Connecting Assembly with Batch Processes Via Basic Pull Systems

By Art Smalley

Pull systems are an integral part of lean manufacturing yet they are frequently misunderstood and considered hard to implement. Specifically, one area of repeated struggle I find is the proper connection of assembly processes with upstream batch processes such as stamping, injection molding, paint, or a machining operation, for example.

There are three basic types of pull system; replenishment pull, sequential pull, and mixed pull system with elements of the previous two combined (see glossary at end of story). In all three cases the important technical elements for systems to succeed are:

1. Flowing product in small batches (approaching one piece flow where possible)
2. Pacing the processes to takt time (to stop overproduction)
3. Signaling replenishment via a kanban signal
4. Leveling of product mix and quantity over time

In order to connect a batch processes to assembly, a replenishment pull system used in conjunction with a specific type of kanban known as signal kanban is used. There are at least three different types of signal kanban and the proper usage depends upon the nature of your process and manufacturing situation. The three types of signal kanban are:

1. Pattern production
2. Lot making with a batch board
3. Triangle kanban

Pattern production establishes a fixed order of production and some basic rules around lot sizes, which can vary somewhat as needed. However, the fundamental pattern or order of production is always maintained. This creates a fixed sequence (i.e. pattern) of production that is continually repeated. Often this style of production is needed in heat treat processes, paint, or other processes with specific changeover sequences. While the sequence of product may be fixed in these cases, the actual amount produced each time in the cycle is unfixed and varies according to customer needs. This is known as a fixed order/unfixed quantity replenishment cycle.

For example, in an eight-hour cycle, part numbers always are run A through F. (The difficulty of your changeovers may dictate this order.) Inventory in the central market is a function of the length of the pattern-replenishment cycle. Thus, a one-day cycle implies a minimum of one day of inventory must be kept in the market; a one-week cycle means one week of inventory. The main disadvantage of pattern production is that the sequence is fixed; you can't jump from making part C to part F.

Pattern production establishes a basic rhythm and sequence of production. As long as the overall production pace is kept close to final assembly takt time and the market governs pull between the two processes, the system works. It's a very basic, first step toward establishing a production control and pull system between two processes separated by a market.

The pros are that it creates a stable order where one did not exist before. It also allows the best sequencing of product to ease changeovers, if they are an issue. The cons are that it is not a perfect just-in-time system as you have to wait until your turn in the order queue comes around and thus some delay is involved. Unfortunately, it also does not tend to place pressure on the system to reduce changeover time, inventory time, or lead time (except by management intervention).

Batch Board

Lot making with a batch board requires creating a physical kanban for every container of parts in the system. As material is consumed from the market, the kanban are immediately detached and periodically brought back to the upstream batch-producing process and displayed on a board that highlights all part numbers and displays an outlined shadow space for each of the kanban cards in the system. This is an example of a fixed quantity/ unfixed order style of replenishment pull.

A returned kanban card placed on the board in the shadow space indicates inventory has been consumed in the market; unreturned cards represent inventory still in the market. As predefined trigger points are reached, the production operator knows to begin making product to replenish the material in the market. If several trigger points are reached at once, then production needs to establish some rules regarding what to produce first (since in this scenario the pattern or sequence is not fixed) and where to produce it. Otherwise, supervisors will decide what to make next based upon the available information.

There are several advantages of this approach. One is that the schedule sequence is more flexible than with pattern production. Second, it provides a visual representation of inventory consumption and highlights emerging problems in the central market.

There are disadvantages to this type of kanban as well. It requires many kanban cards if you have many part numbers, and the cards must be brought back in a timely and reliable manner for the batch board to be accurate. Furthermore there is the inherent problem of multiple part numbers reaching a trigger point as the same time. In these instances human judgment is required to decided what to make next between two or more items that have been triggered. Discipline is required on the part of schedulers and supervisors not to build inventory in advance of when it is needed.

Triangle Kanban

The method that Toyota facilities would follow most of the time in connecting a batch process to a downstream assembly process is called the triangle kanban. Below is an example in rough detail of how to evaluate to use this method for implementation (See LEI Workbook *Creating Level Pull* for a more detailed explanation). The example assumes you have a batch machining department feeding some type of final assembly. The machining department has five machines building 11 totally different product variations. The demand for each product varies significantly, as does the standard pack quantity of each part. The setup time for changeover is about three hours.

1. Take the 11 part numbers and dedicate them to the five machines. (Ideally for example four machines will have two part numbers to run and one will have three part numbers to run, but this depends upon volume and mix, etc.)
2. Now let us take one machine which now has several dedicated part numbers assigned to run on it.
3. Create a single triangle kanban for each part number including the following information (see picture below).
    a. Part number
    b. Part description
    c. Inventory location
    d. Machine to be run on
    e. Space for date triggered to be written
    f. Tool number
    g. Lot size*
    h. Trigger or reorder point*

    *We'll calculate these below

4. Determine the required daily run time for each part number. Calculate your average daily demand for each part number. If you haven't leveled the build in final assembly it's probably wise to add some amount in to cover demand variation. Let's assume that you have two shifts of eight hours available production time which equals 16 hours of production (simplest case). If you have two part numbers on a machine, you can derive required production time based upon average daily demand and the cycle time to run the

part number. For example, let's say, for ease of calculation, that required production time adds up to 10 hours.

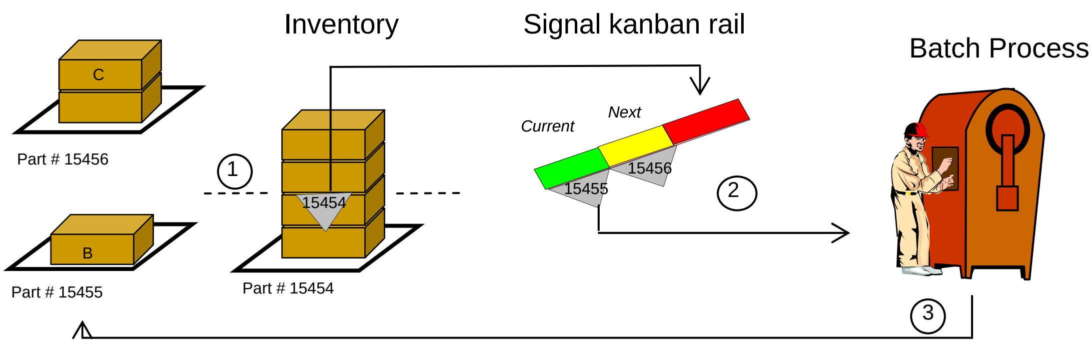
5. Determine the time available for changeovers. In the simple case of 16 hours of production and 10 hours of demand it leaves six hours for changeovers each day. You should incorporate any average down time and scrap you have in the short run – but eliminate this in the long run. Thus, in our simple case (ignoring downtime, etc.) you have available time for two changeovers daily or roughly one per shift under current calculations.
6. Determine your lot size for each part number. There are different ways to do this but for simplicity I will only describe the easiest. In this case your lot size is simply set at one day of production since you probably have two part numbers on this machine and will average two changeover events per day. In essence you are making every part ever day. If you had 10 part numbers on the machine and only two changeovers per day you would have lot sizes of five days.
7. Determine your trigger point for replenishment. To do this, add up the run time for the longer of the other components, add the changeover time, and add the time to make the first container and get it reliably back into the market. This is the minimum level you can establish for a replenishment trigger point for product. (An average trigger point might be 300 pieces.)

The system works by hanging a single triangle kanban at the trigger point in the inventory location. When the trigger point is reached, material handling takes it back to the producing machine. The kanban is hung on a rail at the machine and dictates what to build next and the lot size.

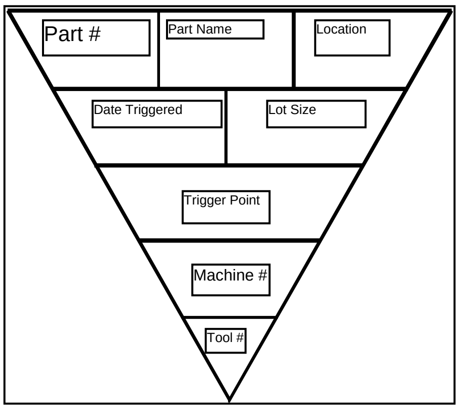
The triangle pull system has many advantages. It is virtually self-running once established as long as average demand does not change. If it does, change your lot sizes accordingly. Also, the triangle helps stabilize quality since product is dedicated to a machine, and it takes out some variability but you can choose to run product on other machines.

The down side is that the inventory in the market is not visible at all times since there is no batch board with cards. You do see, however, how many kanban are hanging on a rail at the machine which is a good indication of inventory depletion and a tool for visual control. Furthermore, there is only one kanban per part number to manage. Other minor challenges include, of course, figuring out the die maintenance schedule (for some types of machine) and what impact this will have on the schedule. Also you must figure out a simple signal to reliably bring any needed raw material to the machine.

Sample Work Flow Using Triangle Kanban for Scheduling



Triangle Kanban Detail



In the end, you must decide what is the main goal is of implementing the pull system. Each of the three described ways works in terms of scheduling a batch process in conjunction with a market but each has a slightly different emphasis in mind. It might be worth the time to calculate the inventory levels, run times, operating rules for each of the three different ways. Construct a simple matrix, evaluate each of them in accordance with your priorities, and select the one that best fits your needs and company ability.

About the Author

Art Smalley is the author of *Creating Level Pull*, a workbook on how to implement a level, pull-based production control system, published by the Lean Enterprise Institute (www.lean.org). While at Toyota he was instrumental in the transfer of equipment and lean manufacturing methods to overseas plants. He subsequently was director of lean production operations at Donnelly Corp., (now part of Magna Inc.), and a leading lean expert in manufacturing at McKinsey & Company. He is a member of the Shingo Prize Hall of Fame.

Glossary (adapted from the *Lean Lexicon*)

There are three basic types of pull production systems:

Supermarket Pull System
The most basic and widespread type, also known as a fill-up or replenishment pull system. In a supermarket pull system each process has a store -- a supermarket -- that holds an amount of each product it produces. Each process simply produces to replenish what is withdrawn from its supermarket. The disadvantage of a supermarket system is that a process must carry an inventory of all part numbers it produces, which may not be feasible if the number of part numbers is large.

Sequential Pull System
A sequential pull system may be used when there are too many part numbers to hold inventory of each in a supermarket. Products are essentially "made-to-order" while overall system inventory is minimized. In a sequential system, the scheduling department must set the right mix and quantity of products to be produced. This can be done by placing production kanban cards in a heijunka box, often at the beginning of each shift. A sequential system requires strong management to maintain, and improving it may be a challenge on the shop floor.

Mixed Supermarket and Sequential Pull System
Supermarket and sequential pull systems may be used together in a mixed system. A mixed system may be appropriate when an 80/20 rule applies, with a small percentage of part numbers (perhaps 20%) accounting for the majority (perhaps 80%) of daily production volume. Often an analysis is performed to segment part numbers by volume into (A) high, (B) medium, (C) low, and (D) infrequent orders. Type D may represent special order or service parts. To handle these low-running items, a special type D kanban may be created to represent not a specific part number but rather an amount of capacity. The sequence of production for the type D products is then determined by the method the scheduling department uses for sequential pull system part numbers.