The OPT system keeps the levels of WIP low by releasing jobs to gateway work centres as late as possible for them to arrive at the assembly line in time for use. This is in keeping with the JIT approach. However, WIP is specifically provided in two areas and this means that WIP levels are higher than would be considered ideal under the JIT philosophy. Firstly, each bottleneck is provided with a 'time buffer' of, for example, x hours worth of parts awaiting processing. This serves to insulate the bottleneck for x hours from problems occuring at other processes upstream, and since the upstream processes are non-bottlenecks, they can presumably 'catch up' once the problems are resolved. Hence this time buffer is a protection against loss of throughput caused by the bottleneck standing idle. The other time buffer is provided at the assembly line in the form of parts that are not produced by bottlenecks. This means that even if there are delays in producing these parts, they will be already available for assembly once the part from the bottleneck arrives. This time buffer protects due date performance since delivery dates are promised relative to the output of the bottlenecks.

These two time buffers undoubtedly make sense in the context of a conventional production system, but they are in conflict with the JIT philosophy adopted by GEC Small Machines. The problem is that in providing time buffers, the continuing existence of wastes in the preceding non-bottleneck processes is accepted as a fact of life.

There is no incentive to make improvements to eliminate the wastes because the factory's throughput is insulated from them by the time buffers. If these are removed, the problems in the preceding processes will begin to receive serious and urgent attention, because they will begin to jeopardise the throughput of the plant. Once these wastes have been eliminated entirely, there is no need for any buffer stock at all.

The OPT system could never be considered seriously for GEC Small Machines because as Haylett [28] points out:

"... OPT is not a cheap option. Fees are on a sliding scale and start off at several hundred thousand pounds."

Also, OPT is a far more powerful software package than the company required because the scheduling of a job shop is extremely complex. This underlines one of the most important benefits of the GT approach used by the company. The change from a job shop configuration to a cell configuration vastly reduces the complexity of the control systems required. The need to consider almost every machine in the factory disappears and is replaced by the need to consider only the bottleneck machine in each cell, as discussed earlier. This approach is in line with the JIT philosophy's emphasis on simplicity, which was discussed at the beginning of this chapter.

The purpose of this discussion of OPT has not been to consider the use of the software package, but rather to consider the applicability of the ideas that the software package is designed to implement. These ideas represent a significant departure from conventional thinking and are, for the most part, applicable within the JIT philosophy. Bicheno [29] concurs:

"The prime requirement for a JIT scheduler is to know where the bottlenecks are and what their throughput rates are. Bottlenecks must be kept busy, but non-bottlenecks must occasionally be idle, and this is the desireable state."

3.1.5 The Periodic Control System at Kumera OY

Kumera OY is a Finish producer of a broad range of gear-driven, speed-reducing power transmissions. The company and its control system are discussed in some detail by Vollmann, et al [30] and the quotes that follow are derived from that source. The company's business appears somewhat similar to that of GEC Small Machines:

"The typical product would consist of a housing, mounting devices, shafts, gears, assembly hardware and perhaps a motor. Some 50 different part numbers would be usual ... (of which) one third to one half would be manufactured (generally the large items such as gears, shafts, housings, etc.), and the rest purchased ..... The large number of potential end items possible means virtually no finished goods inventory can be carried ....."

Kumera's MPC system is in many respects similar to a typical MRP II system and component production is controlled by a material requirements planning (MRP) approach. The applicability of MRP to GEC Small Machines will be discussed later. However, the structure of Kumera's master production schedule is substantially different from that of a typical MRP II system. Several similar ideas were used in the development of a master production schedule for GEC Small Machines. The important concepts are as follows:

i) "Five product groups have been formed on the basis of production process similarities. Within each group, divisions exist for product options. The groupings do not exactly conform to the catalogue product families, but it is easy to translate from customer orders to the groups for order entry purposes .... The product groups provide the basis for master production scheduling."

At GEC Small Machines this approach made sense since there are three distinct machine winding cells and a machine assisted handwinding cell. Each of the four cells winds a specific range of stator sizes and types. Even non-standard, customer specified windings can be assigned to one of these groups. At Kumera the five product groups are produced sequentially, but at GEC Small Machines all these cells operate independantly and simultaneously. Therefore it was logical to use four separate winding schedules at GEC Small Machines and to divide each winding schedule further into groups of products which would be produced sequentially in a particular cell. Thus, the approach is similar to Kumera's except that four schedules are produced instead of one. The rest of the discussion of the Kumera schedule and its groups is thus applied to each of the four schedules at GEC Small Machines and to the groups of products within each of the schedules.

ii) The year is divided into equal periods. "Each of the groups of products will be scheduled <u>once</u> during a period."

Exactly the same approach was applicable at GEC Machines, but the question arises as to how many periods to divide the year into; that is, how long should each period be. The period length 'trade-offs' for Kumera are given as follows (with comments by the author):

<u>Longer Period</u>	<u>Shorter Period</u>
- "Manufacturing efficiency""	- "Quicker customer response"

Perhaps some learning curve effects will occur	A definite advantage and one of the objectives of JIT
- "Larger purchase quantities"	- "Lower inventories (cycle stock)"
Under JIT this is considered a disadvantage	A definite advantage with all the resultant JIT benefits
- "Fewer set ups"	- "Higher potential market share"
This could be an advantage in the short term until the set up reductions envisaged under JIT are implemented.	A possibility due to improved customer response.

Kumera decided on a five week period, for the year under discussion, meaning that each group of products would be scheduled (and therefore produced) 10 times per year. The choice of period length for GEC Small Machines was not immediately apparent and required analysis of the physical constraints involved. This will be further discussed in the implementation chapter. It was clear at this stage, however, that the JIT philosophy would recommend the shortest possible period.

iii) "The sequence of production sets (the specific products to be produced within each group) within a period is based on manufacturing efficiencies in changing from one group to another."

Under GEC Small Machines' current constraints this idea was certainly useful since the machines in the winding cells had not yet received any attention with respect to setup reduction. Stator winding is a bottleneck operation and so the factory's throughput is dependent on its output, which can be improved if less time is spent on setups. Thus it was decided that this simple first approach to reducing setup time, by intelligent sequencing, should be used. It was also evident that if this sequence was fixed and repeated in each period, the demand for component parts would tend to be more level. This is highly desirable from the point of view of controlling component production and so it was envisaged that this fixed sequence concept would not be changed, even after some setup reduction had been performed in the future. The determination of the optimum sequence will be discussed in the implementation chapter.

iv) "An attempt is made to assign each order to the production schedule that will meet the delivery date requested by the customer .... This order promising is done in accordance with the commitment between marketing and manufacturing. No more orders may be entered into a production set than the period quantity."

Kumera limits the number of products that can be loaded into each group in the schedule since certain components have lead times longer than five weeks. These components may therefore be in process before the forecast quantity for the group has been consumed by actual customer orders. This makes it necessary to limit the maximum quantity of orders in a group per period, or else the necessary components will not be available to complete all the orders.

At GEC Small Machines it was expected that a system could be developed to control components which would reduce their lead times to less than the length of one scheduling period. In this way the only constraint on each schedule would be the total capacity available at each winding cell during the period and there would be no constraint on the mix of products loaded. This could mean that in a particular period, due to customer demand, certain products in the fixed sequence would be missed out and others produced in larger quantities than the average. Thus, it was expected that the GEC Small Machines system would be more responsive to customer demand than the Kumera system.

v) "Kumera looks at the timing table as a train schedule. When the launch date arrives, the production orders for the set are released (the train leaves). The sales department can make any allocation of schedule dates among customers until the train leaves for the factory. After that no more changes, they have to wait for the next train."

This principle was considered essential for GEC Small Machines. The new system had to inspire confidence on the shop floor and there is no better way to undermine this than to issue a schedule and then alter is whilst it is in progress in the factory. The future schedules would always be open to changes and revisions by the sales managers, but at a specified time on the day before a production period began, the upcoming schedule would be frozen in its present format and issued to the factory. The sales managers could then concern themselves only with the schedules for the upcoming period and beyond.

3.1.6 Theoretical Model of the Master Production Schedule

The conclusions that can be drawn from the various discussions up to this point lead to a theoretical model of the design of an appropriate MPS for GEC Small Machines which will have the following features:

1. The design of the schedule will avoid unnecessary complexity and its basic mechanism will be simple enough to be clearly understood by shop floor supervisors.

2. It will be a finite capacity schedule.

3. The total workload scheduled will require less than the full capacity that is theoretically available.

4. The schedule will be derived from exact customer requirements for the period in question.

5. The schedule will be designed to load only the stator winding section of the factory. From this MPS an identical, time offset, FAS will be derived.

6. The schedule will thereby drive the entire factory, and the component control system must be designed to respond directly to the requirements of the FAS.

7. There will, in effect, be four separate master production schedules; one for each winding cell and the range of products produced by it.

8. The capacity of each winding cell will be determined by the capacity of the slowest operation in the cell (the bottleneck). The loading of the other machines in the cell will be ignored.

9. The MPS will be divided into fixed periods of time (or time buckets) of equal length.

10. The full range of products will be available for production in each time bucket, but the mix that is actually produced in a particular time bucket will depend on the customers' requirements for that time bucket.

11. The complete range of products produced in each cell will be arranged into a fixed sequence which will be followed by that cell during each time bucket. The sequences will be designed to minimise the total setup times involved. The length of the time buckets will be as short possible but dependant primarily on these total setup times, as well as on marketing considerations.

12. Customer orders will be loaded directly into the soonest available capacity space (unless the required date is specified as a later date) in the MPS for the cell which produces the motor's stator. Once a particular time bucket has been filled to the specified percentage of theoretical capacity, subsequent orders must be loaded into the following time bucket or a later one. This procedure will be controlled by the sales managers and subject to trade-offs at their discretion. For example, they may decide to remove an order from a fully loaded planned time bucket and replace it with one that they consider to be more important. They will have complete flexibility in this regard, right up to the last moment before the MPSs for a particular time

bucket become current. When this launch date is reached, the MPSs for each of the four cells which pertain to that time bucket, will be frozen and issued to the shop floor. Once this point has been reached, no further changes to the schedules will be permitted.

13. The sales force will be given daily updates of the current lead times for each cell and thereby for the range of products associated with it. These lead times will vary according to the current workload awaiting each cell and will be used as a guide in negotiating with customers.

Once a customer's order has been allocated a space in one of the schedules it will be possible to give that customer a delivery date promise. This will be derived from the date of stator winding plus an appropriate allowance for assembly (i.e. the FAS date), plus shipping time. If the order is later rescheduled due to some trade-off decision, the customer will be advised immediately of the new delivery date.

Using the above framework, the company's master production scheduling system was developed and implemented.

bucket become current. When this launch date is reached, the MPSs for each of the four cells which pertain to that time bucket, will be frozen and issued to the shop floor. Once this point has been reached, no further changes to the schedules will be permitted.

13. The sales force will be given daily updates of the current lead times for each cell and thereby for the range of products associated with it. These lead times will vary according to the current workload awaiting each cell and will be used as a guide in negotiating with customers.

Once a customer's order has been allocated a space in one of the schedules it will be possible to give that customer a delivery date promise. This will be derived from the date of stator winding plus an appropriate allowance for assembly (i.e. the FAS date), plus shipping time. If the order is later rescheduled due to some trade-off decision, the customer will be advised immediately of the new delivery date.

Using the above framework, the company's master production scheduling system was developed and implemented.

3.2 Development of the Component and Raw Material Supply System

From the foregoing MPS discussion, three points concerning the component and raw material control system were already apparent:

i) The system had to respond directly to the production requirements of the motors in the FAS and to do so whilst discouraging overproduction.

ii) A procedure would have to be developed to decide what batch sizes should be run in component producing cells. The objective would be to utilise the spare time available at non bottleneck cells for doing extra setups. The batch sizes would be as small as possible whilst still allowing sufficient capacity to meet the demands of the FAS.

iii) The fixed production sequences to be built into the MPS would to some extent alleviate the unpredictability of demand for certain components. Although the quantities demanded would be variable, certain groups of parts would only be required once in each time bucket. This demand would come at fairly regular intervals (approximately the length of a time bucket) depending on the quantities of motors not requiring these parts, which had been scheduled in two successive time buckets. Unfortunately, certain parts are required by all the motors in a production sequence or by more than one group of motors in the sequence and so these patterns would not occur for all components or raw materials.

Thus the demand for components would certainly not be termed level, that is, regular in timing and in quantity; which is the JIT ideal. However, it was expected that there would be some improvement on the totally random component and raw material demands experienced in the past.

3.2.1. The Reorder Point System

The oldest and probably the most widely used inventory system in the world is the reorder point system (ROP). The classic reorder point is a perpetual inventory system because every time an issue is made (perpetually) the stock level is checked to see whether it is down to the ROP. This can be a physical check or, in the more formal situation usually found in larger stores, the stock record balance is checked. This is the basis of the 'Kardex' system that was currently in use at GEC Small Machines to control raw material and component stocks. Another common version of ROP is the two-bin system, where the stock must be drawn from one of the bins first. Once this is empty, an order is placed and the second bin, which contains the ROP quantity, is then used.

The ROP is calculated as that level of stock which will cover the average expected demand over the average replenishment lead time, plus some safety stock to protect against non-average surges in demand or lengthy replenishment lead times. Schonberger [31] points out the disadvantages of this system:

" .... ROP results in high inventories. More parts and raw materials are ordered for the sake of the (reordering) rule rather than because of need. Manufacturers that use ROP do so because of a difficulty in associating parts requirements with the schedule of end products."

GEC Small Machines certainly suffered the high inventory problem with its associated redundant and slow moving stocks, but Schonberger doesn't mention the other major problem. This was that the stock record balances on the Kardex system were almost always incorrect. Consequently, many mistakes were made in ordering parts or raw materials that were not required, even by the ROP system, and failing to order those that were. The stock balances were corrected twice a year after stocktaking, but within weeks would be totally unreliable again. As a result, when planning an important order, the planners would be obliged to go into the factory to see for themselves what was available. This was not usually an easy task since storage areas were disorganised and so the required parts might be held in more than one location, buried under other parts, etc. Thus there was a tendency to avoid physical checking and order "just in case", which led to still higher inventory levels.

It was pointless for GEC Small Machines to try to improve their use of the ROP system since there would, in fact, be no difficulty in associating parts requirements with the FAS once the new MPS system was implemented. Thus, a superior system could be utilised.

3.2.2 Push vs Pull Systems

The American Production and Inventory Control Society (APICS) gives the following definitions of push and pull systems [32]:

"Push System .... In production, refers to the production of items at times required by a given schedule planned in advance. In a material control context, refers to the issuing of material according to a given schedule and/or issued to a job order at its start time."

"Pull System ... In production, refers to the production of items only as demanded for use, or to replace those taken for use. In a material control context, refers to the withdrawal of inventory as demanded by the using operations. Material is not issued until a signal comes from the user."

From the above definitions it is clear that any schedule based system, such as MRP, is a push system. Kanban systems, on the other hand, which are driven by actual consumption and use synchronised replenishment signals for control, are pull systems.

A pull system appears to have the potential for more accurate control since it responds to actual demands and usages, whereas a push system responds to a schedule of expected demands and usages. This schedule could become inaccurate due to deviations in the performance of the factory, or the demands of customers, from what was expected at the planning stage; when the schedule was drawn up. For example, if the usage of parts at a particular work centre has been temporarily halted by a problem, a pull system will automatically adjust to this by not requesting parts from the upstream work centre. A push system however, will continue to push parts forward according to a planned schedule. Thus it is logical that a pull system will result in less WIP inventory than a schedule based push sysem. Hall [33] likens a pull system to a pipeline which connects the work centre where the stock is made with the work centre where it is used. The pipeline can only contain a certain fixed volume of inventory and it synchronises the movement of material with the rate at which it is withdrawn for use at the output end of the pipeline. If the producing work centre has no place to put inventory except in the pipeline, it can only produce when the using process has withdrawn some inventory and created a space.

This concept of a fixed-volume pull system has other advantages as Heard [34] notes:

"When the output end of a production process is tightly controlled and upstream workers are only authorised to produce in response to a pull signal, areas with too few and too many people are highly visible."

Clearly, any other capacity imbalances such as machine capacity bottlenecks or bottlenecks due to quality problems, will also be visible. Other pull system advantages noted by Bicheno [35] are:

" ... the pull system automatically prevents overproduction at non-bottlenecks ... the pull system often allows simplicity in scheduling since only end-items are scheduled with components simply being pulled as required."

The last point, especially, was of great importance to GEC Small Machines, where a simple system was prefered. If the master scheduling system already outlined could be combined with a pull system of component and raw materials control, planning activities would be vastly simplified. However, the search for the most appropriate solution would have been incomplete without a thorough examination of push systems.

3.2.3 Material Requirements Planning (MRP)

MRP is the most widely used, and developed push system in existence. It is also the most sophisticated except for OPT, which had already been ruled out due to its expense. A large number of companies use various manual push systems, but these cannot be expected to equal the performance of MRP. This is due to the need to maintain accurate and extensive data bases and to perform numerous calculations. MRP uses computers to handle these tasks.

There are many different versions of MRP software on the market, but they all have the same fundamental approach. The computer is used to perform thousands of simple calculations in transforming a master schedule of end products into parts requirements via a bill of materials. These gross parts requirements are compared with data concerning the existing stocks on hand and those already on order, to give nett requirements. Orders are then placed for the nett parts requirements. Fixed, predetermined procurement and manufacturing lead times are used to decide when orders should be launched by means of a backscheduling approach.

Work orders, despatch lists and input/output control are the major execution mechanisms. All of these are heavily dependant on the collection of data after every operation so that the progress of work orders on the floor can be monitored and output can be measured. Substantial information processing capability is needed to track work orders and update priorities. Work in progress is limited by controlling the rates of input to gateway operations and attempting to manipulate capacity where possible.

Harrison [36] points out some flaws in the MRP approach:

"To produce a schedule without recognising the bottlenecks is unrealistic, but that is what MRP modules do! The supporters of classic MRP claim that they get around this by subsequently adjusting the demand pattern in the Master Production Schedule (MPS) to avoid a bottleneck that has been recognised by a Capacity Requirements Planning (CRP) module, and starting the process again."

He believes that this can have one of three results. One is that the reallocation of the load will generate a bottleneck somewhere else and the MRP-CRP-MPS route must be followed repeatedly. Another is that the resulting capacity plans will consist of minor over and underloads everywhere. "This is frequent MRP practice and the resultant mess of shop floor loading is left to shop planners and foreman to sort out!" The third option is that <u>all</u> resources can be planned at underload with the resultant loss of potential throughtput. Harrison advocates, instead, the forward finite scheduling of bottlenecks as part of the OPT system discussed earlier.

Mather [37] mentions these shortcomings as a part of his discussion of complete MRP II systems:

"MRPII's scheduling logic is fundamentally wrong and simplistic for all but the simplest of operations; there is nothing in its scheduling process that even attempts optimisation. Infinite capacity scheduling, frozen parameters, and an illogical attention to materials without concern for capacity are just some of its shortcomings."

Hartland-Swann [38] cautions:

"Our own experiences within numerous companies in Europe and overseas suggest that there are inherent problems with the operation of closed-loop systems which make them almost impossible to operate as intended. Moreover, the problems are far worse as a company develops towards more responsive manufacturing, using JIT techniques."

Schonberger [39] notes that MRP correctly calculates part requirements by precisely associating them with the master schedule of end products.

"But what is correct at the time of calculation is subject to error later. The reason is that lots are sizeable, and the production lead time is long, from one to several weeks. During that lead time there will be delays and schedule changes so that the lot being produced no longer is correct in relation to the master schedule of end products. The lot size and lead time erode the close association between parts requirements and end-product schedules."

Due to these long lead times a substantial portion of the MPS has to be based on forecasts; but these don't match the actual customer demand. Then as Hartland-Swann [40] observes:

"In real life, companies respond to customer demand not by re-running the MRP system (usually a complex and time consuming task), but by generating a "hot list" of urgent requirements and expediting them on the shop floor."

This means that certain jobs are sidelined and that planned capacity is used for unplanned activities. The system will tend to stabilise itself eventually if the feedback loops operate correctly, but the result is higher WIP and even more lengthy lead times.

Schonberger critises the guesswork involved in preparing schedules of estimated demand and in planning on estimated lead times. He comments [41]:

"The system allows corrections to be made daily (called shop floor control). Nevertheless, bad guesses result in excess inventories of some parts, ..."

Clearly MRP systems require very accurate inventory records if they are to operate correctly. Figures of 95% or even 99% accuracy are often quoted for so-called "Class A" users. The goal of every MRP user is to become "Class A", but Hartland-Swann asks [42]:

"What have they actually achieved? What value have they added to their manufacturing process by such accurate knowledge? Very little, because in truth, inventory recording is not a value-adding process.

Furthermore, accurate recording is very difficult for humans to achieve and sustain, especially if they have to count large quantities, in a number of different locations or in transit between locations. To sustain accurate recording usually requires a big investment in stores control,movement control, video terminals, bar code readers and so on."

Many MRP systems attempt to keep their records accurate by requiring a computer transaction for each movement material so that the inventory can be charged out of t sender's records and into the receiver's. Other system charge material into stockrooms (whether physically there or not) after every operation; and then out of the stockroom when the work gets to the next operation. The use of such stockroom transactions doubles the number of data preparation and entry steps. All this effort is seen as necessary in order that jobs cannot be lost or misplaced; which is a common problem in plants with little production control discipline (the existing GEC Small Machines included). This system is also supposed to encourage everyone to keep work moving forward so that it will get charged off their work centre's records. Schonberger [43] sums up these methods of control:

"In some plants the system works so well that nothing gets lost anymore. The elaborate system adds no value to the product, however, and it is a system that averts chaos by keeping many fingers in the holes in the dike."

It would obviously be better to avoid the need for all this complication and expense by reducing the inventory levels until there is little to count and by using systems that do not require very accurate inventory records. The ultimate "successful" MRP implementation strives for complete closed-loop control and this is traditionally considered to be one of the prerequisites for <u>any</u> successful MPC system. However, Hartland-Swann [44] challenges this important principle of traditional thinking:

"the systems requirement for accurate inventory levels can only be avoided if closed-loop control is abandoned."

As well as accurate inventory records, MRP requires realistic lead times; but in practice they can vary enormously, according to many factors. These include scheduling priorities, batch sizes, capacity utilisation, quality problems and even motivation of personnel. This means that in the process of guessing lead times, mentioned earlier, the tendancy is to inflate them with safety factors. The result is higher levels of WIP than are necessary or desirable; for the sole purpose of providing the MRP system with accurate data. Even so, these pessimistic lead times often still prove inaccurate and so the resulting manufacturing plan is inappropriate. One of the symptoms of this is noted by Sandras [45]:

"It is not uncommon for production control to spend approximately 75 percent of their time scheduling and rescheduling workorders (i.e. shop orders, manufacturing orders). In each MRP cycle some workorders are pulled up and others are slid out. It is important that workorder dates are maintained so that they can be initiated on the appropriate date ..."

However, processing times for full regeneration of MRP data are often too long to be done overnight. Instead, processing is usually done over a weekend, which means that the MRP cycle has to be at least one week. Some MRP systems allow net change runs, which can usually be done daily, but the changes which can be accommodated are restricted.

The most common result of all the inherent MRP system problems that have been considered up to this point is shortages. This gives rise to a need for separate, often manual, shortage control systems and extra staff to administer them. Manufacturing control becomes dominated by expediting and thus the errors of the MRP system drive the factory, as opposed to the MRP system itself. This makes the information in the MRP system less accurate still, and the discrepancy between the two systems is increased. The increasing level of shortages and "hot lists" can eventually discredit the formal MRP system.

The preceeding discussion highlights some of the problems experienced with MRP in a traditional environment. However, when considering the use of the JIT philosophy, the problems are even more fundamental. Schonberger [46] points out an apparent paradox in the use of MRP systems:

"MRP companies order in lots, rather than piece-for-piece (just-in-time), because American MRP companies have not lowered set up times in order to make small lots economical. If they did so, then simple manual Kanban rather than complex, expensive, computer-based MRP would be the logical choice. The MRP paradox is that if the company removes the set up time obstacle in order to make MRP truely effective in cutting inventories to the bone, then MRP is no longer needed; Kanban is preferable."

The author believes that the above is an understatement and that the use of MRP systems in an advanced JIT environment would probably be impossible. Hartland-Swann [47] seems to suggest this:

"Responsive manufacturing demands much faster response to changes, frequently on an hour-by-hour or minute-by-minute basis. MRP systems are inherently cumbersome and can only be made to respond at this rate by enormous expenditure on hardware and communications facilities. Even then, the complexity of the output reports usually prevents the results from being usable within the time required."

Perhaps even more serious than this is that MRP systems require an order to be generated, printed and issued every time production or purchasing of a batch of parts is required. In an advanced JIT system, with batch sizes approaching one, this should happen many times per day for the majority of part numbers. The consequent opening and closing of shop and purchase orders every minute, by dozens of clerks, would require enormous computer capacity in order to attempt to maintain a real-time system. Shop foremen and buyers would be overwhelmed by piles of computer generated paperwork and the previously described non value adding aspects of the system would be multiplied ten or a hundred fold. In short, MRP with lot sizes of one and lead times measured in hours or minutes is impossible to envisage.

The above was, for GEC Small Machines, an extremely long term view. But, in considering the present, a further fundamental conflict between the JIT philosophy and MRP became apparent. In discussing the MRP system, Shingo [48] states:

"It is doubtful whether that system deeply considers the improvements of the fundamental points of the control system itself .... Considerations must be performed not to lean towards measures of utilising computers extensively, but maintain not to lose the revolution of production system."

What he appears to be saying here is that in designing and implementing a complex computerised control system that will run an existing factory effectively, the present inefficiencies, problems and wastes are being accepted and included. The revolution of the production system is the WCM approach of eliminating waste through better factory layouts, setup reduction, TQC, TPM, etc. MRP <u>accepts</u> the current values of planning parameters such as lead times, lot sizes, routings and defect rates, whereas the WCM/JIT philosophy is primarily concerned with continuously <u>improving</u> these parameters. As Coetzee [49] puts it:

"..... our classical change philosophy is 'unfreeze, introduce change, refreeze' .... This tendency is a negation of basic JIT philosophy and should be fought."

In order to implement an MRP system, enormous amounts of data concerning conditions in the factory must be collected or estimated. After at least a year of hard work, the system might finally operate satisfactorily; once the many planning parameters have been specified with sufficient accuracy. There is then a strong disincentive to make any changes to the factory and thus the system, let alone to make the <u>daily</u> improvements envisaged by the JIT philosophy.

3.2.4 The Selection of a Pull System

Schonberger [50] states that the various inventory systems fall naturally into a continuum and presents them in a diagram similar to Figure 3.1. As it becomes more difficult to associate parts requirements with end product schedules, the inventories carried tend to increase from the theoretical zero on the left to months' worth on the right.

On the inventory scale, GEC Small Machines fell in the MRP bracket with days' to weeks' worth of iventory carried, and in some cases, even in the ROP bracket, with months' worth. However, on the scale of ease of associating parts requirements with end-product schedules, the company should have been in the dual-card or even the single-card kanban bracket. This was a clear signal of the potential to improve and evolve towards a company with hours' worth of inventory, using simple control systems. Since the company could have been described as one struggling to use ROP and a little manual MRP, the advice of Schonberger [51] appeared particularly apt:

"Western ROP companies that are able to streamline their processing and reduce the uncertainty of the coupling between parts and end-product schedules often will want to bypass the MRP stage. MRP is a very expensive undertaking, and its approach is to attack problems with complex solutions, i.e. computer systems .... In most cases, however, money is better spent on JIT/TQC than on computer-based planning and control. If we have learned

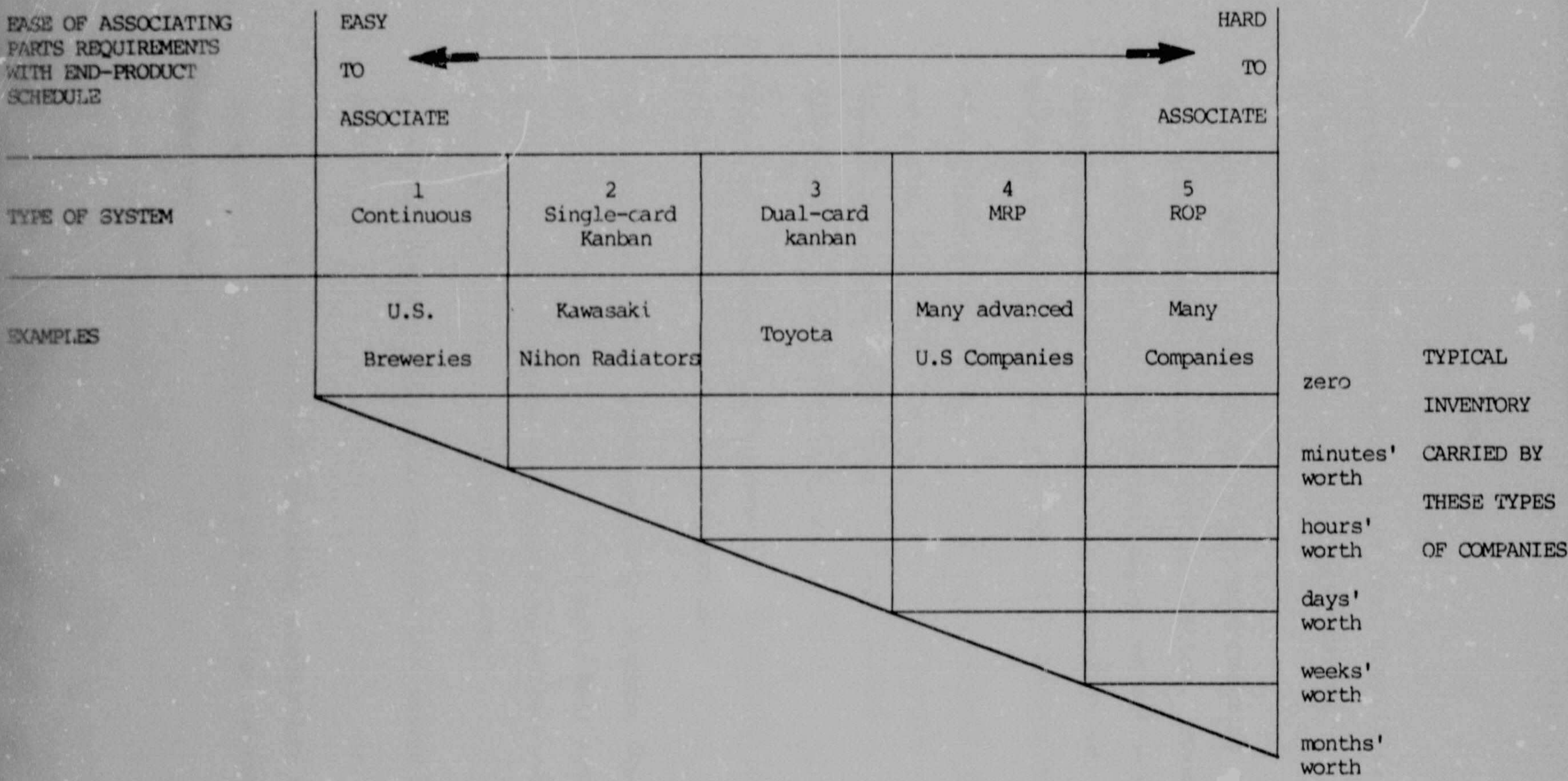

Figure 3.1: - Manufacturing Inventory Systems

anything from the Japanese, it is that simplification is generally the safest path to improvement."

His last point is one which enjoys overwhelming support from other experts too numerous to mention. Professor Jinichiro Nakane has had considerable contact with major Japanese practitioners of JIT and summarises the most important aspects of their approach [52], beginning with:

"The simplest way is best. Study the physical problems carefully and try to solve them by simple means before turning to more complex ones. Simple solutions to problems are more likely to succeed .... "

Ohno [53] condemns push systems in the context of waste elimination and continuous improvement:

"To accomplish this, management must abandon the practice of giving production plans to each and every one of the processes, or of making the preceding process transport its products to the subsequent process."

Hall [54] believes that pull systems are essential for JIT:

"the initial objective of stockless production is the development of a pull system and all that is required to make one work."

He sees the physical conversion of the plant to a pull system of planning and control as a major milestone from which further development can proceed and advises [55]:

"Provide a pull system of production control ... This system should:

- Get the right part to the right place at the right time.
- Prevent the ballooning of throughput times by controlling the amount of work-in-process inventory for each part at each operation.
- Prevent the need to put inventory in stockrooms."

In summary then, the MRP push system approach was rejected by GEC Small Machines in favour of a pull system, for the following reasons:

i) The expense and implementation effort required to adopt MRP and develop into a 'Class A' user.

ii) The many inherent MRP system problems that would remain even when it was fully implemented.

iii) The difficulty of maintaining the JIT philosophy of continuous improvement through waste reduction in a factory controlled by MRP.

iv) The certainty that the MRP system would need to be abandoned before the company could become an advanced JIT producer with lot sizes of one and lead times measured in hours or minutes.

v) The company's relative ease of associating parts requirements with end product schedules.

vi) The overwhelming support of the JIT authorities for simple systems and for pull systems where possible.

vii) The power of pull systems as a tool to facilitate the continuous improvements envisaged by the company's JIT philosophy.

3.2.5 Kanban Systems

The word 'kanban' originally referred to the signboards of a store or shop, but in the manufacturing context it has come to mean a card or any other small sign or signal displayed to a worker. The kanban card system requires that all containers of parts have a card attached to them. Basically, a card is removed whenever a container quantity is used and the card is returned to the preceding process, where the same quantity of parts is manufactured to replace those used. The quantities per container are very carefully and exactly controlled, as are the numbers of containers per part number. No parts may be produced without a kanban 'demand.' With such disciplines, Schonberger [56] observes:

"... inventory control is simple and far more precise than manual or computer-based Western systems."

and Vollmann et al [57] admit:

"One beauty of the kanban system is its simplicity."

Kanban is a mechanism that is used to request parts, authorise production and movement of those parts and control the level of WIP. Thus, it serves as both a production control and a stock control system in one. Kanban cards replace all shop paper such as identification tags, operation instruction tickets, move tickets, despatch lists and input/output reports. Unlike the conventional shop paper, kanban cards can be used repeatedly. Also, the time and effort involved in opening and closing shop orders in MRP is avoided, as is the inventory needed to cover the delays involved in these activities.

Ohno [58] describes the various functions of kanban:

"The kanban contains information that serves as a work order. This is its first function........

The quantity to be produced, the time, methods, procedures and amount to be transported, where it is to be transported, the place to store it, the means of transportation and the type of containers to be used can all be known at a glance through the kanban ..... The second function of the kanban is to move with the actual material."

Thus, kanban facilitates the control of material, prevents overproduction and gives directions based on the actual conditions existing in the workplace. Work is always performed in the same, standard manner since all pertinent information is given on the card. This standardisation of procedures helps to prevent the production of defectives.

One of the most powerful features of kanban is the ability to highlight problems, or wastes, in the production process. When everything is running smoothly between a supplying and a using work centre that are linked by kanban, one or two kanban cards and their containers are removed from the system. This has the effect of reducing the level of buffer stock between the two processes and is done gradually until the producing work centre has difficulty in keeping up.

Numerous problems can be exposed in this way such as poor quality, long setups, unreliable machinery, transportation delays and so on. The deliberate withdrawal of cards focuses the attention of both workers and supervisors on these problems. The easiest or least expensive changes, that will allow smooth operation at the lower level of inventory, should then be made. Thus, kanban is also a productivity improvement technique and attaining such a control system, although very useful in itself, should rather be seen as a method for promoting continuous improvement and developing the accuracy of the production system through waste elimination.

Dual Card Kanban

This is the most well know form of kanban, developed at Toyota. Each part number has a certain calculated number of containers provided for it. The number of containers needed for a part number depends on the rate of usage and the cycle time to replenish what has been used. The number of kanban cards issued determines the number of containers of stock in the system.

For each container of parts there are two cards. One is the withdrawal or conveyance kanban (C-kanban) and the other is the in-process or production kanban (P-kanban). When a using work centre begins to consume the parts in a container from its inbound stock point, that container's C-kanban is detached and sent, together with an empty container, to the stock point at the output end of the supplying work centre. At this stock point there should be at least one full container of the required parts waiting, with a P-kanban attached to it. This P-kanban is detached and placed in a collection box, or hung on a rail, at the supplying work centre. The C-kanban is then attached to the full container, which is taken to the inbound stock point of the using work centre, where the C-kanban cycle begins again.

The P-kanban are the production orders for the supplying work centre and are worked on in the sequence in which they are received; hence the use, in some cases, of a rail where the cards are hung up in a queue. When the parts specified by a particular P-kanban have been produced, they are placed in an empty container. The

P-kanban is attached to this container, which is then left at the outbound stock point to await collection via a C-kanban.

Containers are not dedicated and can usually be filled with any of the part numbers that a work centre produces. The discipline of returning an empty container from the using work centre with every C-kanban ensures that the supplying work centre always has empty containers available.

During the enforced improvement process of removing one or two cards and containers, or due to some other cause, it may happen that when a C-kanban is returned to the stock point, the required parts are not available. In such a case, the C-kanban is simply handed to the supplying work centre, which interrupts all of its other work and devotes itself fully to production of the parts for which a shortage has been noted. As soon as these parts have been made, they are delivered to the using work centre.

The dual-card kanban system cannot function effectively without the disciplines contained in the following rules:

i) Any container filled with parts must always have either a C-kanban or a P-kanban attached to it. The only exception to this occurs when cards are removed from the system. Those containers without cards are used first and then are automatically not replenished because of rule (ii).

ii) No parts may be produced without an unattached P-kanban to authorise them. A work centre will come to a halt rather than produce parts that are not yet required. This means that dual card kanban is a pure pull system with no over-production permitted. Workers do maintenance work or improvement projects when there are no P-kanban to work on.

iii) No parts may be drawn or transported without a C-kanban.

iv) Only standard containers may be used and they must always be filled with the exact (small) quantity of parts specified by the kanban card. All of these parts must be of the correct specification, that is, no defectives are permitted; as this would be equivalent to supplying less than the required quantity.

v) Production must take place in the sequence in which the P-kanban are received.

The dual card kanban system can also be used when the supplying work centre is actually an outside vendor. In such a case, the outside supplier might deliver the parts demanded by the C-kanban, rather than the user fetching them. However, in certain cases the user may find it beneficial to send its own small truck, with

C-kanban, to collect parts from several suppliers in one trip. Alternatively, several suppliers may share a truck making frequent kanban deliveries to the same customer. These are just two examples of the kind of special arrangements with suppliers that are usually necessary in order to implement vendor kanban systems. The kanban cards are often modified to include extra information according to the special requirements. Bar codes on the cards may be useful to record incoming deliveries or alternatively, backflush methods of payment may be agreed upon. In spite of the need for special arrangements with outside suppliers, the basic mechanism and the rules of kanban remain the same.

Single Card Kanban

Single card kanban is more widely used than dual card kanban, probably because it is simple to implement, although the benefits are not as great. This system uses only the conveyance kanban (C-kanban), to control the movement of parts to the user. The method of circulating the C-kanban is essentially the same as for the dual card system. Parts are produced, or bought, according to a daily schedule, and are left in containers at the outbound stock point to await collection by users with C-kanban. Therefore single card kanban could be described as a schedule based push system of production coupled with a pull system for deliveries.

Single card kanban controls deliveries very tightly and the inbound stockpoint found at the using work centre in the dual card kanban system is not required. In fact, the using work centre often has only one or two containers of parts at a time and so confusion and clutter is avoided.

On the other hand, the outbound stock point at the producing work centre is larger than would be necessary under dual card kanban. This is because the scheduled production will be pushed into this stockpoint even if the using work centre has slowed down or stopped due to a problem. However, the amount of overproduction that can occur is limited by some kind of stop signal; for example, a maximum storage space limit, which may not be exceeded, is often painted on the floor at the outbound stockpoint. Also, the stock buildup may not be serious in companies in which it is relatively easy to associate the required quantity and timing of components with the schedule of end products. For this reason, it appeared that single card kanban might be effective at GEC Small Machines.

Unfortunately, there is one major disadvantage of the single card system when compared to the dual card system. The productivity improvement feature of removing kanban cards to expose and solve problems cannot be used without P-kanban because there is no control on the number of full containers of a given part number.