Applying the Theory of Constraints to Supply Chain Collaboration[*]

Togar M. Simatupang[**]
Institute of Information Sciences & Technology
Massey University
Private Bag 11222
Palmerston North
New Zealand
e-mail: t.m.simatupang@massey.ac.nz

Alan C. Wright
Institute of Technology & Engineering
Massey University
Private Bag 11222
Palmerston North
New Zealand
e-mail: a.c.wright@massey.ac.nz

Ramaswami Sridharan
Institute of Information Sciences & Technology
Massey University
Private Bag 11222
Palmerston North
New Zealand
e-mail: r.sridharan@massey.ac.nz

[*] To appear in *Supply Chain Management: An International Journal*, Vol. 9 No. 1, 2004.

[**] Corresponding author. e-mail: t.m.simatupang@massey.ac.nz

Applying the Theory of Constraints to Supply Chain Collaboration

Abstract

Supply chain collaboration amongst independent firms or business units often provides larger benefits from effectively satisfying end customer needs than working in isolation. However, a lack of awareness about the existence of constraints along the supply chain prevents the benefits of collaboration from being fully realised. This paper attempts to apply the Theory of Constraints approach to overcome difficulties in realising the potential benefits of supply chain collaboration. Specifically, it shows how the Theory of Constraints approach can be used to expose an inherent dilemma of collaboration and establish collaborative replenishment policy and collaborative performance metrics so that the chain members can work together to advance supply chain profitability. Several opportunities for future research are recommended.

Keywords: Supply chain collaboration, Theory of constraints, Performance measurement

Introduction

Supply chain collaboration can be defined as two or more independent firms jointly working to align their supply chain processes so as to create value to end customers and stakeholders with greater success than acting alone (Horvath, 2001; Simatupang and Sridharan, 2002). Collaborating firms share responsibilities and benefits by establishing a degree of cooperation with their upstream and downstream partners in order to create competitive advantage (Spekman *et al.*, 1998). When all the chain members in the chain integrate and act as a homogenous entity, performance is enhanced throughout the chain as the matching of supply and demand improves profit (Fisher, 1997). Bowersox *et al.* (2000) estimated that as little as twenty percent of the scope of improvement initiatives is within the responsibility of the focal or individual firm. The balance (80 percent) involves the areas of responsibility of managers from other partners. Thus, joint decision-making is preferable to create competitive advantage through mechanisms such as increased market access, better material sources, and cost-effective transportation.

Reports from real world practice show that supply chain collaboration brings benefits for all participating members (Ireland and Bruce, 2000). Under continuous replenishment programs, Procter and Gamble controls the stock management for its retailers and receives information about retailer inventory levels and demand conditions (Parks, 1999). From the time when Procter and Gamble were able to have higher visibility of actual customer demand, it experienced a five percentage points increase in perfect orders, less variability in retailer orders, and reduced delivery expenses by being able to utilise cube space in transit. Meanwhile, its retailers reaped benefits such as an increase of over 100% in inventory turns, inventory levels were significantly reduced, service levels increased, retail sales went up two percent, and storage and handling costs were reduced. Zara, the fashion retailer, synchronised its global production networks with customer requirements so that it would be able to respond quickly to the changing tastes of fashion-conscious customers (Walker *et al.*, 2000).

Although the main thrust in supply chain collaboration is to achieve a win-win solution for all participating members, there is often a large disparity between the potentials and the practice (Mentzer *et al.*, 2000). For some chain members, supply chain collaboration is limited to volume guarantees (Lambert *et al.*, 1996). The retailer, for example, requires continuity of supplies with lowered wholesale price. Furthermore, it is not surprising that "collaboration" often manifests as a power plays amongst the chain members. Such behaviour has the potential to result in "win-lose" rather than "win-win" outcomes. For example, large retail chains, due to their market power, often shift inventory costs, cycle time, and burdens of information technology to their upstream members (Betts, 1994; Munson *et al.*, 1999; Poirier, 1999). This inherent difficulty presents chain members with the considerable challenge of formulating the appropriate modus operandi that will enable them to realise the potential mutual benefits of supply chain collaboration.

The main objective of this paper is to discuss the potential of the Theory of Constraints (TOC) approach to deliver the benefits of supply chain collaboration.

There are two basic questions that need to be addressed: (i) "What is the main blockage to effective supply chain collaboration and how might this be overcome?" and (ii) "Are supply chains subject to constraints as defined by the TOC and how could the chain members carry out and realise the constraint-based improvement initiatives?" Two frameworks are suggested to answer these questions. First, the evaporating cloud diagram is used to capture a major dilemma in supply chain collaboration, and the second framework proposes strategies for managing collaborative replenishment and establishing collaborative performance metrics that motivate the chain members to improve the supply chain as a whole. It is hoped that this paper will stimulate practitioners and academicians to realise real benefits through applying TOC to supply chain collaboration.

The paper is organised as follows. The next section presents current literature on the application of the TOC to supply chain management. Next, a major dilemma of supply chain collaboration and discussion of how conventional management approaches fail to successfully address it are presented. The succeeding section presents how the constraint-based approach can be used to alleviate the dilemma. Strategies for improving supply chain profitability are outlined in the two separate sections: replenishment policy and performance metrics. Limitations of this paper and discussion of several concepts requiring future study follow. The concluding section summarises the main ideas of this paper.

Literature review

The theory of constraints (TOC) has been widely known as a management philosophy coined by Goldratt (1990a) that aims to initiate and implement breakthrough improvement through focusing on a constraint that prevents a system from achieving a higher level of performance. The TOC paradigm essentially states that every firm must have at least one constraint. Goldratt and Cox (1992) define a constraint as any element or factor that limits the system from doing more of what it was designed to accomplish (i.e., achieving its goal). The owner of a system is assumed to establish its goal. The fundamental goal of most business entities is to make money now and in the future. Other stakeholders may develop necessary conditions that must be met to allow the system to continue operating. The TOC thus encourages managers to identify what is preventing them from moving towards their goals - as well as necessary conditions - and find solutions to overcome this limitation.

The TOC comprises a set of three separate but interrelated areas – namely, logistics, performance measurement, and logical thinking (Cox and Spencer, 1998; Simatupang *et al.*, 1997). The TOC applications to logistics include the drum-buffer-rope scheduling method, buffer management, and the VAT analysis. Measurements are required to determine whether or not the system is accomplishing its goal of making money. Performance measurement includes operating measures (i.e., throughput, inventory, operating expenses) and local performance measures (i.e., throughput-dollar-days and inventory-dollar-days). Logical thinking comprises the five-step-focusing process and the thinking processes.

The TOC solutions initially attempted to resolve core problems in production systems using methods such as the drum-buffer-rope scheduling, constraint-focused performance measurement, and buffer management (Goldratt and Cox, 1992). Further development of the TOC incorporates solutions for marketing and sales (Goldratt, 1994), project management (Goldratt, 1997), and supply chain management (Goldratt *et al.*, 2000). Blackstone (2001) provides an exhaustive review of the latest developments of the TOC applications.

Over the last decade the development of the TOC and accounts of its application have burgeoned with the publication of a considerable number of articles, proceedings, and books based on the TOC approach (Mabin and Balderstone, 1999). Rahman (1998) reviews the TOC approach on manufacturing firms. Siha (1999) applies the TOC approach to addressing problems in different types of service organisations. Beyond business firms, Klein and Debruine (1995) and Dettmer (1998) used the TOC thinking processes to identify core problems in public policies. Womack and Flowers (1999) applied the TOC approach to the healthcare system to improve its performance.

Literature on TOC supply chain solutions deals mainly with managing the supply chain from a single enterprise perspective (Cox and Spencer, 1998; Jackson and Low, 1993). Umble *et al.* (2001), for instance, described how a manufacturing firm applied the TOC approach to direct the implementation of enterprise resource planning (ERP).

Gupta (1997) also recognises that the TOC approach can be used to guide a single firm to concentrate on exploiting resources based on different logistics cost along the supply chain.

Little attention has been given, within the literature, to the application of the TOC concepts to the management of supply chains where collaboration must be fostered between independent firms. Covington (1996) applied the TOC thinking process to identify problems in the apparel supply chain and describes the bringing together of managers from different firms to cooperate in improving the overall supply chain profit. Stein (1997) proposed a conceptual model of locating the time buffer at different positions of participating members to protect actual sales from demand and supply uncertainty. Goldratt *et al.* (2000) conceptualised performance measures to maintain trust amongst the participating members.

This paper follows up the previous research through clarifying how participating members can benefit from applying the TOC to supply chain collaboration and devising strategies to improve the supply chain performance as a whole. The next section presents an inherent dilemma amongst the chain members who want to maximise the benefits of collaboration.

A dilemma of supply chain collaboration

When chain members involve in collaboration, there is a dilemma between accommodating decisions that take into account the interest of the supply chain as a whole and preserving decisions in the interest of an individual firm. A evaporating cloud diagram can be employed to capture and describe the dilemma of supply chain collaboration between taking decisions based on link-centric-measures and taking decisions based on supply chain-wide measures as shown in Figure 1 (see Goldratt, 1994, and Dettmer, 1998, for further explanation of the evaporating cloud diagram).

Take in Figure 1

The upper path of the diagram in Figure 1 can be read as follows. "In order to maximise the benefits of collaboration (O), the chain members must maximise the revenue of the entire supply chain from sales to end customers" (R_1 to O on the diagram). The perceived rationale for this is that a supply chain really only makes a sale when the final customer consumes the product, hence for a supply chain as a whole to benefit there must be an effort to maximise revenues from final consumers. Reading on, "in order to maximise the revenue of the supply chain, the chain members must take decisions tied to supply chain-wide performance measures" (P_1 to R_1 on the diagram), which assumes that maximum revenue (i.e., more sales) can be created only if the chain members perform their efforts in a synchronised fashion and that they will need supply chain focused measures to help achieve this.

The lower path of the diagram can be read as follows. "In order to maximise the benefits of collaboration (O), the chain members must protect their individual profitability" (R_2 to O on the diagram). This is because chain members get benefits from collaboration only if their individual company profitability (i.e., bottom line) is enhanced. Further, "in order to protect the individual company profitability, the chain members must take decisions tied to link-centric-performance measures" (P_2 to R_2 on the diagram). This reflects the fact that chain members have direct control over their individual part of the supply chain and hence they are likely to be focused on this. It also reflects – the predominant indicator regimes in many companies – i.e. very functionally oriented and conveying the assumption that if functions perform well against their link-centric measures then the result will sum to the optimum for the system. This paradigm is, not surprisingly, further extended to the overall supply chain where individual companies assume that the best interests of the supply chain will always be served if they focus on being the best they can be internally – which is measured by link-centric measures. This assumption is also reflected in the conflict presented in P_1 and P_2 of the diagram.

The diagram dilemma can be read: "taking decisions based on supply chain-wide performance measures (P_1) is in direct conflict with taking decisions based on link-centric-performance measures (P_2)". Frequently, individual members tend to make decisions based on their individual link-centric measures rather than considering supply chain-wide measures. The first key assumption is that the chain members often think that supply chain collaboration means a decrease in bargaining power. Often, each individual member focuses its decisions to maximise myopic revenue (i.e., sales from immediate downstream partners) and minimise myopic costs (i.e., buying from immediate upstream partners) rather than to maximise the overall market expansion of the entire supply chain. The source of myopic revenue is the transferring of payments from exchanging inventories among the trading partners, because an inventory transaction before the consumption point of end customers is assumed to be a sale. Furthermore, the existence of cost reduction indicates a compensation policy that encourages self-optimisation. For example, perverse incentives such as periodical promotions encourage the retailer to sub-optimise its performance through diversion and forward buying at the expense of the overall supply chain revenue (Buzzell *et al.*, 1990). If each member takes decisions that maximise only its own performance and ignores the impact of its actions on other links of the supply chain, the maximisation of an individual performance often occurs at the expense of the performance of the entire supply chain (Simatupang and Sridharan, 2002).

Second, there is no clear, quantitative, link between the supply chain-wide measures and the individual performance metrics. Linkages, if any, are based on the assumptions and intuition of individual managers, hence are highly variable in quality and validity. If individual decisions have no strong link to supply chain-wide performance measures, then the chain members do not have effective internal indicators that will show indisputable connection to the results of collaboration and therefore they will be unable to determine whether or not the collaboration is working.

The traditional approach to addressing the dilemma of collaboration remains focusing on cost minimisation at each link of the supply chain (i.e., maintaining a heavy bias towards P_2, with minimal compromise towards P_1). Smith (2000) describes this situation as a zero-sum game in which each chain member bargains to save as much in costs as possible, often at the expense of the other partners. Lee *et al.* (1997) discovered that chain members take “rational” counter-decisions when other chain members bias decisions towards P_2. Typically, such decisions lead to actions such as multiple forecasting, shortage/rationing “games”, price fluctuations, and volume and transportation discounts. In support of the contention that the conventional bias remains on the P_2 side of the dilemma, Poirier (1999) found that retailers with higher bargaining power often require their suppliers to carry out (i.e., cost transferring) improvement initiatives such as time compression, vendor-held inventory, and bar coding. As a result, decisions apparently bring gains to individual members but have a devastating impact by lowering the overall supply chain profit.

To resolve the dilemma of supply chain collaboration, the chain members need to attack the logical connection at the $P_1 \leftrightarrow P_2$ part of Figure 1 (or alternatively the $R_1 \leftrightarrow P_1$ or the $R_2 \leftrightarrow P_2$ arrows). The proposed injection (solution direction) is mainly aimed at the $P_1 \leftrightarrow P_2$ arrow, but also breaks the $R_2 \leftrightarrow P_2$ arrow. In essence it suggests that supply chain focused measures can be used at the individual member level without jeopardising individual company profitability (in fact benefits to profitability would be expected). The proposal is to use Goldratt's measures of "Throughput", "Investment" and "Operating Expense", as applied to the total supply chain, and derivatives thereof, which drive the appropriate synchronisation actions of members.

It is important to note that the requirement that exists for individual members to maximise their profitability is not being challenged, in fact its position in the diagram indicates that it is recognised as a "necessary condition" of successful collaboration. It is also acknowledged that the mechanism of measuring company profitability (the Profit and Loss statement) is not the target of criticism within this paper. The problem captured by the $R_2 \leftrightarrow P_2$ arrow occurs at a more operational level. The level at which prices are negotiated (i.e., prices paid for inputs and price agreed for outputs) and also where decisions on inventory levels, minimum shipment quantities and modes, and scheduling (to name some) are made. At these levels measures contain assumed connections to the overall profit statement, but remain very myopic in focus, with the frequent outcome that the size of potential profit for the supply chain as a whole is reduced. For example, a purchasing manager who wins a cost reduction after a tough negotiation on a purchased part will assume this will result in a company profitability increase; whereas it may in fact lead to reduced product quality, or delivery problems, causing less customers to buy. Similarly, a sales manager may set a higher selling price for product and make the, assumed, connection of the action to a positive profitability increase. The actual outcome may well be that the price increase is passed on down the chain causing an increase in the price to the end-consumer. Such an increase may result in the end-price being pushed beyond a critical price-point in a less than smooth supply and demand curve, thus dramatically reducing the market size.

As the problem is occurring at the operational level it is important that any new measures suggested can be used at that level, that they do give a more holistic view and linkage to overall supply chain maximisation. It is also vital that existing measures and manager performance indicators are reviewed (and removed if necessary) to ensure that manager behaviour is in synch with the new measures.

Many questions need to be answered if the usage of T, I, OE is to be successful. Some of the key questions are:

- How will an individual supply chain member be able to judge the change in supply chain T, I, OE caused by their actions or proposed actions?
- How will the supply chain members decide how to share the overall "cake" (difference between T_{SC} and OE_{SC}) and improvements to it?
- What aspects of the physical supply chain operation and dynamics need understanding and changed in order to better support the use of such measures?
- What aspects of the typical policy and measurement regimes in companies will need change in order to assure no clash of measurement systems occurs and thus synchronisation of behaviour with appropriate objectives follows?
- What additional operational measures are necessary and how can it be determined that these have sufficient causal connection to supply chain T, I, OE to assure the appropriate actions by members?
- What technological changes are necessary, or desirable, to support this activity?

This paper will explore potential answers to most of these questions but it is acknowledged that many details will need to be the subject of further research. Similarly it is acknowledged that the relevance of the ideas to all types of supply chains (e.g., divergent versus convergent versus mixed divergent/convergent product flow) has yet to be explored.

The throughput accounting measures

The constraint-based approach can be defined as a way of realising productive change that alleviates the detrimental impact of the constraint(s) on supply chain profitability. The productive change focusing on actions of managing constraints(s) can directly contribute to profitability. There are two ways in which the constraint-based approach can help managers improve the supply chain: (i) by providing reliable global performance measures that help the chain members to measure the progress of accomplishing the total revenue of the supply chain and (ii) by focusing on improvement efforts that have a dramatic impact on the supply chain performance.

The supply chain can be viewed as a system established for the purpose of accomplishing a system's goal. In order to ascertain whether or not the supply chain is accomplishing its goal of making money, three global (system level) measurements are used – namely, Throughput, Investment, and Operating Expenses (e.g., Dettmer, 1998; Goldratt and Cox, 1992). Throughput (T) is the rate at which the supply chain generates money through sales. This measures all the money coming into a supply chain from the outside. This definition excludes internal transfer pricing. Throughput is comprised mainly of the revenue a supply chain generates through sales of its product, less the truly variable cost of generating the sale. For many supply chains, truly variable costs include the material costs, sales commissions, markdowns, consumable supplies, and so on. Investment (I) is all the money the supply chain invests in things it intends to sell. The larger portion of this investment for a supply chain is in the inventory of various forms. This definition excludes the added value of labour and overheads. In a supply chain, inventory investment comes in three forms: raw materials, finished products not as yet sold, and things somewhere in between (work-in-process). Investment in plant, equipment, and buildings also falls into this category. Operating Expense (OE) is all the money the supply chain spends in turning investment into throughput. This includes direct labour and overheads and other fixed and semi-fixed expenses that would be incurred even if it did not produce a single product. The chain members need to rank the increase of Throughput higher on the priority list than reducing Investment and Operating Expenses because the potential to increase this measure is virtually unbounded, whereas the potential to reduce I and OE is limited (but not to be ignored). In addition to this scale of importance, all other local or departmental metrics are less important than the global, or supply chain-wide metrics.

All these measures are monetary in nature. Throughput is calculated as sales, less any sales discounts, less any expenditures that vary directly with product units, i.e. - any truly variable costs (TVC). In other words, $T_{unit} = Sales_{unit} – TVC_{unit}$ or $T_{total\ for\ an\ accounting\ period} = T_{per\ unit}$ x Number of units sold. To determine this for a supply chain there needs to be agreement on the start and end points of the chain subject to the collaboration. The members must be willing to share information about the volume and costs of raw materials and other truly variable inputs arriving from outside the chain at various locations. The sales price used should be that achieved or proposed at the end-point of the defined supply chain. This information can be extended, through research, to determine the demand probability function at various price-points. Such price versus volume information can then be fed back up the chain to determine raw material prices at different volume levels and thus modelling of the potential profitability of the supply chain under different price regimes can occur. It may well be that a lower price to the end-consumer results in such increases in volume that the total chain profitability is hugely enhanced. This approach is a necessary prerequisite if the cell-centric approach of adding standard margins on top of within-chain transfer prices is to be eliminated. Such an approach can easily result in prices which cause low demand and hence low profitability to all chain members.

The supply chain Net Profit is equal to Throughput minus Operating Expenses for a given period or NP = T-OE. The supply chain Productivity is essentially the ratio between Throughput and Operating Expense or P = T/OE. The supply chain Return-On-Investment is Net Profit divided by Investment or ROI = (T-OE)/Investment ($I_{materials} + I_{others}$). Cash flow is Net Profit plus or minus the change of Investment for the same period or $CF = T – OE \pm \Delta I$. To share this kind of information is a big step for many companies and needs to be initiated and lead by top management of each of the chain members. Sustaining the process will also require the creation of executive-level positions focused on such supply-chain wide decisions and collaboration. These executives will need sufficient authority and a policy framework that enables them to influence the internal functions of their own unit so that decision-making is in synch with the concept of a supply chain “win-win”. The provision of appropriate projections of supply chain T, I, and OE will help quantify the potential benefits and clarify the rationale of decisions.

A key additional requirement for the above to work is for companies within a supply chain to come to agreement on the sharing of the total T for the supply chain. Negotiation in this case would demand the same skills as currently experienced between units of a supply chain, but would have to become more multi-party and involve all contributors to the “throughput chain”. Negotiation power would depend on many factors (as per the current case) but the authors envisage that the most equitable way to share the total T would be in some relationship to the amount of OE and I that each contributor has to bring to the supply chain in order for it to be viable. In this instance the degree of revelation of OE and I may depend on the nature of the negotiation and its participants. Members may also negotiate with one another for revised shares of the T based on actions which they can take to help contribute to reductions in the I and OE of other members. At this stage the operational dynamics of supply chains becomes a factor and hence must be understood. In addition answers to the question of what is appropriate technology to invest in become clearer (and include a value equation expressed in terms of the expected change in T, I, and OE)

Supply chain dynamics and the constraint-based approach

Supply chain dynamics and the interdependence of the logistics processes within the supply chain, means that many initiatives will have only a small impact on supply chain profitability. Only those initiatives that focus on the constraint(s) and result in improvement in its performance will have a significant positive effect on supply chain performance. For a "for profit" supply chain, the constraint would be whatever keeps the chain members from generating more profits. Such supply chain has at least one constraint otherwise infinite profit would result. Thus the dilemma of supply chain collaboration can be resolved if the chain members realise that the real constraint to the overall system may reside outside their “local” control (hence “local” measures won’t help much to elevate their profits). To really elevate their profitability they must take a supply chain-wide view and try to help achieve better performance at the constraint. This approach, with the added coexistence of an appropriate mechanism to measure and share improvements, can lead to real benefits for all supply chain members.

The constraint-based approach recognises the importance of identifying the constraint(s) that prevent the chain members from satisfying a necessary condition or achieving overall profitability. Agreement about the types and locations of the constraint(s) is crucial for initiating supply chain improvement. The nature of a constraint can be either physical or non-physical and its location can be either internal or external. Physical constraints can take the form of raw material shortages, limited capacity resources, limited distribution capacity, and lack of customer demand. Non-physical constraints include obsolete rules, procedures, measures, training, and operating policies that guide the way in which decisions are made. The location of a constraint can be either internal or external. Internal constraints – such as the raw material constraint, the capacity constraint, and the distribution constraint – are located inside the supply chain. External constraints include market constraints and supply constraints which arise outside the agreed boundaries of the collaborative supply chain.

Different types of constraints are interrelated with each other. Physical constraints are often driven by human behaviour (i.e., habits, decisions, and actions). At the same time, human behaviour is affected by metrics and policies around supply chain operations. When the goal of a supply chain is created, the chain members create policies that regulate and guide the behaviour of the supply chain – including performance metrics – to assess the attainment of the goal. However, with changes in the environment of the supply chain and within the supply chain itself, these policies and metrics often become unsuitable for motivating desired decision-making behaviour and thus result in negative impacts on supply chain performance. People are comfortable with regular habits and predictable outcomes, and hence often continue to use outdated policies and inappropriate metrics in making decisions even though supply chain technologies or the business environment has changed (Goldratt, 1990a; Simatupang and Sridharan, 2002). Besides outdated policies and inappropriate metrics, flawed mental models and past training can be constraints that affect human behaviour. Counterproductive decisions and actions cause resources to be used in ways which do not maximise profitability. It is common that physical constraints reflect, and are sustained by, obsolete policies and inappropriate metrics.

The improvement process focused on the constraint can be summarised as follows: (a) the current supply chain profitability is determined and dictated by the constraint that exists within it, (b) the profitability can be improved through better exploitation of physical constraints and, ultimately, if the constraint is permanently removed, and (c) when the constraint is removed, the supply chain moves to a higher level of profitability, and immediately encounters another constraint. This process continues again to find and manage a new constraint. The formal procedure of the focusing process is known as the five-step-focusing process (Goldratt, 1990a). The first step is to identify the constraint in the supply chain that limits supply chain profitability. The chain members must focus on the constraint as the performance of the entire supply chain depends on it. As discussed earlier, many types of constraints – such as raw materials, capacity, distribution, and market – exist along the supply chain. It seems that the constraint moves from one resource to another resource. However, in many cases, the resource constraint is caused by local optima rules that attempt to reduce costs in production, distribution, and marketing. For example, the retailer often operates based on min-max inventory levels for each product. If the product availability reduces to below a minimum level, then the retailer places a big order up to the maximum level. Use of this rule is an attempt to reduce transportation costs because a small order size is not sufficient to fill the truck. In many cases, however, the retailer has the same supplier that delivers various products to the stores. In this case, the retailer and the supplier need to change this replenishment policy to reflect the actual consumption rate in which the supplier replenishes the stores based on number of products sold.

The second step is to decide how to exploit the constraint. This means optimising the existing capacity at the constraint, which is frequently wasted by making and selling the wrong mix of products and by improper rules for scheduling and controlling. The chain members need to ensure that the identified constraint is working on the right products to maximise profits. For example, if immediate product availability is causing a capacity constraint to go idle, the supplier has to observe the constraint and continuously replenish the constraint without delays.

The third step is to subordinate all other activities to the constraint. This means to change traditional rules and metrics that discourage all other activities to support decisions to exploit the constraint. In supply chains, this is an important area for collaborative decision-making as individual parts of the chain may be able to help exploit a capacity constraint through changes to the way they operate – i.e., being subordinated.

The fourth step is to elevate the constraint. This means to increase the capacity of the constraint to a higher level. For instance, a supplier can elevate the capacity constraint by redesigning products to allow postponement of product diversity until points in the supply chain where fast response to changing customer needs can be achieved (Lambert *et al.*, 1998).

The fifth and last step of the focusing process is to prevent inertia from stopping the process of continuous improvement. If the constraint is broken in step four, a new constraint may appear somewhere else in the supply chain. This means that the chain members need to focus on the new constraint, return to step one, and repeat the process. Collaborative sharing of information will ensure that constraint changes are quickly identified and revised tactics, policies, and measures are put in place to refocus the chain appropriately.

This description of the global operating measures of the supply chain and the five-step-focusing process reveals that there are two implications for the chain members. The first implication is the importance of modifying the improper business rules amongst the chain members. For example, the traditional "rule" that says "each member gets sales from other chain members" needs to be altered to "the chain members get paid when they sell products to end customers". This means that once the retailer has obtained payment from the end customer the downstream partners must be paid from their proportion of supply chain throughput (Covington, 1996; Goldratt *el al.*, 2000). This incentive scheme encourages both parties to focus on throughput and speed of flow because their individual profitability and cash flow depends on their collaborative efforts to increase supply chain throughput.

Second, the chain members would be able to identify improvement initiatives that focus on the supply chain constraint and at the same time contribute to supply chain Net Profit (NP), Return on Investment (ROI), and Cash Flow (CF). The chain members will be better able to understand the impacts of operational decisions on the overall supply chain goal because Throughput, Investment, and Operating Expense provide a usable link to NP, ROI, and CF. The potential change in T and OE (and hence NP) that an improvement brings can be determined through the usage of operational level questions such as: "Will it enable more saleable product to flow through the capacity constraint?", "Will it increase sales?", "Will it increase responsiveness and will customers respond to it positively?", "Will it reduce material costs?", "Will it reduce markdowns?", and "Will it reduce fixed expenses?" Similarly changes in I (and hence ROI) can be determined by additional operational questions such as: "Will it shorten the time between product delivery and time of payment?", "Will it increase the volume of sales in the same time period?", "Will it increase the time between order placement and order delivery?", "What impact on inventories within the supply chain will it have?", "Will it need less equipment (i.e., reducing I)?", and so on.

Providing the chain participants attempt to answer these questions collaboratively, then the real potential impact of improvement initiatives on the supply chain can be determined.

Collaborative replenishment policy

Another key question is: "How should a supply chain be changed in terms of its inventory policies and information handling to facilitate better collaboration and the usage of supply chain wide perspectives and measures? The constraint-based approach proposes the need for a replenishment system to better match supply with demand, in conjunction with applying information technology for control and to accelerate the improvement process (Goldratt *et al.*, 2000). In the make-to-stock supply chain, the constraint can be viewed as the market – i.e. end customers who come to the store to buy products. In order to exploit this constraint, the supply chain needs to have the right product in the right place at the right time (Fisher, 1996). The idea is to authorise the chain member with the best knowledge and intuition to make local decisions that improve supply chain profitability. In the retailer and supplier collaboration, the retailer is responsible for taking market initiatives that are beneficial to all members of the supply chain. On the other hand, the supplier is responsible for ensuring fast response delivery to fulfil market customer needs.

In addition to focus on customer service and pricing, the retailer is also able to gather timely information about customer behaviour and product life cycle. Customers need to be segmented along different dimensions such as product features, availability, quantity and price discounts, and credit terms. Customer segmentation helps the retailer to prioritise a variety of services based on profitability, key customers, and operational requirements (Levy, 1999). It also enables differentiated product offerings to be made, with the intent to maximising the total Throughput dollars earned by the supply chain.

The constraint-based approach uses the water supply system analogy as a reference system to refer to the supply chain characterised by on-time delivery that is robust to demand fluctuations, self-managing, and where costs are kept low. The water supply system does not store water at customers' homes; rather, customers draw water only when they need it. Instead, it has several focal points to store water to absorb the fluctuations of water consumption. Similarly, the supplier can strategically place buffers at key focal points to protect throughput so that holistic performance is enhanced significantly. Figure 2 depicts that there are three focal points along the supply chain that require buffers: retailer buffer, warehouse buffer, and production buffer. The store holds just enough on-hand stocks (i.e., the retailer buffer) to cover demand of end customers and replenishment time from the warehouse for a given period. The supplier watches the buffer and replenishes it with what has been sold. The formula for retailer buffer level is Peak Customer Actual Consumption times Reliable Replenishment Time from the warehouse times a Safety Factor of 1.5 (Goldratt, 1994). The warehouse holds only enough inventories to satisfy the expected demands of the stores during the time it takes to reliably replenish the warehouse with what was actually delivered. The warehouse buffer level equals the Peak Monthly Consumption of all stores in the region times Reliable Replenishment Time from manufacturing times a Safety Factor of 1.5. The manufacturer holds just enough inventories to replenish what the warehouse has delivered. The manufacturing buffer level approximately equals the Peak Total Consumption of the whole chain times Reliable Manufacturing Frequency times a Safety Factor of 1.5.

Take in Figure 2

The supplier who has intimate knowledge about product designs and responsiveness is authorised to decide the quantity and timing of its product deliveries to the retailer. The idea of replenishment policy is that sales would be protected if the supplier should replenish what the market has consumed at the stores. Key to this system is to understand that "replenishment" means replacing what was sold by the next step in the chain during a fixed (short) time interval. Replenishment orders quickly pass up the system giving rapid confirmation of what is actually happening at the market place. Such orders are not distorted by localised imposition of "economic order quantity" rules. The system is thus far less dependent on the accuracy of forecasting. The buffer inventories help absorb any forecasting inaccuracies and buffer management enables the system to rapidly learn about demand pattern changes. The requirement for e-technologies is thus not paramount, but these could provide a powerful facilitator of the concept in many circumstances.

Buffer management is used at all replenishment inventory points to control corrective replenishment actions and to enable a "learning system" approach to buffer size adjustments. As the buffer size reflects consumption patterns, it can be closely monitored as the basis to determine appropriate actions. Often the buffer size is divided into three zones: green, yellow, and red. The green zone represents the replenishment level of the buffer (i.e., the up-to policy). The red zone represents when levels have fallen such that a high probability of losing sales exists and therefore requires emergency delivery. The supplier does not need to worry if the green zone has some of its buffer consumed. If the buffer of the yellow zone has begun to be consumed, then the supplier needs to watch the ongoing consumption closely and makes plans for the quantity to order to assure replenishment to the top of the green zone (usually such ordering occurs at the regular time interval). Should the red zone be penetrated, the supplier should take action to expedite the replenishment to the top of the green zone.

There are considerable gains for the supplier who has decision right to manage inventory at its retailer sites. The supplier can focus on compressing the length and variations of total lead-times based on different customer segments that want to pay for faster delivery time (Blackstone, 2001). Second, the supplier may enjoy lowered inventory levels due to demand risk pooling (Goldratt, 1994). Third, the supplier could apply the concept of postponement - that is, to design the products to commit to final product differentiation nearer to the time of purchase in order to reduce costs from risk of inventory and demand uncertainty (Lambert *et al.*, 1998). Postponement enables the supplier to observe more demands before deciding on how much inventory is delivered to any particular retailers. Fourth, the supplier would be able to combine shipments for multiple retailer stores and thereby could dispatch a full truck (Stenzel and Stenzel, 2002).

Collaborative performance metrics

Supply chain collaboration assumes a fair amount of trust amongst the partners. For example, the retailer trusts the supplier to manage its inventory and the supplier trusts the retailer to effectively serve the market with its products. The success of supply chain collaboration thus depends very much on how each party abides by the collaborative scheme. Collaborative performance metrics are required to guide participating members in evaluating whether or not their actions are truly contributing to the global goal (Goldratt *et al.*, 2000). Collaborative performance metrics consist of a set of metrics that specify how the progress of collaboration is evaluated at the level of both the individual link and the overall link.

The authors propose a hierarchy of collaborative performance metrics to assess progress towards a supply chain goal. As shown in Figure 3, the chain members monitor the progress of supply chain profitability, competitive factors, individual performance metrics, and supporting metrics. First, global performance metrics shows the indicators of supply chain profitability such as net profits, return on investment, and cash flow in terms of supply chain T, I, and OE. Second, competitive factors represent product and service features recognised by the market as superior compared to competitors such as service, quality, price, and fast response. These factors support the supply chain goal of continuously increasing profits, return on investment, and cash flow. Third, individual performance metrics are important to allow each participating member to be evaluated as well as to assess other partners. Finally, the individual chain member also monitors supporting metrics such as daily reporting on constraint rate, status of all buffers, customer satisfaction and loyalty, and identification of customer needs.

Take in Figure 3

Individual performance metrics should be able to induce the chain members to follow a global plan of the supply chain as a whole, which is to deliver the right products at the right time to the right place (i.e., creating Throughput) using no more than a specific level of Inventory and a specific level of Operating Expense. Since global operating metrics are measured by Throughput, Inventory, and Operating Expenses and each chain member should control the execution of the given global plan in its particular area, then the chain members should measure their individual performance in terms of Throughput, Inventory, and Operating Expense (Goldratt, 1990b). There are three individual performance metrics that can be used by the individual member to measure how well it executes the global plan: Throughput-Dollar-Days (TDD), Inventory-Dollar-Days (IDD), and Local-Operating-Expense (LOE). TDD measures failures to deliver specific products by specific dates. IDD measures the value of inventory and the time its stay within an area. This metric can be used to indicate excess inventory. LOE measures variances between actual and planned spending.

Besides using the three local metrics to control supply chain operations, the chain member can also use these metrics to judge how other participating members contribute to global metrics. The supplier can use IDD to judge the retailer's inventory performance. IDD equals the sum of the dollars of inventory times the number of days on hand. The supplier can also offer the same measure to its vendors so they can measure the supplier's inventory performance. On the other hand, the retailer can judge the delivery performance of its suppliers by using the TDD that equals the sum of sales dollars times the number of days' delay. An order is considered delayed if the buffer is below the emergency level. Each player attempts to reach zero TDD with as few IDD as possible. In the same way, the supplier can use TDD to judge the delivery performance of its vendors. The supplier will be responsible for the results as measured by TDD and IDD. In this way, IDD and TDD are performance metrics that help participating members to have uniformity of accountability. This means that if the retailer wants fast delivery of certain products, the supplier will fulfil this request and is able to judge how rapidly the retailer will sell the products. At the same time, the supplier can also require its vendors to provide the same delivery service. This chain of accountability stimulates participating members to improve the mutual response to end customers.

Discussion

This paper attempts to apply TOC to supply chain collaboration in order to assist the chain members to realise the benefits of collaboration. Breaking the inherent dilemma of collaboration implies the use of the constraint-based approach. This approach encourages the chain members to manage constraints that prevent them from optimising supply chain profitability. Any improvement initiative should be based on the knowledge of constraints because the act of exploiting or removing a constraint brings a supply chain closer to its goal. Two strategies to realise this concept have been presented: collaborative replenishment policy and collaborative performance metrics.

Collaborative replenishment policy consists of new operational rules that enable the supplier to quickly replenish the stock as the end customer draws it. The retailer buffer used for immediate sales is the inventory interface between the retailer and the supplier. The retailer and the supplier need to determine replenishment and emergency levels. On one hand, the supplier wants to have a large emergency level to minimise Throughput-Dollar-Days. On the other hand, the retailer wants a small replenishment size to minimise Inventory-Dollar-Days. Although the process of the buffer size is an experimental learning situation for them, there is a need for future research to quantify the necessary replenishment and emergency levels.

Collaborative performance metrics are required to ensure that each chain member is doing what it should do to create more Throughput dollars. TDD is used to evaluate the performance of a supplier and IDD can be used to assess the performance of a retailer. The TDD counter for make-to-stock supply chain can be extended to the make-to-order supply chain. The difference relates to how to count the correct number of days to calculate TDD. For the make-to-stock supply chain, the TDD counter ticks once the on-hand inventory is below the emergency level. For the make-to-order supply chain, the count of days starts ticking after the order has been delayed beyond the promised date. Moreover, the new business rule under the TOC approach is that transfer payment (i.e., the supplier portion of Throughput) to the supplier is made immediately following each daily sale by the retailer. The chain members' profit function clearly depends on collaborative actions to eliminate lost sales with minimum on-hand stock. This scheme has a self-enforcing property (Klein, 1996) that is able to encourage a supplier to provide fast response and a retailer to provide excellent service. However, this incentive scheme can be based on outcomes that are observable after the fact and rewards can be granted if the participating members do honour the scheme. Future research is thus required to formalise the self-enforcing property of collaborative performance metrics.

The TOC approach to supply chain collaboration currently focuses on the forward supply chain. Yet, there is no report on the effectiveness of the approach when applied to reverse logistics. Reverse logistics deals with managing both return products heavily driven by customer returns and product packaging that involves recycling product such as plastic and cardboard to reduce disposal costs. Nowadays, reverse logistics is becoming more important because of liberalised returns policies and a growing emphasis on customer service and part reuse, especially for automotive spare parts and perishable products (Carter and Ellram, 1998). The route of the supply chain has been widened to accommodate the backward flow of goods from the customers to the suppliers. The combination of fast response on forward logistics and proactive reverse logistics leads to the unique capability of a supply chain. This is worthy of future research.

Conclusion

This paper has provided the conceptual framework for using the TOC approach to assist the chain members to realise the potential benefits of supply chain collaboration. The inherent dilemma can be broken if the chain members manage the few constraints as their common denominator to achieve the overall goal. The goal of any collaboration is to increase Throughput and at the same time reduce Inventory and Operating Expense. As a prerequisite to ensuring profitability, the chain members must be able to quickly identify and remove the constraint(s) and ensure that they can continue to meet changing customer requirements accurately.

Collaborative replenishment policy has been proposed to authorise a supplier to decide when and how much stock should be delivered to a retailer's site as it is sold. Collaborative performance metrics has also been presented to encourage the individual chain members to contribute to the goal of optimising supply chain profitability. These two strategies help the chain members to ensure the swift and smooth flow of products to end customers and maintain the level of trust amongst themselves. However, the application of TOC should be adopted with care due to its intensive training requirements and radical approach that requires experimental learning. Further research is recommended to refine the TOC approach in dealing with the quantification of replenishment and emergency levels, the evaluation of the self-enforcing property of collaborative metrics, and the inclusion of reverse logistics.

References

Betts, M. (1994), "Manage my inventory or else", *Computerworld*, Vol. 28 No. 5, pp. 93-5.

Blackstone, J.H. (2001), "Theory of constraints – a status report", *International Journal of Production Research*, Vol. 39 No. 6, pp. 1053-80.

Bowersox, D.J., Closs, D.J. and Stank, T.P. (2000), "Ten mega-trends that will revolutionise supply chain logistics", *Journal of Business Logistics*, Vol. 21 No. 2, pp. 1-16.

Buzzell, R.D., Quelch, J.A. and Salmon, W.J. (1990), "The costly bargain of trade promotion", *Harvard Business Review*, Vol. 68 No. 2, pp. 141-49.

Carter, C.R. and Ellram, L.M. (1998), "Reverse logistics: a review of the literature and framework for future investigation", *Journal of Business Logistics*, Vol. 19 No. 1, pp. 85-102.

Covington, J.W. (1996), *Tough Fabric: The Domestic Apparel and Textile Chain Regain Market Share*, Chesapeake Consulting, Severna Park, MD.

Cox, J.F. and Spencer, M.S. (1998), *The Constraints Management Handbook*, Lucie Press, Boca Raton, FL.

Dettmer, H.W. (1998), *Breaking the Constraints to World-Class Performance*, ASQ Quality Press, Milwaukee.
Fisher, M.L. (1997), "What is the right supply chain for your product?", *Harvard Business Review*, Vol. 75 No. 2, pp. 105-16.
Goldratt, E.M. (1990a), *Theory of Constraints*, North River Press, Croton-on-Hudson, NY.
Goldratt, E.M. (1990b), *The Haystack Syndrome*, North River Press, Great Barrington, MA, pp. 144-55.
Goldratt, E.M. (1994), *It's Not Luck*, North River Press, Great Barrington, MA.
Goldratt, E.M. (1997), *Critical Chain*, North River Press, Great Barrington, MA.
Goldratt, E.M. and Cox, J. (1992), *The Goal: A Process of Ongoing Improvement*, North River Press, Croton-on-Hudson, NY.
Goldratt, E.M., Schragenheim, E. and Ptak, C.A. (2000), *Necessary But Not Sufficient*, North River Press, Croton-on-Hudson, NY.
Gupta, S. (1997), "Supply chain management in complex manufacturing", *IIE Solutions*, Vol. 29 No. 3, pp. 18-23.
Horvath, L. (2001), "Collaboration: key to value creation in supply chain management", *Supply Chain Management: An International Journal*, Vol. 6 No. 5, pp. 205-07.
Ireland, R. and Bruce, R. (2000), "CPFR only the beginning of collaboration", *Supply Chain Management Review*, Vol. 4 No. 4, pp. 80-8.
Jackson, G.C. and Low, J.T. (1993), "Constraint management: a description and assessment", *International Journal of Logistics Management*, Vol. 4 No. 2, pp. 41-8.
Klein, B. (1996) "Why hold-ups occur: the self-enforcing range of contractual relationships", *Economic Enquiry*, Vol. 34 No. 3, pp. 444-63.
Klein, D. and Debruine, M. (1995) "A thinking process for establishing management policies", *Review of Business*, Vol. 13 No. 3, pp. 31-7.
Lambert, D.M., Emmelhainz, M.A. and Gardner, T.J. (1996), "Development and implementing supply chain partnership", *International Journal of Logistics Management*, Vol. 7 No. 2, pp. 1-17.
Lambert, D.M., Stock, J.R. and Ellram, L.M. (1998), *Fundamentals of Logistics Management*, Irwin/McGraw-Hill, Boston.
Lee, H.L., Padmanabhan, V. and Whang, S. (1997), "The Bullwhip Effect in supply chains", *Sloan Management Review*, Vol. 38 No. 3, pp. 93-102.
Levy, M. (1999), "Revolutionizing the retail pricing game", *Discount Store News*, Vol. 38 No. 18, p. 15.
Mabin, V.J. and Balderstone, S.J. (1999), *The World of the Theory of Constraints*, Lucie Press, Boca Raton, FL.
Mentzer, J.T., Foggin, J.H. and Golicic, S.L. (2000), "Collaboration: the enablers, impediments, and benefits", *Supply Chain Management Review*, Vol. 4 No. 4, pp. 52-7.
Munson, C.L., Rosenblatt, M.J. and Rosenblatt, Z. (1999), "The use and abuse of power in supply chains", *Business Horizons*, Vol. 42 No. 1, pp. 55-65.
Parks, L. (1999), "CPR investment pays off in many ways", *Drug Store News*, Vol. 21 No. 2, p. 26.
Poirier, C.C. (1999), *Advanced Supply Chain Management*, Berrett-Kochler Publishers, San Francisco, CA.

Rahman, S. (1998), "Theory of constraints: a review of the philosophy and its applications", *International Journal of Operations & Production Management*, Vol. 18 No. 4, pp. 336-55.

Siha, S. (1999), "A classified model for applying the theory of constraints to service organisations", *Managing Service Quality*, Vol. 9 No. 4, pp. 255-64.

Simatupang, T.M., Hurley, S.F. and Evans, A.N. (1997), "Revitalizing TQM efforts: a self-reflection diagnosis based on the theory of constraints", *Management Decision*, Vol. 35 No. 10, pp. 746-52.

Simatupang, T.M. and Sridharan, R. (2002), "The collaborative supply chain", *International Journal of Logistics Management*, Vol. 13 No. 1, pp. 15-30.

Smith, D.W. (2000), *The Measurement Nightmare*, Lucie Press, Boca Raton, FL.

Spekman, R.E., Kamauff, J.W. and Myhr, N. (1998), "An empirical investigation into supply chain management: a perspective on partnerships", *International Journal of Physical Distribution & Logistics Management*, Vol. 28 No. 8, pp. 630-50.

Stein, R.E. (1997), *The Theory of Constraints*, Marcel Dekker, New York.

Stenzel, J. and Stenzel, C. (2002), "ERP system opportunities and limitations: an interview with Eli Goldratt", *Journal of Cost Management*, Vol. 16 No. 2, pp. 5-12.

Umble, M., Umble, E. and von Deylen, L. (2001), "Integrating enterprise resources planning and theory of constraints: a case study", *Production and Inventory Management Journal*, Vol. 42 No. 2, pp. 43-8.

Walker, B., Bovet, D. and Martha, J. (2000), "Unlocking the supply chain to build competitive advantage", *International Journal of Logistics Management*, Vol. 11 No. 2, pp. 1-11.

Womack, D.E. and Flowers, S. (1999), "Improving system performance: a case study in the application of the theory of constraints", *Journal of Healthcare Management*, Vol. 44 No. 5, pp. 397-407.

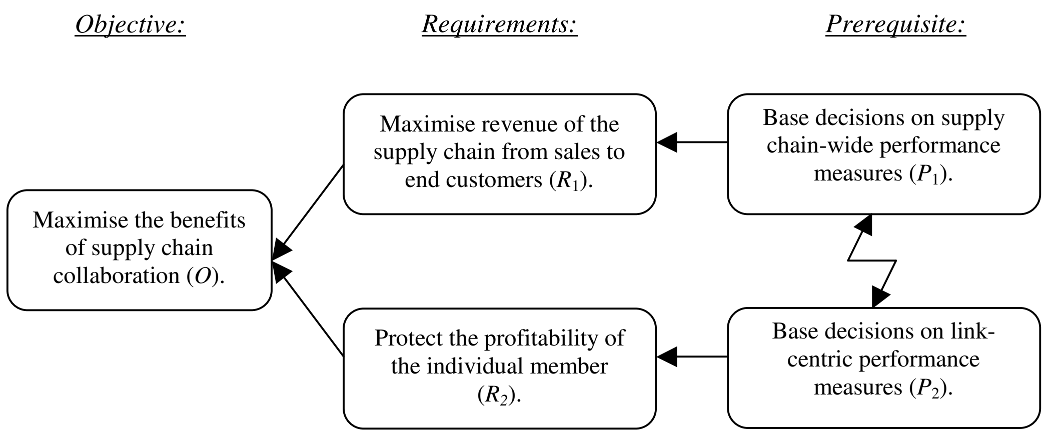

Figure 1. A dilemma of supply chain collaboration

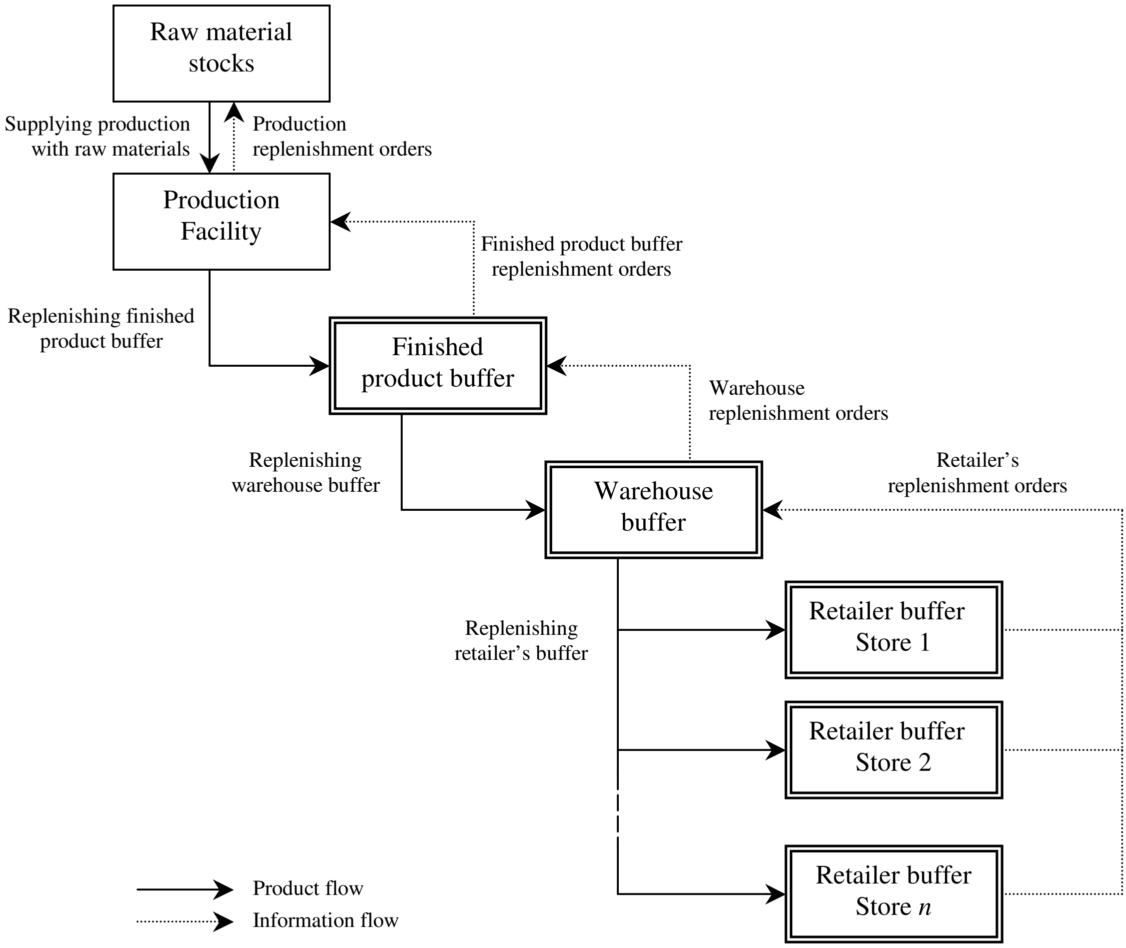

Figure 2. Replenishment policies at different points of the supply chain

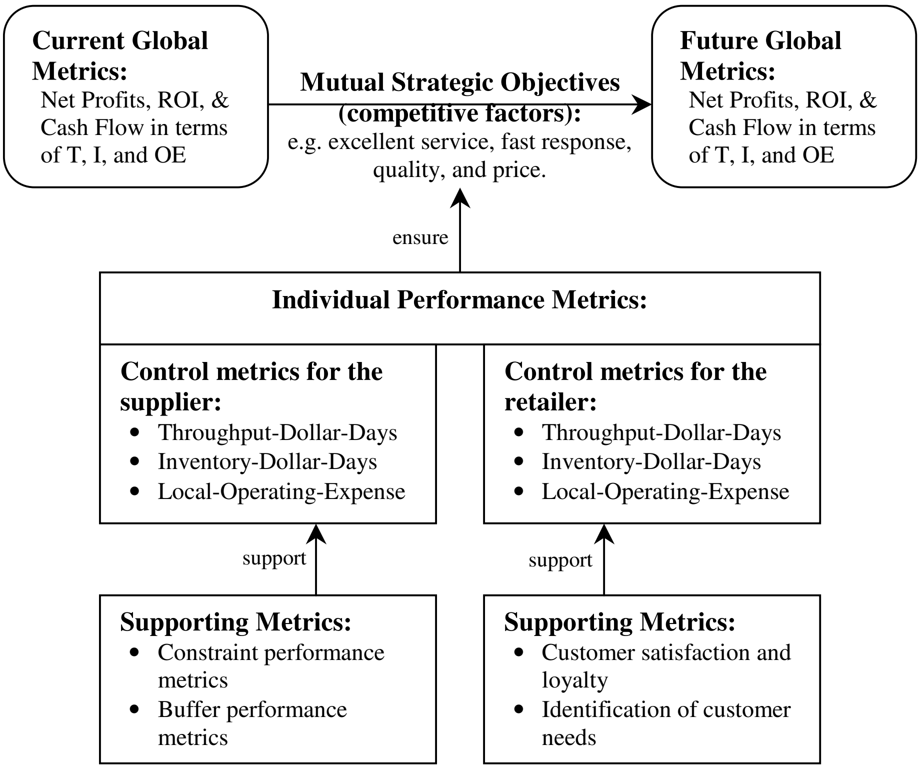

Figure 3. The hierarchy of collaborative performance metrics