*Measuring Performance;*
*Metrics for Time, Rate, Cost, Quality, Flexibility, and the Environment*

*Tim Gutowski, July 8, 2002*

## Introduction

Manufacturing processes and systems operate in a competitive environment. We generally evaluate their competitive performance in six key areas: **time, rate, cost, quality, flexibility, and the environment**. The first four of these attributes can be measured directly, the fifth, "flexibility" can be defined but there is no universal agreement on how to measure it. The sixth, "environment" depends upon the particular issue under consideration. In this paper, we will discuss common units of measure for these attributes. Note that these measures can be made for different sized units of the manufacturing enterprise. In general the relationship between a process level attribute and a system level attribute is not straightforward. For example, even for cycle time, the cycle time for a system is not just the linear sum of the cycle times for the processes within the system. This is because processing can be done in parallel, and because in most systems, parts spend most of their time doing things other than being processed. Examples of these "other things" are; waiting (as raw materials, work in progress and/or finished goods), transport, and inspection. The most important point to keep in mind is that optimizing a <u>process</u> attribute will not, in general, optimize the <u>system</u> attribute.

Comparisons between manufacturing alternatives usually require making trade-offs. That is, often we will find that while one attribute will improve, another is degraded. For example, a common trade-off when increasing production rate by the use of hard tooling, is balancing the positive effect of the increased production rate with the negative effects of increasing the initial cost, and lead time for tooling, and decreasing the flexibility of the process. Technically, these kinds of problems are known as multivariable decision problem. They cannot be optimized without placing some valuation on each attribute. One way around this is to place constrains on most attributes (e.g. make rate greater than or equal to some target etc.) and then optimize one variable, usually cost.

## Competitive Attributes

### 1) TIME

Among the various attributes, time, "t", is one of the easiest to define and measure. Figure 1 shows three important lead times; 1) Customer Lead Time, 2)

Manufacturing Lead Time, and 3) Factory Lead Time.  In the time between an order being made and being released, all necessary raw materials and prefabricated parts need to be procured.  Upon release a part is queued into the manufacturing system, and starts production when all of the appropriate resources (machines, materials and operators) are available.  During the "Factory lead time" the part moves from station to station as specified in the process plan.  The "value added" steps in this phase including all of the processing steps(e.g. casting, machining, assembly).  Usually however, the Factory Lead Time is dominated by "non-value added" steps such as waiting in queues, transportation between stations and inspection.  A final delay is due to transportation and in many cases storage.
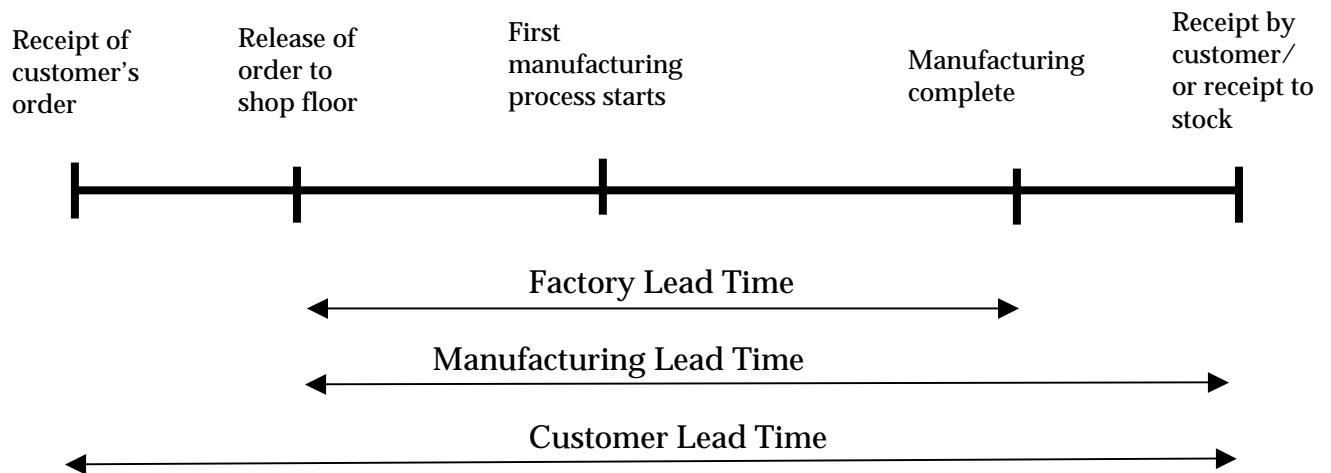


**Figure 1** Lead Times from various perspectives

Unfortunately there is no universal agreement on the different measures of time, and you may hear alternative terms or definitions.  Therefore always take care to define terms before any discussion of manufacturing time. The terms used here are in agreement with those in the APICS Dictionary [ref 1].

At the machine level we can also define a **"machine process time"** as the time to make one part. If parts are made in batches such as when machining on "tombstones" or when injection molding using multiple cavity dies, we must distinguish between the time the parts spend in the machine, say $t_{residence}$, and the average time per part. For a batch of "n" parts that would be $t_{residence}/n$. To make the first part, the total processing time at a specific machine is made up of the **set-up time** and the **process time**.  Set up time is the time required to prepare the equipment to produce a good new part. This usually involves changing tools, checking out software and making final adjustments.

We can also use these same ideas to identify the time a unit of production spends in a manufacturing cell. A manufacturing cell is a collection of manufacturing processes arranged next to each other for the purpose of making a part. (In some cases a cell may be only one process, but usually it is more than one.) Hence a "cell" is usually smaller than a factory and larger than a process.

**2) RATE**

Rate " $\lambda$ ", is the rate of material flow through the system. The "flow in" is raw or unfinished goods, while the "flow out" is the finished or semifinished product.  Rate is usually measured as "units per time". The average time in the system can be related to the average steady state rate and the average number of units in the system by a very useful expression derived by John Little of MIT. [See Ref 2].
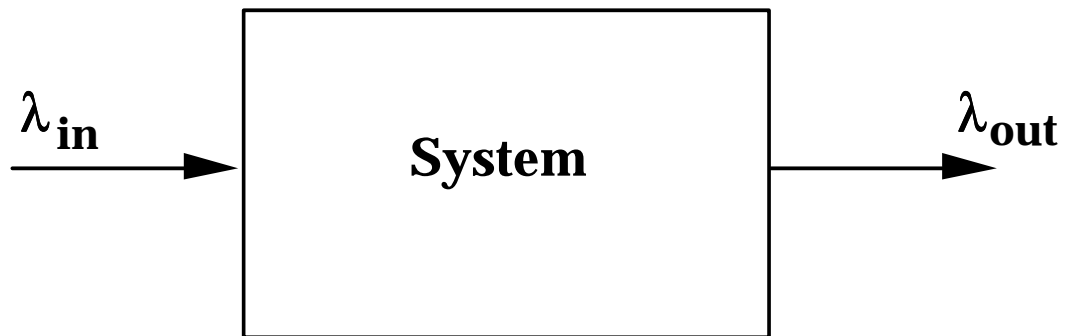
$$\lambda_{in} \longrightarrow \boxed{\textbf{System}} \longrightarrow \lambda_{out}$$

**Fig. 2 Representation of material flow through
a manufacturing system.**

Note that in steady state, the rate of material coming into the system is equal to the rate of material going out, therefore:

$$\lambda_{in} = \lambda_{out} = \lambda$$

Under these conditions, we may state Little's law as,

$$\mathbf{L} = \lambda\, \mathbf{W} \qquad\qquad\qquad \text{eq. 1}$$

where      **L**     =     units in system ( or inventory )

and       **W**    =     time in system

The above law applies to a single process, a cell, or a factory. The important point is to define what is the system and be clear about what is inside, outside, and passing through the boundaries. A simple example would be a machine operating at rate $\mu$, with material supplied at rate $\lambda < \mu$. In a deterministic system this machine would operate in a "starved" mode at a rate "$\lambda$". However, with statistical fluctuations in the processing times, a queue would build up in front of the process. The system would still operate at rate "$\lambda$", and the queue plus the part in the machine would be the inventory "L". This example is similar to the M/M/1 queue, for those of you familiar with queuing theory.

We can us eq 1 to define terms when discussing time and rate for processes and systems. For example "cycle time" for a part is sometime defined as the time spent in the system, Hence "cycle time" = "W". However you will find that some people use "cycle time" to mean the time between parts. Hence if we produce at the rate $\lambda$, then "cycle time" = $1/\lambda$. NOTE THAT THESE TWO DEFINITIONS AGREE ONLY WHEN L = 1.

If one thinks about a factory as a collection of individual processes, then matching the process rates to the factory rate is crucial for smooth flow in the factory. In those cases when the process rate is given, say controlled by the physics of the process, then the factory could be made up of many different processes with many different rates. The unequal rates between processes are then handled by adding additional machines, batching, and by using buffers that smooth the flow through the factory. For the general case then, the factory rate $\lambda_f$ , will be less that the individual process rates $\lambda_i$. That is,

$$\lambda_f \leq \lambda_i \qquad\qquad\qquad\qquad \text{eq.2}$$

The factory rate is equal to the individual process rate only when the process is the "bottleneck".

Some factories try to match the production rates with the demand rate. This is a key feature of the "Toyota Production System". The time required per part to meet demand is called the "Takt Time". And one over the Takt time could be called the demand rate. Takt time is defined, <u>for a given time period</u>, as,

**Takt Time = Time Available/ Number of Required Units**      eq. 3

In order to synchronize production with demand one should design the production rate of the factory such that,

$$1 / \lambda_f \leq \textbf{Takt Time}. \qquad\qquad \text{eq. 4}$$

Note that a factory can meet this production rate requirement by adding parallel processes and by the use of inventory (which follows directly from Little's Law). However, this approach can lead to a sluggish factory with high Work In Progress (WIP) inventory. This inventory makes it hard to respond to change and hard to detect problems.

An alternative approach directs attention to designing processes that can be synchronized in rate with the system. This is a fundamental shift in thinking suggested by the inventors of the Toyota Production System which has profound implications on the design of manufacturing processes. One way to meet the goals of eq 4 then, is to design synchronized manufacturing cells, where each cell individually satisfies eq 4.

### 3) COST

Cost "C", while measured in dollars, or some other unit of currency, is usually estimated in terms of a physical unit ( labor or machine time, weight of material, unit of energy, piece of equipment, etc.) multiplied by the market value of this unit. We will use cost estimates primarily to make decisions about the viability of alternative processing schemes, to justify new equipment expenditures, and as an aid to design for manufacture.

The best place to start a discussion on cost is at the enterprise level. Any local optimization of cost should be reviewed for other possible cost impacts elsewhere in the enterprise. A company must make a profit to stay in business. Profit can be simply defined as,

$$\textbf{Profit = Income - Expenses.} \qquad\qquad \text{eq. 5}$$

For our discussion, income comes from sales of the product, while expenses are due to manufacturing costs. Hence, our goal should always be to increase sales and to decrease manufacturing costs. Sales, of course, depend on the market, which is not well known, and generally not under the control of the company. Manufacturing costs, on the other hand, should be much more within the control

of the company. The key components of manufacturing costs are; 1) **operating expenses**-labor, capital equipment and facilities, materials, and energy and 2) **inventory**. Any strategy to reduce one of these components should be evaluated for its impact on the others. Examples of some commonly used cost reduction strategies are; a) to reduce labor (operating expenses) b) to optimize inventory to reduce set-up costs, and c) to maximize machine utilization. Ironically, the Toyota Production System, which has revolutionized automobile manufacturing does not emphasize any of these strategies.

In some cases, very useful decisions can be made only by considering direct manufacturing costs. Direct manufacturing costs are the costs that are directly related to the processing of a part. These components are usually observable on the factory floor. The common components of direct cost are; **Direct Labor** ("touch" labor hours X hourly rate for all of the required workers), **Material** ("raw" material weight, including scrap X the cost per weight, and prefabricated goods), **Tooling Cost** (this includes cutting tools, jigs, fixtures, dies etc.), and **Equipment Cost.** Other costs, which perhaps are not so readily observable, should be included if they are the direct consequence of a particular process. Overhead charges are those costs that are necessary to run the business but difficult to track down and identify with a process. In several cases, we will use direct costs as a representative measure of total cost to evaluate alternative designs and alternative processes. This type of exercise is very useful because it simplifies a complex problem. However care must always be exercised to ensure that the simple calculation is not leaving out something important.

One simple scheme to represent cost components in order to describe the effects of "economies of scale" is by; 1) Recurring or variable cost per unit "V", which when multiplied by the number of units "N" gives the total variable cost, and 2) Non-recurring or fixed costs "F", which are one time expenditures. By this scheme, the total cost C can be represented as

$$\textbf{C} = \textbf{F} + \textbf{V·N} \qquad\qquad \text{eq. 6}$$

Hence the unit cost C/N goes down as more units are produced and tends asymptotically to the variable cost per unit "V". This result is shown graphically in Fig. 3. "F" may represent tooling cost, set-up cost, and/or equipment cost, while "V" usually represents direct labor and materials. High product specific fixed costs are a major barrier to process and factory flexibility.
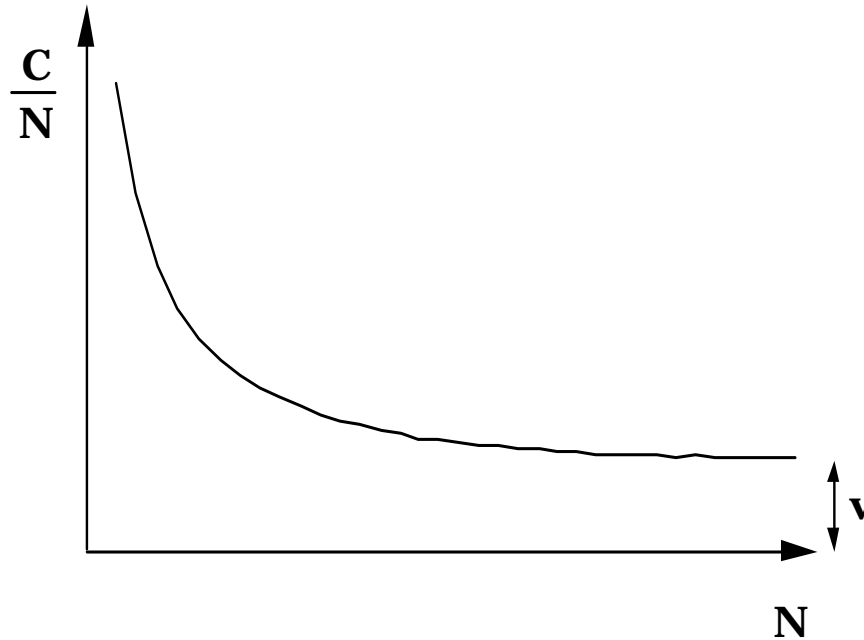
$$\frac{C}{N}$$

N

**Fig. 3  The effect of the number of units of production "N" on the cost per unit.**

### 4) QUALITY

Quality, at the process level can be measured as the ability to hit a specific target.  The target is usually given as the specification or tolerance on the engineering drawing of the part.  Tolerances can be specified for product characteristics such as dimensions, surface finish and physical properties (hardness, void content etc.).  A useful measure of the ability of a process to hit a target is the "Process Capability Index, "$C_p$". This is usually written in terms of the ratio of some measure of the size of the "target" (specification range) divided by some measure of size of the variation in the process. A good measure of the former is the Upper Specification Limit (USL) minus the Lower Specification Limit (LSL). The variation in the process can be captured by the standard deviation "$\sigma$", provided that the process mean is centered on the tolerance mid point. This situation is shown in Figure 4 below.

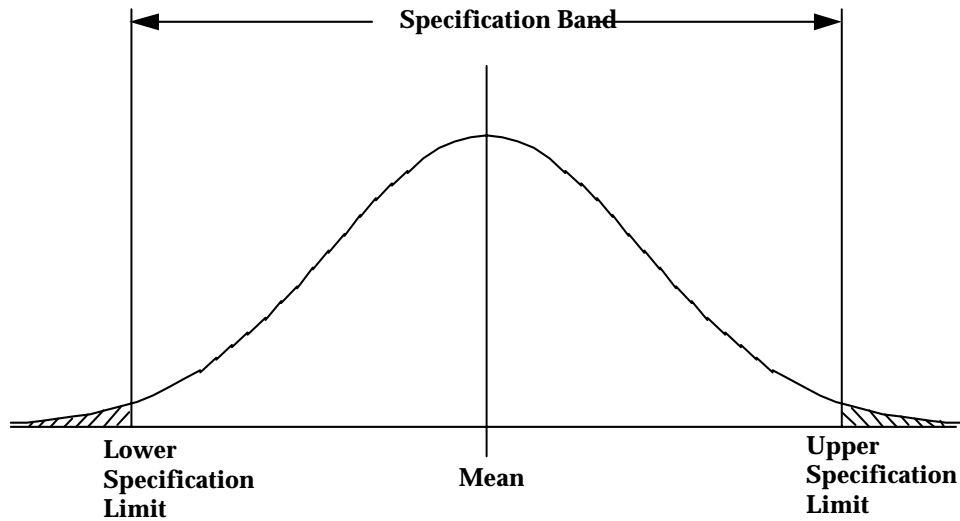$$C_p = (USL - LSL) / 6\,\sigma \qquad\qquad eq.\ 7$$

**Fig. 4   Comparison of the measured values of a part characteristic with the tolerance limits for that characteristic.**

Note that for the case of $C_p = 1$ and with a normal distribution for the measured part characteristic , the percent of out of specification parts (as indicated by the two shaded areas on the tails of the distribution) can be obtained from statistical tables as 0.27% or 2700 defects for every 1,000,000 parts, or 2700ppm. Note that when the process mean $\mu$, is not centered on the tolerance midpoint, then $C_{pk}$ is used.

$$C_{pk} = min( USL - \mu, \mu - LSL)/3\sigma \qquad eq\ 8$$

Quality at the system level is basically a measure of how well a product satisfies a customer.  This involves translating the customers wishes down through the various levels of the manufacturing organization, picking the right targets at each level and then hitting them. The "right targets" are sometimes called "key characteristics" of the product.  Identifying the key characteristics for a product is extremely important in order to focus attention on the critical few rather than on the trivial many.  The product should be designed robustly such that the key characteristics can be met within the natural variation of the process. Note that to get quality in the parts per million (ppm) range, a $C_p$ greater than 1.67 is required.

**5) FLEXIBILITY**

8

Flexibility, is the ability to adapt to change. A major challenge in designing factories is to be able to meet variable demand. Variable demand means changes in the production requirements, e.g. an increase or decrease in the number of parts required, or a change in the mix of products. These changes require a shift in manufacturing resources, either putting resources into production or taking them out. Most of these issues are addressed as higher level factory decisions involving capacity planning and human resource planning. At the process level, we will look into two different aspects of this problem. One is set up time. Shorter set up times allow you to respond more quickly to changes, and therefore shift resources more quickly into and out of production. And the second is in cell design. We will show you how to design manufacturing cells to be able to respond to both increases and decreases in product demand.

A second type of change, which is very important at the process level is change to the part design. Here we mean a change that does not represent an existing product. Changes to a part design may require changes in processing, materials and procedures. In this regard manufacturing processes are vastly different in their ability to accommodate these differences. In general, net shape processes which require hard tooling are less able to accommodate part shape changes, while sequentially subtractive processes like milling, and sequentially additive processes like stereolithography are more easily reprogrammed for the new shape.

Flexibility can be measured in terms of a penalty for implementing the required change. This penalty, in turn, can be measured using the other manufacturing attributes of time, rate, cost and quality. Hence, while a useful notion, flexibility may not be a fundamental attribute, but rather a derived attribute.

### 6) ENVIRONMENT

The environment includes many possible effects, many of these are rather straightforward to account for. They include; energy consumption, waste and toxic or hazardous emissions. Some effects are more difficult to measure such as benzene emissions during sand casting. Still others require a full systems analysis, for example when examining the life cycle impact of a product. Generally speaking we try to normalize all impacts and emissions per weight (kg) of material processed. For example, the production of one kg of aluminum requires approximately 300 MJ of energy and produces about 14.1 kg of $CO_2$. A machining process which produces 50% scrape therefore requires 600 MJ per kg of product and produces 28.2 kg of $CO_2$ per kg of product plus the energy and $CO_2$ consumed during the machining process (yielding approximately 610 MJ and 29 kg $CO_2$ per kg of product).

References

1. APICS Dictionary, 8$^{th}$ edition, American Production and Inventory Control Society, Falls Church, VA 1995.

2. L. Kleinrock, "Queuing Systems, Volume 1 Theory", Wiley 1975.