

Gerenciando a Distribuição de Acordo com os Princípios da TOC

Por Amir Schragenheim, www.inherentsimplicity.com

Traduzido por Adail Muniz Retamal, www.heptagon.com.br

A Prática Atual de Gerenciamento de Cadeias de Abastecimento

É quarta-feira à tarde. Entro numa mercearia e quero comprar algumas pimentas verdes. Porém, eles não têm nenhuma em estoque. Tampouco consigo encontrar tomates bonitos. Continuo até uma loja de eletrônicos. Ouvi boas críticas sobre um novo mouse que a Microsoft lançou e gostaria de obter um. Entretanto, chego até uma estante vazia com apenas uma descrição dizendo “esgotado”. Quantas vezes você foi a uma sapataria, procurou um maravilhoso par de sapatos que você queria comprar, mas eles não tinham nenhum do seu tamanho? Por que as lojas não mantêm os estoques certos para atender à demanda? Por que eles não conseguem fazer nada direito?

As cadeias de abastecimento da nossa era moderna operam de uma maneira que parece fazer muito sentido. Fabricantes possuem maquinário robótico para automatizar os processos; muitos fabricantes que operam hoje em dia já instalaram o estado da arte dos novos sistemas ERP para ajudá-los a gerenciar o chão de fábrica.

Distribuidores e fabricantes possuem software muito sofisticado de previsão de vendas, para prever quantos itens serão vendidos de cada produto ou item de estoque (SKU¹). Portanto, eles deveriam saber quantas unidades eles gostariam de enviar para os pontos de consumo (lojas de varejo) e quando.

Como é que as organizações continuam a experimentar problemas no gerenciamento das cadeias de abastecimento? A tecnologia não é suficiente?

A tendência natural para o comportamento de empurrar

Qual é o ponto de vista do fabricante/distribuidor quando ele está decidindo sobre o quanto manter de estoque em cada localidade? Ele tem duas perguntas principais em mente:

- Quanto manter acima na cadeia de abastecimento?
- Quanto manter abaixo na cadeia de abastecimento?

A tendência natural é manter o estoque o mais próximo possível dos consumidores – se um produto não está no ponto de consumo, então há uma chance (muito) menor de o item ser vendido. Apenas uns poucos consumidores deixarão seus fornecedores enviar o produto para eles em alguns dias, em vez de levá-lo imediatamente – consumo imediato é o nome do jogo. Portanto, é apenas lógico que o fabricante/distribuidor deva manter a maior parte do estoque o mais próximo possível do consumidor – o mais abaixo na cadeia que ele puder gerenciar – geralmente no nível do varejista.

Este é um típico comportamento de empurrar: empurrar os produtos abaixo na cadeia, para aumentar o consumo. Entretanto, o comportamento de empurrar requer um bom modelo de previsão, para prever onde e quando os estoques serão necessários nos locais de estocagem.

Por que é impossível encontrar um bom modelo de previsão de vendas?

¹ SKU: *Stock Keeping Unit*

Os módulos avançados de previsão existentes hoje tentam modelar a demanda e criar uma boa resposta para a pergunta de disponibilidade: O que manter em qual local e quando. Porém, o mecanismo de previsão, não importa o quão bom ele seja, não pode realmente prever como será a demanda. Pesquisas de mercado bem precisas podem dar algumas respostas, mas deve-se sempre considerar alguns fatos da realidade com relação à estatística.

O primeiro fato é que quanto mais estreita é a agregação, pior torna-se a resposta – significando que a pergunta de “quanto será vendido do produto em geral?” produzirá uma resposta bem melhor do que a pergunta: “quanto nós venderemos do produto nesta localidade específica?”. Este fenômeno deriva do fato de que as flutuações se compensam nos eventos agregados (assumindo que eles são eventos independentes). Se predissermos as vendas em 100 diferentes localidades, podemos obter uma resposta que as vendas numa localidade típica variam na faixa de 10 a 25 unidades por dia. Se fizermos a mesma pergunta sobre a quantidade em geral que precisamos fabricar, teremos uma resposta muito mais acurada – provavelmente algo na faixa de 1650 a 1850. Se tomarmos apenas os mínimos e máximos de cada ponto de consumo e os agregarmos, teremos uma resposta muito pior – de 1000 a 2500.

O segundo fenômeno é a interpretação errônea dos dados – pessoas usando a estatística devem ter um bom entendimento do mecanismo de agregação. Existem alguns grandes enganos sendo cometidos diariamente em todo o mundo, por causa da falta de compreensão da estatística. Por exemplo – uma pessoa inteligente, mas não experiente na estatística, pode deduzir do exemplo acima que o consumo será entre 1650 a 1850 para todos os pontos de consumo, então cada ponto terá um consumo entre 16.5 e 18.5 – mantendo 18 unidades em cada localidade e ficando sem estoque num grande número delas, enquanto outras serão deixadas com um monte de estoque que elas não conseguem vender. O fato de que temos uma soma agregada não significa que ela pode ser aplicada aos pontos que perfazem esta soma. Outra pessoa pode sugerir a proteção da disponibilidade pela colocação de 25 unidades em cada localidade – aumentando substancialmente o investimento e o número de pontos de consumo nos quais teremos excesso de estoque – tomando espaço e investimento desnecessários. Quanto mais sofisticado o algoritmo, mais sofisticado precisa ser o usuário final, para poder utilizar corretamente esse algoritmo.

Outro problema é que nenhum modelo de previsão pode levar em consideração a mudança repentina nos padrões de consumo. Um exemplo pode ser um artigo muito entusiástico num jornal (ou vice-versa) que muda repentinamente o padrão de consumo em toda uma região. No mercado dinâmico de hoje tais eventos estão se tornando bastante frequentes.

Conforme a previsão de um único item de estoque (SKU) numa localidade específica está sujeita aos impactos de flutuação e incerteza mencionados acima, ela é uma base muito pobre para a determinação do nível de estoque exigido deste item naquela localidade específica. Está claro que outro mecanismo é necessário para conseguir esta decisão.

O Jeito TOC: Distribuição Puxada

A Teoria das Restrições² analisa o impacto do fornecimento junto com a demanda sobre o gerenciamento dos estoques da cadeia de abastecimento, com uma ênfase no lado do fornecimento. Se for possível responder num instante à demanda, não há necessidade de recorrer a nenhuma previsão! Embora esta situação seja claramente inatingível em quase todos os ambientes de negócio, um passo nesta direção deveria ser considerado. No caso de manter a quantidade correta de estoque na cadeia de abastecimento, o objetivo é conseguir uma disponibilidade muito boa dos itens em todos os pontos de consumo. Este objetivo é limitado pela disponibilidade de dinheiro e espaço, o que significa que é impossível manter altos níveis de estoque de todos os itens em todas

² TOC (*Theory of Constraints*)

as localidades, mesmo quando a obsolescência não é um problema. Não apenas isso, mas também, como será explicado mais tarde neste artigo, manter estoques muito altos de itens (SKUs) de baixa demanda, diminuirá as vendas em geral.

A solução da TOC está baseada na renovação constante dos estoques consumidos e é composta de diversos passos:

- Agregar o máximo possível na fonte – a fábrica ou armazém central – definindo um alto alvo de inventário lá (chamado Tamanho do Pulmão de Estoque)
- Determinar os alvos de inventário em todas as localidades de estoque (Tamanhos dos Pulmões de Estoque)
- Habilitar a transferência dos dados de consumo real de todas as localidades de estoque
- Encurtar o tempo de reposição o máximo possível
- Repor o mais frequentemente possível a partir do armazém principal (da fábrica ou central) para os pontos de consumo – unidades são enviadas apenas para repor o consumo real (ou para reajustar os tamanhos dos pulmões)
- Monitorar os tamanhos dos pulmões de acordo com o consumo e reajustá-los apropriadamente

Agregação: Criação de um armazém de fábrica/central

A parte importante do modelo proposto para gerenciar uma cadeia de abastecimento é manter os estoques no ponto de divergência – onde os estoques podem ser usados para servir muitos destinos diferentes – e usar um mecanismo de puxar a partir do destino para repor. Este método garante a manutenção do menor estoque possível para suportar a demanda dos vários pontos de consumo.

Para ter o produto disponível em diferentes localidades é recomendado agregar os estoques na fonte e criar um armazém de fábrica ou central (AF/AC). Se a organização é um fabricante, a entidade é chamada de Armazém de Fábrica (AF), pois é um armazém de produtos acabados da fábrica. Se a organização é um distribuidor, a entidade é chamada de Armazém Central (AC). Seja qual for o nome, é nesse armazém que mantemos **a maior parte do estoque**. De acordo com os princípios da estatística, esta agregação garante um sistema mais estável do que manter o estoque nos diferentes pontos de consumo. Neles, a quantidade de estoque é bastante limitada. Uma vez que um determinado ponto de consumo vendeu uma unidade, esta será reposta o mais rápido possível a partir do AF/AC.

Quando o tempo de transporte do AF/AC até os pontos de consumo é muito grande, um Armazém Regional (AR) poderá ser necessário entre o AF/AC e os pontos de consumo. Um armazém regional se comportará como um ponto de consumo para o AF/AC e como um armazém central para os pontos de consumo que ele estiver servindo. Isto é apenas uma extensão do modelo da TOC e todas as premissas e considerações permanecem as mesmas – a idéia é continuar a puxar do AF/AC, baseando-se apenas no consumo do AR.

O Tempo de Reposição e como ele deve ser gerenciado

O tamanho dos estoques necessários nas diferentes localidades depende de dois fatores totalmente diferentes:

- Demanda: este é o fator que afeta a taxa na qual o estoque está sendo consumido nos diferentes pontos de consumo
- Fornecimento: este é o fator que afeta o quão rápido as unidades consumidas podem ser repostas

Surpreendentemente, o fator de fornecimento é geralmente ignorado na tomada de decisão tática e estratégica. A maioria dos esforços para melhoria é direcionada para o lado da demanda – especialmente tentando propor algoritmos de previsão mais sofisticados.

O tempo de reposição (TR) é definido como: O tempo que leva desde o momento que uma unidade é consumida até sua reposição a partir do elo anterior na cadeia de abastecimento. O TR é composto de 3 partes diferentes:

- Tempo de Pedido: este é o tempo que leva a partir do momento que uma unidade é consumida até que um pedido seja emitido para repô-la. Em outras palavras, esta é a **freqüência de pedido do mesmo item (SKU)**
- Tempo de Produção: este é o tempo que leva para o fabricante/fornecedor, desde o momento em que ele decide enviar o pedido até que finalize sua produção
- Tempo de Transporte: este é o tempo que leva para realmente enviar o produto acabado, do ponto de fornecimento para a localidade de estoque

A TOC sugere desafiar todos esses três elementos para cortar o Tempo de Reposição a um mínimo possível. Ao cortar o TR, o fator do lado do fornecimento torna-se menos dominante e atingimos o seguinte:

- Os níveis de estoque necessários nos pontos de consumo (e nos armazéns) são menores, já que precisam cobrir menos dias de demanda
- As flutuações no tempo de fornecimento tornam-se menores, na medida em que o tempo de fornecimento diminui
- A previsão necessária para as vendas de novos produtos é muito mais acurada, já que o erro de previsão torna-se maior ao precisarmos prever mais no futuro (a trajetória torna-se mais larga)
- A habilidade de responder muito mais rápido à demanda real é visível

Os princípios da TOC nos orientam a encontrar formas de aparar os diferentes elementos do TR. Estas são as orientações gerais:

- Tempo de Pedido: se possível, cortar até zero – geralmente significa tentar repor diariamente, em cada ponto de consumo, aquilo que foi consumido naquele dia. Outras considerações envolvidas serão cobertas mais tarde neste artigo.
- Tempo de Produção: Tambor-Pulmão-Corda Simplificado (TPC-S)³ – a metodologia da TOC para gerenciar a produção no chão de fábrica – deve ser implementada e a prioridade das partes manufaturadas deve ser amarrada ao seu nível de estoque no armazém da fábrica – isto será elaborado mais adiante neste artigo.
- Tempo de Transporte: tente enxergar alternativas para o transporte. Por exemplo, trens ou navios diários, em vez de semanais, ou despachar alguns produtos por avião. Descobrir fornecedores mais próximos para Matéria Prima (MP) ou peças adquiridas também é uma possibilidade em muitos casos. Geralmente esta é a parte do TR que menos se pode fazer algo a respeito, portanto, todas as possibilidades devem ser verificadas.

Freqüência de reposição versus custo de remessa

Ao aplicar a solução da TOC para gerenciar a distribuição na cadeia de abastecimento alguns fatores são relevantes quando se considera quão alta deve ser a freqüência da entrega.

A prática atual é despachar em grandes volumes. As principais razões são:

- 1) Geralmente um desconto é oferecido para grandes quantidades de cada item pedido. Esse desconto pode ser negociado de forma a ser oferecido para grandes quantidades pedidas durante um período de tempo. Desta forma podemos pedir freqüentemente e ainda assim

³ No original, S-DBR: *Simplified Drum-Buffer-Rope*

desfrutar do desconto, mas isto nem sempre é possível (embora esteja se tornando bem típico hoje em dia)

- 2) Há um certo esforço para listar todos os inventários disponíveis e emitir pedidos mesmo para uma quantidade pequena
- 3) Alguns itens só podem ser despachados em volumes por causa de questões de transporte – itens frágeis algumas vezes podem ser mais bem protegidos se despachados num container inteiro

Há uma escolha a fazer entre o custo adicional ao se investir no aumento da frequência das remessas e o custo de ter menor disponibilidade. O aumento da frequência de entrega cria uma melhor disponibilidade, enquanto o custo das remessas aumenta. Diminuindo a frequência, deve-se pagar ou com a disponibilidade menor ou com níveis de estoque maiores nos pontos de consumo. Na maioria dos casos, o custo extra ficará ínfimo perto da receita adicional produzida.

O que são os pulmões e o consumo dos pulmões num ambiente de distribuição?

A lógica da TOC define uma segurança e monitora constantemente como essa segurança está sendo utilizada. Esta segurança é chamada de **pulmão**⁴. Num ambiente de distribuição, a quantidade que gostaríamos de manter nas localidades de estoque (incluindo o AF e os ARs) é definida como o **tamanho do pulmão**⁵, e este é um tipo de pulmão de **estoque**. O tamanho do pulmão num ambiente de distribuição (tamanho de pulmão para Produção Para Estoque⁶) é o número de unidades que se quer manter em geral na cadeia de abastecimento, para esta localidade de estoque, deste item (SKU). Por exemplo, se o tamanho do pulmão de estoque for 100 unidades e atualmente na localidade de estoque temos 40 unidades, esperamos que 60 unidades já estejam pedidas ou a caminho a partir da localidade de estoque que abastece esta localidade (a localidade de estoque de abastecimento para o AF é a fábrica). Se essas 60 unidades não estão a caminho, um pedido de reposição de 60 unidades deveria ser emitido imediatamente.

Nota: localidades de estoque diferentes terão pulmões diferentes para o mesmo item (SKU), já que o padrão de fornecimento e/ou demanda podem ser diferentes entre elas.

O **consumo do pulmão**⁷ é definido como o número de unidades que faltam do pulmão dividido pelo tamanho do pulmão. Para o exemplo acima, o consumo do pulmão para o estoque no local é 60% ($(100 - 40) / 100$). O tamanho do pulmão é dividido em 3 zonas iguais. O consumo do pulmão define a cor do pulmão, de acordo com as diferentes zonas:

- Consumo menor que 33%: Verde
- Consumo entre 33% e 67%: Amarela
- Consumo entre 67% e 100%: Vermelha
- Consumo de 100%: Preta

A cor do consumo do pulmão fornece uma indicação da urgência de reposição deste estoque:

- Verde: o inventário no ponto de consumo está alto, fornecendo mais do que a proteção necessária no momento
- Amarela: o inventário no ponto de consumo está adequado, não há necessidade de pedir mais unidades para os elos superiores na cadeia de abastecimento
- Vermelha: o inventário no ponto de consumo está em risco de falta – unidades em transporte/fabricação (dependendo de qual é o ponto de consumo) devem ser consideradas para esforços de apressamento e um pedido de reposição urgente deve ser colocado para a fonte abastecedora, se nada estiver disponível a caminho do ponto de consumo

⁴ Buffer

⁵ Buffer size

⁶ Make to Stock buffer size

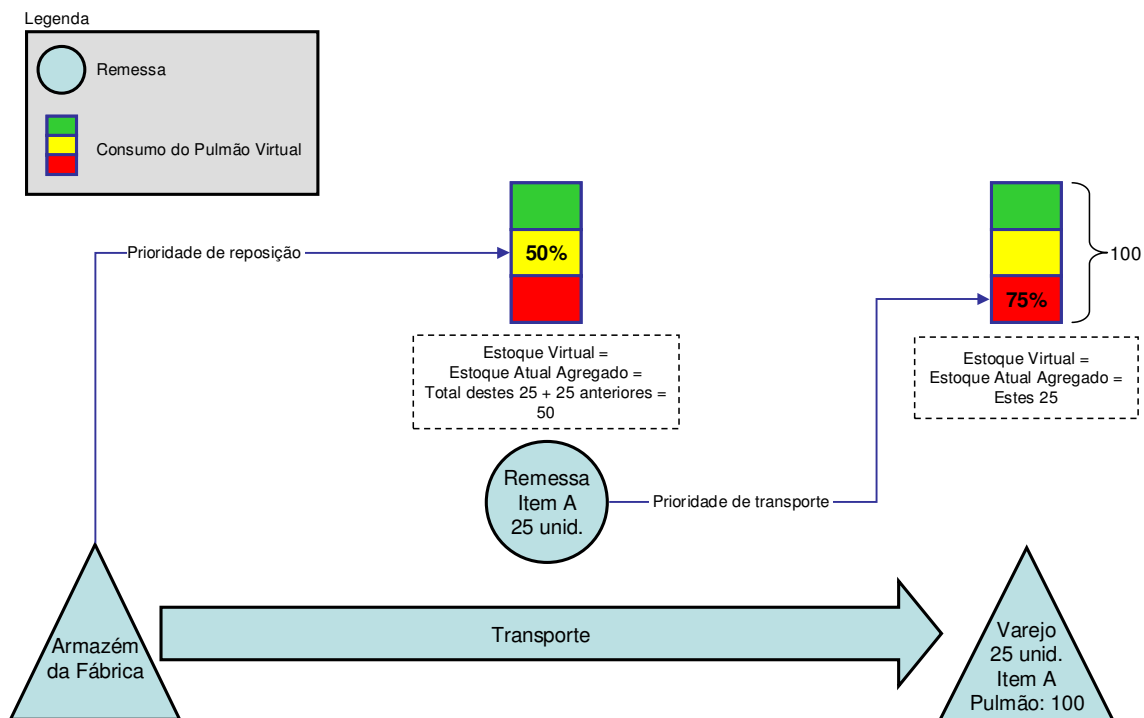
⁷ Buffer Penetration

- Preta: o estoque está esgotado no ponto de consumo, o que significa que a cada hora que passa perdem-se oportunidades de venda – esta situação deve ser resolvida o mais rápido possível, pois representa prejuízo real, especialmente nos elos mais abaixo da cadeia de abastecimento

Podem existir diversas visões dos pulmões no mesmo pulmão. Nós, na Inherent Simplicity, chamamos esse conceito de **Pulmão Virtual**⁸. Vamos examinar o seguinte:

Prioridade para um item mantido numa localidade de estoque

- O estoque agregado atual (estoque virtual) é calculado para todos os elos abaixo na cadeia, para o mesmo item (SKU), e o Consumo do Pulmão Virtual apropriado é calculado com base no que está faltando para o pulmão completo contra o tamanho do pulmão
- A prioridade é determinada pelo consumo do pulmão virtual do próximo elo



O tamanho do pulmão para este item na localidade de estoque é de 100 unidades. Temos 25 unidades em estoque nesse local e temos uma remessa a caminho desde o AF para essa localidade de 25 unidades. Você pode ver acima dos estoques a caminho o seu consumo de pulmão – levando em consideração o estoque agregado dos estoques abaixo na cadeia. As prioridades deles são determinadas pelo Consumo do Pulmão Virtual do próximo estoque abaixo na cadeia.

O consumo de pulmão nos fornece uma ferramenta muito poderosa – nós temos apenas uma medida com diferentes visões, mas todos os decisores envolvidos na cadeia de abastecimento podem obter suas prioridades de acordo com suas necessidades:

- O gerente dos estoques na localidade de estoque pode ver claramente que a prioridade deste item (SKU) é vermelha (75% de consumo de pulmão) – ele precisa descobrir como conseguir mais estoque deste item imediatamente
- O gerente de transportes pode obter a prioridade das remessas – quais despachos precisam ser apressados. Neste caso, a única remessa precisa ser apressada (75% de consumo de pulmão)

⁸ *Virtual Buffer*

- O gerente do armazém da fábrica pode obter a prioridade de reposição deste item nesta localidade de estoque. Neste caso, ele precisa repor 50% do tamanho do pulmão deste item A neste local e a prioridade desta remessa de reposição é amarela (50% de consumo de pulmão)

Gerenciamento Dinâmico de Pulmões

A TOC objetiva o uso de métodos muito simples e diretos, de modo que as pessoas que os utilizem realmente os entendam e, portanto, não queiram usar os módulos muito sofisticados de previsão. Anteriormente, neste artigo, os problemas que tais módulos criam foram apresentados – é preciso um entendimento profundo de estatística para usá-los corretamente. A lógica da TOC mede dinamicamente o uso atual dos estoques e reajusta os níveis de inventário apropriadamente. Este método é referenciado na atual literatura de TOC como **Gerenciamento Dinâmico de Pulmões (GDP)**⁹.

Ao monitorar o consumo do pulmão em cada localidade de estoque para cada produto, nós podemos identificar se o tamanho do pulmão que estamos mantendo deste produto neste local é aproximadamente correto. A abordagem do GDP argumenta que, ao monitorar e ajustar os tamanhos dos pulmões, podemos facilmente chegar ao estoque “real” que precisamos manter no local de modo a cobrir a demanda, levando em consideração o lado do fornecimento (o quão rápido podemos entregar na localidade de estoque).

O GDP observa duas ocorrências distintas: uma é se o tamanho do pulmão está muito grande, e a outra é quando ele está muito pequeno.

Quando tentamos medir se o tamanho do pulmão está muito grande, a indicação é quando o consumo do pulmão no local de um item (SKU), numa certa localidade, ficou Muito no Verde (MVD)¹⁰, significando que esteve na zona verde por muitos dias consecutivos (período de checagem verde – geralmente igual ao tempo de reposição). Isto significa que nós temos um pulmão muito grande para suportar a demanda, pelo menos neste período de tempo, o que sugere diversas alternativas:

- A demanda caiu
- O lado do fornecimento passou por uma importante melhoria
- O tamanho inicial do pulmão foi muito alto
- A demanda flutua severamente (e então o período de checagem verde deveria ser ampliado, em vez de se reduzir o pulmão)

A recomendação padrão para lidar com muito no verde é reduzir o pulmão. O princípio básico diz que reduzimos o pulmão em 33% quando precisamos, mas esta é uma orientação que depende de vários fatores:

- O quão rápido queremos diminuir os inventários, uma vez que observamos que a demanda está caindo
- O quão arriscado/importante achamos que este item (SKU) é
- O quão arriscado/importante achamos que esta localidade de estoque é

Um mecanismo muito similar é usado para determinar se o pulmão está muito pequeno: determinar se este item nesta localidade de estoque esteve Muito no Vermelho (MVM)¹¹. Porém, o algoritmo é geralmente diferente, já que nesse caso gostaríamos que o algoritmo fosse bastante sensível ao esgotamento do estoque, diferentemente do caso muito no verde, no qual gostaríamos de ter tempo

⁹ DBM: *Dynamic Buffer Management*

¹⁰ TMG: *Too Much Green*

¹¹ TMR: *Too Much Red*

e agir com segurança. O algoritmo mais básico para o MVm é determinar se um item (SKU) está no vermelho por muitos dias (geralmente usando o tempo de reposição), o mais avançado que se tenha em consideração e também o quão fundo no vermelho o inventário no local caiu.

As razões para se estar muito no vermelho são:

- A demanda aumentou (a razão preferida)
- O lado do fornecimento passou por uma deterioração
- O tamanho inicial do pulmão foi muito pequeno
- A demanda flutua severamente

O tratamento geral numa indicação MVm é aumentar o pulmão, sendo o padrão em 33%. Novamente, estas são apenas orientações e cada caso é diferente do resto.

Após ajustar o pulmão, o item (SKU) precisa entrar num “período de descanso”¹², no qual nenhuma sugestão de pulmão na mesma direção é dada (até que o sistema se ajuste ao tamanho do pulmão revisado). Este período de descanso deve ser longo o bastante para deixar o ajuste acontecer (as novas quantidades pedidas chegarem ao local de estoque), mas curto o suficiente para que uma mudança real repentina na demanda do mercado não ocorra sem alguém notar. Para o MVm, o período de descanso é um tempo de reposição completo, e para o MVd esse período equivale a deixar o inventário no local passar para o verde a partir de cima (já que a redução do tamanho do pulmão provavelmente fez com que o inventário no local ficasse acima do nível do tamanho do pulmão).

Prioridades da manufatura de acordo com a urgência no AF

Normalmente os fabricantes produzem sob pedido¹³. Isto significa que cada ordem de serviço no chão de fábrica é para um cliente específico, para uma data de entrega determinada. A TOC, para esses ambientes, prioriza as ordens de produção com base em suas datas de entrega (para mais detalhes favor recorrer à literatura publicada sobre o TPC Simplificado).

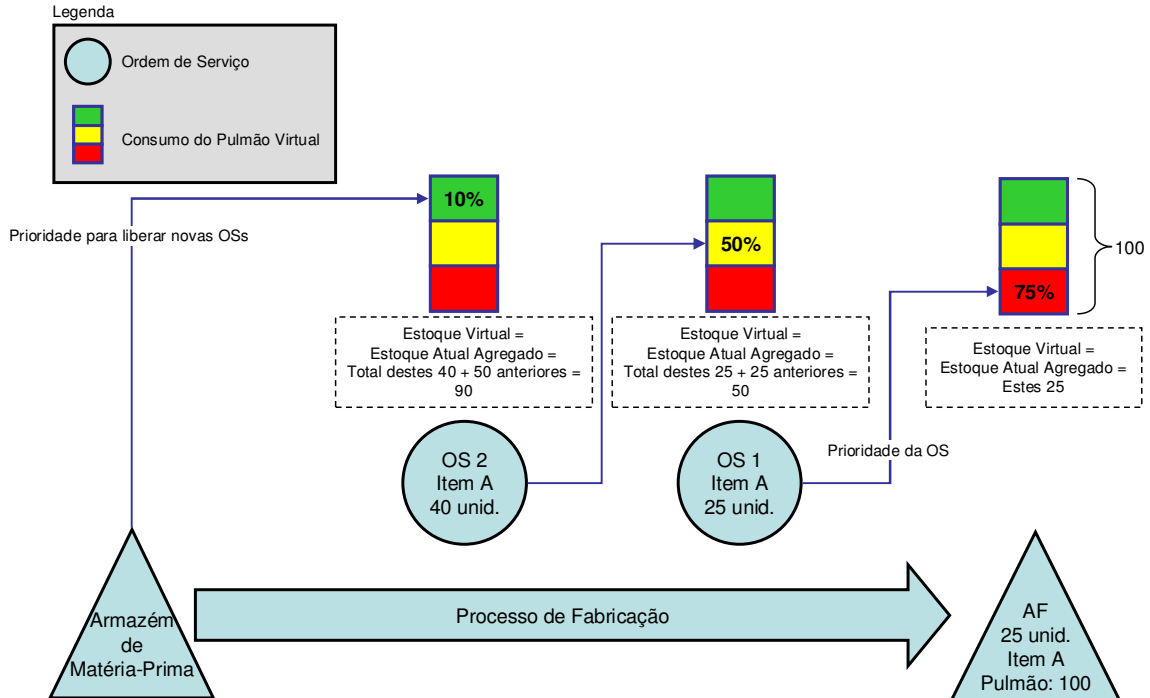
Quando os fabricantes abraçam a solução da TOC para distribuição, outro ângulo deve ser lembrado. Neste caso, as ordens de produção não são para um cliente específico e estão apenas cobrindo o consumo do AF (Armazém da Fábrica). Portanto, a prioridade correta deve ser definida não de acordo com o tempo, mas, em vez disso, da mesma forma que a prioridade para os itens nas localidades de estoque foi definida. O melhor mecanismo de priorização é tomar o consumo do pulmão no local como a prioridade para a ordem de serviço que precisa abastecê-lo. Se há mais do que uma OS para o mesmo item, o melhor mecanismo de priorização é tomar, novamente, o consumo do pulmão virtual da seguinte maneira:

¹² *Cooling period*

¹³ *Make to Order*

Prioridade para um item mantido num armazém de fábrica

- O estoque agregado atual (estoque virtual) é calculado para todos os elos abaixo na cadeia, para o mesmo item (SKU), e o Consumo do Pulmão Virtual apropriado é calculado com base no que está faltando para o pulmão completo contra o tamanho do pulmão
- A prioridade é determinada pelo consumo do pulmão virtual do próximo elo



Cada OS olha para o consumo do pulmão virtual da próxima OS em produção (aquela que foi liberada antes) para obter sua prioridade de produção. Isto garante que a produção está alinhada com o uso real do estoque. Se o estoque se esgota rápido, a OS será apressada através da produção, caso contrário, ela irá flutuar na produção na capacidade em excesso do sistema de produção. Toda entidade na cadeia de abastecimento está plenamente alinhada e sincronizada com a meta do sistema – ser responsiva ao consumo real dos estoques do próximo elo, para criar disponibilidade que, de outra forma, seria inatingível.

Por que uma cadeia de abastecimento baseada em distribuição puxada funciona melhor?

Observemos uma loja de varejo e as diferentes entidades que operam nesse ambiente. Podemos categorizar os itens na loja em 3 tipos diferentes:

- **Chitas**¹⁴: esses itens são vendidos muito rápido, habilitando o varejista a alcançar altos níveis de giro do inventário
- **Elefantes**: esses itens são aqueles que o varejista simplesmente não consegue se livrar – itens que se movimentam muito devagar, com baixos giros de inventário
- **Normais**: itens de movimentação regular, que não se enquadram nas categorias acima

O que na certa vai acontecer com os itens de giro rápido?

Quando os itens são chitas, por definição a demanda do mercado é alta para eles, relativamente à quantidade de inventário que mantemos para eles. Portanto, eles são os itens **mais prováveis de serem vendidos**. Se formos a um varejista e perguntar quantas faltas de estoque ele experimenta, a

¹⁴ Guepardo, leopardo-caçador ou onça africana

resposta mais provável seria: muito poucas, talvez 2-3%. Há muito conceito errôneo aqui. Se a pergunta fosse: digamos que estamos parados em frente à sua loja e perguntássemos às pessoas se elas encontraram o que vieram buscar, em quantos casos teremos uma resposta “não”, mesmo se presumirmos que elas deveriam levar o que procuravam? A resposta mais provável seria: OK, provavelmente de 10-15%. Isto significa que o nível experimentado de falta de estoque nas lojas é muito maior do que os varejistas pensam. Se o padrão de compra típico dos clientes da loja é a compra de mais de um item por vez, este fenômeno é dez vezes maior: qual a chance, tendo apenas 10% de falta de estoque, de um cliente encontrar todos os 8 itens que ele veio procurar na loja? A resposta é: quase zero – afetando a experiência de compra de quase todos os clientes que compram na loja.

Um fator muito interessante aparece ao analisar esses itens em falta: os 10-15% consistem principalmente dos itens chitas! Se o varejista soubesse que esses itens seriam vendidos tão rapidamente, ele teria comprado muito mais. Portanto, o volume de vendas perdidas que ele experimenta é muito maior do que os 10-15% que ele realmente admite! Isto é verdadeiro especialmente no ramo da moda. Os produtos são comprados pelos varejistas uma vez para toda a estação. Dessa forma, os itens de giro rápido estarão em falta durante quase toda a estação! Por exemplo, se um item que vende tão rápido todo o estoque for consumido em 2 semanas numa estação de 8 semanas, perdeu vendas de 3 vezes o tanto que foi estocado dele!

O outro lado da moeda são os itens elefantes. Esses itens não são vendidos como o varejista imaginou quando os comprou, senão ele os teria evitado. O fenômeno que acontece aqui é absurdo: o varejista investirá muito de seus esforços para vender esses itens elefantes e bloquear seu espaço de exposição, em detrimento de outros itens na loja! Este comportamento, embora seja esperado do ponto de vista psicológico, é anti-intuitivo no sentido de negócio – grandes esforços que serão investidos pelo lojista para vender os itens elefantes poderiam resultar em receitas muito maiores nos itens chitas. Este fenômeno algumas vezes torna desprezível o efeito das faltas de estoque nos itens chitas!

Alguns ramos de negócio adotaram até mesmo frases para esconder o fato de estarem operando de maneira anti-intuitiva, porque ficaram desesperados tentando resolver esses problemas. Esses ramos do mercado glorificam as faltas de estoque dos itens chitas (em TOC isto é chamado de perda de vendas) chamando-as de “esgotado”! Eles simplesmente ignoram o fenômeno dos itens elefantes marcando-os como “em liquidação” e investindo grandes esforços em sua venda.

Numa cadeia de abastecimento baseada na distribuição puxada esses fenômenos negativos são cortados ao mínimo. Já que o mecanismo da TOC está baseado na reação à demanda real do mercado e no ajuste dos pulmões à altura, se a demanda do mercado melhora, os pulmões serão aumentados, criando um mecanismo que permite a falta de estoque apenas por períodos de tempo muito limitados. Isto significa que as vendas perdidas de itens chitas por causa da falta de estoque são mínimas com a metodologia da TOC. Devido ao fato de que os inventários de todos os itens são mantidos menores, e as quantidades são diminuídas ainda mais quando o consumo for baixo, os itens elefantes são muito menos problemáticos, pois suas quantidades são mínimas. Assim, o uso da distribuição puxada é muito efetivo na eliminação das perdas de vendas.

Alguns dos Melhores Pontos na Implementação da Distribuição com TOC

Definindo bons critérios para decisões de variedade

Para diferenciar entre itens chitas, normais e elefantes, existe um critério simples: o giro dos inventários – significando quantos itens foram vendidos de um item (SKU) específico, numa localidade de estoque específica, relativos ao nível de inventário desse item. Porém, não é o

bastante conhecer as quantidades nas quais diversos itens são vendidos. É importante saber também seu valor financeiro. Somente saber quais itens são chitas e quais são elefantes não ajudará muito na orientação de qualquer decisão operacional. Há outros critérios que devem ser considerados. É importante saber o valor financeiro de tais itens.

O objetivo de definir tais critérios é obviamente relevante quando o dono da loja precisa escolher quais itens irá manter e quais não manter. Esta é a única coisa importante quando a variedade dos itens é muito grande e a capacidade de cada localidade de estoque para manter um grande volume de itens é limitada. Apenas levar em consideração os giros de estoque não ajudará – alguns itens são vendidos com margens tão baixas que, mesmo que eles sejam chitas, não contribuem muito para o resultado financeiro, e certo item pode ser vendido apenas uma vez por ano (obviamente um elefante), mas a margem é tão importante para o investimento que é um ótimo item para se ter. Para o fabricante/distribuidor, uma medida como essa pode ser usada para determinar quais produtos ele preferiria não ter de forma alguma na cadeia de abastecimento, sinalizando que uma nova concepção de produto é necessária.

A melhor medida para determinar o quanto um determinado item (SKU) é digno de ser mantido numa localidade de estoque é simplesmente o Retorno Sobre o Investimento (RSI)¹⁵ – o quão rápido este item está trazendo valor. Já que a medida do RSI foi criada para ajudar nas decisões relativas à escolha entre diferentes projetos, ela se encaixa perfeitamente aqui. O distribuidor e o lojista estão sempre limitados pelo volume de dinheiro e/ou espaço, então eles deveriam estar focados nos itens que mais contribuiriam para o resultado financeiro.

Nos termos financeiros da TOC, o retorno é medido desta forma: quanto Ganho¹⁶ (em resumo, G – significando margem – preço de venda menos custos verdadeiramente variáveis) se obtém deste item num período de tempo. O melhor período de tempo para se observar é um ano, para levar em conta o efeito de picos repentinos na demanda (normalmente devidos à sazonalidade).

Para calcular o Investimento, considere o seguinte:

- O inventário mantido na localidade de estoque é o que está cobrindo a demanda
- O inventário mantido a caminho também é um investimento para proteger contra as flutuações na demandas
- Existe quase sempre alguma coisa a caminho na distribuição puxada – os estoques da solução de reposição são repostos numa base diária (e algumas vezes, mais freqüentemente)

Levando isso em consideração, o melhor número para representar o Investimento necessário para gerar o Ganho que este item (SKU) gera é o **tamanho do pulmão**. Multiplicando o tamanho do pulmão pelo Custo Totalmente Variável (CTV)¹⁷ deste item, o Investimento real necessário para gerar as vendas deste item é percebido.

Portanto, a fórmula é muito simples: para calcular o RSI, tudo que é necessário é tomar anualmente o Ganho deste item e dividi-lo pelo CTV por unidade desse item, e multiplicar pelo tamanho (médio) do pulmão no ano todo.

A medida do RSI possibilita a diferenciação entre 3 grupos diferentes de itens (SKUs):

- 1) **Estrelas**: estes itens representam um RSI muito rápido, significando que mantê-los é muito bom para o negócio. Para o fabricante/distribuidor este é um tipo de produto que ele gostaria de manter em todas as localidades em que ele presta serviços.
- 2) **Buracos Negros**: estes itens exigem um período muito longo para retornar o investimento feito neles. Para o fabricante/distribuidor, um item neste grupo sinaliza um possível item

¹⁵ ROI: *Return on Investment*

¹⁶ Throughput (T)

¹⁷ TVC: *Truly Variable Costs*

para interromper a fabricação/compra. Entretanto, isto não é conclusivo, pois alguns itens (geralmente referidos como **Estratégicos**) têm que ser mantidos obrigatoriamente, mesmo que sua margem seja tão baixa que os coloque neste grupo.

- 3) **RSI Normal**: estes itens não estão nas outras duas categorias.

É óbvio que existe uma correlação entre os itens Chitas e os itens Estrelas, mas de forma alguma a correlação é 1:1, como foi claramente demonstrado pelos casos extremos discutidos anteriormente.

A decisão de como definir o limite entre os diferentes grupos fica a cargo do ambiente específico, mas as orientações gerais são tomar os maiores 10% como estrelas e os menores 20% como buracos negros. Uma das possibilidades para tratar os itens buracos negros é tentar mudar o preço de etiqueta de alguns desses produtos – tornando-os mais lucrativos se eles puderem ser vendidos por preços maiores.

Regras para definir o tamanho inicial dos pulmões

O primeiro passo ao se mover da distribuição empurrada para a distribuição puxada é instalar o armazém da fábrica (AF) e começar a construir os inventários para encher os pulmões de estoque iniciais.

A decisão do que deve ser o pulmão de estoque inicial parece ser muito complexa. A quantidade de incerteza é enorme, por isso o medo é muito natural – o receio de tomar a decisão errada, como se alguma dificuldade for encontrada, a lógica da TOC será culpada e todo o esforço da mudança para a distribuição puxada seja considerado estúpido.

Não há palavras suficientes no dicionário para enfatizar a diferença aqui entre estar exatamente errado e aproximadamente certo. Na Inherent Simplicity encontramos diversos casos onde a determinação dos pulmões iniciais levou mais de 3 meses! Esta quantidade de tempo seria suficiente para atingir o pulmão correto a partir de praticamente QUALQUER tamanho de pulmão inicial, e nós teríamos atingido alguns resultados em cima disso.

Já que o mecanismo de GDP ajustará os pulmões de acordo com o consumo real, a estimativa inicial só precisa estar na vizinhança do pulmão correto, e mesmo isso não é obrigatório.

A Inherent Simplicity sugere aos seus clientes iniciarem com um chute inicial: tomar o tempo de reposição desde a fonte até o destino e multiplicá-lo pelo consumo médio e por um fator de 1,5.

O tempo de reposição a ser usado deve ser:

- Para um ambiente de produção (armazém de fábrica): tomar o tempo de produção atualmente especificado para esse item (após implementar TOC no ambiente de manufatura o tempo de produção normalmente será cortado pela metade e, então, o GDP vai automaticamente sugerir a redução do pulmão)
- Para um ambiente de transporte (armazém central, regional e pontos de consumo): principalmente o tempo de transporte mais alguma coisa para levar em conta o número limitado de remessas a cada semana

Uma regra simples também pode ser usada para determinar se um item (SKU) tem alguns efeitos sazonais anuais: ao olhar para o consumo do último ano (e do ano anterior, se possível), se as vendas em um mês forem mais de duas vezes maiores que a média mensal das vendas totais (algo entre 15-20%), este item deve ser definido como sazonal naquele mês.

Para os itens sazonais, um pulmão inicial diferente pode ser definido para os meses sazonais e para os meses normais (usando as mesmas regras declaradas acima). A diferença nos tamanhos dos pulmões pode ser calculada e alimentada no modelo de sazonalidade – no qual os pulmões são

definidos manual ou automaticamente pelo software antes que a estação comece/termine. A sazonalidade mensal/semanal pode ser detectada usando um mecanismo similar.

Implementando o Modelo de Distribuição da TOC: Como o software pode ajudar? Ele é realmente necessário?

Para implementar com sucesso a metodologia da TOC para gerenciar um ambiente de distribuição, dois principais requisitos precisam ser satisfeitos:

- 1) Reposição: significa repor as diferentes localidades de acordo com o consumo
- 2) GDP: Gestão dinâmica dos pulmões para mudar o tamanho do pulmão constantemente e mantê-lo no tamanho correto para suportar o consumo atual dos pontos de consumo

Esses requisitos não são os únicos que precisam ser implementados, mas esses dois são os mais básicos – eles serão necessários em qualquer ambiente de distribuição.

Mesmo considerando apenas esses dois requisitos, a conclusão deve ser que nenhuma organização pode gerenciá-la sem software, a menos que ela seja uma cadeia de distribuição realmente pequena (qualquer coisa maior que 50 pulmões para gerenciar requer algum tipo de software). A questão é: que tipo de software pode ser usado?

Primeiro, defina quantos pulmões provavelmente serão mantidos sob o modelo de distribuição da TOC:

- O primeiro número que precisa ser calculado é o número de itens que estão planejados para serem gerenciados – este é o número de itens que a empresa oferece neste instante ao mercado
- O segundo número é o de localidades de estoque nas quais o item será gerenciado – todos os armazéns (regionais, de fábrica), bem como os armazéns de distribuidores e lojas de clientes nos quais o item será mantido em estoque no futuro

A estimativa sobre o número de pulmões que serão necessários para gerenciar deriva da multiplicação dos dois números acima.

Em geral, existem três opções para escolher, com relação ao software:

- 1) Desenvolver os componentes necessários do software dentro de um sistema ERP já existente, usado pela organização
- 2) Desenvolver os componentes necessários do software como planilhas de cálculo (Excel), externas ao sistema ERP
- 3) Comprar um software externo especializado em TOC

A resposta à pergunta sobre qual das três opções deve ser a escolhida depende principalmente da escala. Para qualquer ambiente no qual menos do que 500 pulmões precisam ser gerenciados, o uso de um software interno é uma possibilidade (seja uma planilha de cálculo ou um desenvolvimento do sistema de TI atual).

Para qualquer ambiente que contenha mais do que 500 pulmões gerenciados, a solução recomendada é conseguir um software externo, que seja plenamente focado nos processos da TOC e na tomada de decisão.

Por que não usar um software desenvolvido internamente, em vez de investir um grande volume de dinheiro e esforço num software externo de TOC?

- 1) Garantia de Qualidade: assegurar que o módulo de software desenvolvido internamente está fazendo o que deveria estar fazendo, é muito problemático. Os bons fornecedores de

complementos de TOC estão investindo grande parte dos seus esforços na verificação da validade dos módulos que eles programam.

- 2) **Confiança:** garantir que, tanto agora quanto no futuro, nenhuma mudança ou adição será feita nos módulos (causando ramificações negativas) por pessoas que “acham que sabem”.
- 3) **Desenvolvimento:** o conhecimento sobre a TOC está agora bem no início do processo. Novas percepções são desenvolvidas continuamente por consultores de TOC e empresas de software, as quais estão investindo muito para incorporar os últimos conhecimentos de TOC em seus softwares. Um sistema desenvolvido internamente nunca acompanhará os desenvolvimentos.
- 4) **Experiência apropriada:** existem muitos detalhes finos que não são de conhecimento e domínio público. Ao considerar empresas com necessidades especiais – tais como produtos sazonais, grupo de produtos similares ou grande número de pulmões – apenas uma empresa de software de TOC pode incorporar módulos de software que correspondam a essas necessidades. Desenvolvê-los internamente na empresa tomará enormes esforços e tempo, sem a promessa de quaisquer resultados.
- 5) **Longo prazo de entrega:** a partir de muita experiência, tentando desenvolver os módulos de software de TOC, o tempo necessário excede até mesmo as mais pessimistas estimativas. A Inherent Simplicity possuía muita experiência prestando consultoria para as empresas construírem seus próprios módulos internos para suportar os processos da TOC. Mesmo quando as capacidades internas de TI não eram o problema, enormes quantidades de tempo foram gastas esperando pela solução ser incorporada no sistema, com um monte de incertezas quando os módulos ficarão prontos.
- 6) **Uma planilha de cálculo (Excel),** a despeito do seu uso relativamente fácil, é especialmente não recomendada. Ela é muito fácil de ser alterada e, portanto, não pode realmente ser utilizada para reforçar o uso correto da ferramenta. Além disso, uma planilha é **muito** difícil de ser depurada. Em outras palavras, as duas primeiras razões se aplicam fortemente ao uso de planilhas.

Piloto e Software

Antes de lançar a solução de distribuição em escala total, muitas empresas gostariam de experimentar com um piloto para ver se as soluções de distribuição da TOC fazem sentido e trazem resultados. Uma fase piloto nem sempre é possível (por exemplo, manter estoque num AF apenas para uns poucos varejistas que usam a solução de distribuição da TOC resultará em resultados muito menores do que seriam para uma cadeia de abastecimento completa), mas onde é possível, a questão do software novamente se apresenta.

Um poucas empresas de software de TOC oferecem hoje um modelo no qual o software pode ser usado numa fase piloto por um preço relativamente baixo.

Já que a solução não intencionava ser uma de longo prazo, ela fornece a opção de obter soluções de curto prazo de propósito, significando:

- Para uma solução de curto prazo, se a escolha for prosseguir com um complemento de software de TOC, as interfaces podem ser gerenciadas manualmente e, portanto, a implementação pode começar imediatamente sem muito incômodo.
- Uma planilha de cálculo é utilizada muito facilmente, já que esta é uma solução de curto prazo e o objetivo não é um processo perfeito, mas uma prova de conceito (não há necessidade para um software guiado por processos), esta opção é realmente melhor do que tentar programar o software interno e lidar com os requisitos do piloto.

Resultados do Uso da Distribuição Baseada em TOC

A partir da experiência da Inherent Simplicity na implementação da solução da TOC de distribuição puxada é seguro dizer que os resultados são notáveis. Se forem usados os princípios básicos listados aqui, especialmente para definir os tamanhos iniciais dos pulmões, resultados extraordinários serão encontrados dentro de um período de 3 meses, desde o início da implementação. A média (sim, média) dos resultados da implementação da solução de TOC com o software da Inherent Simplicity (Symphony) é de 40% de aumento nas vendas, juntamente com uma redução de 50% no estoque geral nas localidades de estoque.

Transformando esses resultados em Giros de Estoque, de uma perspectiva puramente matemática, normalmente os Giros de Estoque são melhorados por um **fator de 2,8**.

Os resultados ainda não consideram os critérios que foram introduzidos aqui (itens estrelas e buracos negros). Estes critérios começarão a ser utilizados apenas no segundo trimestre de 2007 (obter mais devido à melhor disponibilidade dos itens estrelas). A expectativa é que este critério trará resultados ainda maiores.

Para conhecer mais sobre o Symphony acesse <http://www.heptagon.com.br/symphony>