Многопоточное программирование

Моделирование и Анализ Параллельных Вычислений

План лекции

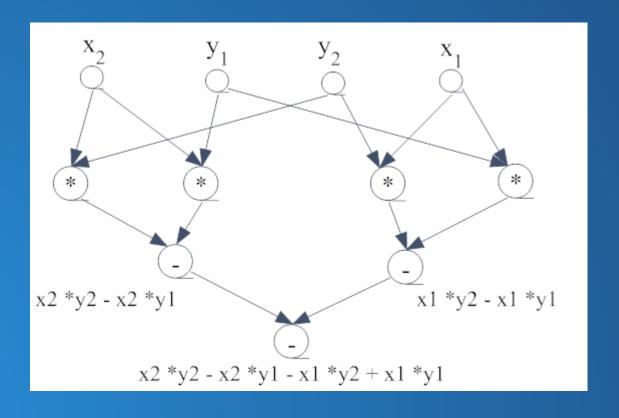
- Анализ внутренней структуры алгоритма
- Модель параллельного алгоритма
- Показатели качества параллельного алгоритма
- Законы Амдала и Густафсона
- Оценка коммуникационной способности

Модель «операции - операнды»

Ориентированный ациклический граф G(V,R)

- V множество вершин, выполняемые операции
- R дуги графа. r(i,j) принадлежит графу, если ј использует результаты операции i

Пример графа



Модель операции-операнды Решение квадратного уравнения

Расписание

- Р количество процессоров
- Расписание

$$H_p = \{(i, P_i, t_i) : i \in V\}$$

Расписание

- Р количество процессоров
- Расписание

$$H_p = \{(i, P_i, t_i) : i \in V\}$$

• Условия реализуемости

$$\forall i, j \in V : t_i = t_j \Rightarrow P_i \neq P_j$$

 $\forall (i, j) \in R : t_j >= t_i + 1$

Расписание

- Р количество процессоров
- Расписание

$$H_p = \{(i, P_i, t_i) : i \in V\}$$

• Условия реализуемости

$$\forall i, j \in V : t_i = t_j \Rightarrow P_i \neq P_j$$

 $\forall (i, j) \in R : t_j >= t_i + 1$

• Модель параллельного алгоритма

 $A_p(G, H_p)$,

Упрощения

- Время выполнения каждой операции 1
- Накладные расходы на передачу данных между процессорами отсутствуют (происходят мгновенно)

• Время выполнения параллельного алгоритма

$$T_p(G,H_p) = \max_{i \in V} (t_i + 1)$$

Время выполнения параллельного алгоритма

$$T_p(G,H_p) = \max_{i \in V} (t_i + 1)$$

• Время выполнения с оптимальным расписанием

$$T_p(G) = \min_{H_p} T_p(G, H_p)$$

Время выполнения параллельного алгоритма

$$T_p(G, H_p) = \max_{i \in V} (t_i + 1)$$

• Время выполнения с оптимальным расписанием

$$T_p(G) = \min_{H_p} T_p(G, H_p)$$

• Время выполнения наилучшей вычислительной схемой

$$T_p = \min_G T_p(G)$$

Время выполнения параллельного алгоритма

$$T_p(G, H_p) = \max_{i \in V} (t_i + 1)$$

• Время выполнения с оптимальным расписанием

$$T_p(G) = \min_{H_p} T_p(G, H_p)$$

• Время выполнения наилучшей вычислительной схемой

$$T_p = \min_G T_p(G)$$

• Минимально возможное время выполнения

$$T_{\infty} = \min_{p>=1} T_p$$
 $T_{\infty}(G) = d(G)$

Время выполнения последовательного алгоритма

• Время выполнения для заданное схемы

$$T_1(G) = |V|$$

Время выполнения последовательного алгоритма

• Время выполнения для заданное схемы

$$T_1(G) = |V|$$

Врем выполнения алгоритма

$$T_1 = \min_G T_1(G)$$

Время выполнения последовательного алгоритма

Время выполнения для заданное схемы

$$T_1(G) = |V|$$

Врем выполнения алгоритма

$$T_1 = \min_G T_1(G)$$

• Время последовательного решения задачи

$$T_1^* = \min T_1$$

Ускорение (speedup)

$$S_p(n) = \frac{T_1^*(n)}{T_p(n)}$$

Эффективность (efficiency)

$$E_p(n) = \frac{S_p(n)}{p} = \frac{T_1(n)}{pT_p(n)}$$

• S=р Идеальный случай

- S=р Идеальный случай
- S < p
 - Последовательные части алгоритма
 - Накладные расходы
 - о Координация

- S=р Идеальный случай
- S < p
 - Последовательные части алгоритма
 - Накладные расходы
 - Координация
- S > p

- S=р Идеальный случай
- S < p
 - Последовательные части алгоритма
 - Накладные расходы
 - Координация
- S > p
 - Сверхлинейное ускорение
 - Увеличения кэша и оперативной памяти
 - Нелинейная зависимость сложности решения от объема входных данных
 - Различные вычислительные схемы

Ускорение и Эффективности

• Часто бывают противоречивыми

Стоимость

• Стоимость

$$C_p = pT_p$$
.

• Стоимостной-оптимальный параллельный алгоритм

Вычисление частных сумм

Закон Амдала

• Доля последовательных вычислений

$$f = \frac{T_{seq}}{T_1}$$

• Время выполнения параллельного алгоритма

$$T_p = fT_1 + \frac{(1-f)T_1}{p}$$

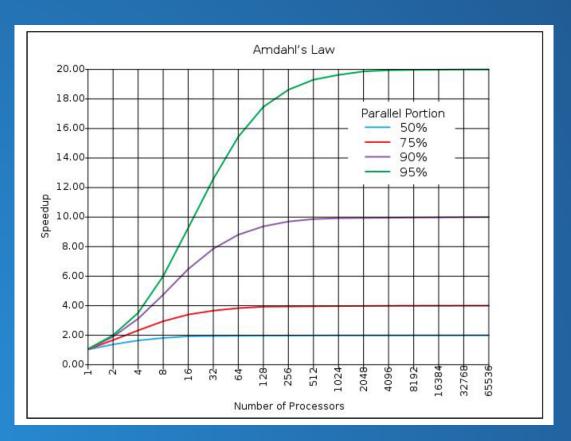
• Ускорение

$$S_p = \frac{1}{f + \frac{1-f}{p}}$$

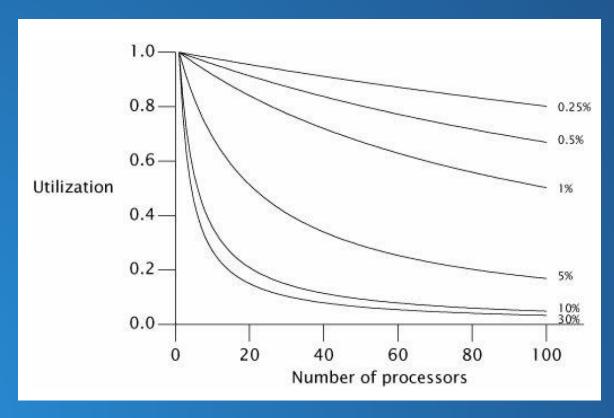
• Максимальное ускорение

$$\lim_{p\to\infty} S_p = \frac{1}{f}$$

Закон Амдала



Закон Амдала



Закон Густавсона-Барсиса

$$g = \frac{T_{seq}}{T_{seq} + \frac{T_{par}}{p}}$$

$$T_1 = gT_p + p(1-g)T_p$$

$$S_p = p + (1 - p)g$$

Вопросы