ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФГОБУ ВПО “СИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ И ИНФОРМАТИКИ”

Кафедра ВС

**РАСЧЕТНО-ГРАФИЧЕСКОЕ ЗАДАНИЕ**

по дисциплине

“Архитектура вычислительных систем”

Вариант 6

|  |  |
| --- | --- |
| Выполнил студент | Рощупкин Никита Геннадьевич |
|  | Ф.И.О. |

|  |  |
| --- | --- |
| Группы | ИВ-221 Факультета ИВТ |
|  |  |

|  |  |  |
| --- | --- | --- |
| Работу принял |  | доцент Каф. ВС к.т.н. А.В. Ефимов |
|  | Подпись |  |

|  |  |  |
| --- | --- | --- |
| Оценка |  | |
|  | |  | |  |

Новосибирск – 2014**Листинг задания 6**

1. Оценить архитектурные возможности вычислительных средств IV и V поколений.

Описать функциональную структуру одной из суперВС (из списка Тоp500).

2. Построить блок-схему p -алгоритма умножения матриц:

G[1:J, 1:K] , B[1:K, 1:H]

обеспечивающего распределение элементов результирующей матрицы по горизонтальным полосам в элементарных машинах ВС.

Отыскать максимум коэффициента ε накладных расходов при реализации p -алгоритма на вычислительной системе, имеющей следующие параметры:

– разрядность *l* = 64;

– полосу пропускания канала между машинами ν = 5 Гигабод;

– время выполнения операции сложения t­c = 0,5 нс;

– время выполнения операции умножения tу = 1 нс.

1. Оценка архитектурных возможностей вычислительных средств……………..3

2. Архитектурные возможности вычислительных средств 4 и 5 поколения...…..5

3. Особенности архитектуры IBM Blue Gene……………………………….....…..6

4. IBM Blue Gene/P…………………………...……………………………….....…..7

5. P -алгоритм умножения двух матриц.………………….…...…………….....…10

6. Коэффициент накладных расходов.………………….…...…………….....……12

7. Приложение……………………….………………….…...…………….....…..…13

**Оценка архитектурных возможностей вычислительных средств**

Параллельные ВС относят к четвертому, пятому и последующим поколениям средств обработки информации. Вообще говоря, параллельные ВС различаются по своей архитектуре и функциональной cтpyктypе, по способам обработки информации и т. п. Существуют классификации параллельных ВС, и в каждом классе выделяются свои поколения ВС; более того, говорят даже o поколениях ВС, выпускаемых той или иной фирмой.

Для количественной характеристики поколений вычислительных средств будем использовать вектор {Ω, *V*, **θ**, ∑}, где Ω показатель производительности, или среднее число операции, выполняемых в секунду всеми процессорами ВС, опер./с; *V* емкость оперативной памяти ВС, бит; **θ** среднее время безотказной работы ВС в целом, ч; ∑ «цена операции», рассчитываемая как отношение цены ВС к показателю производительности (долл./(опер. \* c­­-1)).

Под **производительностью ЭВМ** понимается

ее способность обрабатывать информацию. Как правило, когда говорят о производительности, то понимают под этим потенциальную возможность ЭВМ по обработке информации, a не реальную, учитывающую аномальности в работе ЭВМ, например простои из-за отказов, из-за профилактического обслуживания и т. п. B процессе обработки информации в ЭВМ реализуются те или иные операции из ее набора (или системы) операций. Состав набора операций, бёзусловно, характеризует архитектypу ЭВМ и, следовательно, определяет ее производительность. Для оценки способности ЭВМ производить обработку информации используют количественные характеристик или показатели производительности.

**Емкость памяти**

максимальное количество информации, которое может в ней храниться.

**Среднее время безотказной работы ВС** в целом (для неё нам нужно знать)

Где

* производительность ЭВМ в момент времени
* функция надежности (или вероятность безотказной работы) ЭВМ

Где

* функцией ненадежности (или вероятностью отказа) ЭВМ

Где

* **N** – число работоспособных ЭВМ в начале испытаний;
* –время безотказной работы i-й машины
* **–** среднее время безотказной работы ЭВМ

**Цена операции**

Где

* **-** показатель производительности машины
* **-** цена машины
* константа a > 2 и коэффициент h, имеющий размерность, зависят от технологии производства

Количественная характеристика ЭВМ , которую называют ценой (одной) операции (в секунду).

**Архитектурные возможности вычислительных средств 4 и 5 поколения**

**Особенности архитектуры IBM Вlue Gene**

Для интенсификации работ по проекту Blue Gene разработчики IBM вынуждены были пересмотреть свою архитектурную платформу. B результате ими была декларирована «новая» архитектурная концепция SMASH (Simple, Many, Self-Healing простая, множественная и самовосстанавливаемая). Эта концепция ориентирована на использование таких архитектурных и функциональных решений, которые позволяют интегрировать большое число процессоров и, следовательно, обеспечивают автоматическое устранение проблем, вызванных сбоями.

При разработке данной суперВС принцип масштабирования воплощался на всех уровнях и аппаратурного, и программного обеспечения.

Все модели семейства IBM Вlue Gene относятся к классу распределенных систем c массовым параллелизмом. Несмотря на то, что они ориентированы на решение суперсложных задач, заложенные в них решения позволяют характеризовать их архитектуру как MIMD.

Функциональная структура суперВС рассчитана на использование хост-компьютерной системы множества хост-компьютеров. Эти компьютеры реализуют файловую систему и выполняют следующие функции: анализ функционирования, контроль, диагностику и восстановление суперВС, a также компиляцию и сервисное обслуживание. Выбор хост-компьютеров для IBМ Вlue Gene определяется областью применения.

Важной особенностью архитектуры IBM Вlue Gene является возможность одновременной работы множества пользователей. Это достигается путем выделения каждому пользователю требуемых ресурсов подсистемы из необходимого количества вычислительных узлов. Следовательно, в данной суперВС реализована возможность рaзбиения “спространства” вычислительных узлов на подпростpанства и достyпа к ним пользователей. Выделение подсистем в суперВС выполняется хост-компьютерной системой.

При разработке функциональной структуры суперВС и основных ее элементов значительное внимание уделялось обеспечению надежности (RAS Reliability, Availability, serviceability надежности, готовности, обслyживаемости). При этом свою роль сыграли принципы простоты, однородности и модyльности, a также введение избыточности, средств контроля, диагностики и восстановления.

Фактически разработчики суперВС Вlue Gene реализовали платформу ВС c программируемой структурой, они достаточно полно воплотили и архитектурные принципы, и принципы технической реализации модели коллектива.

**IBM Blue Gene/P (Суперкомпьютер Shaheen (335 место))**

IBM Blue Gene/P — массивно-параллельная вычислительная система, которая состоит из двух стоек, включающих 8192 процессорных ядер, с пиковой производительностью 27,9 терафлопс. <= Производительность может варьироваться от количества процессорных ядер.

IBM Blue Gene/P можно отнести ко второму поколению ВС.

* Она выпускается в разных сериях Blue Gene/L, Blue Gene/P, Blue Gene/Q, финансируются государством и построена на архитектуре, которую можно отнести к технологиям находящимся в разаработке.
* Она применяется для решения универсальных задач, к примеру моделирования поведения лекарств в теле человека.
* Blue Gene/P — проект массово-параллельной архитектуры, направленный на достижение скорости обработки данных, превышающей 1 петафлопс.
* Для неё разработано индивидуальное программное обеспечение
* Емкость оперативной памяти превышает >= 1012
* Также в случаях исключения, она может сама принять решения для её устранения

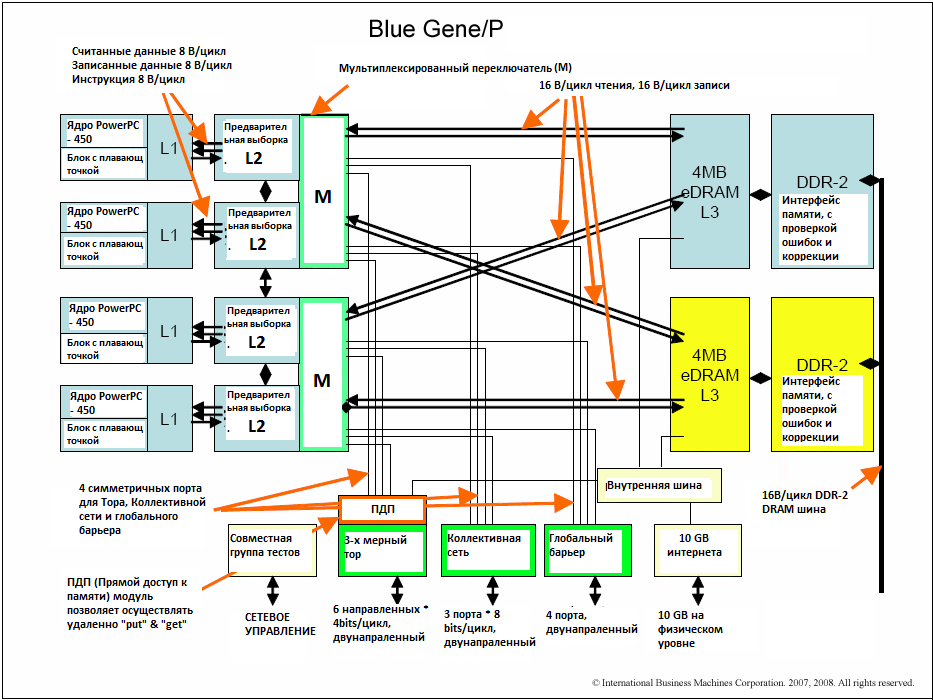


Рис. Применение интегральной схемы к вычислительному узлу

**Вычислительный узел IBM Blue Gene/P**

* четыре микропроцессорных ядра PowerPC 450 (4-way SMP)
* пиковая производительность: 4 cores x 3,4 GFlop/sec per core = 13,6 GFlop/sec
* пропускная способность памяти: 13,6 GB/sec
* 2 ГБ общей памяти
* 2 x 4 МБ кэш-памяти 2-го уровня (в документации по BG/P носит название L3)
* легковесное ядро (compute node kernel, CNK), представляющее собой Linux-подобную операционную систему, поддерживающую значительное подмножество Linux-совместимых системных вызовов
* асинхронные операции межпроцессорных обменов (выполняются параллельно с вычислениями)
* операции ввода-вывода перенаправляются I/O-картам через сеть коллективных операций

**Конструкция системы IBM Blue Gene/P**

* две стойки с вычислительными узлами и узлами ввода-вывода
* 1024 четырехъядерных вычислительных узла в каждой из стоек
* 16 узлов ввода-вывода в стойке (в текущей конфигурации активны 8, т.е. одна I/O-карта на 128 вычислительных узлов)
* выделенные коммуникационные сети для межпроцессорных обменов и глобальных операций
* программирование с использованием MPI, OpenMP/pthreads, POSIX I/O
* высокая энергоэффективность: ∼ 372 MFlops/W
* система воздушного охлаждения

Стойка (rack, cabinet) состоит из двух midplane'ов. В midplane входит 16 node-карт (compute node card), на каждой из которых установлено 32 вычислительных узла (compute card). Midplane, 8 x 8 x 8 = 512 вычислительных узлов, — минимальный раздел, на котором становится доступна топология трехмерного тора; для разделов меньших размеров используется топология трехмерной решетки. Node-карта может содержать до двух узлов ввода-вывода (I/O card). Вычислительный узел включает в себя четырехъядерный процессор,2 ГБ общей памяти и сетевые интерфейсы.

Вычислительные узлы и I/O-карты в аппаратном смысле неразличимы и являются взаимозаменяемыми, разница между ними состоит лишь в способе их использования. У них нет локальной файловой системы, поэтому все операции ввода-вывода перенаправляются внешним устройствам.

**Средства ввода-вывода IBM Blue Gene/P**

* не учитывается при расчете пиковой производительности
* использует сеть коллективных операций для коммуникаций с вычислительными узлами
* подключен к внешним устройствам через Ethernet-порт посредством 10-гигабитный функциональной сети
* операционная система на основе Linux (Mini-Control Program, MCP) с минимальным набором пакетов, необходимых для поддержки клиента сетевой файловой системы и Ethernet-подключений

**Коммуникационная сеть IBM Blue Gene/P**

* [трехмерный тор](http://hpc.cmc.msu.ru/bgp/development/networks#torus) (three-dimensional torus)
  + сеть общего назначения, объединяющие все вычислительные узлы; предназначена для операций типа «точка-точка»
  + вычислительный узел имеет двунаправленные связи с шестью соседями
  + пропускная способность каждого соединения — 425 MB/s (5,1 GB/s для всех 12 каналов)
  + латентность (ближайший сосед):
    - 32-байтный пакет: 0,1 μs
    - 256-байтный пакет: 0,8 μs
* [глобальные коллективные операции](http://hpc.cmc.msu.ru/bgp/development/networks#collective) (global collective)
  + коммуникации типа «один-ко-многим» (broadcast-операции и редукция)
  + используется вычислительными узлами для обменов с I/O-картами
  + каждый вычислительный узел и I/O-карта имеют три двунаправленные связи
  + пропускная способность каждого соединения — 850 MB/s (1,7 GB/s для двух каналов)
  + латентность (полный обход): 3,0 μs
* глобальные прерывания (global interrupt)
  + операции барьеров и прерываний (глобальные AND- и OR-операции)
* функциональная сеть
  + соединяет узлы ввода-вывода с внешним окружением
  + 10-гигабитная оптическая Ethernet-сеть
* сервисная сеть (service/control)
  + загрузка, мониторинг, диагностика, отладка, доступ к счетчикам производительности
  + гигабитная Ethernet-сеть (4 соединения на стойку)

Чтобы разгрузить процессорное ядро от операций, связанных с передачей сообщений по сети трехмерного тора, используется устройство прямого доступа к памяти (direct memory access, DMA). Кроме уменьшения нагрузки на ядро, этот механизм уменьшает вероятность взаимной блокировки процессов, обменивающихся сообщениями, которая может возникнуть вследствие ошибок программиста.

**P-алгоритм умножения двух матриц**

B[1:K, 1:H] \* G[1:J, 1:K] = C[1:J, 1:H]

Пусть требуется построить параллельный алгоритм, вычисляющий произведение двух прямоугольных матриц:

Элементы матрицы-произведения C[1:J, 1:H] вычисляются по формуле. (1)

Минимум емкости памяти будет достигнут, если каждая из исходных матриц будет разбита на n равных частей, и в каждый вычислитель будет размещено по одной такой части матриц B и G. Каждую из матриц B и G можно разрезать на n равных соответственно горизонтальных и вертикальных полос.

Причем в n-м вычислителе строки

столбцы

матриц B и G соответственно.

Через ]x[ обозначено такое ближайшее к x целое число, для которого справедливо неравенство .

Параллельный вычислительный процесс можно организовать следующим способом. Сначала первый вычислитель передает остальным вычислителям первый столбец матрицы B. После этого каждый из вычислителей по формуле **(1)** рассчитывает ]K/n[ элементов первой строки своей полосы для результирующей матрицы C. Затем первый вычислитель рассылает во все остальные вычислители второй столбец матрицы B и производятся вычисления элементов второй строки матрицы C и так до тех пор, пока первый вычислитель не перешлет все строки.

После этого пересылками будут заниматься последовательно второй вычислитель, третий вычислитель и далее до n-го вычислителя. Матрица C получается распределенной по вычислителям, причем в каждом будет своя горизонтальная полоса.



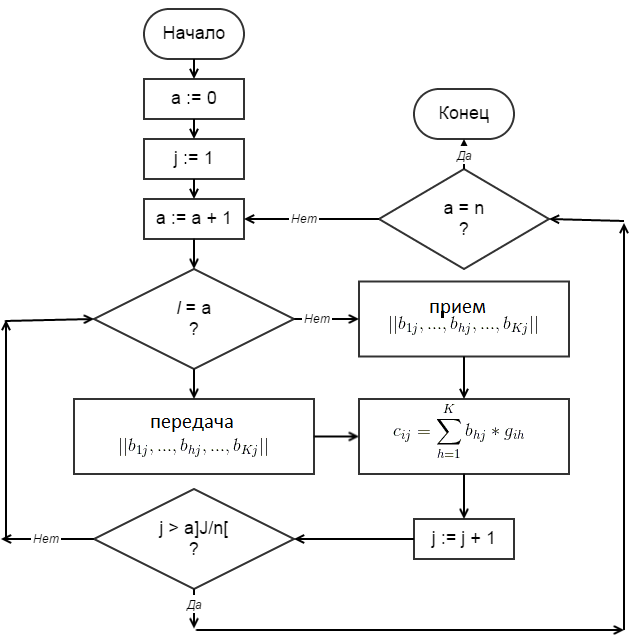


Рис. Схема ветви параллельного алгоритма умножения матриц

а - номер передающего вычислителя;

*l* – ветвь выполняющая вычисления

n - число вычислителей в системе;

**Коэффициент накладных расходов**

Эффективность параллельного алгоритма умножения матриц большого размера можно характеризовать показателями:



Очевидно, что максимум накладных расходов будет при , или, что то же самое, равенство  достигается при  Таким образом, максимум коэффициента ε накладных расходов определяется формулой:

ε = tn / (ty + tc)

tn – время пересылки

tу – время умножения

tс  - время сложения

tn = *l* / ν = 64 / 5 \* 109 = 12,8 нс, тогда

ε = 12,8 / (1 + 0,5) = 8,5333