THOMSON REUTERS
CHECKPOINT™

November/December 2019

# Cost Management

## A Well-oiled Machine

Streamlining Business Practices

Avoiding Risk With Operating Leverage

The Benefits of Work Sharing

How On-site Clinics Can Reduce Costs

THOMSON REUTERS®

# Cost Management




## Editorial Staff

**Executive Editor**
Susan Weisenfeld
susan.weisenfeld@thomsonreuters.com

**Senior Director, Financial Reporting and Management Publications**
Anne-Lise Dorry

**Editor**
Libbey Clayton
(919) 240-5000
l_clayton@technicaeditorial.com

**Desktop Artist**
Vijay Jagdeo

**Advertising Sales**
Thomson Reuters
Terry Storholm
Advertising Sales Director
610 Opperman Drive, D5-S896
Eagan, MN 55123
Phone: (651) 687-7327
Fax: (651) 687-7374
terry.storholm@thomsonreuters.com

**Reprints**
Yurchak Printing, Inc.
(888) 591-9412

**Subscriptions**
1-800-950-1216, ext. 1

**Please Visit Our Website**
tax.thomsonreuters.com/CheckpointJournals

COST MANAGEMENT (ISSN 1092-8057) is published bimonthly by Thomson Reuters/Tax & Accounting, 121 River Street, Hoboken, NJ 07030. Subscriptions: $455.00 a year. For subscription information, call 1 (800) 950-1216; for customer service call 1 (800) 431-9025. We encourage readers to offer comments or suggestions to improve the usefulness of future issues. Contact Libbey Clayton, Editor, (919) 240-5395, l_clayton@technicaeditorial.com

Editorial Offices: Thomson Reuters, 121 River Street, Hoboken, NJ 07030. Web: tax.thomsonreuters.com/CheckpointJournals





Postmaster: Send address changes to COST MANAGEMENT, Thomson Reuters Tax & Accounting, P.O. Box 115008, Carrollton, TX 75011-5008.

THOMSON REUTERS®

# SOME APPLIED MATH FOR MAPPING A LEANER VALUE STREAM

**Value stream analysis helps engage, inspire, and drive individual and organizational learning.**

MARK R. HAMEL AND MICHAEL O 'CONNOR

Prior to the seminal book on value stream mapping (VSM), Mike Rother and John Shook's *Learning to See*, disconnected serial *kaizen* events were the continuous improvement vehicle of choice.[1] Sensei and sensei wannabees alike divined where to next unleash good change based upon a combination of direct observation and their own version of lean "spidey senses." Yet this combination was insufficient in terms of easy transferability, sharing, and organizational learning, because it was often noncollaborative and it did not reliably facilitate true system or flow *kaizen*.

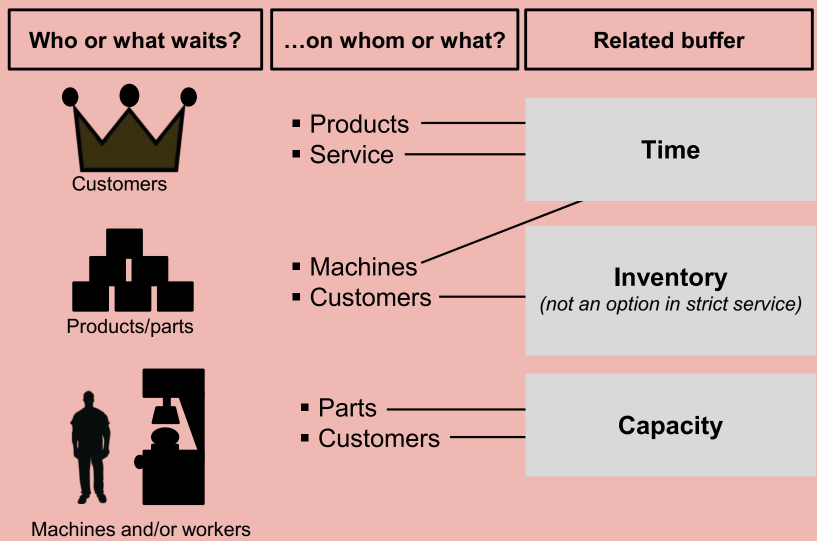
VSM provided context for lean principles, systems, and tools, and gave practitioners the ability to apply them, albeit first on paper, in a qualitative and quantitative manner. This article is primarily about the latter. In other words, we are going to very briefly explore some of the need-to-know principles and formulas to design a leaner value stream. Why? Because bad or missing math can render value stream analysis less than effective.

Before we dive into some of the expected math such as lead time, processing time, and process cycle efficiency (PCE), let us first think about the system from a vantage point that many lean practitioners are rarely exposed to, unless they are secretly risking their lean souls by cavorting with enterprise resource planning and purveyors of *Factory Physics*.[2] Lean machismo often dictates that

MARK R. HAMEL *is partner and COO at The Murli Group, and he is a two-time Shingo award–winning author, speaker, and Lean Enterprise Institute faculty member. In his 19-year pre-consulting career, he held executive and senior positions within operations, strategic planning, business development, and finance. Mark's lean education and experience began in the early 1990s when he conceptualized and helped launch what resulted in a Shingo Prize–winning effort at the Ensign-Bickford Company. He is a CPA in the state of Connecticut and is dual APICS certified in production and inventory management and integrated resource management. Mark was a national Shingo Prize examiner for eight years, has helped develop exam questions for the Association for Manufacturing Excellence/SME/Shingo Lean Certification, and is Juran-certified as a Six Sigma Black Belt. He authored the* Kaizen Event Fieldbook: Foundation, Framework, and Standard Work for Effective Events *and co-authored* Lean Math: Figuring to Improve. *Both SME-published books won a Shingo Professional Publication Award in 2010 and 2017, respectively. Mark can be reached at* mark.hamel@themurligroup.com.

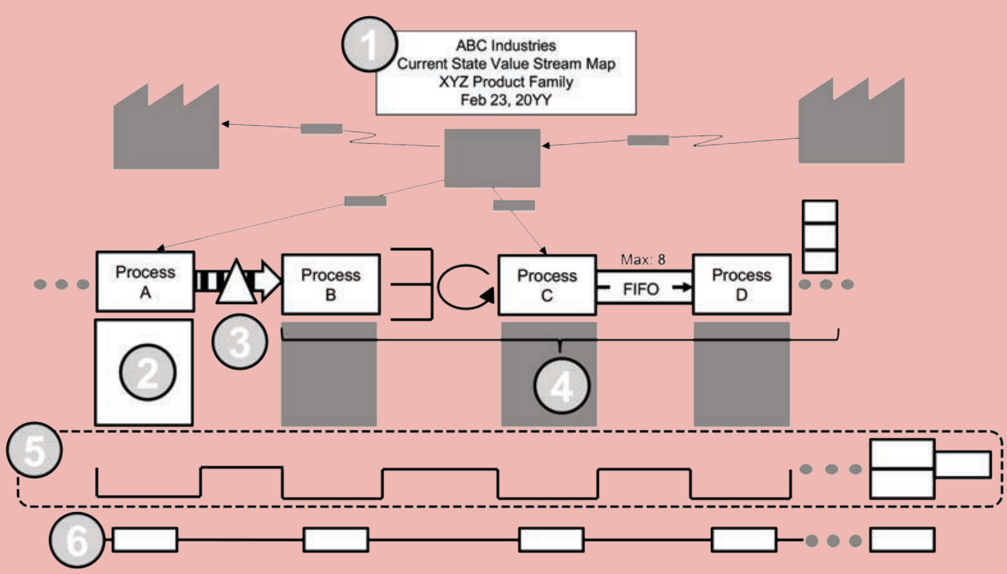
MICHAEL O'CONNOR, Ph.D, *is a learner, teacher, and consultant who helps organizations achieve their goals through the maximal application of their resources. He holds several degrees, including a Bachelor of Science in electrical engineering, a Bachelor of Science in physics, a Master of Science in physics, and a Ph.D. in physics; his enterprise excellence work has been recognized by several organizations. Dr. O'Connor is a Shingo award–winning author, collaborating with Mark Hamel on* Lean Math: Figuring to Improve. *He has also received International Quality and Productivity Center's Master Black Belt of the Year award. He can be reached at* dr.mike@leanmath.com.

**EXHIBIT 1** Buffer Reality




As demand (volume and/or mix) and transformation performance (production or service timeliness) variation increases, one or more buffers must necessarily increase.

**EXHIBIT 2** Six VSM Math Regions




inventory, waiting, and/or excess capacity (see Exhibit 1) is for batch-and-queue newbies.

While we understand that a nonexistent or very low buffer situation is reflective of a longer-term target condition, it is not always realistic in the short term, given the challenges of waste, unevenness, and overburden. As such, the lean practitioner must make purposeful trade-offs to satisfy the customer, while tolerating the least amount of waste possible.

**VSM IS MERELY AN IMPROVEMENT VEHICLE AND NOT THE DESTINATION, SO THE LEAN PRACTITIONER MUST BE THOUGHTFUL ABOUT WHERE AND HOW IT IS DEPLOYED.**

Intuitively, most folks grasp the need for buffers, but what kind of math justifies that intuition? As we know, inventory is the coin of the realm in VSM, by which Little's Law provides a mathematical model of queue time, namely: lead time = work in process ÷ throughput rate, where lead time is the sum of system queue time and processing time. Little's Law assumes a closed queuing network — that is, one in which work in process (WIP) is controlled. As we consider Little's Law and the Kingman equation, which models an open queuing network and where lead time can theoretically grow without bounds, there are a few key points that should make the lean practitioner think about in what, how much, and where they need to pragmatically make improvements within the target value stream:

- Ideal state, meaning zero queue time, is achievable only when there is zero variation in demand and processing time and load never exceeds capacity. This is rumored to happen only in *hei-junka* heaven.
- Zero WIP means zero throughput.
- Optimal standard WIP (and the answer is not always strict one-piece flow) for a given system will yield optimal lead time performance.
- Queues and lead times approach the ideal state when there is significant excess capacity. As such, when utilization of resources approaches 100 percent, lead time grows exponentially.
- The best-case throughput is the bottleneck resource throughput ($R_b$), but $R_b$ will be markedly less than 100 percent of theoretical — more like 80 percent, due to the impact of process variation.

## Why and what are we mapping?

Value stream mappers must, as Taiichi Ohno was credited with saying, "start from need." In other words, VSM is merely an improvement vehicle and not the destination, so the lean practitioner must be thoughtful about where and how it is deployed. Its application should be pulled by the business performance and development needs of the organization.

This should be informed by the company's True North, strategy deployment breakthrough objectives, and the like. VSM entails four, often iterative, fundamental steps:

1. identify and define the very product or service family to map;
2. understand and segment customer demand;
3. characterize the measurable target condition for that value stream (essentially the design parameters for the future state as of a certain date); and
4. formalize the scope that will be treated within the value stream (in other words, determine how far upstream and/or downstream is desirable and appropriate to satisfy the target condition).

Exhibit 2 reflects a simplified and generic map identifying six math regions, the first of which pertains to the four steps. We will ultimately explore all six regions.

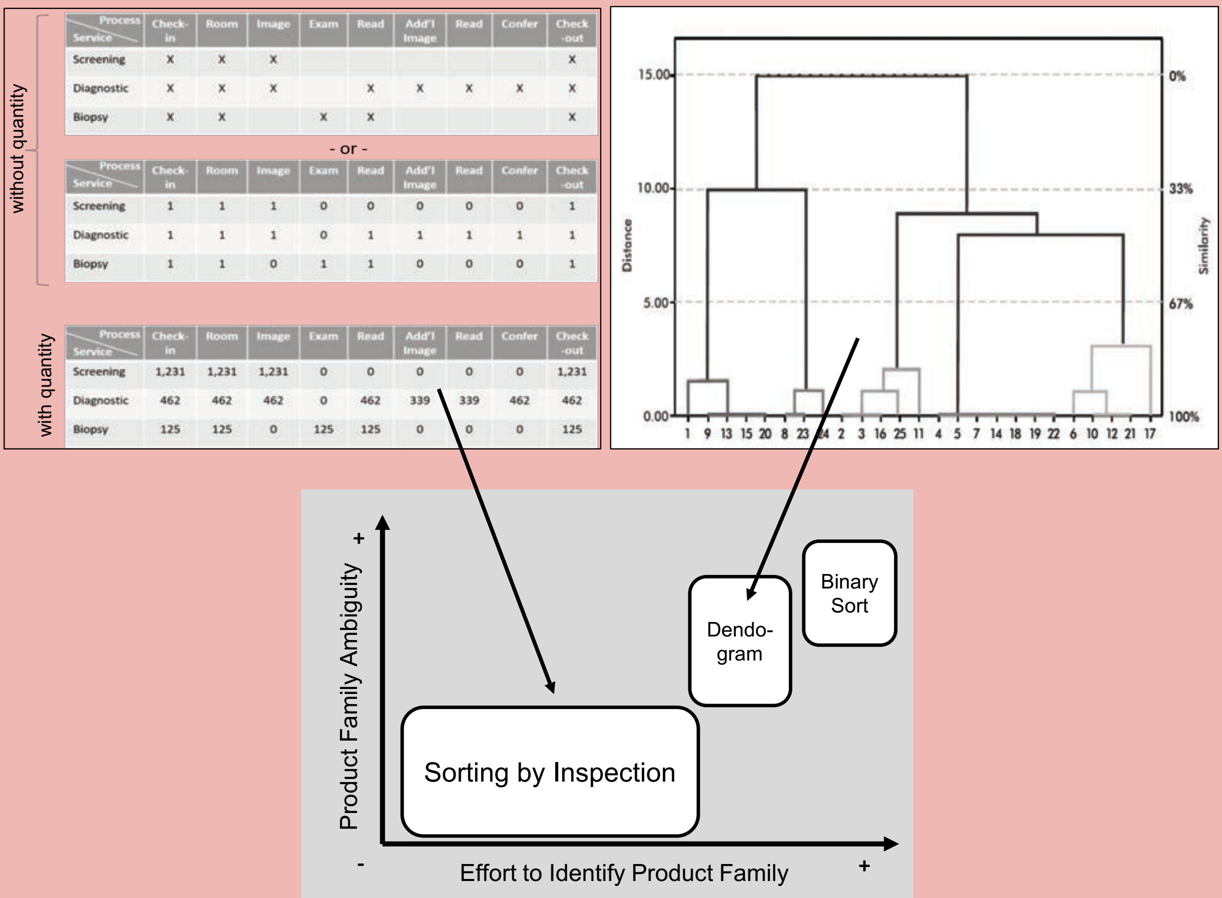
## Identify and define product families (region 1)

A diligent and perceptive practitioner can walk the *gemba* and observe activity (and inactivity), some of which is value-added (but most of which is not), along with the people, inventory, equipment, machines, artifacts, and the like. Often the product families are not readily discernable, even to those who are intimate with the business. Let's be frank: Most folks have never *had* to think in terms of families until they were confronted with VSM. When prompted, they may characterize a family in terms of markets, customers, and product lines — or, even worse, the orientation within a functional silo.

The product or service family analysis (i.e., product quantity process analysis) matrix is a formal way to characterize the intersection of product or service with processes. While not totally unambiguous, the characterization will help the practitioner discern the product family (or families) and converge on the VSM scope. Math can help this exercise. Note that work content variation can be a useful product family discriminator as well, and it also requires some lean math.

As reflected in Exhibit 3, there are options with varying levels of product family analysis (PFA) rigor. Most practitioners are familiar with the preparation of a PFA matrix, with or without demand volumes, and then apply-

**EXHIBIT 3** PFA Method Options




ing some good old-fashioned visual inspection to tease out the product families. However, occasionally the families are not easily identified because of ambiguity or an overwhelming amount of data. Here a dendogram may be an appropriate tool. So, what is a dendogram?

Dendograms provide visual grouping of products (or services) that go through identical or similar process steps. The dendogram reflected in the upper righthand portion of Exhibit 3 is for a relatively simple array of products and processes. As the complexity of the PFA matrix increases, the power of this approach becomes more and more evident.

Essentially, identifying product families is a clustering problem. Clustering by inspection works well for small PFA matrices, and dendograms work well for larger matrices, but if neither is available nor practical, binary sorting can help. In this approach, each product is assigned a binary number based on which process steps that product goes through. The binary numbers are then sorted and like ones are clustered into product family groups.

While PFA can provide the user with insight into demand, it is often prudent to also engage in some demand segmentation analysis. This will characterize historical demand, demand variation, and even inventory levels, which can be analyzed at the stock keeping unit or service level or even aggregated into a family level. Significant demand variation should be understood in magnitude and root cause. See Exhibit 4 for an example of demand segmentation for finished goods. Know that future-state value streams should be designed to accom-

**EXHIBIT 4** Example Demand Segmentation Analysis

XYZ Product Family Finished Goods

| | Historical weekly demand, units | | | | | | | | | | | | Annual Historical Demand | [A] Average Weekly Demand | [B] Standard Deviation Weekly Demand | [B]/[A] Coefficient of Variation ($C_v$) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SKU | 1 | 2 | 3 | 4 | 5 | ... | 48 | 49 | 50 | 51 | 52 | | | | |
| UL490 | 10,012 | 13,747 | 13,389 | 15,431 | 13,948 | ... | 14,741 | 10,125 | 11,041 | 12,146 | 15,499 | 676,411 | 13,008 | 2,070 | 0.2 |
| XP100 | 3,967 | 5,032 | 10,756 | 12,721 | 9,967 | ... | 16,112 | 4,259 | 4,314 | 2,115 | 5,076 | 386,459 | 7,432 | 4,621 | 0.6 |
| ZT250 | 386 | 168 | 1,003 | 417 | 12 | ... | 12,370 | 294 | 56 | 27 | 386 | 78,619 | 1,512 | 3,826 | 2.5 |

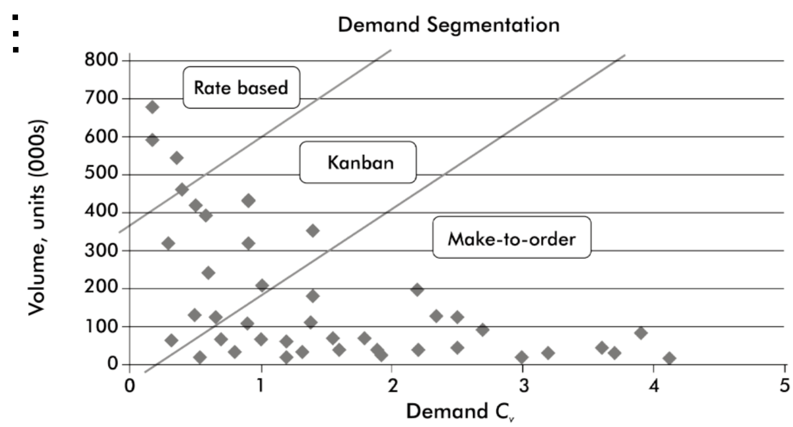



**EXHIBIT 5** VSM Is (a Lot) About Time

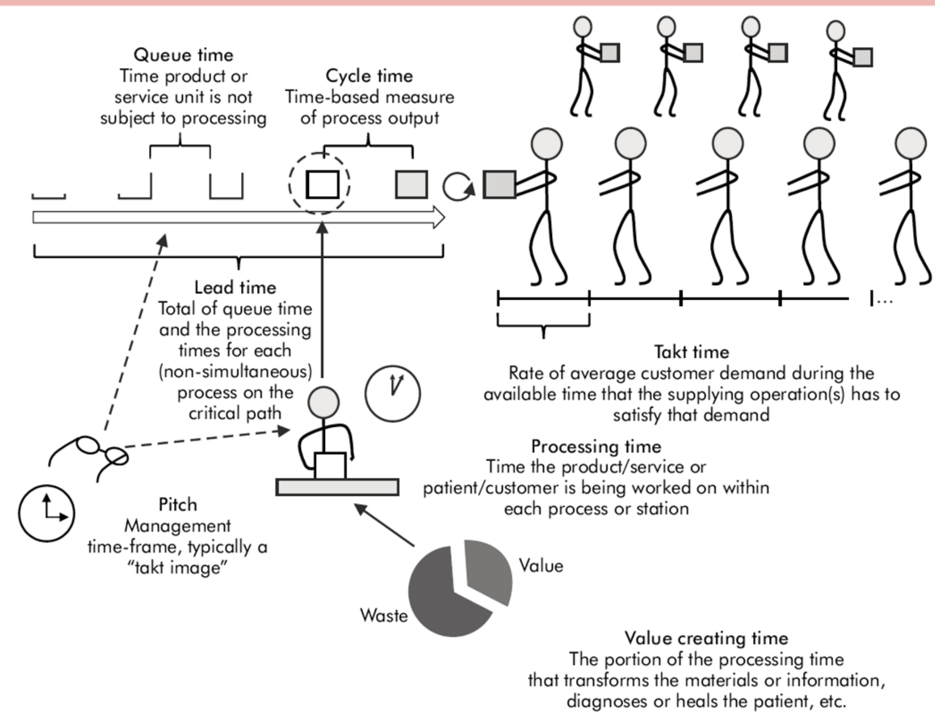


**EXHIBIT 6** Common Data Box Fields

| Data box field | Represents | Value stream mapping notation beyond the data box | Other considerations |
|---|---|---|---|
| **Available time** | Net scheduled time during which the process is available to conduct work | N/A | Used to calculate the local process takt time; provides insight into synchronization between other processes |
| **Average period demand** | Typically average demand per shift or day | N/A | Used to calculate the local process takt time and queue time (see Exhibits 7 and 8) |
| **Takt time** | Local takt time for the process | N/A | Takt time must be ≥ process cycle time or drop-off rate; local takt time may differ versus other processes' takt times and the overall system takt time |
| **Processing time** | Non-simultaneous work content within the process; think of it as the time that the thing is being worked on in the process | Dropped down to the top of the lead time ladder's bottom rung, see Exhibits 9 and 10 | Processing time is usually a very small fraction of overall lead time; the process cycle efficiency calculation, Exhibit 9, provides insight into that reality |
| **Cycle time** | The time, inclusive of manual, walk, and wait times, for the process to complete one full cycle | Cycle time or "drop-off rate" of the process (inclusive of all lines, cells, resources, etc. dropped down to the bottom of the lead time ladder's bottom rung, see Figure 9) | Cycle time or drop-off rate must be ≤ to takt time; helps identify opportunities for cycle time reduction and constraint management |
| **Number of operators** | Count of trained and available workers within the process, usually by shift | N/A | Provides insight into line balance opportunities and optimal staffing |
| **Every part every interval (EPEI)** | Time duration, or interval ("I") for a process to cycle through the production of all material or service types; most relevant in pattern-making environments | N/A | Provides insight into opportunity to reduce upstream inventory and compress lead time using tools like setup reduction and level loading |
| **Uptime** | Percent of the process' regular planned available time during which the process is ready and available to run at standard rate and quality | N/A | Provides insight into process availability and opportunities to address traditional six big losses (breakdowns, setups and adjustments, reduced speed, minor stop and idling, etc.) |
| **Scrap factor** | Percent of parts fabricated, assembled, tested, packed, etc. that cannot or will not be reworked | N/A | Highlights reason for downstream attrition and helps identify process quality improvement opportunities |
| **First-pass yield** | Percent of presented work that requires zero rework or replacement and is complete and accurate in the first cycle | Dropped down to rolled thoughout yield (RTY) line and used to calculate RTY, see Exhibit 11 | Helps prioritize process quality improvement opportunities within the value stream |
| **Inventory** | Count of inventory actually on hand at the time at the current state map or anticipated in the new system and reflected within the future state map | Used to calculate queue time preceding the immediate process, see Exhibits 7 and 8 | Inventory count should be for the relevant inventory proxies; for example, for a cake production value stream, the inventoried/counted items are flour (in the most upstream portion) and NOT eggs, milk, sugar, butter, etc., then cake batter, then baked cake, then frosted cake, etc. |

modate future demand. If future demand is expected to be substantially different than historical demand and can be reasonably estimated, then that should be incorporated in the target condition.

## Ready to map

Now that we understand what we are mapping and the design parameters, at least from a performance perspective (i.e., reduce lead time by 65 percent), it is time to map. Of course, there is requisite pre-VSM preparation (which includes selecting team members, capturing and compiling certain data, and scheduling the kick-off meeting), but let us assume that has been conducted.

As we launch into the VSM activities by which the team(s) will generate a current state value stream map, future-state value stream map, and value stream improvement plan, the practitioners will have to contend with the last five math regions reflected in Exhibit 2. These will intersect in some shape or form with most or all of the eight key future-state questions posited in the book *Learning to See*.[3]

## Material flow and related data boxes (region 2)

Shortly after the big plotter or kraft paper is affixed to the wall and the VSM map title is marked at the top (see "region 1"),

**EXHIBIT 7** Pull System Inventory-related Queue Time (Simplistic Version)

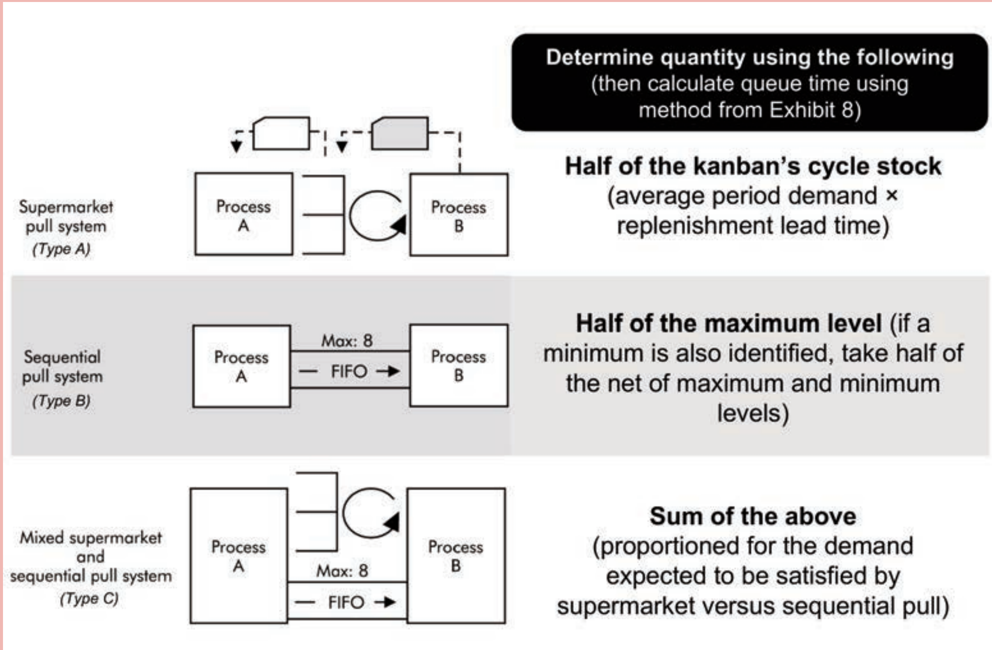



**EXHIBIT 8** Two Methods for Calculating Inventory Queue Time

**Average Daily Demand Method**

$$T_q = \frac{I}{\bar{D}_d}$$

where:

$T_q$ = queue time, in days
$I$ = inventory count, in units
$\bar{D}_d$ = average daily demand, in units/day

*Example.* At Sigma Health Systems' blood draw department, the average inventory waiting for blood draw after check-in is three patients. The average daily blood draw demand is 34 patients.

$$T_q = \frac{3 \text{ patients}}{34 \text{ patients/day}} = 0.09 \text{ days}$$

To convert $T_q$ into minutes (or hours), multiply $T_q$ by the available time for the day, $T_a$. For example, if $T_a$ is 450 minutes, then $T_q$ is equal to 40.5 minutes (39.7 without rounding). This is essentially the same approach as reflected in the following formula:

**Takt Time Method**

$$T_q = I \times T_t$$

where:

$I$ = inventory count, in units
$T_t$ = takt time

*Example.* At Sigma Health Systems' blood draw department, the average inventory waiting after check-in for blood draw is three patients. The $T_t$ is 13.23 minutes (450 minutes/34 patients).

$$T_q = 3 \text{ patients} \times 13.23 \text{ minutes/patient} = 39.7 \text{ minutes}$$

the practitioners can begin populating the map with process boxes paired with their descriptive data boxes. The best way to harvest data for the data boxes is through direct observation. Still, direct observation takes time. That can be especially difficult when the value stream is large and complex, and folks are encumbered with the day-to-day operation of the business. Often there is a lack of important data (such as resource uptime or first-pass yield [FPY]), and there is little time to capture long cycle times or schedule the observations of infrequently used processes. Thus, occa-

**EXHIBIT 9** Lead Time Ladder Math

(A) Lead (queue) time sum $= \sum_{i=1}^{j} T_{l_i}$

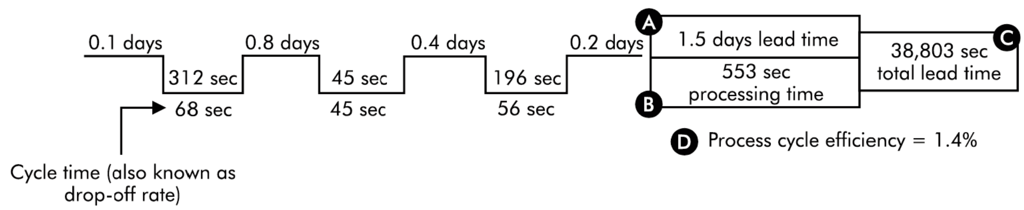
(B) Processing time sum $= \sum_{i=1}^{k} T_{p_i}$

(C) Total lead time = (A) + (B)

(D) Process cycle efficiency $= \frac{(B)}{(C)} \times 100\%$

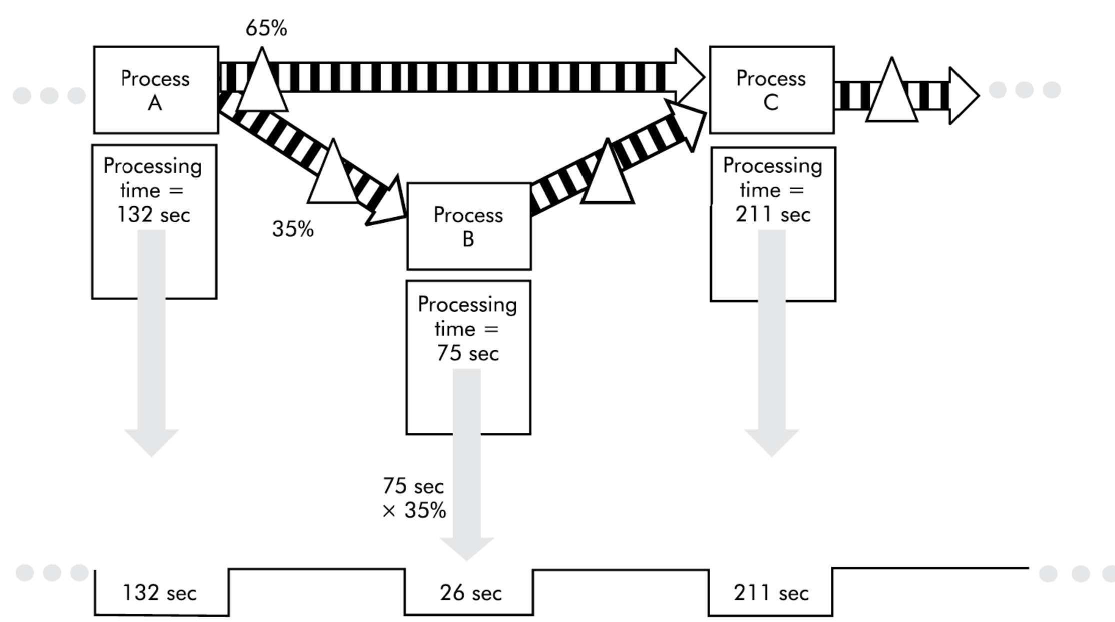
where:
$T_l$ = lead time, typically in seconds, minutes, hours, or days
$j$ = number of queues within the value stream as reflected in the top rung of the lead time ladder
$T_p$ = processing time, typically in seconds or minutes
$k$ = number of processes within the value stream as reflected in the top rung of the lead time ladder

Example




Assumes 425 minutes (25,500 seconds) available time/day

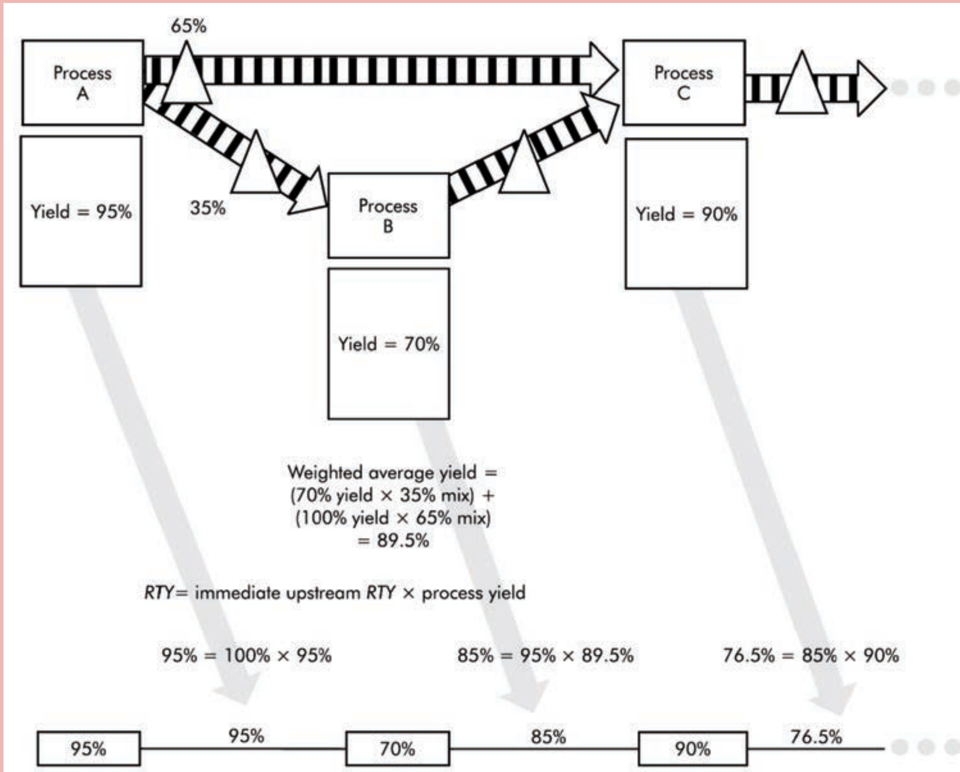
**EXHIBIT 10** Lead Time Ladder and Processing Time, Considering Branches




sionally we must defer to alternate approaches.

Second-best to direct observation is referencing confirmed reliable historical data — for example, accessing machining center programmable logic controller data to understand uptime. Third-best, and surely a last resort (and only for certain metrics such as FPY) are scientific best guesses or estimates by the people who regularly perform the work. These estimations are preferably gleaned with the help of a

**EXHIBIT 11** RTY Line Math




facilitator and done as a team to check for reasonableness. Ultimately, all value stream math should be subjected to common sense questions, like: Does that (fill in the blank — lead time, cycle time, etc.) make sense knowing what we think we know (e.g., patients regularly wait about 35 minutes throughout their visit, or it takes about three weeks to fill a customer order)?

Data boxes represent a sort of critical data laundry list for each process box on the current and future-state value stream map. As is the case with almost anything, there is a hierarchy of need. The short list usually includes cycle time, processing time, inventory queued up before and within the process, uptime, and FPY. While the lean practitioner is somewhat fixated on time (see Exhibit 5), there are other important pieces of data that will help identify opportunities to reduce process and system variation. Without getting very technical, Exhibit 6 provides an overview of the more common data box fields. Note the references to other exhibits (and the associated math).

## Quantify inventory and characterize process linkages (regions 3 and 4)

Inventory, consistent with Little's Law, is the shadow of time. Value stream maps depict or capture inventory in a handful of different modes. Often a single map reflects one or more of the following:

- push (or batch-and-queue), typically represented by a triangle and a zebra "push" arrow;
- type A, B, and/or C pull systems, as reflected in the left side of Exhibit 7; and
- buffer or safety stock to cover demand variation and system instability, respectively.

Here, we know that continuous flow processes are assumed to operate in, or closely approximate, one piece flow, and thus incorporating any queue time is mathematically redundant with the processing time.

So, how do we translate inventory into queue time? For any current state phenomenon described previously, we first *count* the actual associated inventory. Then we have two options: Employ either the average daily demand (ADD) or the takt time (TT) method. Both are mathematically equivalent and illustrated in Exhibit 8.

In our future-state VSM, we hopefully have zero push arrows, but how do we determine the time value of inventory for the other systems? Pull system–related queue times can be determined by applying the ADD or TT method to the relatively conservative — and very quickly calculated — in-use pull system stock levels, reflected in the right side of Exhibit 7.

More appropriately, the A-type pull system stock level could be calculated at half of the cycle stock, plus all the buffer and safety stock. But, in the heat of VSM, that might be unnecessary and impractical due to the underlying data analysis and math (e.g., demand coefficient of variation) needed to properly calculate the buffer and safety stock portions of the *kanban*.

For simplicity and conservativeness, future-state strict buffer and/or safety stock (meaning that there is no cycle stock) should likely be treated by applying the full sizing of the buffer or safety stock within the ADD or TT method. Given the complexity of buffer and safety stock calculations, often the most pragmatic approach is for the lean practitioners to simply peg a buffer or safety stock size in a number of days. For example, we estimate that the buffer will cover two days of demand.

## Create and populate the lead time ladder (region 5)

No value stream map is complete without the ubiquitous lead time ladder (Exhibit 9). It provides practitioners with an extremely important opportunity to check the overall reasonableness of the assembled data and thus the map itself. It also helps quantitatively communicate the opportunity in terms of the current state, and the vision and challenge of the future state.

The lead time ladder compiles all the discrete processing times and queue times and helps flag any cycle time (or drop-off rate) versus TT issues. The ladder provides a summation of queue times and processing times, presents the total lead time, and provides a convenient spot on the map to calculate PCE. PCE represents the percentage of time that the product, service, patient, and so forth is being worked on, delivered, or attended to over the total lead time. Know that it does not necessarily equate to pure value-added time because the processing time reflected on the map probably contains substantial waste.

We would be remiss not to quickly address the phenomenon of branching within value stream maps and the implications it has on lead time ladder math. Branching is reflective of situations where a process can link downstream to two or more processes. For example, a product family is comprised of some products that proceed from a shared subassembly process to a special test stand and then to final assembly. Meanwhile, other products move from the subassembly process directly to final assembly. How is that — specifically, the processing time — reflected on the lead time ladder? The answer is that it is treated as a weighted average to reflect the full target product family (Exhibit 10).

## Create and populate the rolled throughput yield line (region 6)

While it is extremely rare to see a properly prepared VSM without a lead time ladder, the inclusion of a rolled throughput yield (RTY) line is rather novel. (See the bottom of Exhibit 2 for a representation.) This is a missed opportunity.

The RTY line captures the discrete process FPY and then calculates the RTY as it cumulatively (or rather, multiplicatively) changes from upstream to downstream. It enables the VSM creator and reader to quickly identify and target the most significant quality performance opportunities.

Calculating RTY for a value stream with no branches is very straightforward. The RTY is simply the product of the preceding processes' FPYs. Exhibit 11 illustrates the math, with the treatment of one of those pesky material flow branches.

## Conclusion

Value stream analysis is a team-based collaborative flow or system *kaizen* vehicle. Its intended purpose is to help people understand how and why a product family's value flows or does not flow in the current state.

**NO VALUE STREAM MAP IS COMPLETE WITHOUT THE UBIQUITOUS LEAD TIME LADDER.**

By applying lean principles, systems, and tools, it also helps determine how, when, and by whom the product can be improved to achieve a designated, measurable target condition by a certain date. Just as importantly, value stream analysis helps engage, inspire, and drive individual and organizational learning.

The VSM process is an appropriately rigorous but also chaotic experience, replete with sticky notes, 10–30-foot-long pieces of paper, data boxes, weird icons, pencils, erasers, and cold pizza. It employs chainsaw — not scalpel-level — precision.

But while it is about seeing and thinking together in this context, do not let sloppy or wrong math get in the way of creating a reasonably accurate, and thus believable, set of maps. To that end, the regional-based math that we treated in this article should provide a firm foundation for a useful VSM. Incorporate it with the basic math that folks should already have under their belts, like TT, cycle time, and FPY.

What didn't we cover? A lot! Certainly, when you have the appetite to more deeply explore the regional math — or more precisely, size *kanbans* and first-in, first-out lanes, calculate *heijunka* cycles, determine work content variation, calculate process and system capacity, understand performance metric design, characterize and understand variation, and so on — we will be happy to help you dig in. ■


**NOTES**

1 Rother, M. and Shook, J., *Learning to See*. (Cambridge, MA: Lean Enterprise Institute, 2003).

2 Pound, E.S., Bell, J.H., and Spearman, M.L., *Factory Physics for Managers: How Leaders Improve Performance in a Post-Lean Six Sigma World*. (New York: McGraw-Hill Education, 2014).

3 *Op. cit.* note 1.