполнения работы или ее заранее обусловленной части обратно пропорционально размеру выделяемой для этого доли ресурса.

На роль рассматриваемых работ вполне подходят задачи случайного поиска, определенные в [1]. Одна из отличительных черт таких задач заключается в неопределенности местоположения «уникального фрагмента» входных данных, ради поиска которого запускается программа и при обработке которого она завершает назначенную ей работу.

В отличие от шахматных правил аксиомы современного естествознания опираются на интуитивные представления. Современное состояние массово-параллельных вычислительных систем, [2], позволяет полагать, что требования к вычислительной машине являются вполне разумными.

Для простоты дальнейшего изложения предположим, что все задачи имеют одинаковый объем. Кроме того, нет как никаких формальных признаков для установления отношения порядка между работами.

Под планированием будем понимать правила разделения ресурсов вычислительной машины между работами при их запуске и перераспределение ресурсов по факту завершения одной или нескольких работ. Свободных ресурсов не бывает. Совокупность долей ресурсов, назначенных работам в моменты разделения и перераспределения ресурсов, будем называть планом. Понятно, что для всех планов общее время выполнения всех работ одинаково.

Рассмотрим два простейших плана.

Применение *последовательного плана* начинается с выбора из набора одной работы и выделения ей всего ресурса. После ее завершения происходит выбор следующей работы и опять же назначение ей всего ресурса. Так происходит до завершения всех работ. При выборе каждой работы возникает неопределенность, которую администратор может справедливо разрешить, лишь прибегая к случайной процедуре. Например, тщательно перетасовывая работы как колоду карт и выделяя весь ресурс «верхней» работе.

Параллельный план не содержит никакой неопределенности в организации работ. В начальный момент планирования весь ресурс вычислительной машины разделяется между работами поровну. В моменты завершения одной или нескольких работ весь ресурс перераспределяется между оставшимися работами опять же поровну. И так до полного завершения всех работ.

Исполнение таких планов можно реализовать при помощи программного комплекса «Пирамида» [3]. Необходимо лишь предварительно перемешать неделимые фрагменты всех работ и вычеркивать из общего перечня фрагменты, связанные с работой, которая завершилась, например, нашла свой «уникальный фрагмент».

Авторы [3] позиционировали свою работу как способ организации в массово-параллельной вычислительной системе выполнения независимых вычислительных работ (или параллельной программы со схемой вычислений «распараллеливание по данным»). К этому же классу вычислительных работ относятся и задания на выполнение операций случайного поиска из [1].

После выполнения любого из этих планов становится известным упорядоченный набор времен решения (длительностей) всех работ

$$X_1 \leq \ldots \leq X_m$$

при монопольном владении ресурсом с первого момента планирования. Независимо от плана время выполнения всех работ будет равно сумме

$$S_m = X_1 + \ldots + X_m$$

Проведем апостериорное сравнение некоторых характеристик рассмотренных выше планов.

При применении последовательного плана неизвестно, когда будет завершена работа с длительностью X_k Можно лишь утверждать, что среднее ожидаемое время (математическое ожидание) ее завершения (пребывания в системе) при случайной перетасовке колоды будет равно

$$T_k^{\ seq} = \frac{1}{2}(S_m + X_k)$$

При применении параллельного плана первой будет завершена работа с длительностью X_1. Она будет выполнена за время

$$T_1^{par} = mX_1$$

Второй будет завершена работа с длительностью X_2. Она будет выполнена за время

$$T_2^{\ par} = X_1 + (m-1)X_2$$

Далее работа с длительностью X_k будет выполнена за время

$$T^{par}{}_k = X_1 \ + \ldots + X_{k-1} \ + (m-k+1)X_k, k = 2,\ldots,m$$

Сравнивать средние времена завершения работ последовательного и параллельного планов

$$T_{seq} = \frac{1}{m}(T^{seq}{}_1 + \ldots + T^{seq}{}_m) \qquad \text{и} \qquad T_{par} = \frac{1}{m}(T^{par}{}_1 + \ldots + T^{par}{}_m)$$

в условиях полной неопределенности трудно, а в некоторых случаях и не имеет смысла.

В случаях выборочного контроля или применения процедуры «brute force» упорядоченный набор $X_1 \leq \ldots \leq X_m$ можно трактовать как вариационный ряд независимых наблюдений случайной величины и вычислять математическое ожидание разности

$$T_{par} - T_{seq}$$

по формуле из теоремы в [4]. Для экспоненциального семейства математическое ожидание этой разности всегда равно нулю. Для бета-распределения возможны как положительные, так и отрицательные ее значения.

Анализ результатов имитационного моделирования, приведенный в [4], привлек внимание к таким чертам параллельных планов как фильтрация работ по длительности, малое значение среднего квадратичного отклонения времен завершения задач от среднего времени их пребывания в системе и сокращение промежутков времени между моментами завершения длинных работ.

Таким образом, параллельные планы имеют значительные преимущества в тех случаях, когда заведомо имеются многочисленные данные с искаженными результатами наблюдений, которые не найдут фрагментов с уникальными свойствами. Не исключается и тот факт, что познавательная задача, ради которой проводятся вычисления, будет решена уже при завершении работ с малой длительностью.

Для работы положительной длительности xT и каждого плана определим параметр «относительная задержка» работы равенством

$$\pi = par, seq$$

$$\eta^{\pi} = \frac{T_k^{\pi}}{mxT}$$

В том случае, когда имеет равномерное распределение на отрезке [0,1], математические ожидания «относительной задержки» работы положительной длительности xT равны

$$M\eta^{seq} = \frac{1}{m} + \left(\frac{m-1}{2m} \right)\frac{1}{2x} \qquad \text{и} \qquad M\eta^{par} = 1 - \left(\frac{m-1}{2m} \right)x$$

Графики этих функций представлены на рисунке 1.

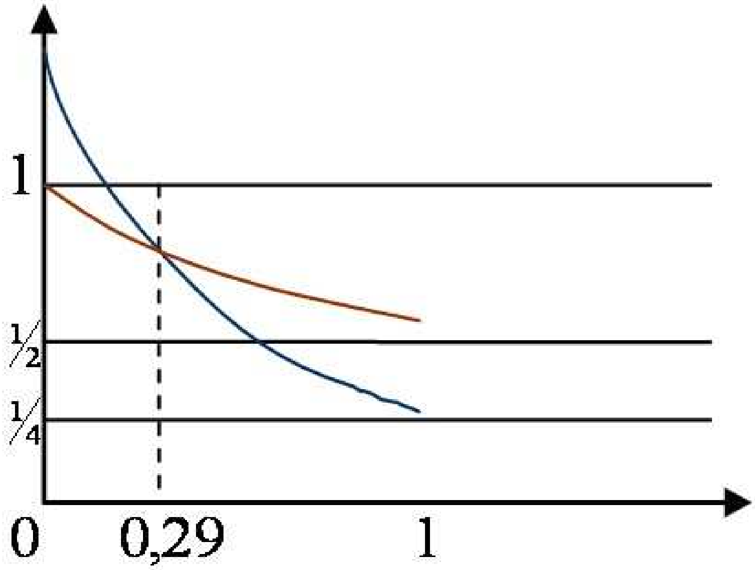

Рис.1. Взаимные характеристики «относительной задержки» работ для последовательного и параллельного планов

ЛИТЕРАТУРА:

1. Ронжин А.Ф., Суриков В.Н. О математических проблемах применения высокопроизводительных вычислительных систем для решения задач случайного поиска. Труды международной научно-практической конференции «Суперкомпьютерные технологии: разработка, программирование, применение». Ростов-на-Дону: Южный федеральный университет, 2010. Т. 2. С.239-243
2. Забродин А.В., Левин В.К., Корнеев В.В. Массово-параллельные системы МВС-100 и МВС-1000.- Научная сессия МИФИ-200. Сборник научных трудов. Том 2. М.: МИФИ, 2000, стр. 194-195
3. Баранов А.В., Киселев А.В., Киселев Е.А., Корнеев В.В., Семенов Д.В. Программный комплекс «Пирамида» организации параллельных вычислений с распараллеливанием по данным. Труды международной суперкомпьютерной конференции и конференции молодых ученых «Научный сервис в сети Интернет» Суперкомпьютерные центры и задачи. Абрау-Дюрсо, 2010, с. 299-302
4. Голосов П.Е., Козлов М.В., Малашенко Ю.Е., Назарова И.А., Ронжин А.Ф. Модель системы управления специализированным вычислительным комплексом. М.: ВЦ РАН, 2010. 46 с
5. Голосов П.Е., Козлов М.В., Малашенко Ю.Е., Назарова И.А., Ронжин А.Ф. Анализ управления специальными вычислительными заданиями в условиях неопределенности. М.: Теория и системы управления. 30 с. В печати.