

CAPACITY BOTTLENECKS : ILLUSIONS AND FALSE GODS

David Crabtree FIOM, GMCS

ILLUSIONS

"I'm sorry, sir", says the order reception clerk on the phone, "There's a 20 week delivery".

The would-be customer gulps in disbelief and protests that the part enquired of could be made from start to finish in two working days.

"That may be so, sir, but I'm afraid we have a production bottleneck on it – 20 weeks is the best we can do".

Really? How can a part that takes two days to make accumulate to itself a 20 week lead time?

To answer the question, we must understand how lead time inflation is created through the vicious circle.

A couple of years ago, the company was quoting to customers a fairly leisurely 3 weeks lead time on the part and fulfilling its delivery promises. Then over a short period, there was a spurt in business. The company released all jobs to the shop floor as it received them, temporarily overloading production and falling behind on its 3 week promises. It then put in motion the first turn of the vicious circle.

Fed up with the complaints of customers and somewhat abashed by its broken promises, it solved its problems at a stroke by increasing the official promised lead time on the part from 3 weeks to 8 weeks. Customers stopped complaining about late deliveries, but they also did something else the manufacturing company never thought of: in their purchasing systems, they naturally amended the delivery lead time on the company to the newly revised 8 weeks.

As a result of so doing, their purchase order systems kicked into action and showed that all the parts that were to have been ordered over the next 5 weeks were now past due, generating a surge of new orders on the manufacturer (and being past due, every one marked 'urgent').

Again, the company immediately released these orders to the shop and again production was overloaded. So then the lead time climbed above the 8 weeks to 12 weeks. More complaints! Another hike in the official lead time. Another surge of apparently past due purchase orders from customers.

And so the vicious circle turned until the company reached the 20 weeks quote, now convinced it had a serious bottleneck.

To beat its bottleneck and reduce its 20 weeks, the company works hard. Operators are on overtime every night and some come in on Saturday morning. Managers dare not be seen to go home before midnight.

But wait! The bright light of knowledge is breaking through the gloom! The company is hearing of the virtuous circle!

Although it is making no inroads on the backlog, the company now announces, loudly and widely, that for all new orders from henceforth the new guaranteed lead time on the part is 14 weeks. The revised lead time replaces 20 weeks in customer's purchasing systems and, for 6 weeks, new orders are reduced to a trickle.

During this respite, production clears 6 weeks of orders mired in the system. Now, with the real lead time down to 14 weeks, a second announcement is made – the official lead time is

reduced to 6 weeks. Again, now, for 8 weeks, new orders dry up and yet further backlog is cleared. Finally, the official lead time is reduced for the last time to 2 weeks.

There is a final act in this drama.

The company invests a very small amount of time and effort to make sure that the shop will never again become overloaded. It installs the simplest and cheapest materials management system in manufacturing – input/output control. Input/output control enables lead times and queues at work centres to be strictly controlled by ensuring that, at least at gateway work centres, no more work goes in than comes out. (A gateway work centre is a starting work centre where jobs first enter the shop). The self-fulfilling I/O equation is as follows:

Queue size at the end of this week (hours)	=	Queue size at the end of last week (hours)	+	Input this week (hours)	-	Output this week (hours)
---	---	---	---	----------------------------------	---	-----------------------------------

The control of lead times and queues at the gateway work centres is rigid, requiring only that an absolutely realistic figure be used to define each work centre's output. Figures are used that have been actually observed, or 'demonstrated' to have been recently achieved. At intermediate work centres, where work arrives from other work centres, queues will take a little longer to control and their complete control will involve discussion between the production controller and the production supervisor.

Our company now cancels its order for the bottleneck busting software it was to have installed and invests instead in a number of simple printed forms. One day it may splash out on a spreadsheet programme.

FALSE GODS

Meanwhile, across the road on the estate, a neighbouring company is experiencing the effects of genuine production bottlenecks. The company makes and distributes a wide range of industrial components – clamping systems, insulators, speciality bearings, and so on – with orders and profits constantly lost due to stock-outs caused by bottlenecked capacity. The reasonable definition of its bottlenecks – two large presses and a horizontal band saw – is that they are work centres which have maximum annual manufacturing capabilities less than the company's annual requirements of their output. Judicious stockpiling in the low season cannot help them and neither can superior scheduling or application of the production supervisor's many shop floor tricks. These really are bottlenecks.

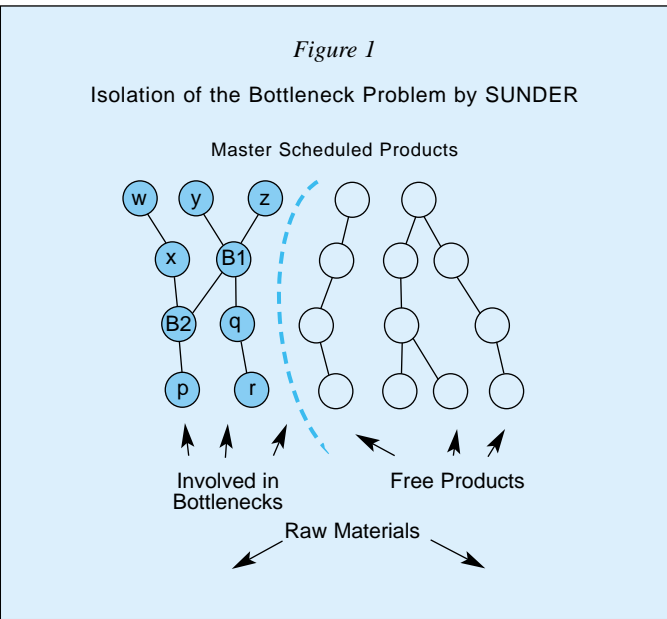
But today, a young man in a smart suit is visiting this company to explain how the TPO computer software package can resolve its difficulties. TPO is an acronym of Transcendental Profit Optimisation, and is indeed regarded almost in a religious light by its many licensees worldwide.

In this presentation, the young man explains to the manufacturing manager, the shop supervisor and other assembled production functionaries the importance in the TPO philosophy of keeping its bottlenecks busy - the arrangement of week-end shifts, ever watchful maintenance, and the provision of buffers of stock both to ensure that they are never

starved of material and that any temporary increases in their capacity can be capitalised on. The drumbeat of the bottleneck must be heard all along the materials chain, he says.

Then at last he begins to reveal the secret ways of TPO. First, he says, it is necessary to agree annual forecasts for all sales products, regardless of present or past capacity limitations as to the actual fulfilment of any such customer demand. Next, this 'ideal annual master plan' must be exploded through the bill of materials to determine the quantities of lower level components and raw materials that would be needed to support it. Thirdly, all of the ideally required production quantities should be assessed against the capacities of the work centres on which they are made to identify which of them are bottlenecked.

Once the bottlenecked products are known, the young man explains how the special TPO module SUNDER will next determine all the bottlenecked products' 'ancestor' materials and raw materials down the bill and all their 'descendent' connected materials up the bill. In effect, SUNDER isolates the bottleneck capacity problem from the remaining materials and work centres in the company not affected in any way by capacity shortages. (These materials are termed 'free products' and are planned by conventional means). At this point in his presentation, matters are summarised and illustrated by Figure 1. This shows the identification of two bottlenecked products B1 and B2; the isolation of seven products involved with these bottlenecks, though not of course bottlenecked themselves, ie. The three ancestors p, q, r and the four descendants w, x, y, z; and twelve free products not connected with the bottlenecks.



Because of the existence of the bottlenecks represented by B1 and B2, not all of the customer demand for the three master scheduled products w, y and z can be manufactured, so that the next step in the proceedings is to decide how the bottlenecked output is to be allocated. It is seen at a glance from Figure 1 that choices are difficult - for example, the more production earmarked for w, the less bottlenecked material B2 is available to make B1, so that the less production will then be achieved of y and z, and so on in endless combinations and permutations.

The young man asks those assembled for suggestions as to the criteria for the allocation of the scarce bottlenecked material. One suggestion made by the production control manager is that it should be directed to sales products yielding the greatest contribution. Contribution is a cost accounting term, and is defined for a product as 'selling price less cost price'. In other words, the suggestion is to manufacture as first preference those products with the biggest profit margins.

To test this strategy, the TPO expert first examines the joint requirements for products y and z, based on a maximum availability of 2,500 units of B1. Since y has the higher contribution (£4 per unit sold), its requirement for B1 is met in full. The remaining availability of B1 to fulfil the demand for z is short by 750 units. The total contribution from this allocation is £3,500, as shown in Table 1.

Table 1
Allocation of Bottlenecked Output to Products with Highest Contribution

Product	Contribution per Unit sold	Units of B1 needed to make 1 unit	Sales Demand	Units of B1 allocated	Units of B1 remaining	Total Contribution
y	£4	3 units/unit	750	2,250	250	£3,000
z	£2	1 unit/unit	1000	250	0	£500
Totals				2,500		£3,500

Unfortunately, the allocation in Table 1 takes no note of the demand of the sales products on the bottlenecked material, albeit at lower levels in the bill. That is, one unit of y absorbs 3 units of B1, whereas one unit of z absorbs only 1 unit of B1. Taking this into account and reversing the order of allocation, so that now the low contribution sales product z receives its full requirement of B1, the total contribution is £500 more than it was before —see Table 2.

Table 2
Allocation of Bottlenecked Output to Products with Lowest Bottleneck Demand

Product	Contribution per Unit sold	Units of B1 needed to make 1 unit	Sales Demand	Units of B1 allocated	Units of B1 remaining	Total Contribution
z	£2	1 unit/unit	1000	1,000	1,500	£2,000
y	£4	3 units/unit	750	1,500	0	£2,000
Totals				2,500		£4,000

In practice, it could never be hoped to cope with a realistic situation by such simplistic considerations as those applied in the two Tables – the different demands on a resource of different products; bottlenecks feeding other bottlenecks; alternative manufacturing sites; and so on. Instead, at the company presentation, it is explained that what is at the heart of the TPO computer package is the mathematical technique of mixed integer programming, an amalgam, as it were, of the well-known operational research techniques of linear programming and integer programming. This valuable family of techniques can be applied when there are constraints and where it is required to deploy them so as to optimise (ie. maximise or minimise) some particular outcome. Optimisation is very widely used in manufacturing industry in such familiar technical applications as sequencing (finding a shortest route), blending (cheapest mix) and scheduling (shortest time).

It seems then that besides isolating the allocation problem, another purpose of TPO's SUNDER module is to reduce to a manageable size the computational task involved in determining the bottleneck allocation by mathematical optimisation. The y.m. explains that the specific TPO module which actually performs the computation is referred to as SMART. SMART is the very centre of TPO – the creation of a 'model' comprising constraints, in the form of plant capacities and product demand, and, above all, an optimisation objective, namely the allocation of production so as to maximise the contribution to profit of the achieved sales product output.

After these wonders are revealed, it is agreed that the young man in the smart suit should take away a computer file of the company's data and return in a weeks time with a draft optimised production plan and a contract for the purchase of the TPO system. There is a buzz of excitement throughout the manufacturing department.

One week later he returns and takes the re-assembled managers through the production plan that has emanated from TPO and the SMART module – bottlenecked output so ingeniously allocated that the profit contribution from the side of the business affected is 50% higher than it previously was. It had not been realised before by those present that discontinuation by the company of its complete range of SuperCherie Components would break the bottleneck holding up full production of its Blague range of clamping systems, or that stopping production of the high end of its range of bearings would enable all demand to be fulfilled for its insulators.

The pen was poised above the contract when, thanks be to God, the company's sales and marketing manager entered the room. He had heard a rumour that a new software model was to be employed to determine the allocation of production to demand. He now looked at the TPO optimised master plan.

He was brought round with a shot of brandy from the manufacturing manager's secret cache. "No! No! No!", he screamed. "Don't you realise that if we discontinue the SuperCherie range, our entire marketing development strategy will fail? Don't you realise that Campbell & Fisher relies on our top bearings, and that it is our biggest customer for the non-bottleneck ranges"?

The contract for TPO was not signed. In fact, all of the literature about it was pitched in the bin. It was realised that mathematical optimisation was fine when it meant determining the mix of nutrients to make up the cheapest mix of animal feed, or when it meant minimising the trim waste of sheet metal in a production schedule. But there are scores of computer packages available to perform these tasks, better focussed for such applications and sold at a fraction of the TPO price – BlueKite and BOSS; Indexin and ILOG; OPUS and OptiRisk...Optimisation is a technical tool to be used where the customer and the marketing manager are indifferent as to the allocation arrived at. Where they are not indifferent, it can never be used to formulate a master plan: with bottlenecked output, formulation of the plan and deciding who gets what is a marketing art calling for sensitivity and the most careful weighting of all interest.

ACKNOWLEDGEMENT

The author wishes to make clear that the Transcendental Profit Optimisation package does not exist; on the Internet, the acronym TPO appears to stand for the Tokyo Philharmonic Orchestra.

About the author

David Crabtree, FIOM is the principal instructor of the UK training company, GMCS and author of "The Manufacturing Manager".