European Journal of Operational Research 237 (2014) 824–835

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Discrete Optimization

Service differentiation through selective lateral transshipments

E.M. Alvarez *, M.C. van der Heijden, I.M.H. Vliegen, W.H.M. Zijm

University of Twente, School of Management and Governance, P.O. Box 217, 7500 AE Enschede, The Netherlands

ARTICLE INFO

Article history:
Received 19 October 2012
Accepted 22 February 2014
Available online 6 March 2014

Keywords:
Inventory
Service differentiation
Lateral transshipments
Spare parts

ABSTRACT

We consider a multi-item spare parts problem with multiple warehouses and two customer classes, where lateral transshipments are used as a differentiation tool. Specifically, *premium* requests that cannot be met from stock at their preferred warehouse may be satisfied from stock at other warehouses (so-called lateral transshipments). We first derive approximations for the mean waiting time per class in a single-item model with selective lateral transshipments. Next, we embed our method in a multi-item model minimizing the holding costs and costs of lateral and emergency shipments from upstream locations in the network. Compared to the option of using only selective emergency shipments for differentiation, the addition of selective lateral transshipments can lead to significant further cost savings (14% on average).

1. Introduction

In the capital intensive industry, a company's operations may fully depend on the performance of certain capital goods, such as radar systems on frigates or MRI and CT scanners in hospitals, with downtime possibly leading to severe consequences. The users of such equipment generally outsource maintenance and spare parts supply to a service provider, with targets on performance measures (such as a maximum response time to failures) formalized in service contracts through service level agreements (Al Hanbali & Van der Heijden, 2013). Because the installed base of capital goods is often geographically dispersed and waiting times for spare parts should be small, the supply chain is typically organized as multiple local stock points, each supporting a subset of the installed base, combined with a centrally located central warehouse. Cheap fast movers are typically stored at local warehouses, expensive slow movers tend to be concentrated in the central warehouse to take advantage of the risk pooling effect.

A complication for managing spare part supply chains is that service level agreements (SLAs) may vary strongly among customers to reflect the value placed on system uptime (Jalil, 2011). A key challenge for the supplier is to satisfy all SLAs at minimal costs. Part of this challenge is to position spare parts in the supply chain such that a target overall system downtime waiting for spares is met at minimal costs. Spare parts suppliers usually handle differentiated service levels by either (i) giving all customers uniform service (also known as "one-size-fits-all", Cohen, Agrawal, and Agrawal (2006)) or by (ii) keeping separate supply chains per customer segment, with premium customers served from stock points nearby and other customers served from the central warehouse. Uniform service is expensive – it must accommodate the tightest service levels – and does not induce standard customers to switch to a premium contract. Separate supply chains, on the other hand, reduce the benefit from risk pooling (Eppen & Schrage, 1981), resulting in higher stock levels in the supply chain than needed.

The literature on differentiation has mainly focused on the use of *critical levels* (Veinott, 1965). That is, a single supply chain is deployed for all customers, but stocks are reserved for premium customers once the inventory drops to a certain threshold, the critical level. Although this approach can lead to large savings in theory, there are practical drawbacks. For instance, service engineers will often not delay a repair if the required part is available, as they are generally accountable for speed of repair. To overcome these drawbacks, Alvarez, Van der Heijden, and Zijm (2013) propose an alternative differentiation approach using *selective emergency shipments*, where demand in out-of-stock settings may either be backordered or satisfied using emergency shipments depending on the customer's class and the item being considered.

In this paper, we extend the selective emergency shipment model by allowing *lateral transshipments for premium customers* as well. That is, a warehouse that is out of stock may obtain the item from a neighboring warehouse that still has the item on hand (see e.g. Kranenburg & Van Houtum, 2009). Such shipments often have shorter lead times than emergency shipments and are also cheaper. At the same time, lateral transshipments facilitate risk pooling, thereby reducing overall stock levels in the supply chain (Paterson, Kiesmüller, Teunter, and Glazebrook, 2011). We limit the use of

* Corresponding author.
E-mail address: e.m.alvarez@alumnus.utwente.nl (E.M. Alvarez).

http://dx.doi.org/10.1016/j.ejor.2014.02.053
0377-2217/

lateral transshipments to premium customer requests only for the following reason. Suppose that a warehouse has only one part remaining on stock and a lateral request of a non-premium customer from another warehouse arrives. If the last item is used to fill this request, an "own" premium customer arriving a bit later will find an empty shelf. Intuitively, it is clear that this is not an adequate way to deal with high priority customers. Table 1 shows the order in which to fulfill a customer's demand at the warehouse by combining selective lateral transshipments and selective emergency shipments.

We can influence the system performance using three types of decisions: (i) the base-stock levels, (ii) the transshipment strategy, and (iii) the shipment strategy. The *transshipment strategy* specifies whether a premium request at a certain warehouse may be met through a lateral transshipment or not. Lateral transshipments cause additional shipment costs, but also result in lower stock levels since implicitly the risk pooling effect is exploited. The *shipment strategy* specifies whether an emergency shipment may be used if demand cannot be filled from stock on hand or through a lateral transshipment. If an emergency shipment is not used, the request is backordered until stock is replenished at the warehouse. If we decide not to use emergency shipments, we avoid high emergency shipments costs, but may need extra spare part inventories to meet the SLAs in terms of downtime waiting for spare parts.

We consider a supply chain of several warehouses supplying multiple spare parts to premium and non-premium customers. In turn, the warehouses are replenished from a central stock point. All locations use a continuous review, one-for-one replenishment policy. Each customer class has a distinct maximum average waiting time that applies over all parts jointly. As high waiting times for some parts might be compensated by low waiting times for other parts, we consider this as a *multi-item* problem, cf. Sherbrooke (2004). An item's waiting time and supply costs will depend on that item's stock levels in the system and on the (trans-)shipment strategy used if the nearest warehouse is out of stock. For simplicity, we assume that the central stock point has infinite stock which is also used for emergency shipments to the warehouses to satisfy customer demand as soon as possible. Shipments aimed at replenishing inventories of local warehouses only will be referred to as regular replenishments.

In Section 2, we give a literature overview and state our contribution. In Section 3, we present our multi-item model and globally describe our approach for solving this model. We then give details on the solution approach in Sections 4 and 5. Section 4 gives the analysis of a single-item building block with lateral transshipments for premium requests, whereas Section 5 details the heuristic solution of the multi-item problem. In Section 6, we discuss our extensive computational experiment. Finally, we draw conclusions and indicate further research areas in Section 7.

2. Literature

Our research is related to literature on service differentiation and emergency supply flexibility, i.e., using the flexibility of lateral transshipments and emergency shipments for meeting demand when the nearest warehouse is out of stock (Alfredsson & Verrijdt, 1999). Below, we focus on literature on lateral transshipments (possibly combined with emergency shipments). For literature on emergency shipments only, we refer to Alvarez et al. (2013). In the *service differentiation* area, we find contributions on both a tactical level (i.e., where the system stock levels are decision variables) and an operational level (i.e., where stock levels are given as input). Most papers on the *tactical level* apply differentiation using the critical level policy, a concept introduced by Veinott (1965). We refer to Teunter and Klein Haneveld (2008) for a literature review. Alternatively, Alvarez et al. (2013) use selective emergency shipments for differentiation with average cost savings of 4.4% over a one-size-fits-all approach. By combining selective emergency shipments with critical level policies considerably larger costs savings are possible; on average savings of 13.9% are found. The above papers only consider a single stock location. In contrast, Alvarez, Van der Heijden, and Zijm (in press) consider dedicated customer stocks as a differentiation tool, with stock possibly kept at customers' sites in addition to a central stock point. The resulting system is a two-echelon supply chain. The literature focusing on the *operational level* is limited to a few multi-location models. Jalil (2011) and Tiemessen, Fleischmann, Van Houtum, Van Nunen, and Pratsini (2012) consider single-item models with multiple warehouses and multiple customer classes, where a request can often be met from more than one warehouse. The supplier may choose to delay satisfying a low priority request or to meet such a request from a warehouse other than its nearest warehouse to reserve stock for premium requests.

The literature on *lateral transshipments* covers two types of models that differ in the way that demand is handled when it cannot be met from stock at either the nearest warehouse or through lateral transshipments from neighboring warehouses: the first model type then backorders demand, whereas the second satisfies it using emergency shipments. Models with *backordering* have initially been considered by Lee (1987) and Axsäter (1990), who consider a two-echelon setting consisting of a depot and various bases which are divided into transshipment pools. Axsäter uses an iterative analysis approach, where each base is analyzed separately over a number of iterations under the assumption that lateral transshipment requests at each base arrive according to Poisson processes. This logic has often been used in other papers, e.g. Alfredsson and Verrijdt (1999) and Van Wijk, Adan, and Van Houtum (2012). Models with *emergency shipments* have initially been considered by Dada (1992) and Alfredsson and Verrijdt (1999), who analyze similar two-echelon models. Some recent contributions are Kranenburg and Van Houtum (2009), where only a subset of warehouses can act as a lateral transshipment source, and Van Wijk et al. (2012), where a lateral transshipment request at a warehouse is only met if the stock level at that warehouse exceeds a so-called hold back level. In these latter two papers, it may in fact be that a lateral transshipment is not allowed even when some warehouses still have stock on-hand. We refer to Paterson et al. (2011) for details.

So far, lateral transshipments have only been considered as a service differentiation tool at an operational level, with contributions limited to single-item models (Jalil, 2011; Tiemessen et al., 2012). In contrast, we consider a multi-item model for which we

Table 1
Overview of order fulfillment options.

Premium customers	Non-premium customers
1. Stock on hand	1. Stock on hand
2. Lateral transshipment from other warehouse	2. Emergency shipment from a central location upstream the supply chain
3. Emergency shipment from a central location upstream the supply chain	3. Backorder, wait for a replenishment order
4. Backorder, wait for a replenishment order	