DISTRIBUTED MANUFACTURING SYSTEMS CONTROL

by

Sencer Yeralan,
Courtesy Professor,
University of Florida
Gainesville, Florida, U.S.A.

and

Helen Emery,
Rigel Corporation
Gainesville, Florida, U.S.A.

Abstract
Recent developments in embedded computing technologies and network communications have shifted the mode of manufacturing systems control from centralized control to decentralized control. We present examples from metal-cutting and consumer electronics assembly industries. Engineering models are proposed to assist in the design and operation of such systems.

Background
High-volume manufacturing often employs dedicated production lines to achieve the lowest possible cost per product. A production line consists of a series of stages, where each successive stage performs the next step in the manufacturing process. Storage space, referred to as buffers, may be provided for work-in-process. For instance, well-balanced production lines are standard practice in most consumer electronics and automotive industries. Balancing the line refers to setting the throughputs of all each stages to the same value, so that no bottlenecks occur. Stages starving for parts are evidence of upstream bottlenecks. Similarly, stages forced down due to a lack of storage space for their output is evidence of downstream bottlenecks. These two conditions are referred to as starving and blocking, respectively. Depending on the length and complexity of the production line, several bottlenecks may develop at any given time. Moreover, over time, bottlenecks may shift position within the production line. This phenomenon is commonly referred to as the “shifting bottleneck(s).”

As stated, bottlenecks cause stages to become idle (starved or blocked), and thus, result in a reduction of productivity. The elimination of idleness or at least the reduction of it is of top priority.

Clearly, bottlenecks, or starving and blocking, are inevitable if the individual throughputs of the stages differ. A slower stage following a faster one will become a bottleneck. Line balancing is thus the first step towards improving the productivity of the line. Sometimes, such as in many automotive industries, a paced line is used. The items move through the line on a conveyor at a constant speed. Each stage then operates on the item within the period the item spends in the stage. Each stage is allocated a length of the conveyor, which in turn determines the time allowed to complete the operation. This is a useful approach, especially if the operations require different times. In other cases, such as high-volume consumer electronics manufacturing, the operations are carefully planned to require the same time. Line balancing through the partitioning of the conveyor, or through the partitioning the set of operations, or a combination of both, is one of the oldest topics of industrial engineering.

Even a balanced line develops bottlenecks, and hence, suffers from stage idle time. These bottlenecks are caused by the temporary deviation of the stage throughputs from the desired line throughput. There are two major causes for such temporary deviations in stage throughput. The first is stage service time variability, and the second, stage breakdown. Variable service times imply that occasionally a downstream stage will be slower than its upstream stage, causing a bottleneck. Similarly, when a stage breaks down, it begins blocking the upstream flow. The term breakdown here does not refer to catastrophic failures of the stage. For instance, a worn or broken tool that must be replaced, or an occasional misalignment that requires resetting the machinery, introduces a delay, and hence a bottleneck. An engineering effort is made in designing the manufacturing tools so that they operate at a constant speed. Typically, the manufacturing tools are rated at a certain capacity, but run at a lower speed to assure constant operation times. Moreover, such tools are designed and built with attention to the reduction of downtimes due to failures. In spite of these efforts, uncertainties in the actual manufacturing processes still introduce variations, which inevitably lead to bottlenecks. If the time-average throughputs of the stages are equal, the bottlenecks occur at random points and shift over time. Shifting bottlenecks imply that the line is stochastically balanced. In this case, the loss in productivity due to the shifting bottlenecks is a monotonically increasing function of the variability of the individual stage throughputs. Therefore, as a second-degree remedy, manufacturing tools are designed and built to have as little throughput variability as possible. In case of breakdowns, this implies that the down times must be as short as possible. For example, some manufacturing equipment use two feeders, or two tools, for instance. If the supply in one feeder is used up, the operator simply switches the other feeder and continues with the operations. The first feeder is replenished while the stage continues its operation. Switching from one feeder to the other takes considerably less time than to re-supply the feeder. Hence, downtime is reduced.

Work-in-process buffers also reduce blocking and starving due to stage service rate variability. In order to be effective, the buffer must contain enough parts to supply its downstream during a typical breakdown of the preceding stage. Similarly, the buffer must have enough capacity for the upstream stage to store its output during a typical breakdown of the downstream stage.

The usefulness of buffers is evident when the extreme case is considered. With no buffers (such as in chemical processing with no holding tanks), the entire line is forced down as soon as one of the stages become unavailable. Assuming a stage availability of A, n identical stages in series have a line availability of A^n. As n increases, Line availability rapidly decreases. The figure below shows line availability as a function of stage availability. Identical stages are considered, with individual stage availabilities of 0.95, 0,96, .., 0.99. With n=20 and A=0.95, the line availability is about 36% -- an undesirable performance measure indeed.

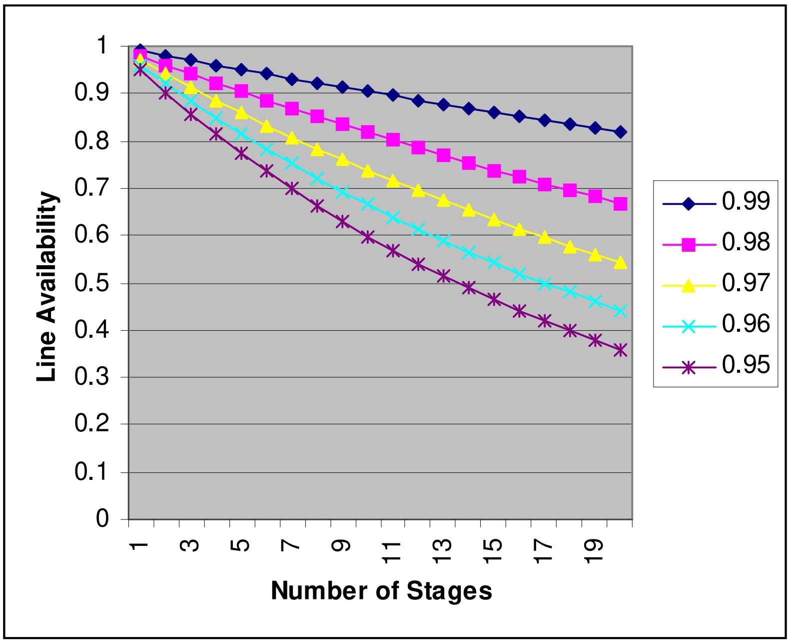

Figure 1. Line Availability as a Function of the Number of Stages and Stage Availability. (Case: no buffers, identical stages.)

Clearly, buffers cannot compensate for line imbalance, but can reduce the strong coupling among adjacent stages.

In summary, in order to increase productivity, a production line must be balanced, its stage service time variability reduced, and stage downtimes lowered as much as possible. In addition, work-in-process buffers need to be implemented to further reduce bottlenecks from forcing neighboring stages to become blocked or starved. Much research has been conducted to quantify line productivity as a function of the stage properties and buffer capacities. These findings help to design and operate lines in the most productive manner possible by selecting the stage service times and buffer capacities.

Modern Production Lines

The preceding section is most applicable to manufacturing tools of the mid 20-th century. Such tools, employing fixed automation, were capable of high-volume manufacturing, but show little flexibility in operating modes. Moreover, such tools implemented an open-loop control scheme, rather than adjusting their operations according to the work-piece being manufactured. Two basic elements of modern manufacturing tools distinguish them from their predecessors. Both of these elements are a result of embedded computer technologies. First, the tools implement programmable automation. This not only allows the operating mode of the tool to be changed upon request, but also allows closed-loop control, where the operations are modified according to the specific variations of each work-piece. This requires the tool to include sensors as well as actuators. Second, the tools, or more specifically, their embedded computers are networked. Information obtained at any tool is readily available to other tools. Similarly, the variations in the operating modes may be relayed to the entire line. The production line control is viewed as a distributed control system, since the operating decisions may be collectively made by various computers on the network.

Embedded computing makes the manufacturing hardware "smart tools." They measure the work-piece attributes and decide on the best mode of operation within the constraints imposed by the line. In light of the discussions above, closed-loop or sensory-based manufacturing sometimes increases stage service time variability. From the viewpoint of production line availability, this is an undesired feature. However, it produces higher-quality products, and thus, is becoming common practice. Often, the tool is programmed to compensate for the work-piece variations as long as the resultant operations will fit within the allocated times. Otherwise, an exception is generated, which in turn is handled as a breakdown. Since tools may also be given information about the upstream and downstream buffer contents, it is not unreasonable to speed up or slow down a stage to bring the buffer levels back to nominal values. That is, the tools monitor and fine-tune their speed to locally balance the lone. This tactic may be employed to reduce or remove the shifting bottlenecks produced even when the line is time-average balanced. Modifying the operating mode based on work-piece or neighboring buffer information may be considered as local control. That is, local information is used to maintain a local balance of operations.

The networking capabilities enhance the effectiveness of the tools. Networking allows global information to be used in local decisions. Typically, past manufacturing systems relied on master schedules to plan line activity. A schedule may depict the target manufacturing volume for the next period (e.g. shift, day, week). Such a schedule is a top-down control approach, designed to guide local operations. With their embedded computers networked, modern tools may undertake such higher-level planning on their own. This should not be regarded as the tools collectively setting their target manufacturing volumes. A desired target still needs to be specified. Rather, such self-control or self-scheduling allows the line to vary the manufacturing output from shift to shift, while still maintaining the desired weekly production. A good example of how global information can be used to improve line productivity is in scheduling maintenance operations. Smart tools can evaluate, say tool ware, and broadcast this information. Rather than waiting for the breakdown condition to change the tool, the stages may collectively decide to stop the line and change all the tools at once. This way, the line is stopped only once, rather than once for each tool.

Research Opportunities

The modern tools present an unprecedented flexibility to manufacturing operations. This flexibility also removes many of the constraints according to which traditional production line operations were optimized. Local as well as global manufacturing information may be used to continuously fine-tune line operations. However, theoretical work has not yet fully classified, let alone resolved, many of the issues involved. Some of the possible research topics are listed below.

<u>Service Time Control</u>

Stage service times may be modified to compensate for local deviations from ideal conditions. More specifically, a stage should increase its service rate if its upstream buffer is above the nominal level, and its downstream station below its nominal level. Usually, service times are reduced or lengthened gradually, without sudden impacts on the system. In case of lengthening service times, the extra time may be used for work-pieces that may need a little extra work to improve product quality. These items may be detected upstream and reported to the stage, in preparation for a slow down. Models that describe the effect of such dynamic policies are required to quantify and balance the productivity improvements and costs. It has been suggested that, similar to simpler inventory control models, trigger levels be set for the upstream and downstream buffers. The stage should then switch to a different operating policy when the trigger levels are reached. In general, a formula may be sought to give the optimum stage service time as a function of the levels of the neighboring buffers.

1. Using Global Information

The approach proposed above may be extended to use more than the levels of just the neighboring buffers. The formula may use all or a proper subset of the buffer levels. Should the information concerning levels of distant buffers be discounted?

2. Using Historical Data

The computer network may store all line activity in a database. This information may be used in future computations to set line control parameters. Currently, the nature of the information embedded in the data is not well understood. In other words, rather than storing raw data, what type of statistically significant performance measures can be obtained and stored. Clearly, such measures would require less effort to store and manage. Moreover, data preprocessed in this manner is easier to use in future computations. This topic is closely related to the relative value of information embedded in raw data. For example, past information may be discarded after it ages a certain amount. Candidates for such performance measures are the up time and down time distributions of the stages, time to blocking and time to starving distributions, blocked time and starved time distributions, etc. For instance, the distributions may be represented by their first few moments.

3. Learning

If a formula, as mentioned in item 1 above, is not available, or not practical, can historical data collected over the operation of the line be used to empirically determine the trigger levels for service time modifications? This would amount to a learning algorithm, where the line learns to pick the best operating mode. It is plausible to expect that the longer the line runs, the better its performance will become. Moreover, since the learning approach requires little prior preparation, the approach is universally applicable to a wide range of lines. In other words, the same control software can be loaded to a diverse range of systems and be expected to perform reasonably well. Wide applicability and scalability are very desirable traits of manufacturing control software. Theoretical work need to be conducted to evaluate the expected time such a learning algorithm would take to converge to an acceptable operating discipline. Can off-line simulation, or even virtual emulation by the same manufacturing tool embedded computers be used to speed up convergence?

4. Global Stability

The term local information refers to the sensory data obtained at the stage, as well as the neighboring two buffers. Many control mechanisms may be devised using this information. However, since the stages are networked, the information is available to any other stage in the line. Intuitively, the use of information from other stages, that is, global information, should allow better control mechanisms. Also plausible is the notion that the further the information comes from, the less marginal value it has. The precise manner in which global information should be used is an open question. Specifically, is it sufficient to achieve optimum production line control with only local information? After all, global information trickles down to the immediate stage, and effects its local information. In practice, it is much easier to conceptualize and realize control based on local information. On the other hand, would the use of only local information make it possible for the entire line to display oscillatory behavior at the system level, continuously cycling through overshoots and undershoots at the local levels. How can system stability be predicted as a function of the control mechanism?

Concluding Remarks

The modern manufacturing systems are several orders-of-magnitude more flexible and sophisticated compared to their fixed-automation-style predecessors. Production lines constructed with such manufacturing tools suggest a very wide range of theoretical research topics. This report provides a background and motivates some of these topics.

Bibliography

1. Fadiloglu, M. and S. Yeralan, “A General Theory on Spectral Properties of State-Homogeneous Quasi-Birth-Death Processes,” European Journal of Operational Research, forthcoming 2001.

2. Fadiloglu, M. and S. Yeralan, “Models of Production Lines as Quasi-Birth-Death Processes,” Mathematical and Computer Modeling, forthcoming 2001.

3. Karabati, S. and Tan, B., “Stochastic Cyclic Scheduling Problem in Synchronous Assembly and Production Lines,” Journal of the Operational Research Society, Vol. 49, No. 11, pp. 1173-1187,1998.

4. Tan, B., "Production Control of a Failure Prone Pull Manufacturing System with Variable Demand," Koç University Working Paper Series 00-1.

5. Tan, B., “State-Space Modeling and Analysis of Pull Controlled Production Systems,” Koç University Working Paper Series 99-5.

6. Tan, B., “A Three Station Continuous Materials Flow Merge System with Unreliable Stations and a Shared Buffer,” Mathematical and Computer Modeling, forthcoming, 2001.

7. Tan, B. and Karabati, S. "Modeling and Scheduling of an Asynchronous Cyclic Production Line with Multiple Parts," Journal of the Operational Research Society, Vol. 51, No.11, pp. 1296-1308, 2000.

8. Tan, B. “Asymptotic Variance Rate of Production Lines with Limited Buffer, ”Annals of Operational Research, Special Issue on Performance Evaluation and Optimization of Production Lines, Vol. 93, pp. 385-403, 2000.

9. Tan, B., “Variance of the Output as a Function of Time: Production Line Dynamics,” European Journal of Operational Research, Vol. 117, No.3, pp.470-484, 1999.

10. Tan, B., “Asymptotic Variance Rate of the Output of a Transfer Line with No Buffer Storage and Cycle-Dependent Failures,” Mathematical and Computer Modeling, Vol. 29, pp. 97-112, 1999.

11. Tan, B. “Agile Manufacturing and Management of Variability,” International Transactions in Operational Research, Vol. 5, No. 5, pp.375-388, 1998.

12. Tan, B., "An Analytic Formula for Variance of Output from a Series-Parallel Production System with No Interstation Buffers and Time-Dependent Failures," Mathematical and Computer Modeling, Vol. 27, No. 6, pp.95-112, 1998.

13. Tan, B., "Effects of Variability on the Due-Time Performance of a Continuous Materials Flow Production System in Series," International Journal of Production Economics, vol. 54, pp.87-100, 1998.

14. Tan, B., “Variance of the Throughput of an N-station Production Line with No Intermediate Buffers and Time Dependent Failures,” European Journal of Operational Research, Vol.101, No.3, pp.560-576, 1997.

15. Tan, B. and Yeralan, S., “A Decomposition Model for Continuous Materials Flow Production,” International Journal of Production Research, Vol. 35,No.10, pp. 2759-2772, 1997.

16. Tan, B. and Yeralan, S., “Analysis of Multistation Production Systems with Limited Buffer Capacity, Part II,” Mathematical and Computer Modeling, Vol. 25, No.11, pp.109-123, 1997.

17. Tan, B. and Yeralan, S., "A Fuzzy Decomposition Method for Multistation Production Systems Subject to Blocking," International Journal of Production Economics, Vol. 42, No. 3, pp. 245-262, 1995.

18. Tan, B., Yeralan, S., Babu. S, and Osborn, B., "Product Quality/Tool Reliability Modeling of Complex Manufacturing Tools," Endustri Mühendisligi (Industrial Engineering Journal of Turkish Chamber of Engineers), Volume 5, Number3, pp. 3-11, 1995.

19. Tan, B. and Yeralan, B., "A Unified-Boundary Interior Model for Continuous Materials Flow Production Lines with Station Breakdown," Yöneylem Arastirmasi Dergisi/Transactions on Operational Research, Volume 6, Number2, pp. 95-120, November 1994.

20. Tan, B., Yeralan, S., Babu. S, and Osborn, B., “Computer-Aided Reliability Modeling and Applications in Semiconductor Manufacturing,", Computers and Industrial Engineering, Vol. 23, Nos 1-4, pp.169-172, 1992.

21. Yeralan, S., M. Fadiloglu., G. P. Cengiz, "The Intelligent Control of Productions Lines - A Framework for Experimentation," Proceedings of the International Workshop on Performance Evaluation and Optimization of Production Lines, pp. 281-291, 1997.

22. Yeralan, S. and Tan, B. “A Station Model for Continuous Materials Flow Production,” International Journal of Production Research, Vol.35, No.9, pp. 2525-2541, 1997.

23. Yeralan, S. and Tan, B., “Analysis of Multistation Production Systems with Limited Buffer Capacity, Part I,” Mathematical and Computer Modeling, Vol. 25, No. 7, pp.109-122, 1997.