计算机系统结构 第二讲流水线的性能指标(1) 冯丹

武汉光电国家研究中心







1.吞吐率的定义

 吞吐率通常使用缩写TP表示,是指在单位时间 内流水线所完成的任务数量或输出结果的数量。

$$TP = \frac{n}{T_k}$$

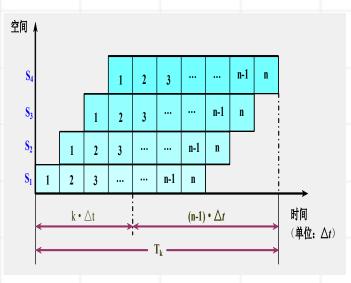
n: 任务数

T_k: 处理完成n个任务所用的时间





2-1.吞吐率分析: 各段时间均等



- ho (k段线性) 流水线完成n个连续任务所需要的总时间 $T_k = (k+n-1) \Delta t$
- > 实际吞吐率

$$TP = \frac{n}{(k+n-1)\Delta t}$$

> 最大吞吐率

$$TP_{\text{max}} = \lim_{n \to \infty} \frac{n}{(k+n-1)\Delta t} = \frac{1}{\Delta t}$$



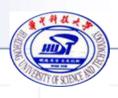


2-2.最大吞吐率与实际吞吐率

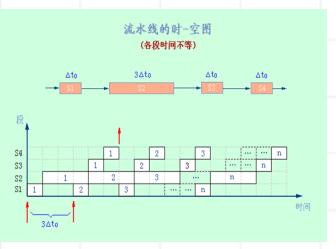
$$TP = \frac{n}{k+n-1}TP_{\text{max}}$$

- 流水线的实际吞吐率小于最大吞吐率,它除了与每个段的时间有关外,还与流水线的段数k以及输入到流水线中的任务数n等有关。
- 只有当n>>k时,才有TP≈TP_{max}。





3.吞吐率分析: 各段时间不全相等



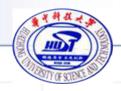
- (k段线性)流水线完成n个连续任务所需要的总时间为
 - $T_k = \sum \Delta t_i + (n-1) \max(\Delta t_i) \qquad (i=1, \dots, k)$
- > 实际吞吐率

$$TP = \frac{n}{\sum_{i=1}^{k} \Delta t_i + (n-1) \max(\Delta t_1, \Delta t_2, \dots, \Delta t_k)}$$

> 最大吞吐率

$$TP_{\max} = \frac{1}{\max(\Delta t_1, \Delta t_2, \dots \Delta t_k)}$$





4-1. 流水线瓶颈问题

- 流水线中执行时间最长的段称为流水线的瓶颈段
- 瓶颈段使得流水线性能下降
 - 各段时间均等的流水线

$$TP_{\text{max}} = \lim_{n \to \infty} \frac{n}{(k+n-1)\Delta t} = \frac{1}{\Delta t}$$

• 各段时间不全相等的流水线

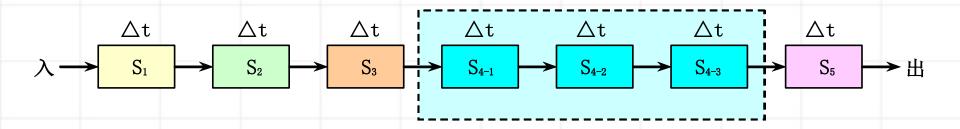
$$TP_{\max} = \frac{1}{\max(\Delta t_1, \Delta t_2, \dots \Delta t_k)}$$

- 解决流水线瓶颈的方法
 - 细分瓶颈段
 - 重复设置瓶颈段





4-2. 细分瓶颈段

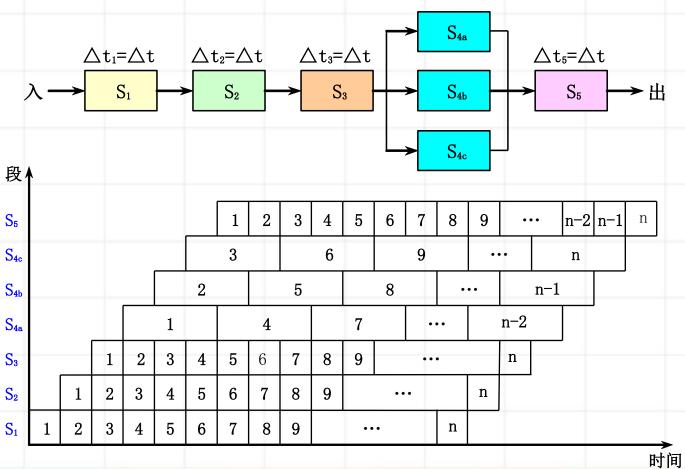






4-3. 重复设置瓶颈段









5. 加速比

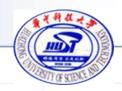
- -加速比(S)
 - 完成同样一批任务,不使用流水线所用的时间与使用流水线所用的时间之比。

$$S = \frac{T_s}{T_k}$$

Ts: 不使用流水线(即顺序执行)所用的时间

T_k: 使用流水线后所用的时间





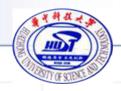
6.加速比分析: 各段时间均等

- ▶ 一条k段流水线完成n个连续任务所需要的时间
 - $T_k = (k+n-1)\Delta t$
- » 顺序执行n个任务所需要的时间
 - T_s= nk∆t
- > 实际加速比

$$S = \frac{nk}{k + n - 1}$$

$$S_{\max} = \lim_{n \to \infty} \frac{nk}{k + n - 1} = k$$

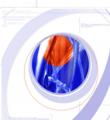


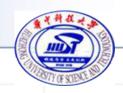


7.加速比分析: 各段时间不全相等

> 实际加速比

$$S = \frac{n\sum_{i=1}^{k} \Delta t_i}{\sum_{i=1}^{k} \Delta t_i + (n-1) \max(\Delta t_1, \Delta t_2, \dots, \Delta t_k)}$$





8.思考

> 段数越多越好?





谢谢大家

