О быстродействии подсистемы памяти узлов вычислительного кластера

Д.В. Биматов, Д.С. Сущенко, С.П. Сущенко
Томский государственный университет

Введение

В современных вычислительных кластерах в качестве узлов широко используются компьютеры общего назначения, в которых для обеспечения быстрого доступа к наиболее часто используемым данным применяется многоуровневая организация памяти. Основным узким местом современных вычислительных систем является тракт процессор–память [1,2]. Разрыв скорости обработки данных в центральном процессоре и скорости доступа к операндам в оперативной памяти современных компьютеров составляет 1-2 порядка.

Главными требованиями, которым должна удовлетворять подсистема памяти, являются достаточно большая емкость, высокое быстродействие и экономическая эффективность с точки зрения технической реализации [3]. Желательно также, чтобы при минимальных физических размерах память обладала как можно большей информационной емкостью. Удовлетворить все эти требования одновременно в одном устройстве невозможно, поэтому обычно комбинируют несколько запоминающих устройств с различными параметрами, добиваясь создания комплексного решения с требуемыми характеристиками. Обычно оперативная память и механизмы, реализующие виртуальную память, являются средством обеспечения большой информационной емкости, а буферная память (кэш) – средством повышения быстродействия подсистемы памяти.

Таким образом, многоуровневая организация памяти позволяет сгладить разрыв между скоростью обработки данных центральным процессором и скоростью доступа к адресуемым объектам в оперативной памяти вычислителя. Эффективность доступа к иерархической памяти определяется не только быстродействием и емкостью отдельных уровней подсистемы памяти, но и набором архитектурных параметров, важнейшими среди которых являются коэффициент ассоциативности, длина межуровневого интерфейсного блока и глубина неблокируемости кэш-памяти.

Ассоциативность определяет способ отображения пространства адресов оперативной памяти в кэш-память и, наряду со стратегией вытеснения кэш-строк при конфликте адресов, в значительной мере

влияет на степень локализации прикладных задач в памяти верхнего уровня. Степень локализации в, свою очередь, определяет вероятность «попадания» адресуемого объекта в кэш и время доступа к нему.

Как правило, различные уровни иерархической памяти работают асинхронно, поэтому между ними предусмотрены буферные устройства, которые хранят адреса востребованных вычислителем элементов данных и позволяют параллельно выполнять операции доступа к различным уровням. Емкость межуровневого буферного устройства определяет параметр глубины неблокируемости кэша и задает потенциально достижимую степень параллелизма выполнения совокупности транзакций доступа к различным уровням иерархической памяти.

Эксплуатационные характеристики вычислительной системы в значительной мере определяются не только указанными параметрами иерархической памяти, но и распределением востребованности процессором рассеянных в оперативной памяти адресуемых объектов решаемых пользователями прикладных задач.

Набор важнейших вероятностно-временных характеристик подсистемы иерархической памяти включает вероятность попадания в кэш, вероятность блокировки кэш-памяти, среднее время доступа к адресуемым объектам и пропускную способность подсистемы памяти. Кроме того, при разработке приложений, их настройке на конкретные применения и условия эксплуатации важными являются знание и прогноз пределов изменения операционных показателей пропускной способности и средней задержки доступа к адресуемым объектам в различных нагрузочных условиях. Чрезвычайно полезными являются также оценки времени заполнения кэш-памяти блоками памяти при переключении процессора на выполнение нового приложения.

Влияние ассоциативности на вероятность попадания в кэш

Основным фактором влияния на этот показатель является распределение вероятностей востребованности блоков памяти прикладной задачи центральным процессором вычислительного узла. Кроме того, частоту попадания в кэш в значительной мере определяет стратегия вытеснения блоков из группы кэша при возникновении конфликта адресов [3]. Очевидно, что самой предпочтительной является идеальная стратегия вытеснения самого неиспользуемого блока в группе (имеющего наименьшую вероятность быть востребованным). По оценкам специалистов, наиболее близка к идеальной стратегия LRU [3], имеющая умеренную трудоемкость реализации.

Процессы перемещения адресуемых объектов между уровнями иерархической памяти для идеальной стратегии вытеснения хорошо описываются цепями Маркова. В предположении упорядоченности по убыванию вероятностей востребованности блоков памяти процессором вычислителя с ростом адресов блоков, отображаемых на заданную группу кэша ассоциативности A, динамика процесса замещения блоков данных в кэше моделируется случайным Марковским процессом в A-мерном пространстве. При этом переходные вероятности цепи Маркова, описывающей факты вытеснения наименее вероятных к использованию в группе кэша и подкачки востребованных вычислителем блоков памяти в кэш, имеют вид перемещений вдоль координатной оси с максимальным исходным значением координаты.

В стационарных условиях цепь Маркова, описывающая функционирование группы кэша, распадается на совокупность замкнутых множеств состояний, образующих $A!$ неприводимых подцепей Маркова. Вероятности обнаружения блоков памяти в заданной группе кэша при этом имеют следующий вид [4]:

$$\Pi_a = 1,\ a = \overline{0, A-2};\Pi_m = \frac{f(m)}{1-\sum_{k=0}^{A-2} f(k)},\ m = \overline{A-1, M-1}.$$

Здесь $f(m)$ – условная вероятность востребованности вычислителя в m-м блоке памяти, отображаемом на одну группу кэша; M – количество блоков памяти, отображаемых на одну группу кэша.

Отсюда следует, что в каждой группе кэша все строки кроме одной содержат самые востребованные вычислителем блоки памяти, а одна строка используется для попеременной подкачки недостающих вычислителю адресуемых объектов. Вероятность попадания адресуемого объекта в кэш с идеальным вытеснением определяется зависимостью [4]

$$\Pi = \sum_{g=0}^{G-1}\left[\sum_{m=0}^{A-2} p(g,m) + \frac{\sum_{m=A-1}^{M-1}\left[p(g,m)\right]^2}{\sum_{m=A-1}^{M-1} p(g,m)}\right],$$

где G – число групп кэша; $p(g,m)$ – вероятность потребности вычислителя в m-м блоке памяти, отображаемом на g-ю группу кэша.

Из данного соотношения нетрудно видеть, что вероятность попадания достигает максимального значения при полностью ассоциативном кэше, в котором только одна строка единственной группы используется для замещения блоков памяти, в то время как в других находится

самая востребованная процессором информация. В целом для широкого спектра распределений востребованности адресуемых объектов прикладных задач в оперативной памяти зависимость вероятности попадания от коэффициента ассоциативности имеет характер кривой с насыщением в области $A > 10$. В тех случаях, когда приложение полностью локализовано в кэше или же востребовано процессором в соответствии с равномерным законом, вероятность попадания инвариантна к ассоциативности кэша.

С ростом размера приложений вероятность попадания монотонно снижается и при достижении объема прикладных задач до 5-10 размеров кэш-памяти стабилизируется на определенном уровне. Следует отметить также, что время заполнения кэша самыми востребованными процессором блоками памяти для идеальной стратегии вытеснения распределено по геометрическому закону со снижающимся параметром распределения по мере загрузки наиболее используемых адресуемых объектов. При этом параметр распределения снижается на значение, равное вероятности востребованности загруженного блока памяти, если его номер лежит в пределах от 0 до A–2.

Влияние глубины неблокируемости на среднее время доступа к подсистеме памяти

Для получения аналитических зависимостей среднего времени доступа к адресуемым объектам от параметров вычислителя подсистема памяти описывается стохастическим конвейером с нарастающим временем обработки в последовательных фазах, случайным количеством последовательных фаз обработки операции доступа к адресуемому объекту, определяемым вероятностями промаха на каждом уровне иерархической памяти, и длиной, равной числу уровней иерархии.

Стохастический конвейер описанной структуры моделируется многоэтапной системой массового обслуживания с поступлением транзакций доступа к заданному уровню подсистемы иерархической памяти в темпе промахов в предыдущий уровень иерархии, детерминированным обслуживанием и дискретным временем. Время между поступлениями заявок в систему массового обслуживания распределено по геометрическому закону с параметром, определяемым вероятностью промаха в кэш первого уровня, а число этапов обслуживания – временем обработки заявки на следующем уровне иерархии, вычисленным в длительностях циклов доступа к первому уровню подсистемы памяти.

Среднее время доступа к адресуемым объектам для двухуровневой подсистемы памяти блокирующего типа определяется линейной зави-

симостью от вероятности промаха в кэш и складывается из времени обращения к кэш-памяти и, с вероятностью промаха, времени выборки объекта из оперативной памяти. С ростом глубины неблокируемости зависимость среднего времени доступа принимает нелинейный вид от вероятности попадания. При этом для низких уровней попадания средняя задержка превышает данный показатель для подсистемы памяти блокирующего типа, а для высоких – оказывается ниже.

Область значений вероятности попадания, обеспечивающих выигрыш в среднем времени доступа для каждой следующей величины коэффициента неблокируемости, сужается, из чего следует, что большие значения этого параметра следует применять для сильно локализованных задач, имеющих низкие вероятности промаха. Отметим, что рост глубины неблокируемости при прочих равных условиях всюду приводит к увеличению пропускной способности подсистемы памяти до потенциально возможных значений.

Литература

1. **Корнеев В.В., Киселев А.В.** Современные микропроцессоры. М.: Нолидж, 1998.
2. **Корнеев В.В.** Параллельные вычислительные системы. М.: Нолидж, 1999.
3. **Кохонен Т.** Ассоциативная память. М.: Мир, 1980.
4. **Сущенко М.С.** Анализ производительности множественного ассоциативного кэша // Вестник ТГУ. 2002. № 275. С. 218-223.