SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

TECHNICAL REPORT NO. 659

May 1985
Revised April 1986

# PRODUCTION-DISTRIBUTION SYSTEMS INVENTORY PLANNING (SIP): RATIONALE, ECONOMICS & REALITIES

By

W.L. Maxwell, J.A. Muckstadt,
P.L. Jackson, R.O. Roundy

This research was partially supported by National Science Foundation Grants MEA-81-21429 and ECS-84-04641.

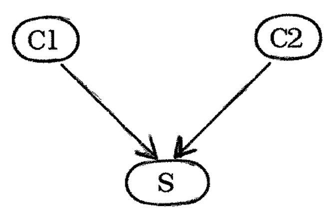
ABSTRACT

We have developed a set of algorithms and associated software for systems inventory planning (SIP) of production-distribution systems. The algorithms have been published in a set of technical reports that are quite detailed and complex in their mathematical developments. The basis for these algorithms, however, is very intuitive. The purpose of this paper is to explain the modelling basis of the problem addressed, including the reasons for certain key assumptions, and to discuss implications of our results for process and facility design.

In previous technical papers we have developed a method for determining the reorder intervals for all of the parts of a complex product. This paper has several purposes:

1. explain the rationale for the model that is used and its assumptions
2. give an intuitive basis for the method
3. point out the uses for the method.

Consider a very simple product structure, conisting of two components, C1 and C2, and one subassembly S,




C1, C2, and S are <u>parts</u> that may be stocked. For simplicity assume that one each of parts C1 and C2 are needed to make subassembly S.

Let

$K_i$ = the setup cost for part $i$, in dollars

$D_i$ = the annual demand for part $i$

$h_i$ = the annual holding cost for part $i$, in dollars

$T_i$ = the reorder interval for part $i$, in years.

The standard textbook approach says to choose the Economic Reorder Interval, $ERI_i$, by minimizing $K_i/I_i + h_iD_iI_i/2$, yielding,

$$ERT_i = \sqrt{K_i/(h_iD_i/2)} \ .$$

This approach, in which the economic reorder intervals are chosen independently of one another, has two serious drawbacks. For example, consider the data and corresponding Economic Reorder Intervals given in Table 1.

| part $i$ | setup cost $K_i$ | annual demand $n_i$ | holding cost $h_i$ | $ERT_i$ |
|---|---|---|---|---|
| C1 | .625 | 10000 | .05 | .05 |
| C2 | 74.25 | 10000 | .10 | $\sqrt{.1485}$ |
| S | 12.5 | 10000 | .25 | .10 |

Table 1

The inventory plots of part C1 and part S are given in Figure 1.

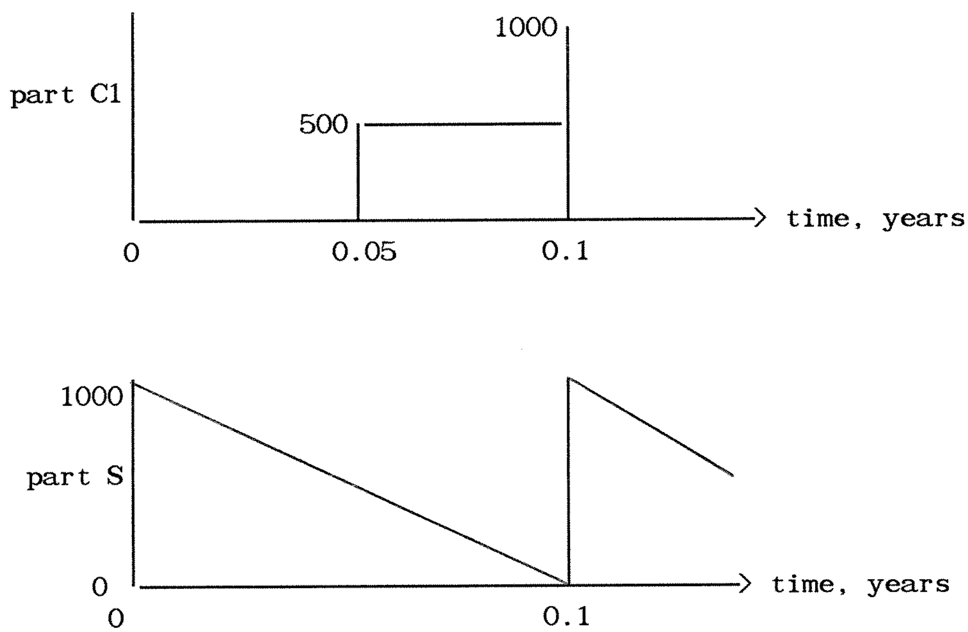



Figure 1

Every 0.1 years part S consumes 1000 of part C1. Since part C1 is produced twice and used only in this interval, there is a buildup of part C1. An obviously better solution is to produce part C1 only when it is needed--this will involve only one setup every 0.1 years rather than two and will not involve holding inventory of part C1. Clearly this is less expensive, so there is something amiss in the way the calculations were done. As we shall see it is the <u>model</u> that is wrong.

Second, part C2 is produced every $\sqrt{.1485}$ years while part S, which needs part C2, is produced once every .11 years. The inventory plots of part C2 and part S are given in Figure 2.

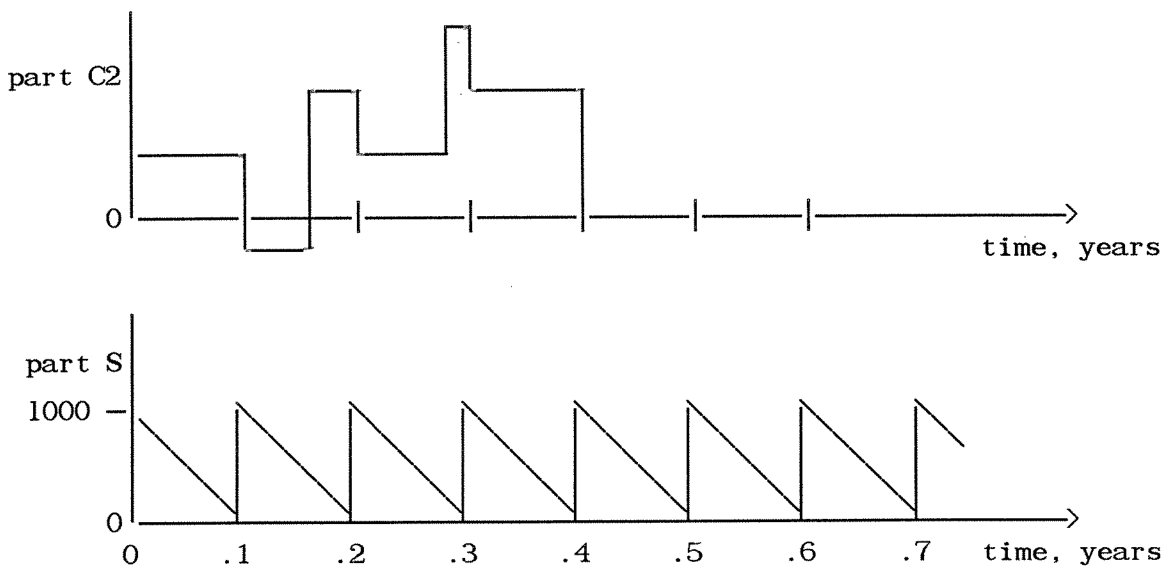



Figure 2

Suppose that these economic reorder intervals are to be used indefinitely. One can show that if shortages of part C2 are to be avoided, the inventory level of part C2 after orders at time zero are placed cannot be less than 1000. Also note that the inventory level of part C2 has a very complex time path. The problem is that the <u>model</u> for the average inventory level of part C2 is incorrect.

The problem is that the cost model

$$\text{cost} = \sum_{i=C1,C2,S} (K_i/T_i + h_i D_i T_i/2)$$

ignores the fact that parts C1 and C2 are needed to make part S.

The first drawback can be overcome by requiring that the reorder interval of a part should not be less than the reorder interval of a part that it goes into. Thus we should have the restrictions

$$\begin{aligned} T_{C1} &\geqq T_S, \\ T_{C2} &\geqq T_S. \end{aligned} \tag{1}$$

The second drawback was that in order to avoid component shortages we needed an initial buffer inventory of a component of size equal to the economic reorder quantity of the part using the component. In a system that (in theory) is completely reliable and has infinite capacity, this seems excessive. This drawback is not eliminated by requiring (1) to hold because in the example we had $ERT_{C2} > ERT_S$.

The second drawback arose from the fact that $ERT_{C2}/ERT_S$ was an irrational number. We can eliminate the need for an initial buffer inventory of a component and insure that components are not produced earlier than they are needed by modifying the restrictions to

$$\begin{aligned} T_{C1} &= M_{C1} T_S, \\ T_{C2} &= M_{C2} T_S \end{aligned} \tag{2}$$

where $M_{C1}$ and $M_{C2}$ are positive integers.

There is a final, and subtle point, which is the interpretation of the holding cost. Consider component C2 and subassembly S and assume that $T_{C2} = 3T_S$, and $T_S = 1$. The inventory plots of part C2 and part S for this situation are given in Figure 3.

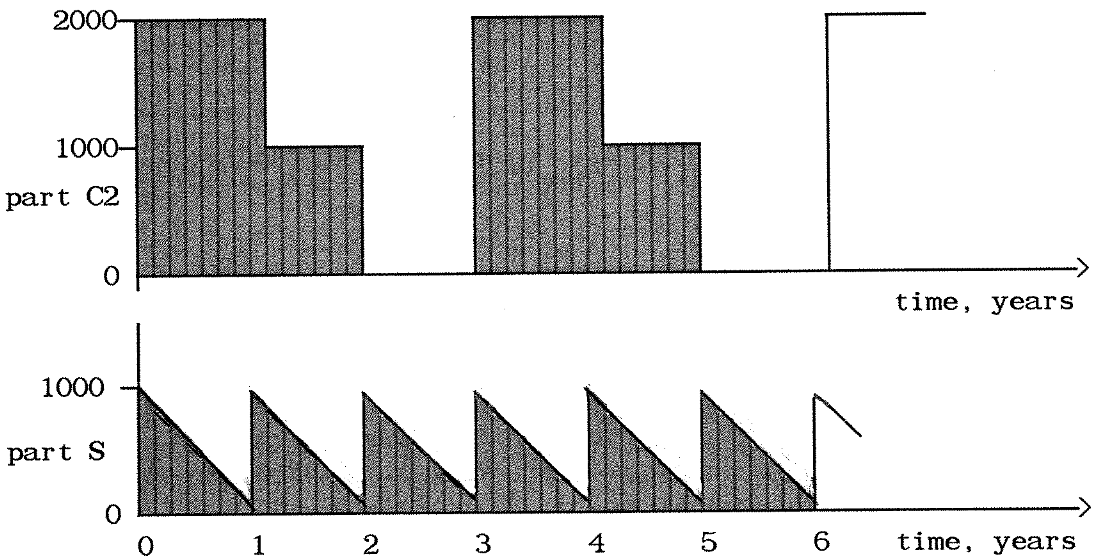



Figure 3

The average inventory cost for part S is $h_S D_S T_S/2$. The average inventory cost for part C2 depends upon <u>both</u> $T_{C2}$ and $T_S$; it is $h_{C2}D_{C1}(T_{C2}-T_S)/2 = h_{C2}D_S(T_{C2}-T_S)/2$ because $D_{C2} = D_S$. The total inventory cost is

$$h_S D_S T_S/2 + h_{C2}D_S(T_{C2}-T_S)/2 = (h_S-h_{C2})D_S T_S/2 + h_{C2}D_S T_{C2}/2.$$

The subtle point is that the inventory holding cost for a subassembly is not proportional to its cumulative value but to its value added. If we had considered both components C1 and C2, we would have found the effective holding cost for part S to be $h_S-h_{C1}-h_{C2}$. This concept of "value-added" holding cost is discussed in Caie and Maxwell [ ]. The closely-related concept of echelon stock, has been in use for decades in the literature on inventory in distribution systems [ , ].

Using this concept of "value-added" holding cost, define

$g_{C1} = h_{C1}D_S/2$, the value-added holding cost if $T_{C1} = 1$ year

$g_{C2} = h_{C2}D_S/2$, the value-added holding cost if $T_{C2} = 1$ year

$g_S = (h_S-h_{C1}-h_{C2})D_S/2$, the value-added holding cost if $T_S = 1$ year.

The problem we wish to solve is

$$\begin{aligned} \min \quad & \sum_{i=C1,C2,S} (K_i/T_i + g_iT_i) \\ \text{subect to} \quad & T_{C1} = M_{C1}T_S, \\ & T_{C2} = M_{C2}T_S, \\ & M_{C1}, M_{C2} \text{ positive integers.} \end{aligned} \tag{3}$$

The variables are $T_S$, the reorder interval for the subassembly, and $M_{C1}$ and $M_{C2}$, the multiples of $T_S$ for the reorder intervals of components C1 and C2. This problem has a non-linear objective function with a mixture of continuous ($T_S$) and integer variables ($M_{C1}$ and $M_{C2}$). We are not aware of any computer codes that can handle such problems.

The optimal solution to this problem using the data in Table 1 was obtained by an exhaustive search procedure. The solution is

$T_S = .1297$

$M_{C1} = 1$

$M_{C2} = 3$

Cost = 583.845

In [ ] we propose a solution technique that is based upon some additional model restrictions. These restrictions are:

1. Every reorder interval will be a multiple of the length of some <u>basic period</u>.
2. This multiple will be a non-negative power of 2, the so called <u>powers of two</u> restriction.

The <u>base period</u> restriction is used to ensure that the economic reorder intervals are tied to the natural planning cycle that one often finds in effect in material planning systems. For example, at GM and Ford the natural planning cycle is one week; in other firms we have found the cycle to be one fiscal quarter. without this restriction it is possible to have reorder intervals which are "unnatural" (such as $\sqrt{2}$ years) for planning purposes.

We have found that the <u>power of two</u> restriction is often used in practice, and for good reason. The power of two is based upon interval lengths, not lot sizes. If the length of the base period is a week than all production planning activity is for one, two, four, or perhaps eight weeks. The actual scheduling of production, setup activity, tool room activity, maintenance activity, etc. is much easier with this sort of interval planning. When demand for the final product varies the lot sizes are varied in this type of planning according to the variation in demand. If lot sizes are fixed then the timing of activities must vary as demand varies; such changes in timing of scheduled activities is difficult to adapt to.

The second reason for the power of two restriction is based upon experimental work that has been done on a classical scheduling problem called ELSP-the Economic Lot Scheduling Problem (see Elmaghraby [ ] for an excellent analysis of work in this problem). This problem requires that one consider not only the economics of a production plan but also the problem of scheduling in time the production time of a single machine on which several parts are produced. The experimental work on sample problems indicates that the best solution is usually a power of two solution.

To illustrate the reason for this phenomenon, consider a three-part example in which the base period is a week, $M_1 = 2$, $M_2 = 3$, and $M_3 = 4$. We must have the capacity of producing one lot each of parts one and two in a single week, as this will need to be done once every six weeks. Similarly, we must have the capacity to produce one lot each of parts two and three. However the average number of lots produced per week is $1/2 + 1/3 + 1/4 < 1.1$. Clearly the capacity of the machine is not being utilized effectively. It seems more appropriate to set $M_1 = 2$ and $M_3 = M_4 = 4$. This allows us to find a schedule in which one lot is produced each week.

The final reasons for the power of two restriction are that such a solution is very easily computed and that there is a guarantee that the cost of the solution will be within 6.1% of the cost of the best possible solution (details and proofs are given in [ ]).

The efficiency of the computations required and the fact that the structure of the solutions computed facilitates the generation detailed schedules and is often compatible with current planning activity make the base period and power of two restrictions extremely attractive.

The technical work we have done [ ] had its original focus in obtaining the actual numerical values for the economic reorder intervals for all parts in complex assembly and distribution systems. The focus was on the system as described by the Bill of Materials, including assembly stages and stages which produce parts used in more than one assembly. Much of what we have said so far has focused on how such a system should be modelled and how to otain the economic reorder intervals simultaneously considering all part interactions.

We now realize that the work we have done can be explained in economic terms that professors of production management have been groping with for years. We also realize that the potential payoff of the work may be more in helping product and process designers understand the economic impacts of production design rather than in simply calculating economic reorder intervals.

## Economic Interpretation

Consider a serial production process with $N$ stages; each stage is represented by a circle with its stage number within

$$(N) \to \ \dots \ \to (i+1) \to (i) \to (3) \to (2) \to (1)$$

Suppose stage $i$ adds value $v_i$. The value-added profile plot is constructed by showing cumulative value, $c_i = \Sigma_{j=1}^{i} v_j$ as a function of stage number, $i$. We suggest drawing this as in Figure 4.

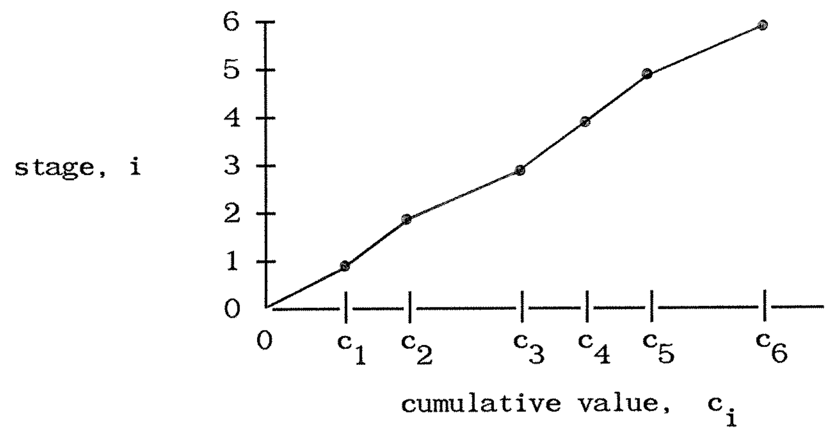



Figure 4

Production economists have argued that the forms of such a plot gives a "feel" for the nature of the entire production system and perhaps guidelines as to how the production system should be managed in terms of positioning of inventories. If this plot were convex, as say in Figure 5, then the late stages add most of the value to the product, whereas if this plot were concave, as say in Figure 6, then the early stages add most of the value to the product.

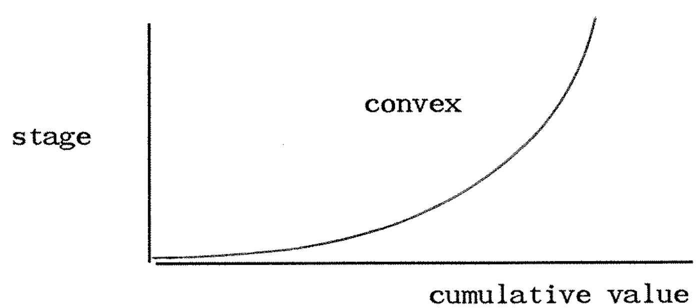



Figure 5

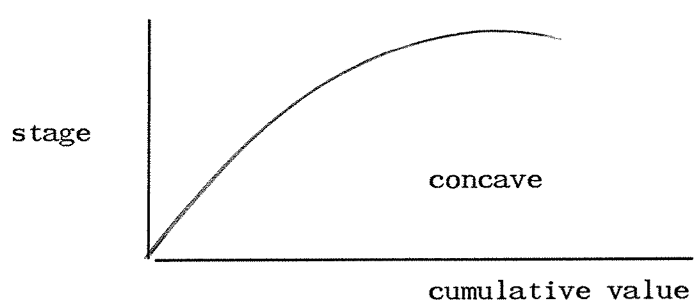



Figure 6

In the convex case one can argue, as economists do, that inventory should exist in work-in-progress since it has relatively small cumulative value added. In the concave case there should be no work-in-progress inventory.

These arguments ignore the processing nature of a stage. As shown in Figure 4, the ordinate of the plot is stage number. Suppose stage $i$ has a setup cost $K_i$. Let $k_i = \Sigma_{j=1}^{i} K_j$ be the cumulative setup cost up to and including stage $i$. Consider the plot of $k_i$ versus $c_i$ as given in Figure 7.

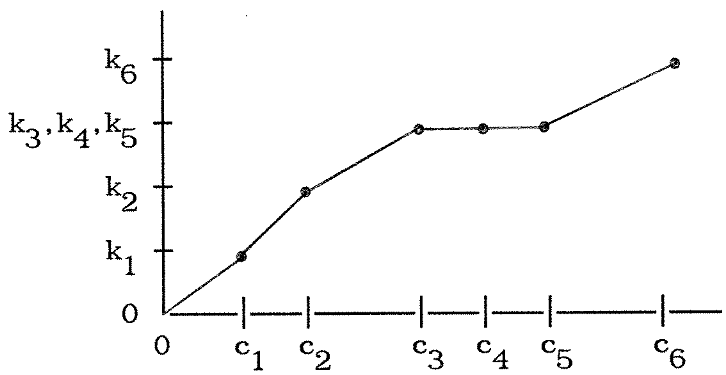



Figure 7

Define $c_0 = 0$ and $k_0 = 0$. Then,

$$K_i = k_i - k_{i-1} \qquad i = 1,\ldots,N$$

$$v_i = c_i - c_{i-1} \qquad i = 1,\ldots,N.$$

The plot of Figure 7 is obtained from the plot of Figure 4 by making the ordinate scale jump not by one (the stage number) but by the stage setup cost. It turns out that $\sqrt{K_i/v_i}$, the square root of the slope of the line segment in Figure 7 for stage $i$, is proportional to the reorder interval one would use for stage $i$ if that were the only stage. Recall,

$$\min \sum_{i=1}^{N} (K_i/T_i + g_iT_i),$$

where $g_i = \frac{1}{2} h_iD_i$, and $h_i$ is value added holding cost. Then,

$$T_i = \sqrt{K_i/g_i} = \sqrt{2K_i/h_iD_i} = \sqrt{2K_i/Rv_id} = \sqrt{K_i/v_i}\ \sqrt{2/Rd}$$

if $R$ is the percentage holding cost factor and $D_i = d$.

We have discussed the need for the restriction $T_i \geq T_{i-1}$, implicit in (2). This means that we cannot allow the slope of a segment in Figure 7 to be steeper than that of a segment further to the right. Thus instead of using the slopes of the solid lines in Figure 7 to compute economic reorder intervals we use their convex hull, the dotted lines in Figure 8.

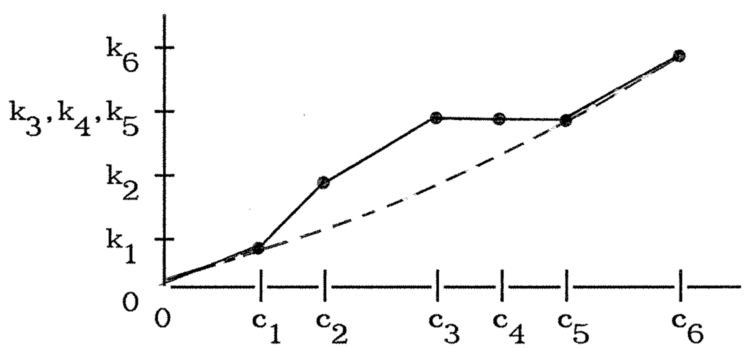



Figure 8

Note that segments 2, 3, 4 and 5 have the same slope for the dotted line. This is a <u>cluster</u>, a group of operations whose costs make simultaneous ordering more economical. The cost of the policy is proportional to the sum of the square roots of the shaded areas in Figure 9.

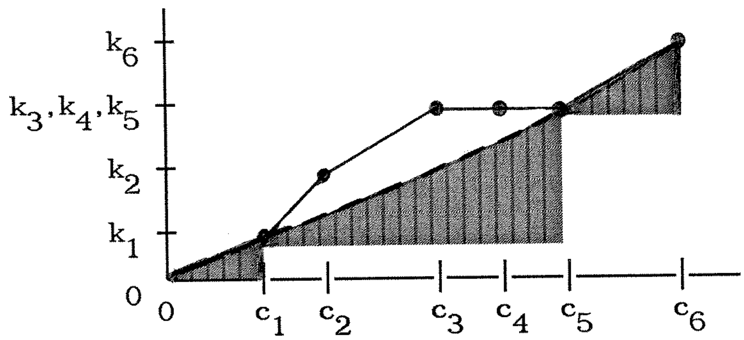



Figure 9

The foregoing analysis has been simplified in one very important respect. In a manufacturing environment the major component of the cost of a setup is usually not the direct cost of performing the setup, but the value of the production time lost during the setup. One of the valuable byproducts of our algorithm for solving this problem is a reliable estimate of the dollar value of an hour of setup time on each machine.

Design Implications

The economic interpretation above has implications for both process design and facility design. We begin with facility design.

All operations within each cluster are performed at the same frequency. In the example of Figure 8 operations 2, 3, 4 and 5 are in the same cluster. If these operations are scheduled in immediate succession, or are performed in parallel with transfer carried out either continuously or in small batches, the inventories of parts 2, 3 and 4 are for all practical purposes eliminated. Inventories of parts 1 and 6 will be held, however, because of the relatively small setup cost of part 1 and the relatively large setup cost of part 6.

The implications of the clusters on facility layout are evident. An effective layout should facilitate efficient transfer between 2, 3, 4 and 5 and should facilitate storage and retrieval of parts between stages 1 and 2 and stages 5 and 6.

With regard to process design, the gold is clearly to engineer the manufacturing process in such a way that the cost is minimized. Usually this is best accomplished by reducing selected setup costs and/or setup times. The question is, which ones should be reduced, and how much should one be willing to pay to reduce them.

The clusters provide insight into the selection of setup costs/times for potential reduction. The total cost can be written as

$$c = \Sigma_j 2\sqrt{K^j g^j}$$

where $j$ indexes the clusters and $K^j$ and $g^j$ are the total setup cost and value-added holding cost in cluster $j$. The partial derivative of the cost with respect to $K_i$ is

$$\frac{\partial c}{\partial K_i} = \sqrt{g^j/K^j} = 1/T_i, \quad i \text{ in cluster } j. \tag{4}$$

Thus the partial derivative is simply the reciprocal of the economic reorder interval, <u>assuming the economic reorder interval has been chosen correctly</u>.

Note that the unit of $K_i$ is dollars per setup. If the major component of the setup cost is lost production time, the dollar value of machine time is required to convert $\partial c/\partial K_i$ from (\$/day)/(\$/setup) into

(\$/day)(days/setup). In either case a reduction of $X$ dollars per setup in $K_i$ is usually easier to achieve for large setup costs than small ones. The larger setup costs in a cluster are usually found among the last operation in the cluster.