UCR-CS217-HW4: Performance Analysis

Infomation

- Name: Yue Zhu
- SID: 862547137
- eMail: yzhu303@ucr.edu

Prerequisite

Notation

- Problem size: n
- Number of processors: p
- Fraction of potentially parallel tasks: f, s
- Inherently sequential computations: $\sigma(n)$
- Potentially parallel computations: $\phi(n)$
- Communication operations: $\kappa(n,p)$

Amdahl's Law

$$\Psi \leq \frac{1}{(1-f)+\frac{f}{p}}$$

Gustafson-Barsis' Law

$$\Psi \leq p+(1-p)s = s+(1-s)p$$

Karp-Flatt Metric

$$\Psi(n,p) \leq \frac{\sigma(n)+\phi(n)}{\sigma(n)+\frac{\phi(n)}{p}+\kappa(n,p)}$$

- Corollary:

$$e = \frac{\frac{1}{\Psi}-\frac{1}{p}}{1-\frac{1}{p}} = f + \frac{\kappa(n,p)[\frac{p}{p-1}]}{\sigma(n)+\phi(n)} \overset{\lim_{p\to\infty}}{\approx} \frac{\sigma(n)+\kappa(n,p)}{\sigma(n)+\phi(n)}$$

Isoefficiency Metric

$$\begin{aligned}\varepsilon = \frac{\Phi(n,p)}{p} \le \frac{\frac{\sigma(n)+\phi(n)}{\sigma(n)+\phi(n)/p+\kappa(n,p)}}{p} &= \frac{\sigma(n)+\phi(n)}{p\sigma(n)+\phi(n)+p\kappa(n,p)} \\ = \frac{\sigma(n)+\phi(n)}{\sigma(n)+\phi(n)+(p-1)\sigma(n)+p\kappa(n,p)} &= \frac{1}{1+\frac{(p-1)\sigma(n)+p\kappa(n,p)}{\sigma(n)+\phi(n)}}\end{aligned}$$

Let

$$T_o(n,p) = (p-1)\sigma(n) + p\kappa(n,p), T(n,1) = \sigma(n) + \phi(n)$$

then, the formula becomes

$$\varepsilon \le \frac{1}{1+\frac{T_o(n,p)}{T(n,1)}}$$

Assume efficiency is constant(to effectively utilize parallelism), it can be re-write as

$$T(n,1) \ge \frac{\varepsilon}{1-\varepsilon} T_o(n,p) = CT_o(n,p) \qquad \text{(Isoefficiency Relation)}$$

- Corollary: Scalability function

The isoefficiency relation can often be simplified as $n \ge f(p)$

Let $M(n)$ denote memory required for problem of size n

Scalability function: $M(f(p))/p$ shows how memory usage per processor required at least to maintain same efficiency

Q1

It's obvious that $f = 95\% = 0.95, p = 10$. According to the Amdahl's Law, the maximum speed up is

$$\Psi \le \frac{1}{0.05 + 0.95/10} = \frac{1}{0.145} \approx 6.9$$

Q2

Let us assume $\Psi = 10$. To get the minimum number of processors, assume that other parts can be parallelized.

Thus $f = (1-0.06) = 0.94$.

$$\Psi = 10 \le \frac{1}{(0.06 + \frac{0.94}{p})} \implies 0.06 + \frac{0.94}{p} \le \frac{1}{10} \implies p \ge \frac{0.94}{0.04} = 23.5 \implies p = 24$$

Q3

As stated in the problem, $\Psi = 50$. According to the Amdahl's Law,

$$\Psi = 50 \le \frac{1}{(1-f)+\frac{f}{p}} < \frac{1}{(1-f)}, f \neq 0 \implies f' = 1 - f < \frac{1}{50} = 0.02$$

Q4

According to the problem statement, $\Psi = 9, p = 10$. According to the Amdahl's Law,

$$\Psi = 9 \le \frac{1}{(1-f)+\frac{f}{10}} \implies (1-f) + \frac{f}{10} \le \frac{1}{9} \implies \frac{8}{9} \le \frac{9}{10} f \implies f \ge \frac{80}{81} \approx 0.988$$

Q5

According to the problem statement, $p = 16, t_{parallel,scaled} = t \times p = 233 \times 16 = 3728$.

According to the Gustafson-Barsis' Law,

$$s = s_{scaled} = \frac{t_{sequential,scaled}}{t_{total,scaled}} = \frac{9}{3728+9} = \frac{9}{3737}$$

$$\Psi \le p + (1-p)s = s + (1-s)p = \frac{9}{3737} + \frac{3728}{3737} \times 16 = 15.96$$

Q6

According to the problem statement, $p = 40, s_{original} = 1 - 99\% = 0.01$.

Thus,

$$s = s_{scaled} = \frac{1 \times t}{1 \times t + 99 \times t \times 40} = \frac{1}{3961} = 0.000252$$

According to the Gustafson-Barsis' Law,

$$\Psi \le s + (1-s)p = \frac{1}{3961} + \frac{3960}{3961} \times 40 = 39.99$$

Q7

No.

According to the problem statement,

$\Psi_1 = 9, p_1 = 10; \Psi_2 = 90, p_2 = 100.$

According to the Karp-Flatt Metric,

$$e_1 = \frac{\frac{1}{9} - \frac{1}{10}}{1 - \frac{1}{10}} = \frac{\frac{1}{90}}{\frac{9}{10}} = \frac{1}{81}$$

$$e_2 = \frac{\frac{1}{90} - \frac{1}{100}}{1 - \frac{1}{100}} = \frac{\frac{1}{900}}{\frac{99}{100}} = \frac{1}{891}$$

The calculation results show that as the degree of parallelism increases, the cost e actually decreases.

According to the Karp-Flatt Metric Corollary,

$$e = \frac{\frac{1}{\Psi} - \frac{1}{p}}{1 - \frac{1}{p}} = f + \frac{\kappa(n,p)[\frac{p}{p-1}]}{\sigma(n) + \phi(n)} \overset{\lim_{p\to\infty}}{\approx} \frac{\sigma(n) + \kappa(n,p)}{\sigma(n) + \phi(n)}$$

If n remains constant, the equation can be simplified to

$$e = f + \frac{\kappa_n(p)}{K}$$

For this question, $e_1 = \frac{1}{81}, e_2 = \frac{1}{891}$.

It can be concluded that $\kappa_n(p)$ decreases as p increases, which is unreasonable (contradicting the model assumptions), and therefore the case is impossible.

Q8

$$t_{total} = t_{sequential} + t_{parallel} + t_{communication}, t_{parallel}(p) = \frac{t_{parallel}(1)}{p}$$

In order to achieve minimum time, assume there are no communication time for the task, thus

$$t_t(p) = t_s + t_p(p) = t_s + \frac{t_p(1)}{p}$$

According to the problem statement,

$$t_t(1) = t_s + t_p(1) = 1000$$

$$t_t(4) = t_s + \frac{t_p(1)}{4} = 500$$

Therefore,

$$t_s = \frac{1000}{3}, t_p(1) = \frac{2000}{3}$$

Thus,

$$t_t(16) = t_s + t_p(16) = \frac{1000}{3} + \frac{\frac{2000}{3}}{16} = \frac{1125}{3} = 375$$

Q9

According to the Isoefficiency Metric and the Scalability function,

a.

$n \geq f(p) = Cp, M(n) = n^2 \implies M(f(p))/p = C^2 p$

b.

$n \geq f(p) = C\sqrt{p}\log p, M(n) = n^2 \implies M(f(p))/p = C^2 \log^2 p$

c.

$n \geq f(p) = C\sqrt{p}, M(n) = n^2 \implies M(f(p))/p = C^2$

d.

$n \geq f(p) = Cp\log p, M(n) = n^2 \implies M(f(p))/p = C^2 p \log^2 p$

e.

$n \geq f(p) = Cp, M(n) = n \implies M(f(p))/p = C$

f.

$n \geq f(p) = p^c, M(n) = n \implies M(f(p))/p = p^{c-1} \quad where\ 1 < C < 2$

g.

$n \geq f(p) = p^c, M(n) = n \implies M(f(p))/p = p^{c-1} \quad where\ C > 2$

Thus,

$g > d > a > f > b > c > e$

Q10

According to the problem statement, we can assume there is no time that can be only executed sequentially for matrix-matrix multiplication. Thus,

$$\sigma(n) = 0, \phi(n) = 2n^3, \kappa(n, p) = 16n^2 \log_2 p, p = 1024$$

According to the Karp-Flatt Metric

$$\Psi(n, 1024) \leq \frac{2n^3}{\frac{2n^3}{1024} + 16n^2 \log_2 1024} = \frac{n}{n + 81920} \times 1024 = 1024\left(1 - \frac{81920}{n + 81920}\right)$$

The above equation is an increasing function of n, so the larger n is, the greater the speedup.

Let's check the memory usage. According to the problem statement, $M(n) = 24n^2$ bytes.

To get the maximum speedup, we have to make full use of memory(the maximum allowable n can be obtained), therefore

$$M(n) = 24n^2 \leq 1024 \times 2^{30} \implies n \leq \frac{128 \times 2^{30}}{3} = 16 \times 2^{15} \times \frac{\sqrt{3}}{3} \approx 302697$$

Thus, the maximum speedup is

$$\Psi(302697, 1024) = 1024(\frac{302697}{302697 + 81920}) \approx 806.6$$

Additionally,

$$\Psi(n, 1024) = 256 \leq 1024\frac{n}{n + 81920} \implies n + 81920 \leq 4n \implies n \geq \frac{81920}{3} \approx 27307$$