

表1-11 几种现代处理器的制造成本因素

芯 片	晶片尺寸 (mm ²)	估测缺陷率 (每cm ²)	制程 (nm)	晶体管数 (百万个)
IBM Power5	389	0.30	130	276
Sun Niagara	380	0.75	90	279
AMD Opteron	199	0.75	90	233

1.1 [10/10] <1.6>表 1-11 给出了影响几种当前芯片成本的相关芯片统计数字。在下面几个练习中，我们将研究 IBM Power5 的不同设计决策所产生的影响。

- [10] <1.6>IBM Power5 的成品率是多少？
- [10] <1.6>为什么 IBM Power5 的缺陷率要低于 Niagara 和 Opteron？

解：a. 成品率 = $(1 + \frac{0.30 \times 3.89}{4.0})^{-4} = 0.36$

b. 因为 IBM 的产品规格更大，更好优化，所以缺陷率会比较低。

1.12 [20/20/20] <1.1、1.2、1.7>在 Amazon 或 eBay 使用的服务器场中，一个故障不会导致整个系统崩溃，而是减少在任意时刻能够满足的请求数目。

- [20] <1.7>如果一个公司有 10 000 台计算机，每台计算机的 MTTF 为 35 天，而且只有当 1/3 以上的计算机发生故障时才会经历灾难性故障，系统的 MTTF 为多少？
- [20] <1.1、1.7>如果一台计算机的 MTTF 加倍，需要另加 1000 美元，这是不是一个很好的业务决策？证明你的结论。
- [20] <1.2>表 1-2 给出了宕机的平均成本，假定在一年的所有时间内，该成本不变。但对于零售商来说，圣诞节是盈利水平最高的（因此，如果因为宕机造成无法销售，损失也最大）。如果目录销售中心在第四季度的通信流量是其他任一季度的两倍，那第四季度每小时的平均宕机成本是多少？其他时间的宕机成本又是多少？

解：a. $35/10000 \times 3333 = 11.67$ 天

b. 在目前的系统中，大约每 5 分钟就有一台电脑出现故障。5 分钟不太可能是足够的时间来隔离计算机，让计算机重新上线。然而 10 分钟有可能。无论如何，这将大大延长 1/3 的计算机同时停机的时间。由于停机的成本非常巨大，能够延长停机时间是非常有价值的。

- $$\begin{aligned} \$90,000 &= (x + x + x + 2x)/4 \\ \$360,000 &= 5x \\ \$72,000 &= x \\ 4\text{th quarter} &= \$144,000/\text{hr} \end{aligned}$$

1.15 [15/10] <1.9>假定我们对一台计算机进行了升级，使某种执行模式提升为原来的 10 倍。升级模式的使用时间占总时间的 50%，这一数值是在使用该升级模式时测得的执行时间百分比。回想一下，Amdahl 定律需要的是能改进但还没有改进的原执行时间比例。因此，在使用 Amdahl 定律计算加速比时，不能使用这个 50% 的测量值。

- [15] <1.9>从快速模式获得的加速比是多少？
- [10] <1.9>转换为快速模式的原执行时间比例是多少？

解：a. 旧的执行时间 = $0.5 \text{ new} + 0.5 \times 10 \text{ new} = 5.5 \text{ new}$

b. 在原代码中，未增强部分的时间等于增强部分提高 10 倍，因此：

$$\begin{aligned} (1 - x) &= x / 10 \\ 10 - 10x &= x \\ 10 &= 11x \\ 10/11 &= x = 0.91 \end{aligned}$$

2.10 [10/12] <2.2>有一种新的微处理器，需要研究分组 L1 数据缓存与流水化 L1 数据缓存的相对性能。假定有一种 64 KB 两路组相联缓存，其块大小为 64 字节。流水化缓存由三级流水构成，类似于 Alpha 21264 数据缓存的容量。分组实现方式由两个 32 KB 两路组相联组成。使用 CACTI，并假定采用 65 nm (0.065 μm) 工艺，回答以下问题。Web 版本的周期时间输出表明缓存可以在什么样的频率下正常工作，不会在流水线中产生气泡。

- a. [10] <2.2>该缓存的周期时间与其访问时间相比为多少？该缓存将占用多少个流水级（精确到小数点后两位）？
- b. [12] <2.2>对比流水线设计与分组设计的每次访问的面积及总动态读取能耗。说明哪种设计占用的面积较少，哪种需要的功率较多，解释其原因。

解：a. 访问时间是 1.12ns，而循环时间是 0.51ns

流水级 = $1.12/0.51 = 2.2$

b. 流水线设计(不包括锁存面积和功率)面积为 1.19 mm²，每次接入能量为 0.16nJ。存储缓存的面积为 1.36 mm²，每次访问的能量为 0.13nJ。分组设计使用的面积略大一些，因为它有更多的感应安培和其他电路来支持这两种设计，而流水线设计消耗的能量略大一些，因为活跃的内存阵列比分组的情况下更大。

2.16 [12] <2.3>你正在准备一台服务器，它采用八核 3 GHz CMP，在执行某一工作负载时的总 CPI 为 2.0(假定 L2 缓存缺失填充没有延迟)。L2 缓存行的大小为 32 字节。假定该系统采用 DDR2-667 DIMM。如果有时需要的带宽为平均带宽的 2 倍，那么应当提供多少个独立存储器通道才能使系统不受存储器带宽的限制？该工作负载平均第 1000 条指令导致 6.67 次 L2 缺失。

解：核心将执行 $8\text{cores} \times 3\text{GHz} / 2.0\text{CPI} = \text{每秒 } 120 \text{ 亿条指令}$ 。这将产生每秒 $120 \times 0.00667 = 8000 \text{ 万 L2 缺失}$ 。

$80 \times 32\text{B} = 2560\text{MB/秒}$ 。

如果有时需要的内存带宽是平均带宽的 2 倍，则为 5120MB/秒。这几乎不在 DDR2-667 DIMM 提供的带宽范围内，因此只需一个内存通道即可。

3.9 [5] <3.4>如果你不清楚寄存器重命名程序必须做哪些工作，可以回过头来看看正在执行的汇编代码，问问自己，必须具备哪些条件才能获得正确结果。例如，考虑一个 3 路超标量机器，同时对下面这 3 条指令进行重命名：

ADDI R1, R1, R1

ADDI R1, R1, R1

ADDI R1, R1, R1

如果 R1 的值在初始为 5，那么执行这一序列之后，它的值应当是多少？

解：ADD R1, R1, R1; $5 + 5 \rightarrow 10$

ADD R1, R1, R1; $10 + 10 \rightarrow 20$

ADD R1, R1, R1; $20 + 20 \rightarrow 40$

所以结果是 40