

[美] DOMENICO FERRARI
[意] GIUSEPPE SERAZZI
[意] ALESSANDRO ZEIGNER



郑衍衡 王春元等 译 李润斋 校

计算机系统的测量和优化

水利电力出版社

计算机系统的测量和优化

郑衍衡 王春元等 译 李润斋 校



水利电力出版社

内 容 提 要

本书比较全面地、系统地介绍了测量和提高计算机系统性能的途径和方法。

全书分十章，主要内容有：工作负荷，测量原理，测量数据的表示方法，测量仪器，系统的调准，程序的调准，解析模型及其应用，投资分析。每章都附有参考文献。

本书是一本面向实际应用的参考书，可供计算机系统引进、科研、设计、生产、应用部门的管理人员和科技人员参考。本书又是美国电气与电子工程师协会(IEEE)推荐的计算机专业研究生有关课程的参考书，可供于我国大专院校计算机专业的教师、研究生和高年级学生参考。

JK451/13

DOMENICO FERRARI
GIUSEPPE SERAZZI ALESSANDRO ZEIGNER
Measurement and Tuning of Computer Systems
Prentice-Hall Englewood Cliffs, 1983

计算机系统的测量和优化

郑衍衡 王春元 等译 李润奇 校

水利电力出版社出版、发行

(北京三里河路6号)

各地新华书店经售

水利电力出版社印刷厂印刷

*

787×1092毫米 16开本 26.25印张 596千字

1988年10月第一版 1988年10月北京第一次印刷

印数0001—3210册 定价8.50元

ISBN7-120-00370-4/TP·15

译 者 序

我国计算机系统的数量正在日益增加，投资已相当可观。为了更好地取得投资效益，如何合理选择系统规模和配置，如何有效地发挥系统的功能和能力，以及如何对系统进行扩充，这些都是需要有科学依据，属于计算机系统性能评价的范畴。开展计算机性能评价方面的研究，对提高我国计算机的应用水平有着重要的作用，是一项值得重视而且极有意义的工作。

系统性能评价研究的重要目标之一，是在系统运行过程中合理调整工作负荷，使之达到最佳的性能指标。而评价的基础则是系统工作负荷的测量。因此，管理与维护计算机系统的科技人员必须掌握系统的测量与调准技术。

本书比较全面地、系统地介绍了测量与提高各类计算机系统性能的途径和方法。它吸取了该领域中大量研究成果，列举了大量事例予以阐述。现将其译出，希望有助于这方面研究工作的开展。

本书第一章、第二章由王春元翻译，第四章、第五章由刘政、王春元合译，第六章由郑衍衡翻译，第三章、第九章由孔繁华翻译，第七章、第八章由郑衍衡、孔繁华合译，第十章由廖万清翻译。全书由郑衍衡统稿。李润斋对全书作了校订。

由于我们的经验和能力所限，本书一定存在不少错误和缺点，欢迎大家批评指正。

郑 衍 衡

1987年6月

前 言

虽然还不能把性能评价看成是一门已经完全成熟的学科，但是在它的不太长的历史中，已经得到了显著的发展。近年来在构造计算机模型技术，特别是在解析技术，以及在工作负荷的特性和预测、容量规划、配置设计方面取得了很大进展。在测量技术和工具、调准过程和治疗技术方面也有一定进展。

近年来，在性能分析和评价方面出版了不少书籍，其中大部分的读者对象是研究工作者和学生，而不是面向从事这个领域的实际工作人员和管理人员。还有几本书，虽是有关这一领域的专著，却没有包含一些取得最重大进展的领域，例如解析模型方面的进展。

一些见多识广的实际工作者在几年以前可以对某些课题视而不见，他们认为这些课题对解决他们日常提出的问题没有任何帮助，但是今天这些课题对实际工作已经变得非常重要，而且他们的知识也立即成为在性能评价领域内不可缺少的技术基础。和其他各类计算机专家一样，一个成功的性能评价专家要不断取得成就，不仅要了解硬件和软件工艺技术的进展，而且要掌握性能的测量和模型构造技术的不断发展状况。

上述考虑是编写本书的出发点，目前还没有一本简便的、易懂的、面向实际的性能评价的入门方面的、让读者能了解本领域最近的重要进展的书。学生或者研究工作者，也可以从以实际经验和坚实的基本概念出发，对课题进行较为实用的探讨中取得收获，在本书中不可能要求讨论性能评价这门学科涉及的所有领域，我们选择测量技术（我们认为测量确实是基础。没有测量，对系统的评价是不可能进行的）和调准技术作为重点。绝大部分的讨论和事例研究的内容都是针对着已经存在的和正在运行的系统的性能评价。在这里，评价活动的主要目的之一是提高计算站的性能价格比。

第一章介绍一些基本概念和问题，并对在进行提高性能的研究时要用到的基本性能指标进行了广泛的讨论。我们把测量技术作为本书的重点之一，反映在第二章中是讨论工作负荷特性。这章的绝大部分叙述了构造和验证用于性能测量实验中的可执行的工作负荷模型，在这一章中也提出了在这个领域中若干新的研究成果和方向。

第三章研究测量计算机系统和计算机系统模型的基本技术。第四章讨论表示测量结果的一些最有效的方法，这些方法使结果易于解释。第五章描述主要的几种测量工具和它们的特性。第六章从方法学的角度描述将这些技术和工具用于解决调准一个系统的问题。第七章提出大量事例研究，包括各类不同的系统、配置和应用环境。第八章讨论一些用于改进程序性能的特殊方法和测量工具。程序性能的改进对程序员和系统调准人员都是一个重要的问题，因为系统工作负荷的变化将会使系统失调，工作负荷的改进常常是改进系统性能的一个有效途径。

解析模型在调准研究中起着越来越重要的作用。在对某些性能方面的问题的诊断上，解析模型将提供有力的帮助。并且在系统要进行一项费钱而又没有十分把握的改动以前，

能够预测这项改动的效果。在第九章中讨论了运行分析法和排队模型技术，其中几个实例研究解释了在性能改进研究中模型技术的若干应用。一些实例说明了模型如何去解决在第七章中只能靠经验进行处理的问题。

最后一章，即第十章，主要涉及一个重要问题，这个问题在进行性能评价技术方面的讨论时是很难涉及的，就是在性能改进中将获得多少经济效益。这问题是一个实际问题，因为并不是每一项调准研究都能产生经济效益。首先，不是每一项调准研究都能成功，其次，即使这项措施在技术上是成功的，在经济上也不一定产生效益，一项改进性能的研究应该看成为一项投资，并且应在价格效益分析的基础上进行评价。几个实例研究解释了如何完成这项分析，并同时讨论了它的一般原则和方法。

本书的每一章节都用二位或三位数来表示。用二位数表示的章节通常是对某一方面进行定性的描述，是面向要求对问题及其解答进行一般性了解的读者。用三位数字表示的章节对某一方面进行较为深入和详尽的研究，或讨论已经描述过的一些技术应用的某些特殊实例。

作者感谢 G.C.Baldovini, Syntax S.P.A. 的主要负责人，在本书准备初稿和写作过程中给予的鼓励和支持。还感谢意大利的 Honeywell Information Systems 的 F.Filippazzi 教授的支持。本书的构想和极大部份写作是 D.Ferrari 作为 Pavia 大学的访问教授时完成的，并得到加州大学和自然科学基金会 MCS78-24618 和 MCS80-12900 号拨款的支持。Pavia 大学的信息和系统讲习会的 I.Delotto 教授和数学分析讲习会的 E.Magenes 教授促成了这次访问并提供了本书写作的适宜环境。G.Serazzi 二次对 Berkeley 的旅行完成的初稿，这是由 Consiglio Nazionale delle Ricerche 的 79.0233.62 和 80.02334.01 号合同所支持的，它并为 Berkeley 和 Pavia 合作的 NSF-CNR 性能评价研究项目提供经费。作者很感激合同的主要主持者 M.Italiani 教授在经济上和精神上的支持。

谨以本书献给作者的妻子和儿女们，以报答他们失去无数的丈夫和父亲的时间。

多梅尼科 费拉里
吉乌斯比·塞拉齐
亚历山德罗·齐格纳

目 录

译者序

前 言

第一章 概述	1
1-1 定义和基本概念	1
1-2 评价的内容	1
1-3 参考系统	3
1-3-1 单道程序批处理参考系统 (UBRS)	4
1-3-2 多道程序批处理参考系统 (MBRS)	4
1-3-3 多道程序交互式参考系统 (MIRS)	6
1-3-4 多道程序交互式虚拟存贮参考系统 (MIVRS)	7
1-4 性能指标	7
1-4-1 解题时间	12
1-4-2 响应时间	16
1-4-3 吞吐率	19
1-5 评价技术	27
参考文献	28
第二章 工作负荷	30
2-1 工作负荷特性	30
2-2 工作负荷模型的代表性	34
2-2-1 模型代表性评价举例	37
2-3 测试工作负荷	40
2-3-1 实际测试工作负荷	41
2-3-2 综合测试工作负荷	43
2-3-3 人工测试工作负荷	47
2-4 工作负荷模型的实现技术	55
2-4-1 分布采样	60
2-4-2 实际工作负荷的采样	62
2-4-3 分类	63
2-4-4 联合概率分布采样	70
2-4-5 主成分分析	71
2-4-6 一种工作负荷模型的实现	75
2-5 几种工作负荷模型的实现方法	78
2-5-1 事例2-1: 一种批处理工作负荷的资源和功能模型	79
2-5-2 事例2-2: 以面向性能准则为基础的一种交互式工作负荷模型	83

2-6	工作负荷预测系统能力计划	89
2-6-1	一种新应用的负荷的估计	94
2-6-2	事例2-3: 一种实时应用负荷的预测	102
	参考文献	110
第三章	测量原理	115
3-1	概述	115
3-2	事件检测	116
3-3	采样	118
3-3-1	样本的选择和精度	119
3-4	模拟	123
3-4-1	一个简单模拟器的构成	127
	参考文献	134
第四章	测量数据的表示方法	137
4-1	引言	137
4-2	表和图	137
4-2-1	利用率(或Gantt)分布图	148
4-2-2	Kiviat图	152
4-2-3	Kiviat图的标准形状	156
	参考文献	157
第五章	测量仪器	160
5-1	测量工具的特性	160
5-2	软件监测器	161
5-2-1	采样监测器	163
5-2-2	时间测量、时钟和定时器	167
5-3	硬件监测器	169
5-3-1	硬件监测器应用举例	174
	参考文献	177
第六章	一种系统调准的方法学	180
6-1	概述	180
6-2	测量工具的选择	183
6-3	一次测量活动的设计	188
6-4	瓶颈的检测	190
6-4-1	脱机瓶颈检测	191
6-4-2	联机瓶颈检测	194
6-5	瓶颈的排除	196
	参考文献	197
第七章	系统的调准	200
7-1	引言	200

7-2 平衡一个多道程序系统	201
7-2-1 事例7-1: 在磁盘和通道上的瓶颈	202
7-2-2 事例7-2: 主存容量不足	206
7-3 改善一个交互式系统	208
7-3-1 事例7-3: 超负荷通道	209
7-3-2 事例7-4: 双机系统之间不合适的负荷分配	211
7-4 改善一个虚拟存贮系统	217
7-4-1 事例7-5: 降低一个双机系统中的页面调用率	225
7-5 控制数据库管理系统的性能	233
参考文献	243
第八章 程序的调准	247
8-1 一般准则和选择程序	247
8-2 程序优化的方式	253
8-2-1 程序的性能指标	255
8-3 减少运行时间	259
8-3-1 代码的改进	263
8-3-2 事例8-1: 循环的优化	270
8-3-3 I/O操作的改进	274
8-3-4 降低页面调用率	282
参考文献	287
第九章 解析模型及其应用	290
9-1 引言	290
9-2 一个解析模型的实现	291
9-3 运行分析法	293
9-3-1 事例9-1: 一个虚拟存贮系统的最佳多道程序数	300
9-3-2 事例9-2: 一个交互式系统的饱和状态的分析	305
9-4 排队模型	310
9-4-1 事例9-3: 在宏观模型级上的性能分析	323
9-4-2 事例9-4: I/O 负荷平衡	329
9-4-3 事例9-5: 一个交互式系统的性能分析	332
9-4-4 事例9-6: 改善一个虚拟存贮系统的配置	336
9-4-5 事例9-7: 一个实时应用的瓶颈预测	347
9-5 具有与负荷特性有关的资源	351
9-5-1 事例9-8: 具有与负荷有关的资源的系统	354
9-6 采用等效流量聚合技术的近似解	357
9-6-1 事例9-9: 在一个交互式系统中估计由于配置变动产生的影响	364
9-7 平均值的分析	374
9-7-1 事例9-10: 将平均值分析法应用于事例9-8	375
9-8 本章中的符号一览表	377

参考文献	378
第十章 投资分析	382
10-1 概述	382
10-2 一种调准研究的成本和效益	386
10-2-1 事例10-1: 推迟采购一个新系统	389
10-2-2 事例10-2: 一个系统程序的优化	390
10-2-3 事例10-3: 改进用户程序	391
10-2-4 事例10-4: 选择一种用于控制数据库工作的工具	396
参考文献	399
附录 一个简单的排队模型分析器	402

第一章 概 述

1-1 定义和基本概念

性能这个术语，是指人或机器向需要者所提供的服务。一个信息处理系统是一种硬件和软件组成的装置，它能按用户所写的程序处理数据。因此，信息处理系统的性能这一术语，是指该系统能够提供用户的各种设施。这些设施包括：与系统交往的编程语言、为程序设计与开发所提供的工具、处理数据恢复故障的性能、所提供的数据安全的级别等等。

然而，本书所使用的性能这个术语，在含义上有较多的约束。这也是在考察其它工程系统时常用的。譬如汽车，虽然在汽车的性能指标中，应该考虑使用汽车的方便、舒适和平稳（即使按购买者的需要，对它们会有相对的权衡，但通常都是这样考虑的），但是对于术语来说，常常是指汽车在全负荷下的最大速度、达到给定速度所需的时间、燃料的消耗量、以及其它以某些方式表示汽车效能的量。汽车是按这些效能实现其功能的。

如汽车那样，计算机系统的选择与许多因素有关，其中之一就是对性能的要求。这个因素的重要程度随着人和应用的类型而变化。设计者、购买者、计算站的管理人员、程序员、临时性的用户、以及维护工程师，由于他们的要求和目的不同，对同一系统的看法就不相同。这一点也同样适用于系统的性能。对于表示性能的各种指标，不同的人，即使不同环境下的同一个人也会给出不同的估价。每个指标是可定量表示的，因而能作为评价的项目。性能指标可以用各种方法评价：可以测量、计算或估计。这些评价方法总是定量的。但是应该记住，在选择系统时有许多考虑的因素是定性的，因此是难以作定量的评价。

本书讨论性能定量评价的最一般的技术。性能是系统价值的重要组成部分（在经济的含义上）。一旦计算机系统已经被选择和安装，性能就显得非常重要。由于性能与负荷的类型及系统的使用方式有关，作相当简单的干预，性能就能经常保持在满意的水平，那么在控制之下，要保持处理效率总是很方便的。若此时性能值低于要求，或者尚存在改进系统性能的可能，则必须对低效的原因及其如何改进进行仔细的分析。这种分析会用到下而将要介绍的性能评价工具。

1-2 评价的内容

值得进行性能评价的理由，已在上节作了概括的说明，本节要作更详细的讨论。所有工程系统都是要作性能评价的。在系统设计、装配、销售及使用时，系统的性能由目的不同的人用不同的观点来评价，以证实该系统在效能上是否能满足给定的要求，是否能在某种范围内使用。

这种规划也适用于而且也应该适用于计算机系统。显然，对于投资较大或者性能要求

较高的系统，证实系统能否满足要求就更为重要。因此，在小型计算机评价中所要投资的资源数量，远远小于大型机性能改进的研究或者为批量销售的产品选择用何种微处理器所要投资的数量。

计算机系统的性能评价不仅在系统的生产过程中需要，而且在设计过程中，在用户选择时，以及在现场安装中都是需要的。评价技术的应用可以分为四种主要类型：

1. 采购：从已有的和现成的各种系统中选择一种系统，或选择系统一部分的所有评价问题均属此类。例如，计算站的设计及硬件、软件的采购。

2. 改进：对已有的运行系统，有关提高其性能的问题均属这一类。这方面的内容，本书将作详细的讨论。

3. 能力计划：这类问题是预计已有系统的未来能力，何时能以给定的性能水平处理计算站的工作负荷。

4. 设计：设计者在建立新系统时所面临的有关问题均属于这一类。

评价技术可以用来解决上述各类问题[FE 78]，[KO 78]。不过，本书主要讨论改进方面的研究。

因此，本书是以系统效率的改进作为评价研究的主要目标。这种改进经常产生十分重要的效果，但又难以度量。例如，提高用户的兴趣常常导致生产率的提高，加速新应用的开发，以及增加用户的数目。然而，有些效果从经济上来看，也还是比较容易度量的，诸如：

1. 减少系统每日运行的时间，由此可减少操作员的人数或工作时间。
2. 取消或减少由于负荷高峰所需的后备。
3. 减少系统的配置成本（例如，取消多余的通道或外部设备）。
4. 推迟因工作负荷增加而使系统达到饱和的时间。因系统进入饱和，势必要对系统扩充或替换。

显然，这些研究的实际目的是要改善价格-性能比，而不仅是性能上的改进。只要不导致性能降低太多，降低价格就是有价值的研究成果。经济问题总是首要的，即使在后面我们专门讨论技术问题（即系统性能）时也是这样。

要从一项评价研究中得益，总得考虑到成本。在承担一项研究以前，利益和成本二者都是未知数。因此，即使常常是很困难的，也必须对其利益与成本作出估计。期望在价格-性能比方面达到最大的收益，这与系统偏离理想运行状况的程度有关。但是，由于此时对最佳值这个概念还很模糊，也由于对偏离的程度往往未能了解，因此上述提法并不十分有用。关于近期的预测工作，也许通常只能作出部分估计。

改善价格-性能比的典型步骤是一种迭代过程，如图1-1所示，它的周期中包含诊断阶段和治疗阶段。诊断阶段的目的是确定价格-性能比不满足的原因。这些原因常常用瓶颈来表示。既然这样，若能消除或“扩大”限制系统性能的瓶颈，则治疗是有效的。显然，存在多种治疗方法可以达到所要求的改善，最后选用最低价格-性能比的一种治疗方法。因为有些治疗方法的费用很大（例如，扩充配置，修改操作系统，改变工作负荷），在使用这些治疗方法以前，最好预测其对价格-性能比的效果。

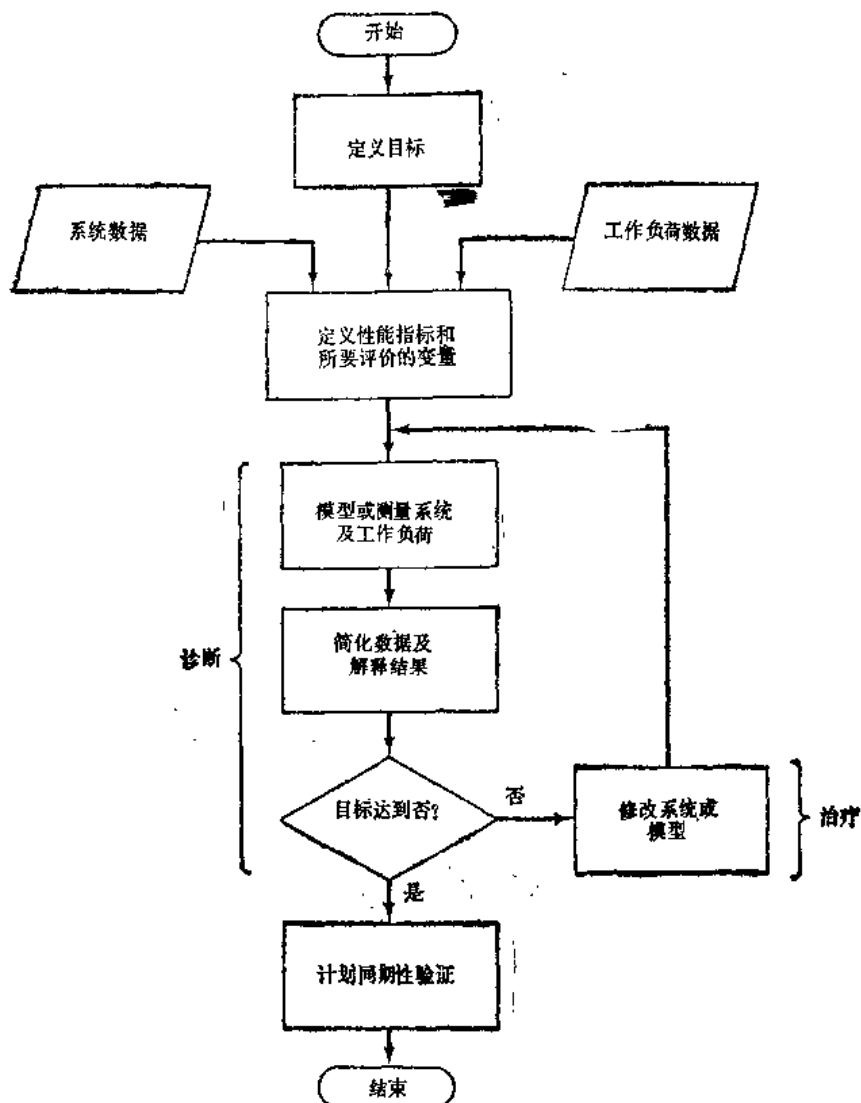


图 1-1 性能评价研究的流程图

在诊断阶段能使用的一些技术，同样也能用于性能预测。在上述过程中，诊断-治疗周期有时必须反复重复，因为有些瓶颈一旦消除，其它瓶颈将出现。不过从这些反复过程中所取得的性能的改善，一般是越来越小。所以最好将反复次数限定在几次。使用了模型方式的评价技术，也有可能将反复过程延长到过程结束，才采纳所建议的治疗方法。当然这种做法并不总是适用的。

1-3 参 考 系 统

本书所考虑的评价研究对象，已经在1-1中定义，是指信息处理系统。为了给读者更具体的定义，以及初步了解在性能评价中观察此类系统所持的观点，我们要讨论几种基本

的系统配置。这些系统的配置在一定范围内是可变的。譬如，对它们的I/O（输入/输出）数目和大容量存贮器未作硬性规定。它们在特征上的区别主要在于使用方式和操作系统所用的资源管理形式。本节选择了四种类型的参考系统（简称为RS）作为例子。它们的配置可以从已有的计算机系统中见到。现按其复杂程度，从简到繁予以介绍。

1-3-1 单道程序批处理参考系统（UBRS）

在此系统中，使用批处理方式，用单道程序方式运行管理计算机的主要资源。图1-2表示一种可能存在的系统配置。此配置的特征是有二个通道，一个（ C_0 ）联到系统控制台C和外部设备（卡片读入器CR、行式打印机LP）；另一个（ C_1 ）联到大容量存贮器（磁盘D和磁带T）。用户将程序、数据以及到操作系统的命令打在卡片上，用卡片读入器CR读入，对系统进行访问。通道 C_0 将打在卡片上的信息传送到主存贮器M，由中央处理器（CPU）进行处理。每个程序总是包含若干程序步，例如：编译、目标程序的加载、执行、以及在文件系统中存入程序。从这些程序步得到的输出被送往 C_0 ，只要程序需要输出，就由 C_0 控制输出打印。甚至读卡也是按步进行，只有在第一程序步完成后，第二程序步的卡片才读入，如此继续下去。

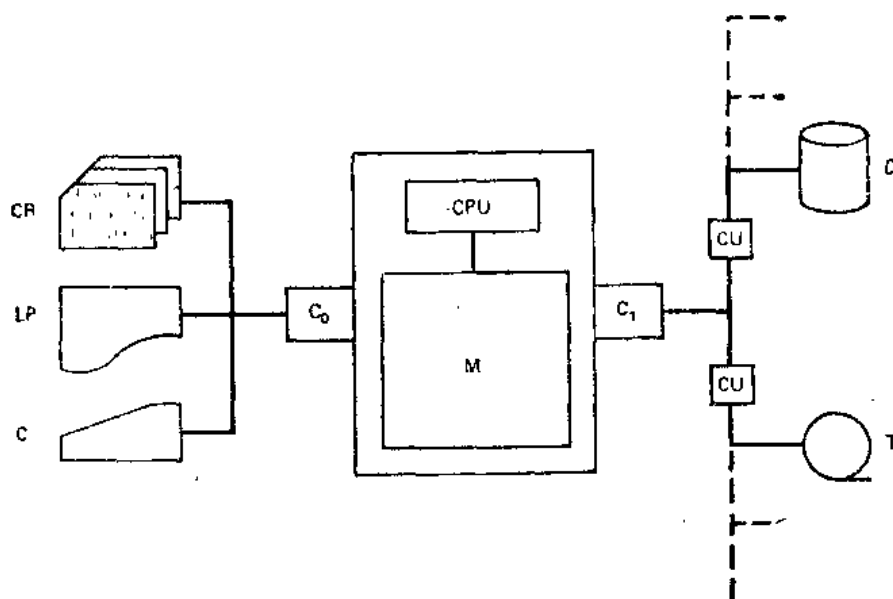


图 1-2 典型的UBRS系统配置

因此，在这类系统中，操作系统未用到的所有存贮空间，都可以为执行中的作业步所使用；在主存贮器与外部设备或大容量存贮器之间传送数据时，CPU并不工作。图1-2中，以缩写CU（控制器）表示一个控制单个I/O设备活动的硬件部件，用通道请求方式使它们之间同步。

1-3-2 多道程序批处理参考系统（MBRS）

历史上，多道程序最初是以相当有限的假脱机方式（所谓假脱机即连机同时操作）开始的。这种技术代表了一种最简单的重叠活动型式。在假脱机系统中，通道的活动可以与

CPU的活动相重叠。因此，在程序执行时，可以读入一组卡片或者打印另一程序的结果。为此，必须在主存储器M中，为来自卡片读入器的数据，或者由程序产生的数据分配缓冲区，作为暂存之用。这些数据以后将积累在大容量的存储器中。

假脱机系统依旧可认为是一种单道程序系统，因为在任何给定的时间内只有一个用户程序（或者更确切地说，只是一个程序步）驻留在主存储器中。不过这种类型的系统允许某些活动互相重叠，这一点是单纯的单道程序系统所不能做到的。

除假脱机外，若与各用户作业步相对应的程序与数据在同一时间内能够存在，则就是实际的多道程序系统。多道程序的概念的引出是为了增加CPU的利用率，从而提高系统生产率。每当挂起一个程序步（例如：因为它要访问外部设备或大容量存储器），CPU就将该挂起程序步的状态保存起来，并开始执行一个早已准备执行、且已装入存储器的程序。若主存储器中没有装好可以立即开始执行或继续执行的程序，CPU才停止活动。图1-3表示一种可能存在的MBRS配置。

此时， C_0 与 C_1 所起的作用与在UBRS中相同， C_1 控制几台固定头磁盘的信息传送（固定头磁盘常常又称为磁鼓，因其操作与鼓相同）。

通常通道有三种型式：多路转换器、选择器与块多路转换器。

多路转接通道（有时称为字节多路转接器）可以对I/O设备传送信息。传送的时候，可以把来自各种设备的信息，以字符为单位混合进行。此种通道是将所有慢速设备与存储器相联，因此，举例来说，从卡片读入器来的字符可能后随着一个从终端来的或者从控制台来的字符。图1-2和1-3的通道 C_0 是典型的多路转接通道。

选择通道是与快速设备联用的，快速设备的信息可以按记录（例如一个磁盘记录或一

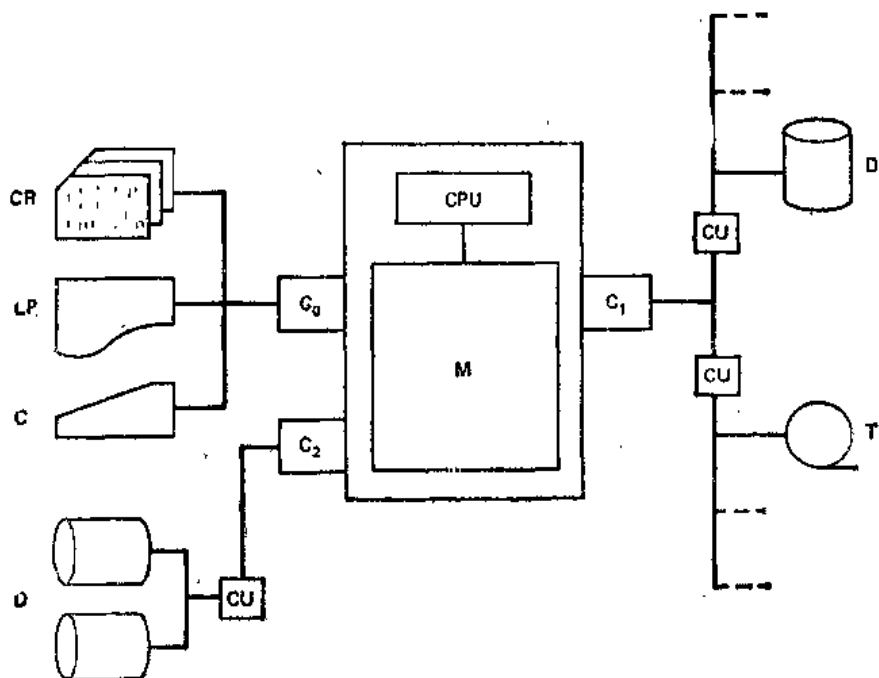


图 1-3 典型的MBRS系统配置

个磁带记录) 传送, 这是因为没有足够的时间将通道控制从一种传送转为另一种传送, 或者没有足够的缓冲空间使之不丢失信息。因此一个选择通道 (若它有数个子通道, 则是每个独立子通道) 在一次给定的传送中完全由它控制的一台设备所占用。不过, 读-写头的运动或者磁带的倒带操作, 包括某些旋转设备的等待时间, 可以与同一通道上其它设备的数据传送相重叠。

块多路转接器能够将来自多个设备的信息组成块的形式 (一块由多个字符组成) 进行传送, 而不是以单个字符形式传送, 这些通道所控制的设备, 比通常联接到多路转接通道的设备要快的多。图1-2和图1-3中的 C_1 与 C_2 可以是选择器式, 也可以是块多路转接器式。

1-3-3 多道程序交互式参考系统 (MIRS)

此类系统的使用方式是混合的, 因为除了批处理方式外, 它还支持交互式的处理方式, 其特征是存在着用户与系统对话的交互式终端。交互式事务处理包括一系列用户-系统的相互作用。相互作用由两部分组成: 用户通过终端的输入设备 (通常是键盘) 向系统送入命令与信息; 系统在同一终端给出回答信息。

在图1-4的配置中, 终端与多路转接通道 C_1 相连。为了实行对话, 系统对负荷轻的请求必须立即回答 (例如, 删除文本的一行, 替换一个字符, 或者命名一个文件)。引入分

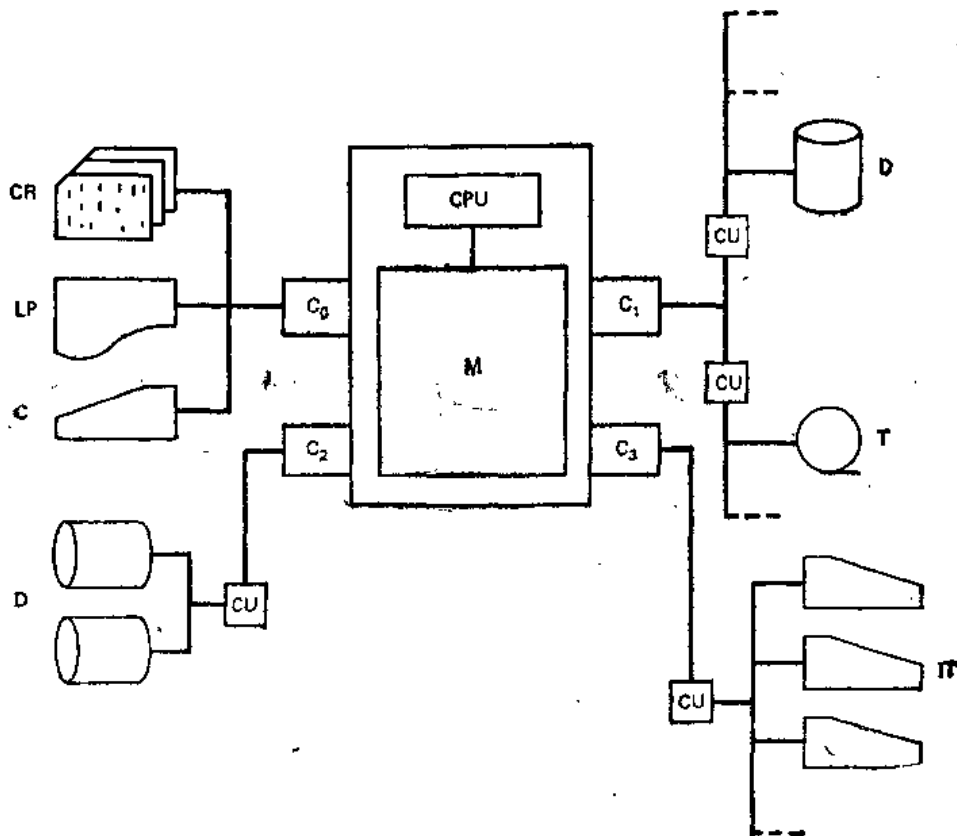


图 1-4 典型的MIRS和MVRS系统配置

分时技术是为了将负荷轻的请求的响应时间减到最小程度。分时技术对一个请求独占某资源（尤其是CPU）的时间间隔规定了一个最大值。若有任何请求在等待，则目前正在得到服务的请求在限时（分配给该技术的时间间隔，例如100ms）一到即被中断。这种方法可避免负荷重的请求严重地推迟较轻负荷的请求，许多负荷轻的请求都可在第一次限时内完成。分时已普遍应用于交互式的系统，以致这些系统常常称为分时系统。不过，这种技术也用于非交互式系统中。例如，它可用于批处理系统，以减少较短程序的解题时间。

1-3-4 多道程序交互式虚拟存贮参考系统（MIVRS）

这类系统具有图1-4所示的配置，在此系统中，用户编程的地址空间（虚拟存贮空间）是与实际主存贮器M的地址空间不相同的。一般这种编程空间非常之大，不用担心因程序不能适合所分配的存贮空间，而要分块（覆盖）轮流装入M。

在具有虚拟存贮的计算机中，存储器的管理完全由系统掌握，换句话说，存储管理是完全自动的。图1-4中所示的设备功能与MIRS相同，只是与通道C₁联接的磁盘有特殊的例外，磁盘是作为页设备使用的，即在系统自动的存储器管理层次中作为第二层使用的。每一个程序只有一部分存放在M中，这部分程序又全部存放在页设备中，并且被分成称为页面的相等部分。磁盘中存贮着全部程序，当程序所需要的某些指令或数据不在M中时，则含有这些指令和数据的页面自动从磁盘传送到M。在实际存贮器中，经过程序修改的页面，当M中不再有空间存放时，由操作系统重新写入磁盘。

1-4 性能指标

计算机系统的性能评价是各方面人员感兴趣的问题，每方面的人员都给“性能”这个词以特定的含义。1-1中指出用系统效率来鉴别性能，这只是给出的第一步。我们将较详细地描述各类人员和他们的观点，以完善我们对性能的定义。

对于参与评价系统的人员可以进行分类，一种可能的分类在图1-5中表示。虽然他们有同一目标，都要使系统有较高的运行效率，但图中的三类人员都以完全不同的观点来考虑问题。

硬件和软件的设计者非常关心系统可能达到的应用范围，而计算站的管理人员一定要满足特定环境的要求。最后，使用者只是对程序感兴趣。

从设计者到使用者，即从第一层到第三层，所需满足要求的范围逐层缩小，左右性能的可能性（即选择如何使用系统的自由度）也更加有限。这三类人员用特定的变量来表示他们对性能的观点；而在有些情况下，他们对相同的变量所给的估价也并不相同。系统设计者经常关心的是系统部件的利用率，如存贮器访问时间、硬件速度、程序库和数据库的组织、存贮分配算法等等变量都会对利用率产生影响。但对其它变量的值，如程序的处理时间，则兴趣较淡。计算站的管理人员最关心系统部件的均衡使用和成本的效益；为大部分使用者（或者对一些最重要的使用者）提供满意的服务；对设备的使用有公平的收费。单个使用者可改变资源（CPU、外部设备、主存贮器和通道）的负荷来影响性能，并按执行时间和费用来评价处理效率；面对其它类型用户有关的重要变量，却并不关心。