# OPTIMAL PROFIT SCHEDULING OF A COMPUTING CENTER

Walter L. Whipple
The University of Michigan, Ann Arbor, Michigan

## INTRODUCTION

The current means of pricing computer services is an ad hoc arrangement by which the owner attempts to recover costs and achieve a profit. Virtually no attempt has been made to solve the concomitant peak loading problem. Where some pricing method of regulating demand exists, it is generally too weak to be effective. The University of Michigan Computing Center offers a 20% discount after midnight and on weekends, but this does not really motivate users who are not really paying for service anyway. Control Data Corporation has long offered a range of priorities for batch jobs but not for time-sharing. The batch priority prices cover a 3 to 1 range, but when various restrictions are taken into account on the high and low priorities, this drops to 1.4 to 1 in three steps. This amounts to a decision as to whether one wants results today or tomorrow. The result is a peak loading problem which can be quite severe. These two examples are typical of the pricing methods for computer services: a priori pricing by time of day or a posteori service by bid price. In either case the price ranges are narrow and peak loading problems exist.

This paper models the economic demand curve as a straight line with the equation

$$P = P_0 - Q \cdot E$$

where $P_0$ is a price at which it is assumed no demand will occur and E is an elasticity. The electrical engineer will recognize this as a voltage source with series resistance. The assumption of linear demand is necessary to ease the analytic solutions. A nonlinear curve such as

$$P = P_0 - Q^{\frac{1}{2}}E \quad \text{or} \quad P = P_0 - Q^2E$$

could be used with the resulting solution difficulties. However, a real problem of determining the correct functional form will arise. An obvious extension of this paper would be to evaluate the changes when nonlinear demands are present. Other analogous electrical elements are used to construct a circuit model of the general demand/supply interface. Even a simple model will require unidirectional elements to represent constraints on the solution. The resultant network will be solved in a piecewise linear fashion for several possible pricing methods. The results will then be compared to yield estimates of rates of returns on sales. The case of non-profit centers and the point of view of consumers is briefly discussed before presentation of the optimal policies.

This paper supports the concept of an absolute auction for computer services based on maximizing revenue based on consumer bids subject to maximum processing rate and a changing job mix. The policy is an adaptive one (a closed loop control) in which prices drop when demand drops and increase when it rises. By permitting prices to vary over a wide range in this adaptive manner, prices are always optimal. A linear programming problem must be solved repetitively in order to determine specific job mixes. However, most operating system schedulers already devote considerable time to selecting the job mix and performing swapping. All that has changed is the objective function. The indicated revenue increase computed in this paper is almost 100% over the constant price schemes which are in current use. The interested reader is referred to the author's thesis for a detailed treatment of this subject(1).

## CIRCUIT ANALOGY

In economics, price and quantity are the basic variables. Price has the characteristics of a potential although the underlying forces and fields may be difficult to identify. Nevertheless, it is clear that an electrical voltage is the analog of price. Similarly, a quantity of goods or services has all the essential characteristics of electrical charge. As is usual in electrical circuits, this paper will consider an economic current flow to be the negative time rate of change of quantity. As such, economic current represents orders for goods and services. This reversal is necessary for consistency with the polarity of price. With these variables in mind, it makes sense to speak of a price source or demand source in a manner analogous to voltage and current sources in electrical networks. These sources may in general be time varying and dependent. The price source demands that quantity of a product or service which will maintain a specified price. The demand source pays that price which will maintain a specified demand.

An elasticity in economics represents the slope of a price-quantity curve in just the same way that an electrical resistance represents the slope of a voltage-current curve. However, in the economic case these curves are generally represented as a curved rather than a straight line due to the effects of diminishing returns and returns to scale. In this paper each elasticity will be assumed constant to preserve piecewise linearity in the economic circuits. At the expense of computational complexity, these linear elasticities could be modified to be non-linear if appropriate values could be found. In this study, the constant elasticity assumption seems justified due to the widespread choices available to the consumer. Other elements of linear circuit analysis which are relevant to a study of pricing are the capacitor and inductor. A capacitor has its counterpart in the warehouse for products or queue of demand. In order to be linear, the price of a warehouse must vary proportionally to the negative of its contents. An inductor has its counterpart in the tendency of demand to continue flowing after the need has passed. Its effect is that of sales resistance.

Many constraints exist in an economic system. For this reason alone, change is traumatic. For instance, consumers cannot usually act as producers and vice-versa. Therefore, if prices rise, consumers suffer; when prices fall, producers suffer. This is true even if prices make it favorable for a consumer to produce. Exceptions to this rule are found in pumped storage hydroelectric stations and in the various middleman situations. In an electrical circuit this kind of constraint is represented by a diode. When applied to economics, a diode is an element which has infinite elasticity when current flows in the forward direction and zero elasticity in the reverse direction. The diodes used in this study are all perfect switches in either an on or off state.

Sources, elasticities, and diodes will be used in the next section to aid in the analysis of the supply/demand situation for computing.

## COMPUTER SIMULATION METHODS

The problems outlined in the previous section may be analyzed by digital computer using one of a wide variety of techniques.

1. Programming languages such as FORTRAN, COBOL, PL/1, ALGOL, BASIC and the like. These will require the user to do his own integration of the profit function and in the case of energy storing elements will require the integration of differential equations as well (2). All equations must be written by the user (for large networks this can be a formidable task).

2. Simulation languages such as MIMIC, DYNAMO (3,4, 5), CSMP (6), and other differential equation solvers. These relieve the user of the problems inherent in integration and solution of differential equations. However, the network equations must still be written.

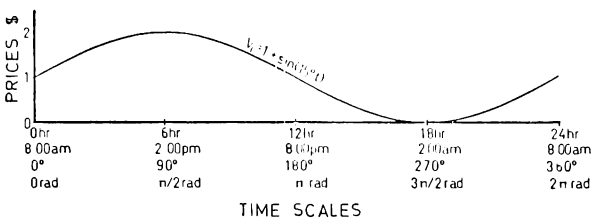


Figure 1. User price source

3. Nonlinear network programs such as ECAP (7) (piecewise linear), SCEPTRE (8), SYSCAP (9), and TESS (10). These programs solve a wide variety of nonlinear electrical circuit problems and permit solution of large networks without requiring the analyst to generate all the simultaneous differential equations. Their ability to handle exceedingly nonlinear problems and their extensive numerical techniques particularly suit them to this problem.

Three examples of pricing problems are shown in this chapter solved with the TESS program. Only a few samples of the output are shown for the sake of brevity.

More complicated problems involving energy storage elements may also be solved using the TESS program, but they were not solved for this study due to limitations of time and funding. Inclusion of energy storage elements in this problem results in time constant problems and difficulties with nonlinearities which make a program such as ECAP or SYSCAP more suitable than TESS. If extensive simulations of this type are to be performed, it may be necessary to write a program which will better handle the difficulties to be encountered.

Average Market Price Model

Suppose that the user community can be represented by a time-varying price with constant but unidirectional elasticity. Further suppose that the computer services are priced at the average of the time-varying price and that any unused services or unsupplied demands are lost forever. The processing rate is limited to a given maximum. Let another elasticity be defined equal to the difference between the price divided by the quantity actually processed. Figure 1 is a typical price function which corresponds to the system loads of many computing services. Figure 2 is the circuit diagram of this average pricing scheme. The following values are assumed:

$V_0 = 1.$

$V_1 = 1. + \sin(15°t)$

$I_0 = 1$

$R_1 = .1$

$R_V$ = selected to maximize $J = \int i V_0 dt$ subject to the constraint $i \leq I_0$.

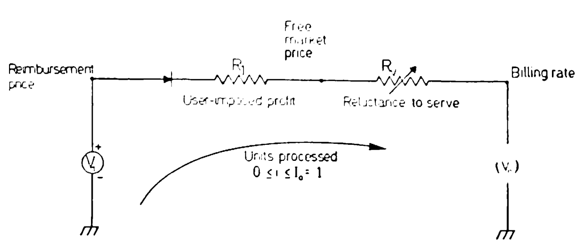


Figure 2. Economic circuit

The program input data are shown in Figure 3. A FORTRAN function subroutine used to define the price is shown in Figure 4. While a complete description of the input data is not within the scope of this paper, some features will help relate the listing to the problem. The network is described under the heading "ELEMENTS." Each element begins with a letter which defines the type of equation to be used:

E - potential or price
J - current or production
R - resistance or profit margin
L - inductance or sales resistance
C - capacitance or storage.

The characters up to the first comma constitute the unique name of the element. Following are the names of the network nodes between which the element is connected. After the equal sign is the value of the element. Values beginning with Q(equations) and T(tables) are defined under "FUNCTIONS" while "FUNCTION" refers to a user supplied FORTRAN function (in this case, OPT1).

The "FUNCTIONS" heading introduces specifications of functions which will be implemented as FORTRAN arithmetic statement functions. Constants, auxiliary variables, and auxiliary differential equations appear under the heading "DEFINED PARAMETERS." Each parameter begins with the letter P and may have a variable value or constant. Derivatives begin with DP. Therefore the state-

```
TESS-TR SYSTEM INPUT    5/1/73      VERSION 2.1
::
CIRCUIT DESCRIPTION
CONSTANT PRICE EXAMPLE
ELEMENTS
E1,G-1=Q1(TIME)
R1,1-2=.1
RV1,2-3=FUNCTION OPT1(E1,EØ,R1,PIØ)
EØ,G-3=1
FUNCTIONS
Q1(T)=(1+SIN(PI*T/12))
Q2(T)=(AMOD(1ØØ*T+8ØØ,,24ØØ.))
Q3(X,Y)=(X*Y)
DEFINED PARAMETERS
PIØ=1
PTIME=Q2(TIME)
PI=3.14156
PROFIT=Ø
DPROFIT=Q3(IEØ,EØ)
OUTPUTS
E1,RV1,EØ,PROFIT,DPROFIT,PLOT(PTIME)
PTIME,PLOT
RUN CONTROLS
STOP TIME=24
MAXIMUM PRINT POINTS=1Ø
RERUN DESCRIPTION(S)
ELEMENTS
EØ=.5,.75,.9,.95,1.Ø5,1.1Ø,1.25,1.5
END
```

Figure 3. Constant Price Example Input Data

```
RUN VERSION 2.3 --PSR LEVEL 363--

                FUNCTION OPT1(E1,EØ,R1,PIØ(
          C...  SUBROUTINE TO CALCULATE RV1 FOR CONSTANT PRICE
ØØØØØ7          IF(E1.LT.EØ)GO TO 1Ø
ØØØØ11          IF(E1.LT.(EØ+PIJ*R1))GO TO 2Ø
          C...  FULL LOAD CASE
ØØØØ14          RV1=(E1-EØ)/PIØ-R1
ØØØØ17          GO TO 1ØØ
          C...  REVERSED LOAD
ØØØØ17      1Ø  CONTINUE
ØØØØ17          RV1=1.E2ØØ
ØØØØ21          GO TO 1ØØ
          C...  PARTIAL LOAD
ØØØØ21      2Ø  CONTINUE
ØØØØ21          RV1=Ø
ØØØØ22     1ØØ  CONTINUE
          C...  ADJUST FOR ZERO RESISTANCE
ØØØØ22          OPT1=AMAX1(RV1,1.E-2ØØ)
ØØØØ26          RETURN
ØØØØ27          END
```

Figure 4. FORTRAN Function Subroutine for Computing RV1

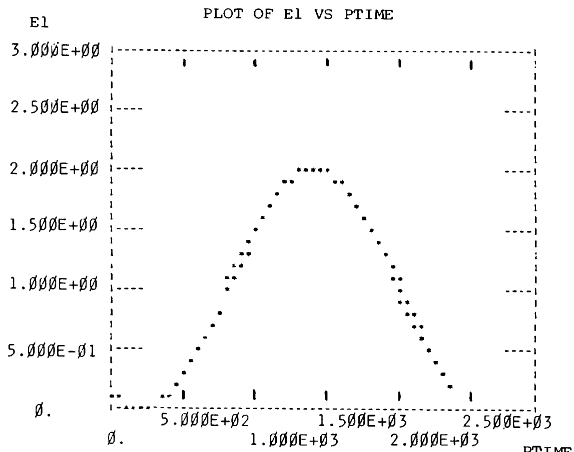

Figure 5 - User Demand, E1.

ments PROFIT = 0 and DPROFIT = Q3(IEO,EO) result in integration of profit with a zero initial value. All outputs are plotted against the parameter PTIME which is the military representation of time of day (0000 - 2400 hours). Various solution parameters may be introduced under "RUN CONTROLS." The Rerun Description section specifies eight repeat analyses at different prices.

The demand curve for this example is shown in Figure 5. This function is not far from the usual situation where there are many users of a computer in a commercial environment. The program permits specifying such curves by explicit equations or tables with ease.

For the case with E0 = 1, the cumulative profit is shown in Figure 6. Note that the final time is 0800 in this case and through a sign reversal the profit is shown negative. Final profit was calculated as 11.592 at the final time. This agrees well with analytic results (1).

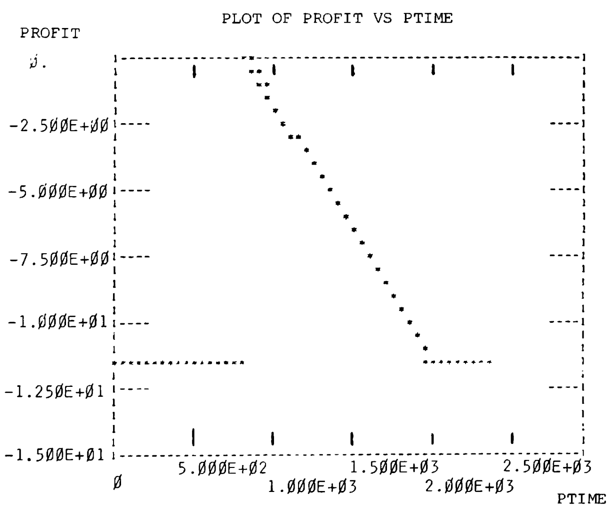

Figure 6 - Plot of Profit

The values of $R_V$ are shown in Figure 7. Due to the scaling of the plot, only two values appear, zero and infinity (represented by $1.0 \times 10^{200}$). Actually where $R_V$ appears to be zero, it takes on values from zero (represented by $1.0 \times 10^{-200}$) to one. The boundary of values for $R_V$ of $1.0 \times 10^{-200}$ and $1.0 \times 10^{200}$ is necessary in TESS for numerical reasons.*

Figure 8 shows the marginal profit throughout the day. Since most of the time the system is either saturated or turned off, the marginal profit is non-zero during a restricted period of time equal to twelve hours (8:00 a.m.

*The presence of these extreme values of $R_V$ is a serious impediment to analysis of networks with energy storage in TESS since they give rise to widespread time constants.

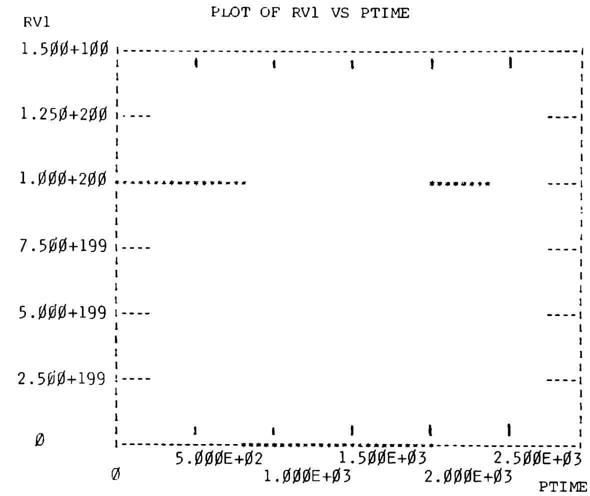

Figure 7 - Plot of RV1.

to 8:00 p.m.). Note that the marginal profit is limited. This is in complete agreement with the analytic results (1).

The total profit (total $iV_0$ less costs) for this scheduler is J = -C + 11.592. The value of $R_V$ is zero until $V_1 = I_0R_1 + V_0$ and then increases linearly with $(V_1 - V_0)/R_1$ to maintain the constant $i = I_0$. The power dissipated in $R_V$ may be interpreted as an economic loss, perhaps due to the congestion at the counter or uncertainty of service.

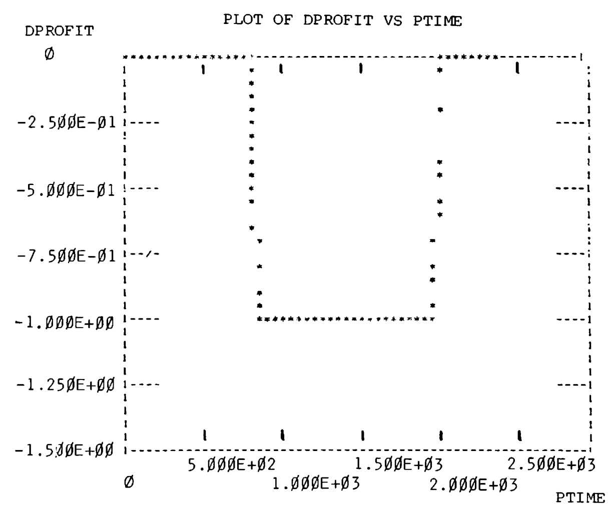

Figure 8 - Plot of Marginal Profit.

Optimal Constant Price Model

The same input data used before were used with reruns for price values of .5, .75, .9, .95, 1.05, 1.1, 1.25, and 1.5. The final profit from each rerun is plotted as a function of price level in Figure 9. Note that the profit reaches a peak at $V_0 \doteq 1.25$ and then declines. Again this result agrees with analytical results (1). The peak is relatively broad and indicates that as long as a constant price is being considered, it does not make a great deal of difference what constant price is chosen if it is somewhere near the optimum. Attempts to modify the price to find the peak would in practice be difficult since the results would be masked by random variations in the market situation.

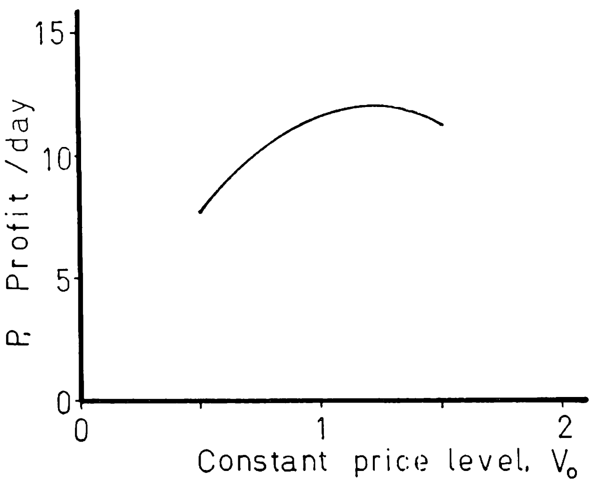


Figure 9. Profit Versus Constant Price Level.

## Demand Price Model

Now suppose that the problem is modified by removing the constraint on the price except that it be positive. The circuit is shown in Figure 2, and the same values apply. Again, $R_V$ is selected to maximize $J = \int iV_o dt$ subject to the constraint $i \leq I_o$. Analysis shows that the value of $R_V$ is zero and the total profit is J = -C + 22.08.

The demand price model is somewhat more complicated because two elements must be varied, $R_V$ and $V_o$. Therefore, two entries to FUNCTION OPT are provided. The profit function was also changed to display positive values. Otherwise the input data in Figure 10 is the same as that

```
TESS-TR SYSTEM INPUT    5/1/73     VERSION 2.1

*
CIRCUIT DESCRIPTION
DEMAND PRICE EXAMPLE
ELEMENTS
E1,G-1=Q1(TIME)
R1,1-2=.1
RV1,2-3=FUNCTION OPT1(E1,R1,PIØ)
EØ,G-3=FUNCTION OPT2(E1,R1,PIØ)
FUNCTIONS
Q1(T)=(1+SIN(PI*T/12))
Q2(T)=(AMOD(1ØØ.*T+8ØØ.,24ØØ.))
Q3(X,Y)=(-X*Y)
DEFINED PARAMETERS
PIØ=1
PTIME=Q2(TIME)
PI=3.14156
PROFIT=Ø
DPROFIT=Q3(IEØ,EØ)
OUTPUTS
E1,RV1,EØ,PROFIT,DPROFIT,PLOT(PTIME)
PTIME,PLOT
RUN CONTROLS
STOP TIME=24
MAXIMUM PRINT POINTS=1Ø
END
```

Figure 10. Demand Price Input Data.

in Figure 3. Similarly, FUNCTION OPT IS shown in Figure 11. The same computations are performed for both entry points and the appropriate value returned. The output plots versus PTIME are shown in Figures 12-15.

The user demand is the same as before. Since user demand is positive semi-definite and overdemand cannot occur with demand pricing, $R_V$ is always zero. The price level is shown in Figure 12. As long as a price equal to half the demand results in partial load, the price appears to be

```
               FUNCTION OPT(E1,R1,PIØ)
          C...SUBROUTINE TO CALCULATE RV1 AND EØ FOR DEMAND PRICING
ØØØØØ6         ENTRY OPT1
          C... ENTRY FOR RV1
ØØØØ16         IOPT=1
ØØØØ17         TO GO 1
ØØØØ2Ø         ENTRY OPT2
          C... ENTRY FOR EØ
ØØØØ3Ø         IOPT=2
ØØØØ31       1 CONTINUE
ØØØØ31         IF(E1.LT.Ø)GO TO 1Ø
ØØØØ33         IF(E1.LT.2*PIØ*R1)GO TO 2Ø
          C... FULL LOAD CASE
ØØØØ37         EØ=E1-PIØ*R1
ØØØØ41         RV1=Ø
ØØØØ42         GO TO 1ØØ
ØØØØ42      1Ø CONTINUE
          C... REVERSED DEMAND CASE
ØØØØ42         RV1=1.R2ØØ
ØØØØ43         EØ=Ø
ØØØØ45      2Ø CONTINUE
          C... PARTIAL LOAD CASE
ØØØØ45         EØ=E1=.5
ØØØØ46         RV1=Ø
ØØØØ5Ø     1ØØ CONTINUE
          C... ADJUST ZERO RV1
ØØØØ5Ø         RV1=AMAX1(1.E-2ØØ,RV1)
          C... RETURN FUNCTION ACCORDING TO ENTRY
ØØØØ54         IF(IOPT.EQ.1)OPT=RV1
ØØØØ57         IF(IOPT.EQ.2)OPT=EJ
ØØØØ62         RETURN
ØØØØ64         END
```

Figure 11. Demand Price Fortran Function.

zero. Actually it is zero only at PTIME = 0200. Similarly, the marginal profit shown in Figure 13 appears to be zero during the period 0000-0400, but it is actually greater than zero. Since the system is saturated, during the time period 0400-2400, the marginal profit is equal to $V_1$ less the constant potential across $R_V$. Therefore marginal profit is also approximately sinusoidal. The prices are shown in Figure 14. Cumulative profit is shown in Figure 15 and reaches a value of 21.9363 at the final time. This agrees well with analytic results (1).

It is interesting to note in the latter problem that if $I_o$ is permitted to vary, then a value of $I_o$ of 10 is required which produces J = 360 - C. If the cost of providing additional service is constant, it probably will be supplied in this case. However the selection of optimum $I_o$ is very sensitive to the exact form of $V_1$ and will not be discussed further in this paper. Demand pricing results in a ratio of price to production equal to the elasticity $R_1$. Now suppose one fixed the demand pricing by using an elasticity. Then the optimal value of $R_1$ from the point of view of maximizing $i^2R_1$ (the user's profit) is also equal to the fixed elasticity. Therefore, demand pricing satisfies both buyer and seller.

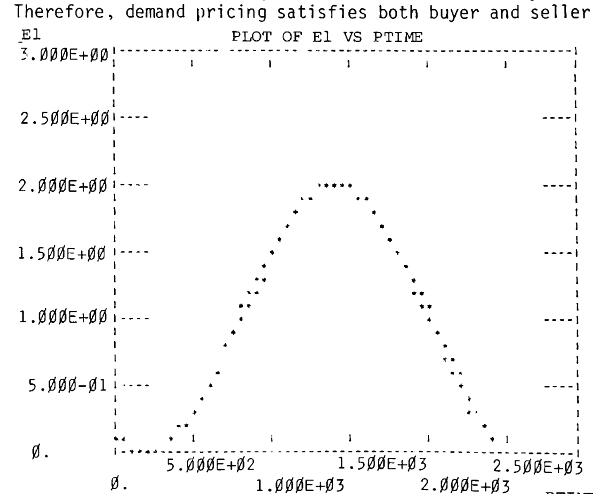


Figure 12. Demand for Processing, $V_1$.

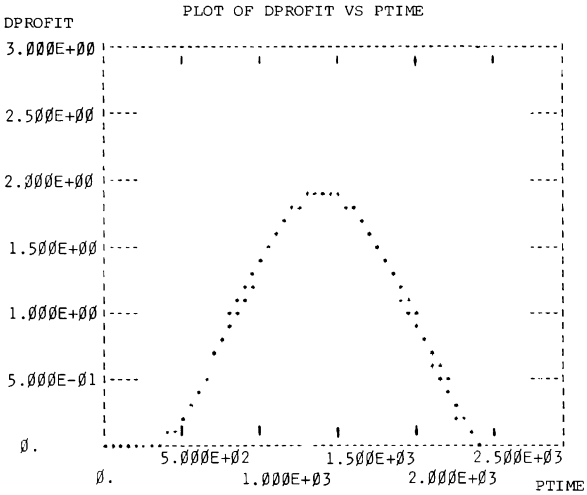


Figure 13. Marginal Profit.

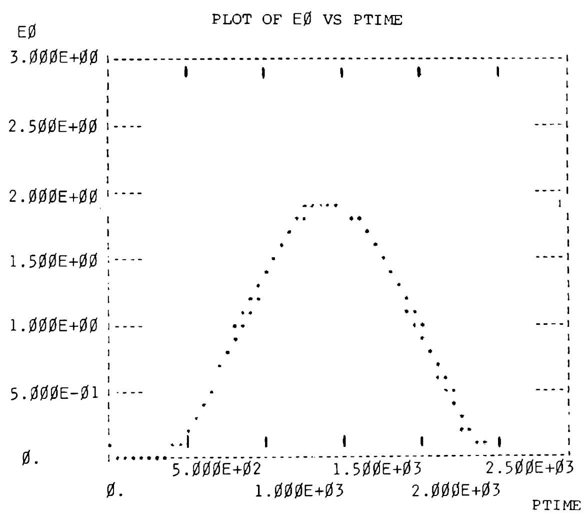


Figure 14. Price Level, $V_0$

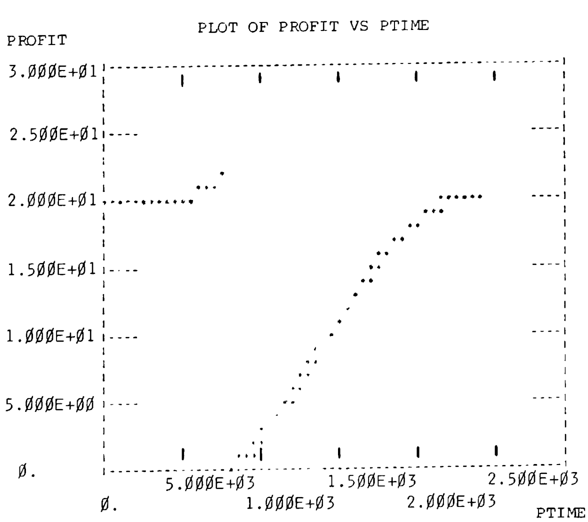


Figure 15. Cumulative Profit.

The demand price is exactly $V_1/2$, and the buyers and sellers exactly share the profit. But there tends to be a greater number of consumers than producers so that the profit to a single consumer is less than to a single producer. However, this argument is only applicable when the resulting demand is less than $I_0$. When processing becomes scarce, the optimal profit cannot be achieved by either producer or consumer.

## APPLICATION

Since most current computer centers are operating with nearly a constant price, their profit is computed as 11.59 - C. Assume that the net profit of such a firm is zero. Then the costs must be 11.59. Now with demand pricing, costs are not appreciably increased (in fact, with a steady load, costs may even decrease and overall demand will be less due to the higher prices under demand pricing) and yet the profit increases to 22.08 - C = 10.49. This is a profit of 47.5% of revenues and 99% of costs. If one assumes an annual cost equal to 25% of machine investment (representing lease cost of equipment, space, personnel, etc.), then this would be a profit of almost 25% of machine investment. The payoff for efficient pricing can be impressive. The reason that demand pricing is so attractive is that it adapts to an unpredictable and varying demand. The more unpredictable and more variable the demand, the better demand pricing can be expected to work.

Since the producer's profit is being optimized by demand pricing, one would intuitively suspect that the consumer must therefore suffer. But as was pointed out previously, in the absence of monopoly on the part of producer or consumer, the demand price is optimum for both. Even if the producer has no interest in profit (such as a captive center operating on a nonprofit basis as part of a company, university, or government agency), the welfare of the consumers would be served by demand pricing. Not only would demand pricing tend to optimize the user's utility, but the added efficiency of the operation due to the unwelcome profit would help justify new equipment and more useful service.

This pricing approach capitalizes on a short-run monopoly situation in that the long-run possibilities of free supplier entry into the market cannot help the short-run congestion. Firms can charge higher prices during such times because it is the long-run profit history which gives rise to the entry of new firms, competition, and lowered prices. To the extent that this pricing method yields higher long-run profits, it will encourage entry of new firms. However the decision for any one firm must be made assuming that the profitability of the best scheme will determine the decision of new entrants. There can be little incentive for a firm to minimize its profits (note carefully that there is much incentive for a firm to understate its profits) in order to avoid entry of competitors when confronted by short-run congestion. The best short-run response to congestion is higher prices in accordance with this paper while the demand lasts. Investment in new physical capacity should be a response to long-run profitability analyses assuming the best short-run decisions.

The implementation of a demand pricing strategy requires some care. Certainly it is clear that the scheduler must solve a linear programming problem repetitively to ensure the best mix of nontrivial jobs to optimize profit. But a fairly complex system of user bids and limits would be necessary for both batch and time-sharing users. Since the profit obtainable through linear programming is a function of the number and variety of user tasks available, it makes sense to obliterate immediately any distinction between batch and time-sharing. Each user would be called upon to provide a bid for each resource used. In the case of files some provision is needed for moving them to less expensive storage when the user's bid for disk space is not high enough. With his bid, the user would have to specify the maximum length of time involved. Various aids to the user should be provided such as default bids and meaningful classes of service with assigned defaults. Perhaps a program could be provided to supply bids as a function of required turnaround. But no guarantee of turnaround can be given by the center. Perhaps an independent agent could, for a fee, guarantee turnaround. But any conflict of interest can result in lowered profit when the agent declines to take the loss when demand unexpectedly increases. An alternative approach would be to have the user specify a completion time and a price range bid. Whenever the price drops below the range, the low bid is taken. When the price is in the range, resources are provided at a rate to maintain linear progress toward completion. When the price is above the range, no resource is allocated. In this way a user may achieve approximate turnarounds when feasible at an average price. This is modeled on the noncompetitive bid for government securities.

The essential quality which must be maintained is the ability to adapt to changing conditions and highly variable load. The rewards for doing so can be substantial.

REFERENCES


(1) Whipple, W.L. "Optimal Profit Scheduling." M.S.E. Thesis, Computer, Information, and Control Engineering, The University of Michigan, Ann Arbor, 1974.

(2) Gordon, Geoffrey. System Simulation. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1969.

(3) Forrester, Jay W. Industrial Dynamics. Cambridge: MIT Press, 1961.

(4) Forrester, Jay W. World Dynamics (Cambridge: MIT Press, 1971.

(5) Jensen, Randall W. and Mark D. Lieberman. IBM Electronic Circuit Analysis Program. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1968.

(6) Bowers, James C. and Stephen R. Sedore, SCEPTRE: A Computer Program for Circuit and Systems Analysis. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1971.

(7) SYSCAP (TRACAP) User Information Manual. Minneapolis: Control Data Corporation, 1971. Publication Number 86602000.

(8) TESS User Information Manual. Minneapolis: Control Data Corporation, 1972. Publication Number 86615900.