Technical Tutorial:
Optimal Level of WIP in a Production System

H. James Choo
Chief Technical Officer, Strategic Project Solutions, jchoo@spsinc.net

ABSTRACT

In the previous edition of the Journal, we featured a tutorial article on Little's Law, which is a fundamental relationship between Throughput (TH), Cycle Time (CT) and Work-In-Process (WIP). These core variables are found in all production systems, including those that are contained within capital projects. A naïve interpretation of Little's Law frequently leads those new to operations science to infer that one need only increase WIP to arbitrarily high levels in order to increase throughput to whatever target level is desired. While Little's Law is generally true under very broad assumptions, it cannot automatically be treated as if independently selecting and altering any two of the variables (Throughput, WIP and Cycle Time) will set the third variable at a desired target. Real production systems always have other physical constraints that place upper limits on Throughput and lower limits on Cycle Time. Using some simple examples, we will explain how physical constraints manifest themselves in limiting the range of feasible values that Throughput, Cycle Time and WIP can achieve. We will also discuss the important concept of Critical WIP, the minimum WIP level necessary to achieve the maximum Throughput in a production system, wherein there is no variability. We then conclude with a qualitative discussion about how variability affects system performance, as well as its effect on the optimum level of WIP needed to achieve desirable Throughput and Cycle Time performance.

Keywords: *Little's Law; Throughput; Cycle Time; Work-In-Process; Optimal Level of WIP; Critical WIP*

INTRODUCTION

The previous article in this Technical Tutorial series [1] discussed Little's Law, relating the concepts of Throughput (TH), Cycle Time (CT) and Work-In-Process (WIP). In this Technical Tutorial, we will explore these relationships further.

The Project Production Institute regularly conducts introductory seminars to explain the concepts underlying Project Production Management (PPM), including the fundamental principles of operations science, as articulated by Institute partner Factory Physics. Occasionally, the Institute co-hosts seminars with Factory Physics, and it is eye-opening to see how wide the variation in understanding and intuition is regarding predicting the behavior of a production system; the results are frequently varied and incorrect.

PPI commonly conducts an exercise during these and other public seminars, providing participants with blank graphs (Figure 1 below) and asking them to draw the relationship between Work-In-Process (WIP) and Throughput on one graph, and the relationship between Work-In-Process (WIP) and Cycle Time on the other.

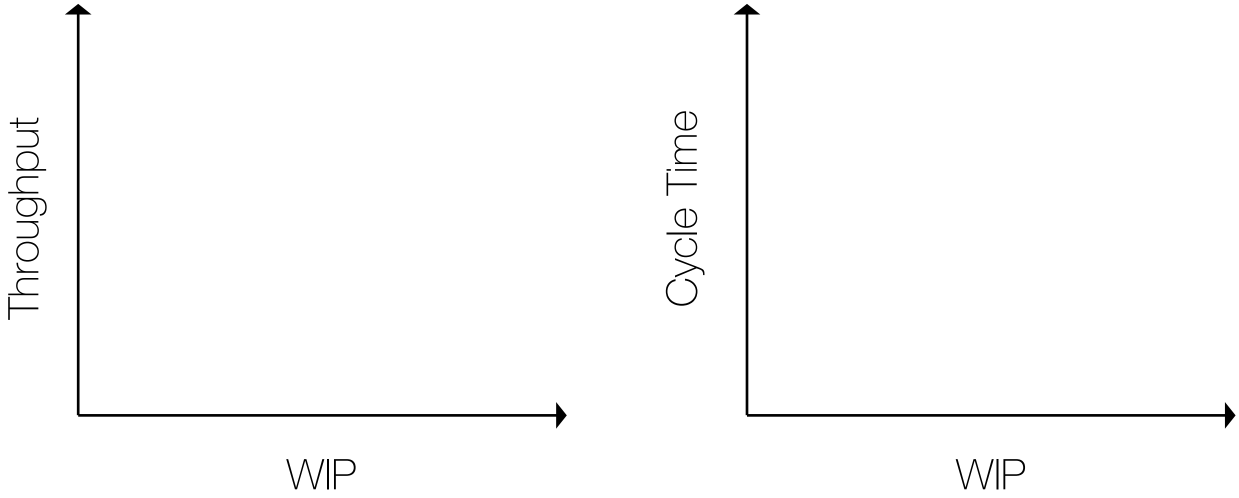

Figure 1: What are the relationships among Throughput, Cycle Time and WIP?

Surprisingly, we receive all sorts of answers, some showing relationships that are the opposite of correct relationships. This variation in understanding and intuition helps explain why we often see people working on the wrong problems, or trying to implement the wrong solutions in their attempts to solve the right problems.

It is therefore critical to fully understand these relationships, which leads us to the purpose of this technical tutorial.

A SIMPLE EXAMPLE

Using a simple example, we will examine the following relationships using Little's Law [1-3]:

- Relationship between Throughout (TH) and Work-In-Process (WIP)

- Relationship between Cycle Time (CT) and Work-In-Process (WIP)

In Institute seminars, participants sometimes suggest that Little's Law implies that Throughput can be made arbitrarily large, simply by increasing the Work-In-Process:

$$TH = WIP/CT$$

This is sometimes used as justification for the common practice in project management of accumulating large amounts of inventory in a project with the belief that "having all the materials available throughout the production system ensures optimum project execution with minimum schedule." Multiple project performance surveys demonstrate that such a belief is not actually supported by the project outcomes achieved. Little's Law does indeed constrain the three variables (TH, WIP and CT) under very general conditions. However, in any real production system, TH and CT usually have other physical constraints that place upper limits on the values TH can have, and lower limits on the values CT can have.

The example shown in Figure 2 below illustrates why this must be true. Note: this example is based on earlier material (the Penny Fab) produced by our partners at Factory Physics, albeit slightly modified, [1, pp 232 – 255].

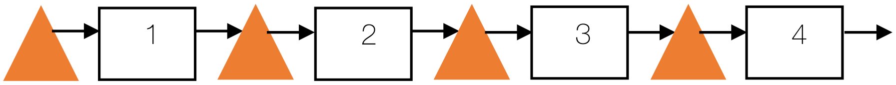

Figure 2: Example Production System with 4 distinct operations

Figure 2 illustrates a generic production system. For the sake of concreteness, let's suppose this is a description of the steps involved in drilling a well and tying it into pipeline infrastructure to produce oil and gas. The high-level sequence of steps is:

1. Secure permits to drill and complete the well

2. Set rig and associated drilling equipment in place and drill the well

3. Move out drilling equipment, set completion equipment and assets in place, and complete the well

4. Connect producing well to pipeline infrastructure

For simplicity, we will assume that each operation cannot handle any more than one well at a time, i.e. the team working on securing permits can only work on the permits for one well, and cannot start on the permits for the next well until the permits for the preceding well have been secured. Similarly, the drilling rig can only be drilling one well at a time, and cannot start drilling the next well until the drilling of the preceding well is completed. The situation is the same for the teams doing the work of well completion and tying the producing well to the pipeline – each team can only work on one well at a time. In other words, we are specifying the capacity of the resources in each step, one well at a time.

Also, for the sake of simplicity, we'll assume there's absolutely no variability in the system. Each step takes a predictably constant processing time and there are no delays anywhere in the system. Let's also suppose each operation takes 10 days, i.e. each individual step has its own cycle time of 10 days. Then, assuming there is no waiting time in between any of the steps, the total Cycle Time (CT) is 40 days, which represents the total time a well resides within the entire process from start (securing the permits) to finish (producing well is connected to pipeline). Each step therefore has its own individual throughput of processing one well every 10 days or 0.1 wells/ per day. And if a well is handed off from one step to the succeeding step without any delay, then it's easy to see that the Throughput of the entire system is also 0.1 wells/per day, or 1 well every 10 days. That is, after initial startup, with the very first well taking 40 days to appear, the subsequent wells will come online at the rate of 1 every 10 days. The Throughput (TH) of the entire production system is therefore 0.1 wells/ day.

It now follows from Little's Law that the total WIP in the system is TH*CT = 0.1 * 40 = 4 wells. It is not surprising, as it's quite clear from our initial assumptions, that there is a well being worked on at each of the 4 steps in the well production system in Figure 2.

Can one increase the overall Throughput (TH) by increasing WIP? Any production system is simply the aggregation of individual operations, perhaps arranged sequentially in a *line* or *routing*, or perhaps a parallel arrangement of multiple lines, or perhaps an even more complex network. Looking at any of the individual steps involved shows the reader why this can't be the case.

For any individual operation, it is easy to see that there will generally be an upper limit on the throughput of that individual step and a lower limit on its individual cycle time. For instance, in our example, one might conceivably get the team in Step 1 to be simultaneously securing the permits for more than one well at a time. One might also be able to expedite the processing of permits to reduce the overall time needed to get a permit. However, one will never be able to get the cycle time to zero, nor will one be able to get the number of wells being permitted to be arbitrarily high. There will be some physical or procedural limit – perhaps the state will not permit more than 50 wells at any one time because of hard constraints on their capacity or perhaps no permit can be executed in faster than 24 hours. Similarly, there will always be a finite upper limit of one on the number of wells that can be drilled simultaneously, until someone develops a rig that can drill multiple wells simultaneously. And it is also true that because there is a lower limit on the cycle time for drilling, it too won't come arbitrarily close to zero.

Within the overall production system, there will always be an individual step or steps limiting the overall throughput. This is the *bottleneck*, constraining overall system performance. Increasing WIP arbitrarily will not overcome the limiting of throughput imposed by the bottleneck.

Let's examine our example production system as it starts up, to generate the graphs in Figure 1 for our example production system in Figure 2. The unit of WIP in this case is a "well," which must undergo 4 steps – permitting, drilling, completion and tie up – to pass through the production system. We can see that as the production system with wells is being put into the production system at the beginning, the values for WIP, TH and CT evolve, as shown in Table 1 below.

WIP	TH	CT	TH*CT
Wells	Wells Per Day	Days	Wells
1	0.025	40	1
2	0.050	40	2
3	0.075	40	3
4	0.1	40	4
5	0.1	50	5
6	0.1	60	6

Table 1: Values of Throughput TH, WIP and Cycle Time CT for the system in Figure 2

We can now produce the graphs in Figure 1 from the tabulated data in Table 1 to produce Figure 3:

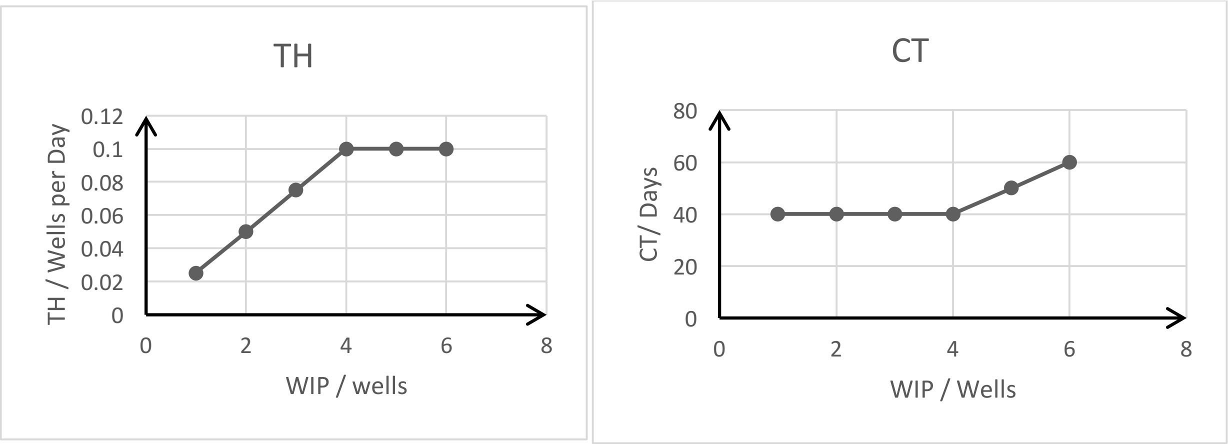

Figure 3: Plots of Relationships among TH, CT and WIP

As seen in Figure 3, TH grows as WIP grows until it reaches an upper limit. This upper limit is known as the Critical WIP [2, pp 231 – 235], defined as the WIP level that achieves maximum throughput (bottleneck rate, or BNR [2, pp 231 – 235]) with minimal cycle time (raw process time, or RPT [2, pp 231 – 235]) in a process flow or production system with no variability. For our very simple example, BNR is 0.1 wells per day and RPT is 40 days. CWIP can be calculated using Little's law: CWIP = BNR x RPT.

What is interesting is that TH does not grow beyond BNR, even if WIP increases. Any WIP above the CWIP level does not result in additional throughput, but rather increases CT due to the need for additional wait time, as is obvious in our example production system in Figure 2. As soon as there are 4 wells being worked on in the production system, any additional wells must wait for permitting, until permitting finishes processing the well they are currently working on. Any WIP level below CWIP, however, results in lost throughput. Although every production system suffers from variability, CWIP provides a critical benchmark for evaluating current production system performance.

Our example system in Figure 2 is very simple – there is no variability and the numbers have been chosen for ease of illustration. But despite these gross simplifications, it still illustrates some very general and foundational principles. The system in Figure 2 is also sufficiently generic – the tasks could be the activities in a variety of different projects. For instance, the operations labeled 1 through 4 in Figure 2 could be engineering, fabrication, delivery and site construction – the basic elements of a construction project. And while the sample values of throughput and cycle time for the individual steps were artificially kept the same for simplicity of calculation, they need not be. One could generate the equivalent of Table 1 and Figure 3 with a different set of numbers, albeit with a bit more arithmetic. The system, even one which is more complex than the sequence of 4 steps we considered, would exhibit the same general characteristics: there would be a bottleneck rate, there would be a critical WIP level and there would be a raw processing time, all deterministically obtained in the absence of variability.

Even in the absence of variability, our example illustrates some fundamental considerations in production system design and management in projects. Greater WIP than necessary results in more capital tied up without return and a decreased ability to take advantage of technological advancements or regulatory changes. Therefore, a system will always perform better if the desired throughput can be achieved using smaller WIP and smaller CT.

THE IMPACT OF VARIABILITY ON PRODUCTION SYSTEM PERFORMANCE

Unfortunately, the type of performance illustrated in Figure 3 is unattainable in the real world due to variability in the flow of work (how often work is provided to a production unit – a single step or operation in the production system) and variability in the process (how long it takes to complete / process a unit of work). Much of the variability that we deal with each day can be categorized by whether it affects the flow or the process, or both.

What is the effect of variability on Figure 3? We're not quite ready to be that quantitative, but we can have a qualitative discussion, referring to some other examples illustrated in Figure 4 below that show the results of simulating a production system with and without variability.

The relationship between WIP and CT is shown below. When there is no variability (best case performance), CT is equal to Raw Process Time (RPT) until it reaches Critical WIP (CWIP), and then it increases as WIP increases. When there is variability (marginal case performance), CT increases as WIP increases. The maximum TH deteriorates as variability increases. Therefore, as WIP increases, the increase in TH becomes smaller and smaller (marginal case performance). The detrimental effects of increased WIP can be further understood by examining the relationship between WIP and CT.

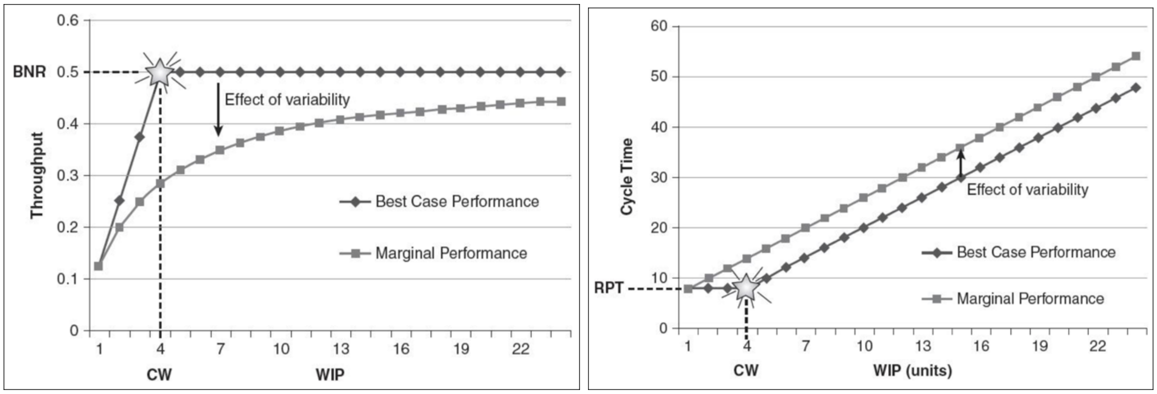

Figure 4: The Impact of Variability
(Source: Factory Physics for Managers)

If you were to put the relationships between WIP, CT and TH together, it is very easy to see that as WIP increases more and more beyond CWIP, the increase in TH becomes minimal but CT increases tremendously. This can be translated to more capital tied up and more delays in the delivery of projects. Understanding these relationships allows us to identify the need to reduce variability and choose carefully how much WIP to carry in any production system.

We have discussed how variability degrades production system performance, causing longer cycle times and resulting in either less TH or higher WIP. In the next tutorial article for this series, we will take a closer look at the quantitative impact of variability on cycle time, throughput and utilization.

REFERENCES

1. H. J. Choo, "Little's Law: A Practical Approach to Understanding Production System Performance", *Journal of Project Production Management*, Vol 1., 2016, pp. 61 – 65.

2. W. Hopp, M. Spearman, *Factory Physics,* 3rd ed. Waveland Press Inc., 2008.

3. J. Little, "Little's Law as viewed on its 50th anniversary", *Operations Research*, Vol. 59, No. 3, May – June 2011, pp. 536 – 549.

4. E. Pound, J. Bell and M. Spearman. *Factory Physics for Managers: How Leaders Improve Performance in a Post-Lean Six Sigma World*. McGraw-Hill Education, 2014.

ABOUT THE AUTHOR

H. J. James Choo is Chief Technical Officer of Strategic Project Solutions, Inc. and has been leading research and development of SPS knowledge, processes and systems to support implementation of SPS offerings since 2001.

Since joining SPS, Choo has worked with various organizations in oil & gas, heavy industrial, civil infrastructure, aerospace & defense and other industries around the world implementing project production management solutions in support of delivering desired project objectives.

Prior to joining SPS, his experience included roles as a construction site engineer, research associate at research institutes, teaching assistant at universities and software developer. He has also been developing computer systems for implementation of Lean Construction since 1997 during his PhD studies at UC Berkeley.

Choo earned a Master's Degree in Civil Engineering from Yonsei University, Korea, and a PhD in Construction Engineering & Management from the Civil and Environmental Engineering Department of University of California at Berkeley.