第7卷 第2期 华侨大学学报 Vol.7 No.2
1986年5月 JOURNAL OF HUACHIAO UNIVERSITY May 1986

实时性能监督程序设计

王 博 文

〔计算机科学（电脑）系〕

摘 要

本设计的最终目的是提供性能分析人员一个灵活、实时的工具，以及时掌握效率的有关数据，从而能分析其效率问题，提出改进的措施。实时性能监督程序是OS管理下一个独立的程序体。它负责搜集计算机系统的性能数据，並进行统计。性能监督程序与用户的交接界面是一个查询语言。本设计方案是针对非虚存的中小型机提出的。

一、前 言

多用户的计算机系统的资源由每个用户共享。这些资源的利用是否平衡？某资源或某些资源的负载是否饱和从而引起瓶颈现象的产生？或者某些资源是否多余，可以加以撤除以减少开支？这是对现存计算机系统进行分析的主要目的。发现现存计算机系统的性能问题，增加某些设备或调整软件系统，有可能使系统性能大幅度地提高，从而避免或拖延购买新的昂贵的计算机系统，在经济上有较大的意义。

计算机系统性能的标准主要归纳为三个方面：

1.流通量 在单位时间内服务完了的用户数。

2.响应时间 用户在终端打入命令的结束符到系统给出响应信息的第一个字符的时间间隔称为响应时间。

3.可预见性 多个用户的响应时间的差异较小时，则响应时间的可预见性就好；当系统忽略对于某个用户的服务时可能造成可预见性差的结果。

通常对系统性能的评价是相对的。这是因为：

1)性能分析员的出发点不一样。批处理系统可能强调流通量，而分时多用户系统则可能强调响应时间。

2)工作负载的特点的变化引起系统性能的变化。比如以需要CPU为主的工作负载和以

本文1985年9月2日收到。

需与I/O为主的工作负载所测量到的系统性能不一样。

3)系统通常是比较复杂的，很难对某个性能指标定义出最佳值。即使系统简单到可以定义出某个性能的最佳值，当其性能达到某个值以后，要进一步地改善性能付出的代价可能非常高，而未达到该值前的性能改善通常只俚付出相对小的代价就可获得。

二、影响系统性能的主要因素

本设计主要用于监督CPU和I/O设备的性能。也兼顾对内存及外存性能进行监督。

(一) CUP利用率

一个系统的CUP时间可分为忙碌时间T_b和等待时间T_w。CPU处于忙碌态即是CPU正在执行某程序段。对于一个非虚存机器CPU的等待时间可分为：

空闲时间：系统中无任何活动的用户。

I/O等待时间：虽然CPU处于等待状态，但系统中存在等待I/O的活动用户。

$$T_w = T_{id} + T_{io}$$

其中T_{id}为空闲时间，T_{io}为I/O等待时间

在时间间隔T中CUP的利用率(%CUP)定义为CUP忙碌时间T_b与T之比，

即$\%CPU = T_b/T$。

%CPU经常可以作为衡量系统性能的重要标志。但并不是说%CPU越高越好。%CPU高于或低于正常值可能说明系统存在CPU或I/O的瓶颈现象，当%CPU接近100%时而其他系统部件运行正常说明CPU存在瓶颈现象。过低的%CPU可能说明I/O等待时间长即I/O存在瓶颈现象或系统中缺少活动用户。比如交互系统的用户们正花费很多的时间在其终端前思考而没有打入任何命令。

有时需要对个别用户的CPU利用率进行分析、从而发现系统或用户潜在的问题。用户占用CPU的时间与T之比。在T时间范围内，用户除了占用CPU外，可以处于下列状态之一：

1.等待CPU(%RUN)：用户外于CPU的等待队列。如果%RUN大于正常值，说明CPU的瓶颈现象可能存在。

2.等待I/O：用户程序暂停，等待I/O完成。

当某用户处于运行态(占用CPU)，或等待CPU、等待I/O时，都称该用户为活动的。

3.空闲态：用户除了处于活动态外，只能处于另一状态，即空闲态。比如交互系统的用户在终端前的思考时间很长或忘记打入"BYE"命令而离开，使空闲等待时间很长。

(一) I/O系统的性能监督

I/O等待时间长短是衡量I/O系统性能的重要标志。非正常的等待时间可能标志I/O系统中的瓶颈现象。除了用I/O等待时间%IOW来衡量I/O系统的性能外，还用设备的利用率%UT，设备的I/O服务率SEC(每秒钟设备服务的I/O请求数)，对I/O设备的I/O请求数IOREQST来衡量。

设备的利用率与设备的特性有关。对于磁盘来说，%UT = 100*R(sk + rs + dt)/T 其中 T 为时间间隔、R 为时间间隔中得到服务的 I/O 请求数。sk、rs、dt 分别是平均寻找时间，平均旋转时间和数据传送时间。

SEC 可用来衡量系统的流通量。非正常的 SEC 低值说明 I/O 系统的服务速度慢或 CPU 的效率低。

(三)对内存外存的性能监督

设内存空间除去系统软件常驻的区间后其余部分为 m，用户能占内存总量为 u，则定义内存利用率 %MEM = n/m。

设各用户占用内存区的大小为 UMEM。性能监督程序提供各用户占用内存量的大小的信息。供性能分析员参考。

性能监督程序也提供各用户占用磁盘区域大小的信息 DSTO。

三，特点和查询命令

本性能监督程序有如下特点：

1.对性能问题进行联机交互的查询：计算机系统性能分析员可根据提供的查询命令联机查询及分析系统性能问题，包括分析整个系统、各别用户和 I/O 设备等资源的工作情况。查询语言可在屏幕上显示也可在打印机上输出以供事后分析。

2.异态的自动检测：可对某些性能参数设置门槛值。当某性能参数(如 CPU 利用率 %CPU) 达到设置的门槛值时，系统自动生成报告。这个特点使分析员及时发觉系统某部件的瓶颈现象，并可能进一步采取对策。门槛值的设置是通过 "SET LOGMSG" 命令实现的。可通过其它命令获取异态信息和询问所设置的门槛值。可动态对门槛值进行修改。

3.可对性能数据的收集，统计时间间隔（也是自动显示，打印时间间隔）的大小进行控制。比如可选择 5 分钟或 60 分钟的时间间隔对 %CPU 及其它性能参数进行收集、统计。

4.提供计算机系统和各个用户的运行情况。分析人员可人工或自动地在屏幕上或打印机上得到系统和各用户运行情况，如 CPU、内存、外存、I/O 设备等资源的利用情况。

5.提供各个 I/O 设备的使用情况。用设备利用率 %UT，设备 I/O 请求数 IOREQST，设备的 I/O 服务率 SEC，设备上处于等待队列的用户数 MQ 来反映 I/O 设备使用情况，这些数据对调整 I/O 系统及检测 I/O 系统的问题是有用的。分析人员可人工或自动地在屏幕上或打印机上得到 I/O 设备的使用情况信息。

查询命令。

1.GENERAL

命令格式为：

GENERAL
G

本命令不带操作分量，G 为缩写型命令形式。

本命令在屏幕上显示最近结束的"总的运行情况"。显示的最后一行以 "<—" 开头，该行结出系统在一个长的时间区间中的性能平均值。这个对间区间最大24小时。区间的起点可以是：a)打入 "RESET" 命令或

“PRINT ALL”命令的时间；b)上次“自动打印时间间隔”的结束点，c)性能监督程序开始运行的起始时间。区间的结束点是打入命令时最近结束的“显示时间间隔”的结束点。称这个时间区间为“当前时间区间”。自动打印时间间隔和显示时间间隔由INT命令设置。

显示的最后第二行是最近结束的显示时间间计算机系统的性能数据。这一行以“->”开头。

当屏幕容纳不下“总的运行情况”时，可用NEXT命令让接下去的数据在屏幕上出现。总的运行情况可用“PRINT G”命令打印出来供事后研究。

总的运行情况的每行是某用户的性能数据。每一行的先后次序可用“ORDER”命令规定。先后次序的内定值是%CPU，即花费CPU最多的用户排在最上面。

输出格式如下：

```
   PFRM GENERAL Start:  end:
   %CPU %ACT %TOW %IELE %RUN CPU UMEM DSTO
用户 1 ···
用户 2 ···
   ⋮
PERM ···
   ⋮
用户 n ···
->
<-
```

%CPU表示CPU利用率；%ACT表示活动态时间的百分比；%IOW表示等待I/O动作完成的I/O等待时间比；%IDLE表示处于空闲态的时间比；%RUN表示等待CPU的时间绝对值；UMEM表示能占的内存空间(以KB为单位)；DSTO表示能占用的磁盘区域(以KB为单位)。

Start表示显示时间间隔开始时间。

end表示显示时间间隔结束时间。

PERM代表性能监督程序。它也是计算机系统的一个用户，故在显示表格中占一行。

USERIDi是多用户帐号。即有登记了的用户(LOGGEDON)都有对应的一行性能数据。

2.DEVICE

当命令带INT操作分量时，它被用来在屏幕上显示最近结束的显示时间间隔计算机系统的I/O设备的性能数据。当命令不带操作分量时，用来显示当前时间区间的显示时间间隔为时间单位的I/O设备性能数据。

命令格式：

DEVICE	ITN
DE	

（INT为INTERVAL的缩写）

输出格式：

PERM DEVICE Start end

%UT IOREQST SEC MQ TYPE

设备地址 1

设备地址 2

⋮

设备地址 n

用“DEVICE”命令产生的显示由多幅上述显示组成，只是每一幅显示对应的显示时间间隔不一样。而且，在显示整个包含在“当前时间区间”的显示时间间隔的I/O设备性能数据以后，有独立的最后一行以<-”为开头的数据，给出“当时间区间”的I/O设备性能数据的平均值。

%UT是设备利用率、IOREQST是对应显示时间间隔中对该设备发出的I/O请求数。SEC是每秒钟该设备服务的请求数。MQ是排队等待该设备能务的请求队列长度。TYPE是设备类型。

用设备地址对系统中多I/O设备进行区分。令及“PRINT DEVICE INT”命令打印出来。

可用“INTERVAL PRINT nnnn nnnn nnnn”命令指定自动打印时间间隔。自动打印的内容包括I/O设备性能数据。

3.HOLD

命令格式为：

HOLD	
HO	

HOLD命令用于让出现在屏幕上的显示画面固定值，以便有足够时间观察。在画面固定期，虽然性能监督程序仍然继续收集，统计性能数据，但自动显示功能受到控制。

4.FREE

FREE	
FR	

恢复自动显示功能。

5.ORDER

命令格式：

ORDER	%CPU
O	%IDLE
	%IOW
	%RUN
	UWEW
	DSTO

ORDER 命令使输出到屏幕或打印机上的各用户性能数据按照某指定的标准按次序排列。这个标准可以是%CUP，%IDLE，%IOW、%RUN、UMEM、DSTO。比如若指定%IOW，则出现在“总的运行情况”显示里最上面的用户是等待 I/0 最久的用户。先后次序的内定值是%CPU。

6.INTERVAL

命令格式

INTERVAL	GD nn
INT	PRINT nnnn…nnnn

本命令用于指定自动显示的“显示时间间隔”和“自动打印时间间隔”。在显示时间间隔结束时，性能监督程序自动显示“总的运行情况”；在“自动打印时间间隔”结束时，性能监督程序自动打印该时间间各种性能数据，并且重置(见 RESET 命令)性能监督程序，开始新的自动打印时间间隔，重新收集，统计性能数据。

nn 是指定的时间间隔，它的范围是 1 至 60(分钟)。nn 也是 I/O 设备性能数据收集、统计时间间隔。

nnnn 表示自动时间间隔的端点。

INT 是命令缩写形式。

例如：命令“INT GD 10”使每隔 10 分钟在屏幕上自动产生“总的运行情况”显示。“INTPRINT 0830 1715 0000”命令将在 8：30、17：15；00：00 三次产生时间区间为〔00：00,08：30〕、〔08：30,17：15〕、〔17：15,00：00〕期间的性能数据。

INTERVAL 命令行长度不超过显示长度。

抑制或不抑制自动打印输出可用命令“SET PRINT OFF”和“SET PRINT ON”实现。

6.QUIT

本命令用于停止性能监督程序的运行，将控制回交操作系统。性能监督程序的启动用命令“PERM”实现。

命令格式：

QUIT	

8.RECOMPUTE

命令格式：

RECOMP	
REC	

本命令用于强行结束当前显示时间间隔。当显示时间间隔较长，而性能分析员又想知道当前系统运行情况，可用这个命令。新的显示时间间隔从打入“RECOMP”命令后开始。REC 是命令缩写形式。

9.RESET

本命令对性能监督程序所收集和统计的性能数据，包括所有的计数器(如 IOREQST)和非正常情况的登记表都置为初始态，并开始新的显示时间间隔和新的自动打印时间间隔。

命令格式：

RESET	
RES	

10.SET

SET 命令用于控制性能监督程序的运行及定义门槛值

SET	PRINT ON/OFF MSG ON/OFF LOGMSG n ON/OFF LIMIT nn

“SET PRINT ON/OFF”用于不抑制或抑制自动打印输出。

“SET MSG ON/OFF”命令用于决定是否进行门槛值的监督。比如“SET MSG OFF”命令抑制对门槛值的监督。即性能监督程序并不对性能参数是否达到其门槛值进行检验。

“SET LOG n ON/OFF LIMIT nn”用于定义各门槛值(LIMIT右边的nn表示门槛值)及决定是否对各别的性能参数进行门槛值监督(当命令中出现的是ON时，对对应的性能参数监督其是否对各别达到门槛值；当出现的是OFF时，不对该性能参数是否达到门槛值进行判断)。

命令中的LOGMSG右边的n是0至7的整数。意义分别为：

◎任何I/O设备每秒服务的I/O请求最大数。

①用户使用的磁盘区域的KB数。

②用户的空前绝对时间。

③用户使用内存的KB数。

④计算机系统的CPU利用率%CPU。

⑤计算机系统的I/O时间等待比%IOW。

⑥计算机系统的内存利用率。

⑦在一秒钟内各别用户的I/O请求得到服务的数目。

例：“SET LOGMSG 2 ON LIMIT 120”命令使能能监督程序将空闲时间超过120分钟的用户登记下来，以供性能分析人员采取措施。

11. LOG

本命令产生最近结束的显示时间间隔异态情况显示。异态指性能参数达到其门槛值。显示内容包括与异态最有关的用户的标识码(USERID)

命令格式：

LOG	

输出举例：

```
   PERM  LOG  -->  10 : 16 : 00
CPU  Utilization  100%  USER 10  39%  USER 01  7%
```

显示说明结束时间为10点16分00秒的显示时间间隔中，计算机系统的CPU利用率达到满载，其中两个用CPU最多的用户是USER 10和USER 01，分别占用CPU的时间为39%和7%。

12. ALOG

命令格式：

ALOG

本命令提供“当前时间区间”中异常情况的屏幕显示。以提供分析。显示的内容与LOG显示一样，但所涉及的时间间隔不一样。当一个屏幕显示容不下所有信息时，可用“NEST”命令显示接下去的信息。

13. SLOG

命令格式：

SLOG	

本命令提供“当前时间区间”中以显示时间间隔为时间单位的计算机算统运行情况信息。其每行显示格式与用 G 命令显示的最后第二行的格式一样。当一个屏幕容纳不下所有信息时，可用“NEXT”命令显示接下去的信息。

输出示例：（假定显示时间间隔 10 分钟）

```
        PERM  SYSTEM  LOG  start  160 end 1710
        %CPU  %ACT  %IOW  ···  DSTO

1610
1620
 ⋮
1710
<—
```

以“<-”记号开头的最后一行是“当前时间区间”中系统性能数据的平均值。

14. ULOG

命令格式：

ULOG	

本命令提供“当前时间区间”中以显示时间间隔为时间单位的“最突出用户”的性能数据。判定“最突出用户”的标准是由 ORDE 命令指定的，比如消耗 CPU 最多的户。

输出示例：（假定显示时间间隔为 10 分钟，ORDER 命令指定的参数是 %CPU）

```
        PERM  USER  LOG  start  ···  ···  end
        USERID  %CPU  %ACT  ···  DSTO
1610    USER 02    10    ···
1620    USER 02    12    ···
 ⋮
1710    ···
```

上面显示说明在最初的两时间间隔中，用户 USER 02 占用 CPU 显间最长，分别达到 10% 及 12%。

15. NEXT

命令格式：

NEXT N	

当所显示的信息多到一个屏幕容纳不下时，可用本命令让接下去的信息出现在屏幕上。

16. PRINT

命令格式：

PRINT	G
PR	LOG
	ALOG
	SLOG
	ULOG
	DEVICE
	DEVICE INT
	ALL

产生打印输出，以供事后分析是必要的，自动打印时间间隔由 INTERVAL 命令指定。自动动打印输出的内容包括用 G,LOG,ALOG,SLOG,ULOG,DEVICE 命令显示在屏幕上的内容。

临时打印由 PRINT 命令实现。它的操作分量可以是 G,LOG,ALOG,SLOG,ULOG,DEVICE,INT 或 ALL。打印结果分别是对应命令（LOG,ALOG,…,DEVICE,INT）在屏幕上显示的内容。如果操作分量是 ALL，则打印由 LOG,ALOG,SLOG,ULDG 和 DEVICE 命令显示的全部内容。

PRINT ALL 命令执行以后，性能监督程序开始一个新的自动打印时间间隔。

17. QUERY

本命令用来查询用 SET 命令所设置门槛值的大小及 ON/OFF 状态。也用来查询显示时间隔及自动打印时间间隔的大小。还用来查询用 SET 命令所置的 PRINT、MSG 的 ON/OFF 状态。

查询结果在屏幕上给出。

命令格式：

QUERY	LOGMSG
Q	INTERVAL
	NSG

输出示例 1

```
PERM QUERY LOGMSG
LOGMSG STATUS LIMIT
  0      ON     50
  1      ON    200
  2      ON    120
  3      ON    100
  4      NO     98
  5      ON     40
  6      ON     95
  7      ON     50
```

输出示例 2

```
PERM QUERY INTERVAL
GD     1      MINUTES
PRINT ON 0800 1600 0000
```

四，后　　言

性能分析员通过查询命令得到的数据只供性能分析时使用。是否存在值得改善的性能问题，通过什么方法去改善，则是一个困难复杂的问题。目前并不存在一个现成的某一张对照表，可根据所得到的数据能从该对照表查出性能问题的所在。

能否成功地寻找出性能门题并加以改善，依赖于对系统的理解，也依赖经验和直觉。比

如磁盘出瓶颈现象的原因可能是下列因系之一或其复合，也可能是其他因系：

1.缓冲区太小导至内外存交换次数的增加。

2.磁盘调度策略不合适。

3.本磁盘负载太大。

实例：

资源	CPU	盘1	盘2	盘3
利用率	0.62	0.26	0.14	0.90

对得到的利用进行分析。盘3的利用率远大于其他部件利用率，且接近1。估计盘3存在瓶颈现象。了解系统的构成之后，知道盘3是用于缓冲往返于慢速设备和程序之间的信息，负载大。调整的方现可以是将此缓冲功能交给盘2。调整后再测量其利用率。

CPU	盘1	盘2	盘3
0.74	0.29	0.27	0.68

调整后的系统性能得到提高，因而验证瓶颈现象存在于盘3的假设及使瓶颈现象消失的方法是合理的。

为了提高系统的性能，可增加硬件部件，多增加磁盘的容量或个数。也可以对软件进行调整。增加硬件通常容易延误时间且比较昂贵。同样，软件调整也要花费人力物力和金钱。软件调整可以是：

1.操作系统参数：多道程序的道数，缓冲区的大小，优先数的定义等等。

2.资源管理算法：作业、CPU，盘的调度算法；存取技术(ACCESS METHOD)；内存管理算法(分配，置换等)；外部设备管理算法等。

3.内部连接：通道——设备连接；设备——总线连接。

4.信息存放：文件在存贮器层次的存放；操作系统常驻内存模块存放的改变。

5.程序性能：对OS的开销及诸如编译程序、编辑程序，数据库管理系统等常用程序所需要的执行时间和空间进行调整。

6.收费标准：对轻负载时的用户如晚上的用户优惠价；对低效率的程序及重负载时的用户加价。

7.作业接受标准：拒绝某种作业或用户。

无论是硬件的改变或软件的调整，事先都应慎重地考虑。

采用软件性能监督的方法的好处是：它适用于任何程度的复杂性。即使系统的构造是很复杂的，总可以建立相应的性能监督程序。由于是软件，故其灵活性是显然的。设计者可根据不同的系统，设计出不同的性能监督软件。如果用硬件方法实现对系统性能的监督，则很少有修改的余地。

但是系统性能监督程序需要CPU时间去执行及需要空间去存放其程序体和存放所收集到的数据，所以在一定程定上影响了系统性能分析的准确性。为了减轻对分析准确性的影响，可考虑：

1.设计性能监督程序时，单独统计出该软件所需的CPU使用时间和空间。

2.如果时间比空间重要，则设计者可将观察到的数据存放起来，等到所定义的时间间隔终了时再进行统计。反之，如果空间更重要，可将每次观察到的数据马上进行统计。

3.系统性能分析员可加大显示时间间隔或自动打印时间间隔，以减少性能监督程序本身的开销。

参 考 文 献

[1] Domenico Ferraari, Computer Performance Evaluation, Prentice-Hall, INC, (1978).

[2] James. A. Morris, Computer performance Management-A Structured Appraach, North-Holland Publishing Company, (1978).

[3] Boris Beizer, Micro Analysis of Computer System Performance, Van Nostrand Reinhold, (1980).

[4] H. M. Deitel, A Introdu ction to Operating Sgstem, Addison-Wesley Publishing Company, (1983).

A Design of Real-Time Performance Monitor

Wang Bowen

Abstract

The article presents a design of performance monitor for mini or high-end micro computer system. The aim is to give a software approach which is in charge of eollecting and analysing information regarding system performanee for existing system in a flexible and real-time manner.