

УДК 004.94

О ВЛИЯНИИ ХАРАКТЕРИСТИК СЕТИ ПЕРЕДАЧИ ДАННЫХ НА ОРГАНИЗАЦИЮ РАБОТЫ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

Коваленко С.М., к.т.н., доцент, МГТУ МИРЭА

E-mail: kovalenko@mirea.ru, Москва, Россия

Аннотация. Рассматриваются оценки времени, расходуемого на вычисления и на обмен данными в параллельных вычислительных системах (ВС). Выводится выражение для оценки оптимального числа вычислительных модулей в ВС, обеспечивающего минимальное время выполнения программы.

Ключевые слова. Параллельные вычислительные системы, обмен данными, вычислительные модули.

ABOUT THE INFLUENCE OF THE CHARACTERISTICS OF DATA TRANSMISSION NETWORK OF ORGANIZATION PARALLEL COMPUTING SYSTEMS

Kovalenko S.M., PhD., MSTU MIREA

E-mail: kovalenko@mirea.ru, Moscow, Russia

Annotation. Discusses the assessment of the time spent on computing and data exchange in parallel computer systems (CS). Displays the expression to evaluate the optimal number of modules in the CS, providing minimum run time.

Keywords. Parallel computer system, data exchange, compete modules.

Решение многих актуальных в настоящее время задач, возлагаемых на компьютерные системы, требует от таких систем обеспечения высокой производительности. К таким задачам относятся и моделирование в САПР сложных изделий и систем и задачи из таких областей науки как ядерная физика, биология, фармакология, климатология и других.

Основой современной высокопроизводительной вычислительной системы (ВС) является сложнейшее изделие современной цифровой микроэлектроники – современный многоядерный микропроцессор (МП), - обеспечивающий производительность при выполнении вычислительных операций в 100 – 1000 миллионов операций в секунду. Однако требуемую производительность вычислительной системы для решения сложных задач могут обеспечить лишь собранные в систему десятки, сотни или даже тысячи подобных МП.

Современные высокопроизводительные параллельные ВС, включающие десятки и сотни МП, создаются на стандартном оборудовании: серийно выпускаемых МП и

сетевом оборудовании. Такие ВС строятся по иерархическому принципу: на первом уровне иерархии несколько МП (от 1 до 8) вместе с необходимой оперативной памятью и адаптерами сетей объединяются в вычислительный узел или модуль (ВМ); на следующем уровне несколько ВМ объединяются в конструктивно законченный узел: блок или стойку, причем функционально ВМ объединяются путем их подключения к портам коммутаторов или маршрутизаторов сетей / 1 /. Чаще всего в ВС используются две сети: сеть передачи данных и служебная сеть (сеть, поддерживающая работу сетевой операционной системы). В качестве сети передачи данных сейчас чаще всего применяют сеть InfiniBand с малой латентностью и высокой пропускной способностью, в качестве служебной сети – Fast- или Gigabit Eht.

Рост производительности ВС с увеличением числа МП и ВМ идет нелинейно, поскольку с увеличением числа ВМ в ВС растут «накладные расходы», связанные с временем, необходимым для реализации обмена данными между ВС.

Для анализа влияния таких «накладных расходов» рассмотрим некоторый идеализированный вариант выполнения хорошо распараллеливаемой программы, имеющей N параллельно выполняемых фрагментов на K ВМ, причем N больше K . Время выполнения программы в рассматриваемом случае можно оценить как:

$$T_{np} = T_{выч} + T_{обм}, \quad (1)$$

где $T_{выч}$ – время, затрачиваемое на вычисления в ВМ,

$T_{обм}$ – время, затрачиваемое на обмен данными между ВМ.

$T_{выч}$ можно оценить с учетом производительности одного ВМ $P_{вм}$ и их числа K как:

$$T_{выч} = \frac{G}{(K \cdot P_{вм})}, \quad (2)$$

где G – оценка числа операций в программе.

Значение G можно получить при анализе алгоритма программы, используя, например, метрики Холстеда.

Затраты времени на обмен пакетом данных $T_{п}$ в сети передачи данных можно получить как:

$$T_{п} = T_{н} + T_{з} + \frac{Q}{V}, \quad (3)$$

где $T_{н}$ – задержки формирования пакета передачи данных в адаптере сети,

$T_{з}$ – задержка передачи пакета в сети, связанная с задержками в коммутаторе,

Q – объем передаваемых данных в пакете данных,

V – скорость передачи данных в сети.

Рассматриваем сеть передачи данных в составе ВС, поскольку объем трафика в ней много больше, чем в служебной сети. Анализируя время обмена данными как «накладные расходы» в процессе вычислений, в первом приближении $T_{обм}$ можно оценить с учетом ограниченной пропускной способности сети как:

$$T_{обм} = K \cdot \frac{Q}{V} \quad (4)$$

Рассмотрим получение выражения (4) несколько подробнее на примере системы, включающей K ВМ. Объем трафика каждого ВМ будет снижен пропорционально числу ВМ, однако считаем, что каждый ВМ пересылает пакеты остальным по окончании выполнения своего фрагмента программы. Объединяя трафик всех ВМ, пересылаемый по сети, получаем оценку времени, приводимую в выражении (4).

Объем данных в байтах пакета обмена связан с числом операций в выполняемой программе G (вычислительной сложностью алгоритма) как:

$$Q = C \cdot G \quad (5)$$

где C – коэффициент, характеризующий класс исполняемых алгоритмов в отношении связности по данным [2], алгоритмы с большей связностью по данным характеризуются большей интенсивностью обменов между ВМ и большим значением коэффициента C .

С учетом выражений (2)-(5), выражение (1) можно переписать в виде:

$$T_{np} = \frac{G}{(K \cdot P_{вм})} + \frac{(G \cdot C \cdot K)}{V} \quad (6)$$

Как следует из приведенного выражения (6) с ростом K снижается время на вычисления и растет длительность времени на обмен данными в ВС. Таким образом, характеристики сети передачи данных будут сильно влиять на работу ВС.

Из представленного выражения (6) следует, что можно провести оценку оптимального значения числа ВМ K , обеспечивающего минимальное значение времени выполнения программы T_{np} . Значение k определяется из условия $dT_{np}/dk=0$ и оценивается как:

$$K_{оп} = \sqrt{\frac{V}{C \cdot P_{вм}}} \quad (7)$$

Для современных ВС значения $K_{оп}$ находятся в районе единицы для V – 1 Гбайт/с и $P_{вм}$ – 1000 миллионов оп/с и только для значений C – 0,01 байт/оп, характеризующих алгоритмы со слабой связностью по данным, $K_{оп}$ составит величину порядка 10. По

мере повышения пропускной способности сети передачи данных число ВМ в ВС можно повышать и далее без опасения потерять производительность ВС из-за увеличения времени на обмен данными.

Список литературы

1. Корнеев В. В. Вычислительные системы. – М.: Гелиос АРВ, 2004. – 512с.
2. Воеводин В. В., Воеводин Вл. В. Параллельные вычисления. – СПб.: БХВ-Петербург, 2002. – 608с.