Operations Research Letters 44 (2016) 186–190

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Monotonicity of base stock policies

Nha-Nghi de la Cruz [a,*], Hans Daduna [b]

[a] Faculty of Business Administration, University of Hamburg, Hamburg, Germany

[b] Department of Mathematics, University of Hamburg, Hamburg, Germany

ARTICLE INFO

Article history:
Received 13 October 2015
Received in revised form
10 December 2015
Accepted 5 January 2016
Available online 11 January 2016

Keywords:
Queueing
Stochastic processes
Inventory
Stochastic ordering

ABSTRACT

We analyze monotonicity of base stock levels in multi-product inventory-production systems where arriving demand triggers production of a new unit. In a paper from 1996, Rubio and Wein used open Jackson networks to describe such integrated models. They conjectured that the base stock level should increase with the utilization in the production system. We present a basic analytical proof of the general presumption for single- and multi-product systems.

1. Introduction

For a manufacturing system which produces products of different types on a make-to-stock basis, we consider a queueing network model to determine an optimal base stock policy. A base stock policy prescribes a target total inventory position for each product to balance backorder and inventory holding costs. In the system, products are stored in a finished good (FG) inventory to serve exogenous demand. If there are finished goods at stock when demand arrives, it is satisfied immediately, otherwise it is backordered (negative FG). In both cases, an order for producing a new unit of the required product is placed instantaneously, which is immediately counted as work-in-process (WIP). Consequently, the sum of FG and WIP inventory is maintained at a fixed level for each product in this CONWIP-like system. Related models are base stock systems with multiple production stages each holding its individual buffer stock (see [2,8]).

Our starting point is the paper of Rubio and Wein [10] whose base stock model consists of a multi-product inventory and a replenishment network of the Jacksonian type.

We focus on monotonicity behavior of optimal base stock policies within the parameters of the network and discuss the following important monotonicity property for optimally setting the target base stock level: Whenever the demand intensity for a product increases and/or somewhere in the network the production capacity decreases, the optimal base stock level increases (we use "increasing" for "non decreasing", and similarly "decreasing").

Although this property seems to be intuitive and natural, it is not easy to prove. In fact, Rubio and Wein provided a proof only for the case of a single-product inventory-production system where all stations have the same utilization (in their terminology: a balanced network). For the single-product case in an unbalanced network, they conjectured that such monotonicity should hold as well, relying on numerical experiments (see [10], p. 263). The main objective of the present paper is to prove the monotonicity property even for a more general system, i.e., the manufacturing system does not have to be balanced nor reduced to the single-product case.

2. Manufacturing queueing system with backordering

We consider a multi-product manufacturing system with products u, $u = 1, \ldots, U$. In the system, we distinguish between the FG and WIP inventory which contain the finished and unfinished goods, respectively. Unsatisfied demand is captured by the backorder level for each type of good.

If an exogenous demand for a specific product arrives, the corresponding WIP inventory increases in the same amount as the FG inventory for this product decreases. This ensures that the demanded product is eventually reproduced. On the other hand, finished products are transferred from the manufacturing network to the FG inventory, i.e., are converted from WIP to FG units.

* Correspondence to: University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany.

E-mail addresses: nha-nghi.de.la.cruz@uni-hamburg.de (N.-N. de la Cruz), daduna@math.uni-hamburg.de (H. Daduna).

http://dx.doi.org/10.1016/j.orl.2016.01.003
0167-6377/

The WIP processes $(N_u(t) : t \geq 0)$ for $u = 1, \ldots, U$, are determined by the manufacturing system, which will be described in terms of stochastic networks of generalized Jacksonian type in Section 3.

Let $Z_u(t)$ denote the FG inventory level of product u, where negative levels count backordered units. Without loss of generality, the initial level is set as the base stock level z_u, $u = 1, \ldots, U$.

Demand reduces the FG inventory whereas finished produced units increment it. Therefore, the base stock level is at all times equal to the total inventory position: $N_u(t) + Z_u(t) = z_u$. The WIP levels $N_u(t)$ can grow arbitrarily high because it also includes the backordered products whereas $Z_u(t)$ is bounded from above by z_u.

We are interested in the long run overall costs per time unit. Constructing a Markovian system process which relies on standard queueing network models, these costs are computed via the stationary expected costs by exploiting the ergodic theorem of the Markovian joint queue length processes in standard Jackson networks.

Let N_u denote a random variable distributed according to the total number of WIP of type u under the stationary WIP distribution, which will be given explicitly below.

It turns out that the main decision variables are N_u, $u = 1, \ldots, U$ (see [10], p. 261). Control variables are stock sizes $z_u \geq 0$, for product type $u = 1, \ldots, U$. Rubio and Wein showed that the cost minimizing base stock levels z_u^*, $u = 1, \ldots, U$, can be expressed in terms of the WIP of product u and the respective inventory holding cost h_u and backorder cost b_u per unit of product u and time unit only. These are the smallest integers $z_u^* \geq 0$, such that for the stationary WIP in the replenishment network holds

$$P(N_u \leq z_u^*) \geq \frac{b_u}{b_u + h_u}, \quad u = 1, \ldots, U. \tag{1}$$

3. Multi-class Jackson network with exponential service times

To describe the behavior of the WIP in the replenishment system, we use a standard multi-class Jackson network in which the WIP inventory N_u is the total population size of product u. We adapt the notation of [4, chap. 5] and consider a multi-class Jackson network with J stations, numbered by $j = 1, \ldots, J$. We distinguish in the network customer classes ($\equiv$ product types) u, $u = 1, \ldots, U$. Set n_{ju} the number of class u customers at node j ($\equiv$ WIP of product u at j), the total population size at node j as $n_j = n_{j1} + \cdots + n_{jU}$, and define $\overline{n_j} = \left(n_{j1}, \ldots, n_{jU}\right)$ and $\overline{n} = \left(\overline{n_1}, \ldots, \overline{n_J}\right)$, the joint class occupation vector. Arrivals of class u at node j (triggered by external demand for product u) from outside of the network follow a Poisson process with rate λ_{ju}. $r_{ju,ku}$ is the probability that class u customers departing from node j join node k, and $r_{ju,0}$ is the probability that a class u customer leaves the system after finishing his service at node j (and enters the FG inventory); it holds $\sum_{k=1}^{J} r_{ju,ku} + r_{ju,0} = 1$, $u = 1, \ldots, U$.

Note that since demand of product u triggers a class u arrival at the replenishment network ($\equiv$ initiates production of a new unit of product u) and leaves as an output of product u ($\equiv$ is shifted to the FG inventory after complete production), customers in the network do not change their class which would be possible in the general setting of [4, chap. 5].

α_{ju}, the total arrival rate of class u at node j, is obtained as the solution (assumed to be unique) of the so called traffic equations

$$\alpha_{ju} = \lambda_{ju} + \sum_{k=1}^{J} \alpha_{ku} r_{ku,ju}, \quad j = 1, \ldots, J;\ u = 1, \ldots, U.$$

At node j all customers require an amount of service which is exponentially distributed with rate μ_j. The server at node j provides service at rate $\Phi_j(n_j)$, which is non-decreasing in n_j and $\Phi_j(n_j) > 0$ if $n_j \geq 1$. This results in a versatile construction of service regimes. For example, if the network consists of first-come-first-served (FCFS) multi-server queues with $s_j \geq 1$ servers, we can write $\Phi_j(n_j) = \min\left\{n_j, s_j\right\}$. We define the class u utilization of node j by $\rho_{ju} = \alpha_{ju}/\mu_j$.

The stationary distribution π of the multi-class Jackson network for the joint class occupation process with values $\overline{n} = \left(\overline{n_1}, \ldots, \overline{n_J}\right)$ is with normalization constants $B_j = \sum_{n=0}^{\infty}\left(\left(\sum_{u=1}^{U} \rho_{ju}\right)^n / \prod_{\ell=1}^{n} \Phi_j(\ell)\right) < \infty$ (see [4], p. 127)

$$\pi\left(\overline{n}\right) = \prod_{j=1}^{J} \pi_j\left(\overline{n_j}\right) \tag{2}$$

with $$\pi_j\left(\overline{n_j}\right) = B_j^{-1} \frac{n_j!}{n_{j1}! \ldots n_{jU}!} \prod_{\ell=1}^{n_j} \Phi_j(\ell)^{-1} \prod_{u=1}^{U} {\rho_{ju}}^{n_{ju}}.$$

Remark 1. With the standard independence assumptions for inter-arrival times, service time requests and the Markov chain property of the customer traffic process, it can be seen that a process, which records the detailed sequence of the products present (waiting or in service) for all stations of the multi-class Jacksonian replenishment network, is a Markov process [4, chap. 5]. The departures from the network immediately turn into FG, while the arrivals at the network indicate arrival of external demand at the FG inventory. This makes our system accessible to a Markov process modeling. The joint class occupation process (which does not count for positions in the queue) with values $\overline{n} = \left(\overline{n_1}, \ldots, \overline{n_J}\right)$ is in general not Markovian. However, it suffices as state description for the relevant quantities for our investigations.

Example 1 (*Deterministic Replenishment Schedule*)**.** Our model encompasses the important case where each class (product) has its dedicated fixed sequence of service stations to visit in the replenishment procedure: Each customer (product) class follows its own route inside the network. Such networks with class dependent fixed routing are called Kelly networks [7], which exhibit stationary distributions with product form structure as introduced before.

Remark 2. So far, we have only considered networks in which customers have exponential service requirements. For arbitrary service requirements, we refer to BCMP networks (see [4, chap. 6]), where the service requirement of a customer u at station j may have an arbitrary distribution, depending on class and node. Denote by S_{ju} a random variable distributed as class u customers' service request at node j with mean $E(S_{ju}) = \mu_{ju}^{-1}$. In this situation, we need to abandon the FCFS-property and employ the so called symmetric service disciplines (see [4, chap. 6, p. 150]).

In the context of our replenishment manufacturing system, we could think of a server network with $\Phi_j(n_j)$ increasing in (n_j), where all arriving customers (products) have the same priority and the server's effort is equally shared among all the customers present at the same node (processor-sharing regime). The service completion rate is $\Phi_j(n_j) \cdot \frac{n_{ju}}{n_j}\mu_{ju}$ for customers of class u at node j, $u = 1, \ldots, U$. The utilizations in the network are $\rho_{ju} = \frac{\alpha_{ju}}{\mu_{ju}}$, $u = 1, \ldots, U$, $j = 1, \ldots, J$. The stationary distribution of this BCMP network with single servers under processor-sharing with $\sum_{u=1}^{U} \rho_{ju} < 1$ and unique solutions of the traffic equations α_{ju}, $j = 1, \ldots, J$, $u = 1, \ldots, U$, has the identical formula displayed in (2) (see [4], Theorems 6.1 and 6.2). But the cautious reader should recall the difference of $\rho_{ju} = \alpha_{ju}/\mu_j$ and $\rho_{ju} = \alpha_{ju}/\mu_{ju}$.