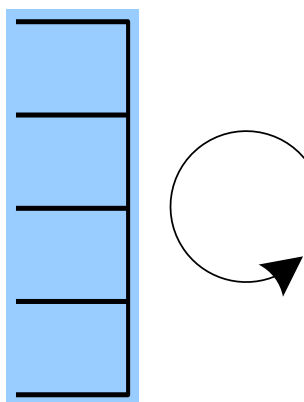


拉式排程系统概要

(Pull Scheduling Systems Overview)

David Hallett

(Copyright © David Hallett 2009)



作者序

阅读本文您将看到能永久转变营业想法的有力概念。拉式排程（Pull Scheduling）确实是种需要思维转变的作法，一种不同于当今最常见之排程方式。这样的系统已被证实绝对优于传统方法，能对企业经营带来超出想象的良性影响。

我将以对您而言或许是新想法来挑战您的思考，一开始似乎是『与直觉相反』，即是如依照经验与直觉的反应，你是不会采取的方式。然而，当你知道了这样的概念，并且应用了这样的概念后，你的思考构面将扩展至不同的层次。

于本文，我分享普遍的拉式排程原理，无论公司或场景为何，其概念的正确性是通用的。我也将分享一些多年实务所得的实践技巧，这些将协助您针对您的情况操作正确的方法，同时提供一些应该从何处开始应用拉式系统（Pull Systems）的见解。

关于作者

David 应用拉式系统于实际工作已有二十年以上的历史。过去十多年以来，他从事管理顾问工作，全力投入协助企业改善营运管理，主要在于应用精实生产（Lean Manufacturing）与限制理论（Theory of Constraints）于改进企业的营运模式。

Copyright © David Hallett 2009
www.PullScheduling.com

David 的求学背景为工业工程，在日本学习与精进其精实生产的技能，而且直接参与日本的精实 Sensel 在美国的开发工作。他在 Goldratt Institute 学习并取得限制理论的认证。

David 致力于成为顾问与讲授，过去十多年以来，参与过超过百家的顾问工作，并参与许多国家级会议、专业社群与国际企业研讨会。他为顾问公司及各重生产延伸伙伴中心（Manufacturing Extension Partnership，MEP）提供专业能力与服务。

David 目前受聘于位于纽约 Buffalo 之 MEP 中心的 Insyte Consulting 机构。

目录

I	拉式系统之境遇 (The Case for Pull Systems)	四页
II	拉式系统的类别 (Types of Pull Systems)	十五页
III	实施拉式系统 (Implementing Pull Systems)	二十页
IV	结语 (In Conclusion)	二十二页
	附录 A – 拉式系统比较表 (Appendix)	二十三页
	专门术语与概念之词汇表 (Glossary of Terms and Concepts)	二十四页

I. 拉式系统之境遇 (The Case for Pull Systems)

本章回顾拉式系统的基础知识，与应用拉式系统对企业经营的利益。

故事从此开始

我第一次接触拉式系统是在 1980 年代中期，当时我服务的公司是一家地铁车厢再制造商 (a subway car re-manufacturer)。地铁车厢再制造商的工作是，重新加工与升级二手的产品使它们像是新 (或更好) 的产品。以地铁车厢产业而言，车厢还在行驶服务时，运输管理局便委托再制造商车厢重整的合约。重整车厢必须从行驶服务的轨道移走，运输管理局不喜欢这样做 (显然地)。如果真的将车厢移走，他们想要尽可能在最短的前置时间 (the shortest lead-time) 内放回去。

地铁车厢产业中，大多再制造商的合约有三项必要条件：

- 在制品 (Work in Progress, WIP) - 在任何设定的时间内，运输管理局容许移走的地铁车厢数量，即是再制造商能加工的车厢数量。
- 前置时间 (Lead-time) - 地铁车厢被从行驶服务移走，直到回归行驶的时间长度。
- 数率 (Rate) - 每个时间单位地铁车厢被送交运输管理局的数目。

在这段期间，我负责两个任务：再制造程序的排程，及监督和分析关于这个程序的成本。一段时间后，我察觉到上述三个变量间的关系：

- 前置时间 = 在制品 / 数率 [Lead-Time = WIP / Rate]

在后来的工作里，我才知道这是被称为『Little's Law』的程序，这个小程序却相当有用，这是拉式系统的关键要素之一。[译者注：Little's Law http://en.wikipedia.org/wiki/Little's_law;
<http://xquality.blogspot.com/2007/12/littles-law.html>]

来看看两个例子。假如合约上的在制品数量是 10 个车厢，而我能每周完成 2 个车厢，那么前置时间是多少呢？

- 前置时间 = 在制品数量 / 比率 = 10 个 / 每周 2 个 = 5 周。

假如在制品数量是 20 个车厢，而必须在 4 周的前置时间完成，那么数率必须是多少呢？

- 首先转换程序，『前置时间 = 在制品数量 / 数率』 转为『数率 = 在制品数量 / 前置时间』。
- 数率 = 在制品数量 / 前置时间 = 20 个 / 4 周 = 每周 5 个。

后来，当我正式更深入理解拉式系统时，我察觉到实际上我们的客户迫使我们采用一种拉式系统。拉式系统的广泛定义是：

- **拉式系统：**当库存受限于某方式的情况下，使用的一种（拉式）排程系统。[以某方式限制库存量的情况下，使用的一种排程系统。]

(Pull System – A scheduling system where inventory is limited in some way)

[注：库存可能是原物料，在制品，成品等等。]

大多经营模式并未强制限制库存数量。在没有限制的情况下，在任何时间点的库存数量与库存所在的位置都相当随意，并倾向随时间而增加。这被视为一种推式系统（Push System），定义为：

- **推式系统：**当不正式限制库存的情况下，使用的一种（推式）排程系统。

(Push System – A scheduling system where inventory is not formally limited)

如何落实推式系统呢？

自年少起，我们就被教导一些基本的工作观念，像是：

- 努力工作
- 赶快
- 超前

在经营管理上常见的绩效考核与奖励都是支撑这些观念，典型的员工与部门都有『每天几个、每人几个、符合计划时程等』的绩效考核及奖励。

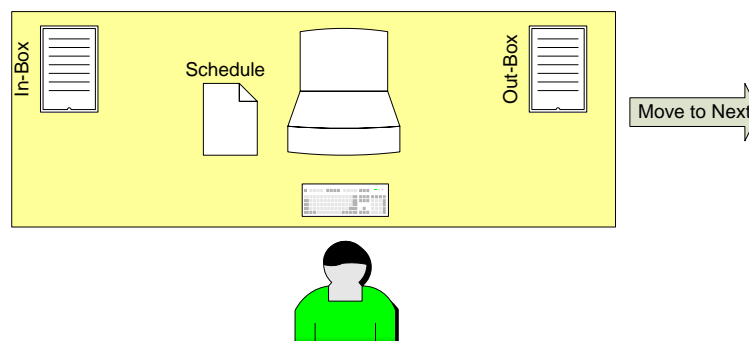


Figure 1

观察一个典型的业务情境（上面的图一，Figure 1），一位工作人员在计算机前工作，并被给予一份哪些数据在何时应该完成输入的工作时间表。数据到达他的桌上，进到『收件箱（In-Box）』。这位工作人员看看工作时间表，从收件箱拿出下一份需输入的数据，用计算

机完成数据输入，然后将数据放到『结束箱（Out-Box）』。一段时间后，他会站起来将结束箱中的数据拿到下一个处理步骤的收件箱。

每一周，他的主管会检查完成的件数，看看是否依照工作时间表完成工作。假如不符合工作目标，主管追踪工作人员后续的修正行动。

根据上述的绩效考核到位的情况，而工作人员的注意力落在『努力工作、赶快、超前』，工作人员保持忙碌的状态。一旦输完的数据放入结束箱，就从收件箱拿出下一份数据开始工作。如果时间表上的下一份不在收件箱，就拿其他后面才得输入的数据去做。

推式有什么问题呢？

有两个假设常见于营运管理：

1. 我们的业务操作是各自独立的。改变一个操作很少会对其他操作造成不利的影响。
2. 我们的业务操作具有相加的效果。要知道业务的表现情况，就看每个工作的操作效果，基本上将这些操作加总即可。要想改进绩效与获利，要求每个操作尽可能发挥效率，一个操作的改善就会改进公司的利润。

根据这些假设，我们使用某种方式来管理业务。我们试图使每个部门、区块、个别程序优化，认为这样做会使绩效与获利最大化。因而经常是各个操作个别排时间表 – 如果他们全都遵守个别的时间表，就能符合整个时程。我们并且加上对每个操作的绩效考核，以能知道操作是否『具有效率』。

这里的问题是，这些假设不正确！下列的原则取代上述的错误假设：

1. 我们的业务操作是具相依性的（Our business processes are dependent）。改变其中的一个操作，会对某些或全部的其他操作造成某种影响。
2. 我们的业务操作不具加总的效果（Our business processes are not additive）。单独改善一个操作通常不会对业务的整体绩效或获利情况造成良性的作用。事实上，根据上面第一项原则，有可能某个部分的改善，却造成整个系统表现的退步。

看看下列的例子。

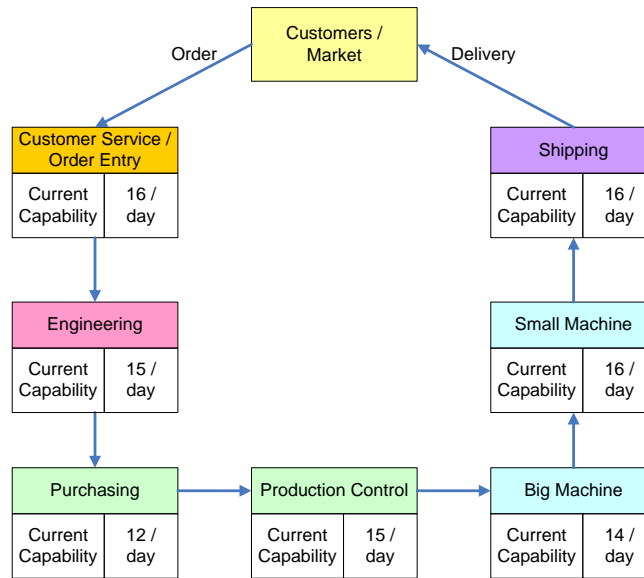


Figure 2

图二（Figure 2）的图形呈现一个『业务循环』。每个业务有一系列程序，最后链接成一个整体的业务循环。每个业务筐内的名称与机能都不一样，但是大体的概念总是相同的。

假如『工程（Engineering）』筐代表上面描述之推式系统中的工作人员，要符合每天 15 个工作日的目标，且准确遵照排程来工作，以整体业务循环来看，每天能完成多少个？答案是 12 个工作。在这整个循环中，工作进行的速度最快取决于最慢的业务程序，目前是采购（Purchasing）每天 12 个工作。那么另外三个工作到哪里去了呢？在采购筐前面等待。对于前置时间会有何影响呢？

- 前置时间 = 在制品数量/数率 [$LT = WIP / Rate$]，在制品数量上升，而数率保持不变，因此前置时间拉长！前置时间拉长是不好的结果吗？绝对不好。

每次 MEP 社群的精实生产的简报都引用下列亨利福特的名言：

『维持福特产品低价的努力中，有一项最值得注意的成就之一是，逐步缩短生产的周期。一个对象在制造过程中越久，被移来移去的次数越多，最后的成本就越高。』

亨利福特所说得重点是，前置时间越长，成本越高，即是下列的公式：

- 前置时间 = 成本 [$Lead-Time = Cost$]

这个公式是精实生产的精髓，专注于消除没价值的附加活动。一般而言，产品前置时间的 95%+ 是没价值的附加时间（non-value-added time）。因此，如果我们投入降低前置时间，本质上会降低没价值的附加时间，成为『较精实（Leaner）』。

回到前面的工程师例子，在年底主管对他符合产出与排程的表现很高兴，他获得加薪。接下来的挑战是，提升他每天的产出从 15 个到 17 个工作。主管认为有其必要性，因为所有的主管都被大老板要求改进今年的效率。

透过许多分析工作，超时加班及辛苦工作，这位工程师改进他的计算机程序，这样一来能每天产出 17 个工作。这时，其他部门在改进上却不怎么成功，事实上，采购还是一天 12 个工作。

对于整体业务产出，工程部的改善有什么作用呢？

- 前置时间 = 在制品数量/数率 [$LT = WIP / Rate$]，在制品数量增加，数率还是一样。因此，前置时间拉长！
- 前置时间 = 成本，前置时间拉长，因而，成本增加！

等一下，到底发生什么了。工程改善了，但成本增加！为何如此？

回到那两个原则，我们的程序不是独立的（Our processes are not independent），工程部的改变影响到我们系统中的其他程序。同时，我们的程序不具加总的效果（our processes are not additive），所以工程部的改善并未提升整体利润。

因此，这些原则表示什么呢？

1. 我们必须管理整体系统，才能使经营优化。
2. 一般而言，在考虑系统限制下，一个程序会限制系统的有效性。藉由改善该系统限制，我们能提升整体系统的表现。而其他程序必须搭配系统限制的运作需求。
3. 藉由降低在制品，我们能改善前置时间及成本，只要不使系统限制闲置。

取代制造更多不需要的东西，工程师是不是该只产出采购需要的数量，而用多余的时间去协助采购的工作呢？

为何降低前置时间会降低成本呢？

这并非一般我们看待成本的方式。事实上，当今 95% 企业使用的成本会计系统没有如此的呈现。这些在 1900 年代早期开发的系统，营运方式与今日的方式相当不同，根据劳工及/或材料分摊经常费用。在传统成本会计上，只有当劳工及/或材料被用于制造产品的时间，才累计工作的成本。这是典型之加工价值的步骤，占整个前置时间的 5%。而步骤间的时间是没做加工的事情，成本会计系统将之被视为『free（没费用）』。

整个营业周期的前置时间，其累计之产品成本图（Figure 3）如下（传统成本会计的呈现）：



Figure 3

然而，应用亨利福特的前置时间等于成本的道理，产生之图形（Figure 4）如下（精实成本会计的呈现）：

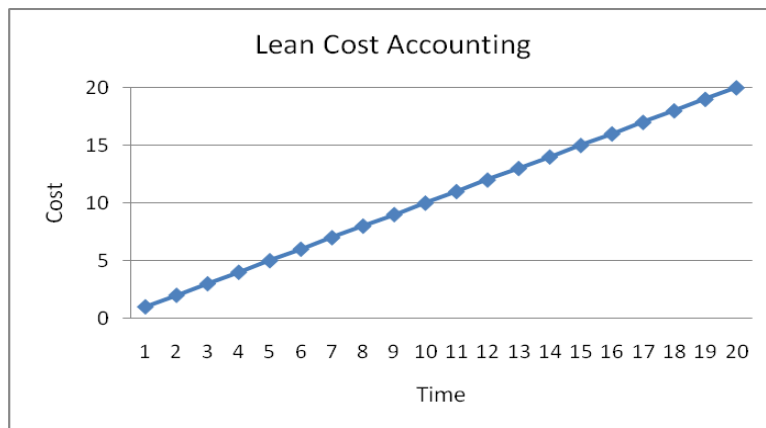


Figure 4

高德拉特，限制理论（Theory of Constraints）的创始人，曾说过：

『告诉我你会如何考核我，我就告诉你我会如何表现。』

假如我们的绩效衡量制度会影响我们的行为表现，传统的成本会计造成什么行为模式呢？取代专注于整体的前置时间，传统成本会计使我们将注意力放到一小部分的时间，就是使用劳力与材料于制造产品的时段。

降低作业的前置时间有很多利益难以在此一一列出。Insyte 顾问公司是位于美国纽约州水牛城的 MEP 中心，在他们的网页上有一份列出 31 项缩短前置时间的利益（网址：

<http://www.insyte-consulting.com/Home/Resources/Tools/BenefitsofReducingLead-Time>)。这些利益大致分为六类：

1. 增加销售 (Increased sales)
2. 改善质量 (Improved quality)
3. 降低营运成本 (Reduced operating costs)
4. 增加产能与有效产出 (Increased capacity and throughput)
5. 降低投资的资财 (Reduced invested assets)
6. 增加员工的满意度 (Increased employee satisfaction (精神/心理上 morale))

以制造公司的一个生产工作中心为例，该生产中心位于整个产品制造路径的一半位置，假如说该中心有一个月的在制品数量，价值是 100,000 元，放在该中心前面等待（见下图，Figure 5）。

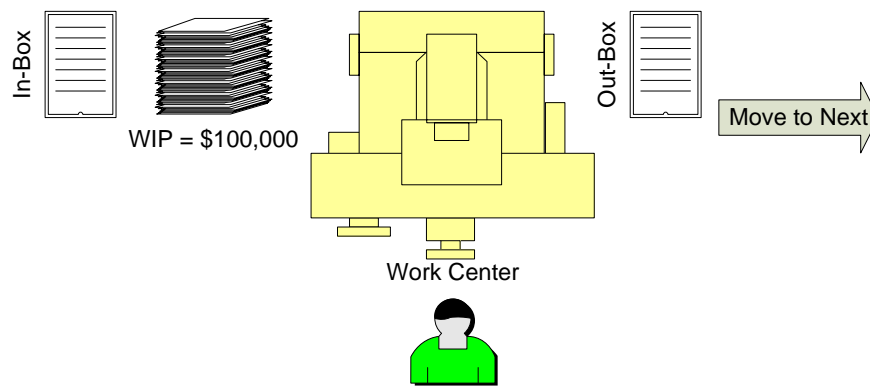


Figure 5

根据传统成本会计系统，对公司而言这些在制品并无花费。事实上，在资产负债表上被列为财产。同时，假如总计任何一年的在制品，在制品增加在会计的损益表上，则显得公司获利更多。

根据精实成本会计 (Lean Cost Accounting)，更正确的看待方式是，这些在制品花用公司的金钱，为何如此？几项理由列举如下：

- 资金的成本 (Cost of capital) – 一家典型公司有信用贷款和/或希望租借或购买设备。假如 1000,000 元没压在那里，公司能用这笔钱以降低信用贷款的额度，或投资新设备或其他财产。
- 空间的成本 (Cost of space) – 价值一个月的在制品需要放置空间。制造商需要投入金钱方有空间，还有与暖气、冷气与照明相关的费用。

- 处理的成本（**Cost of handling**） – 在这个月中，这些在制品等待加工，通常需要搬动、存放、排列、分开、综合等等许多次的作业。相关的成本不只是劳工，还有可能的产品损坏及潜在的人工受伤。
- 质量的成本（**Cost of quality**） – 假如我们发现在制品中有异常情况，有多少不良品呢？可能全都不良！这样就必须重工或丢弃及重做。由于之前的制造步骤是一个月或更久之前发生的，我们会去找出问题的根本原因，及采用有效修正的变更作法的可能性低。
- 流失有效产出（**Lost throughput**） – 到处都有随意置放的在制品会掩盖最慢的作业程序。不知最慢的程序何在，无法有效专注，则流失系统整体的有效产出。
- 丧失销售机会（**Lost sales opportunities**） – 许多顾客订单是视供货商的前置时间而定。一般而言，顾客是差劲的规划者，因为前置时间=在制品数量/数率，供货商的在制品时间会拉长前置时间，这样就可能时间太长而丧失销售机会。
- 降低准时交付（**Reduced on-time delivery**） – 制造时间越长，越难准时交付。还有顾客很可能改变设计、数量、交期等，而增加无法准时交付的可能性。

完整的精实或 TOC 成本会计是超出本文的范围。然而，一般做出营运上好决策的好规则是，专注于考虑决策在现金定位及现金流（**cash position and cash flow**）对增值利益（**accrual accounting profit**）的效应。改进整体现金定位及/或现金流的决策，最后能改善营业成效。评量实践任和改善方案的成效也可使用这个方法。

上述种种与拉式系统有何关系？

有许多事情会造成成长制造前置时间。不过，80/20 法则建议系统中少数的政策造成大多数整体之前置时间。在我的经验中，最容易改善前置时间，又能产生大效果的是：

- 排程系统（**Scheduling systems**）
- 批量大小（**Batch sizing**）

拉式系统针对这两项，使用几个相当简单的技巧，能很快降低 50%或更多的前置时间。

许多精实实践者认为拉式系统是高等技巧，在整个精实的实施中到后面才用上（如果有有用的话）。TOC 实践者认为拉式系统是基础的步骤，通常在导入改善方案先期便采用。假如前置时间是主要的标的，而拉式系统的主要贡献是处理长前置时间，那么前置时间的改善工作，则该以应用拉式系统为起点。

你会发现，一旦导入拉式系统，你的前置时间会降低，系统中其他的关键问题会更显而易见，更难忍受，但更容易入手处理。你也会有工具能将有限的改善资源，投到能造成效果的方案，而不被浪费于无（或者负）投资报酬之处。

拉式系统的类别 (Types of Pull Systems)

有许多能用来实践拉式运作的可能技巧，再次，我们的定义是：

拉式系统 -- 一套排程系统，其中库存受某种方式所限制。

(Pull System – A scheduling system where inventory is limited in some way)

注：库存含原物料、在制品、成品 (raw materials, WIP, finished goods) 等。

根据我们的定义，任何技巧能正规地限制库存，则能造成拉式操作。尽管如此，大多数的技巧大致上可归入下列的主要拉式方法：

- 超级市场补货 (Supermarket Replenishment)
- 设限的先进先出动线 (Capped FIFO Lanes)
- 鼓-缓冲-绳 (Drum Buffer Rope)
- 在制品限制 (WIP Cap)

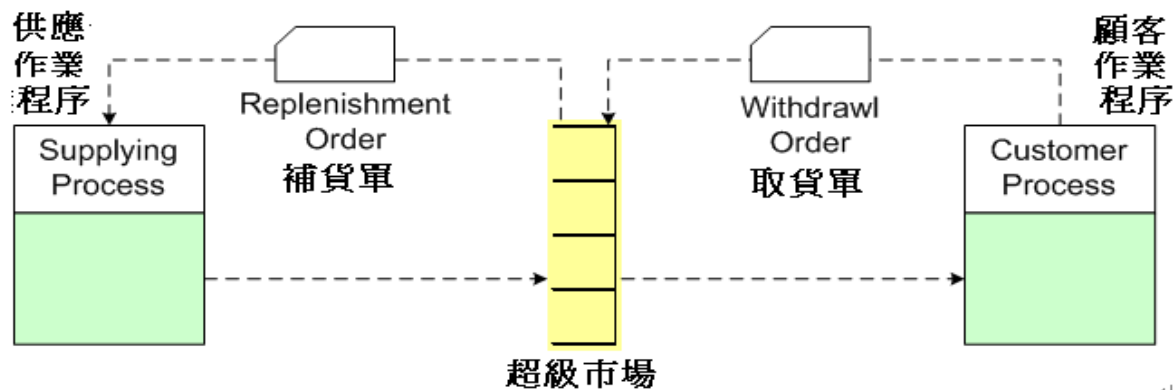
本章将回顾每个作法及适合的应用。

超级市场补货 (Supermarket Replenishment)

如我们所见，传统的制造排程系统要求大量的库存资金。当丰田在第二次世界大战后重建时，他们现金很少，无法依照传统的制造方式投入库存。他们观察并提出：『什么产业最善于运用库存？』答案是，超级市场！

今天，以典型的超级市场来看，他们的库存转数每年超过 50 次。使用推式系统 (Push system) 的典型制造商每年的库存转数超过 1-10 次。使用拉式系统及其他精实 (Lean) 制造技巧，丰田的制造环境中能达每年 50-100 的库存转数。

我们来看超级市场补货系统的运作。



图六

图六呈现一个基本的超级市场补货系统，过程如下：

- 顾客作业程序从超级市场的货品展示架上取得所要的产品。
- 每个物品号都设定一个数量为再订购点。只要是库存量加上已发出之订货量，低于设定的数量，则发出一张新的订单给供应作业程序。一个容器（container）、广告牌卡（kanban card），灯志信号，清空的空间等等，皆能用于代表需要订购（纸张订单不必然需要 – 简单来看，就是『清空了，补足它』）。
- 每个物品号都有一个再订购数量（re-order quantity），新订单的数量等于某个物品号的再订购数量。
- 供应作业程序完成该订单。
- 购得的物品摆上超级市场的展示架。

在此例子中，只有顾客作业程序能设计一套正式的订购运作方式。供应作业程序受带动而自动发生，这是拉式系统非常重要的部分。不管一个拉式系统中有多少程序，只有一个排程点（one scheduling point）！

单一计划点（Single Schedule Point） – 在拉式系统中，于某个单一程序独立计划整个拉式系统，而所有其他的程序由拉式系统自行计划。（Single Schedule Point – Single process in a Pull System where the entire Pull System will be independently scheduled. All other processes are scheduled by the Pull System itself.）

想想相较于分别计划系统中的每个程序，这是多简单与有效用。大多数公司试图不是手动计划每个程序，就是使用传统的 MRP/ERP 系统。

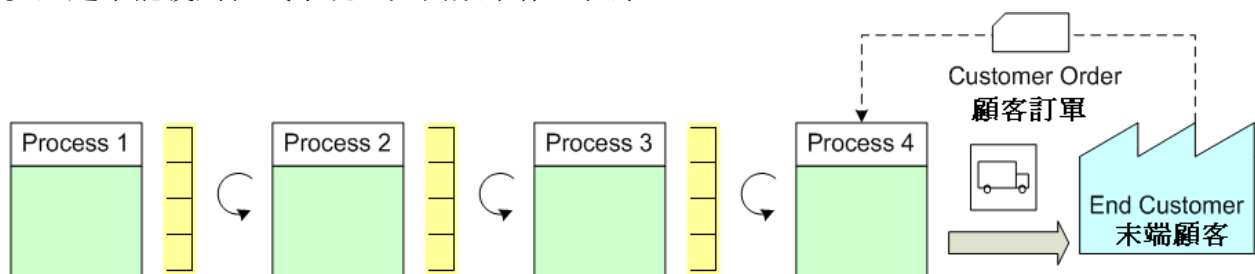
下面的变量会影响到，计算系统中每个物品号的再订购点（Re-Order Point）和整体潜在的库存（Total Potential Inventory）：

- 一段时期的平均需求量（Average Demand per period）
- 一段时期的需求变化性（Variability in Demand per period）
- 从发出一张订单给供应作业程序，直到完成该订单，即是货品补回超级市场的展示架之前置时间（Lead-time from when an order is released to the Supplying Process until the completed order is received at the Supermarket Shelf）
- 再订购数量（Re-Order Quantity）
- 服务程度（Service Level），是超级市场希望在多少时间比率上，展示架上至少有某些产品，可进行顾客作业程序。注：因刻意设计，绝不是 100%，由于费用太高或统计上就是不可能。

上述任何一项的增加会使的系统的整体潜在库存提高。

当使用这套系统时，在超市架上的库存很少等于整体潜在的库存，其中包含拿走而不达再订购数量，等待供应端处理的订单，供应端处理中的订单，及供应端送出在运送途中的订单。在一个设计良好的系统中，超市架上的实际平均库存经常是整体潜在库存的 10-50%。

多重超市能使用拉式系统，如同数个作业程序：



图七

这个例子中的各个超市都是独立的，如图六所示的运作方式。单一计划点（Single Schedule Point）在此系统中是在程序四（Process 4）。一个简单的方式来决定这个位置是：

单一计划点的位置就在整个拉式系统中，紧接在最后一家超市之后的程序。（The Single Schedule Point Location will be located at the process immediately after the last Supermarket in the entire Pull System.）

所以，如果你比对每个拉式系统中接续的作业步骤，从收到原物料和信息，到将产品交给顾客，在超级市场的情况，下一个作业程序就是单一计划点。这指出有多种方法我们能用于单一的拉式系统，这是正确的。事实上，超级市场补货方式通常是最不想要使用的拉式作法。

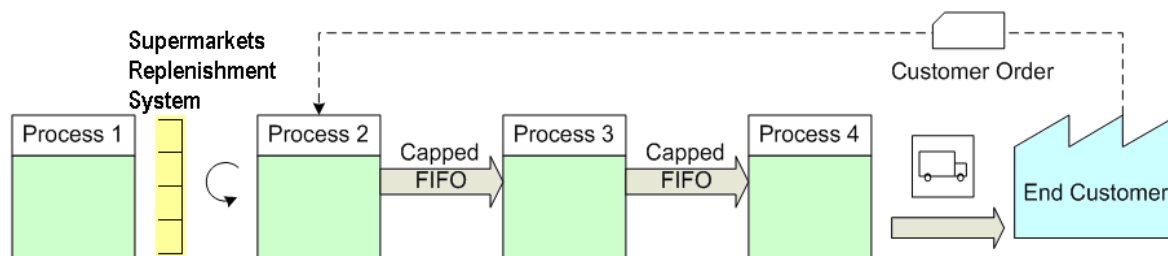
当顾客作业程序需要能选择多样物品号时，采用超级市场补货技巧。假如顾客作业程序是个人计算机的组装程序，这家公司按订单组装个人计算机。在收到订单之前，组装工作无法知道顾客想要的

配件选择。因此，超市建立库存，持有每种想卖的物品，可能有四种不同的声卡。在此例中，为每张不同的声卡需要建立一个特殊的超市补货系统。

II. 拉式系统的类别 (Types of Pull Systems)

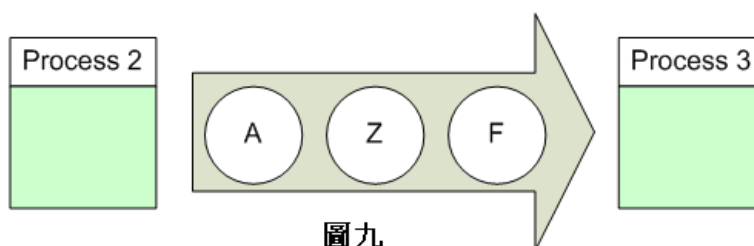
设限的先进先出动线 (Capped FIFO Lanes)

顾客操作的程序不需多种选择时，除了超市 (Supermarket) 以外的方式来看，先进先出动线 (FIFO Lane, First In First Out) 是更好的方式。



圖八

图八呈现一种拉式系统，链接超市的补货系统 (Supermarket Replenishment System) 与设限的先进先出动线。单一计划点 (Single Schedule Point) 的位置在程序二 (process 2)，因为这个程序在系统中超市位置的后面。在程序二与三之间是一种设限的先进先出动线。最易想象的设限先进先出动线的样子是，网球通过一条管子。



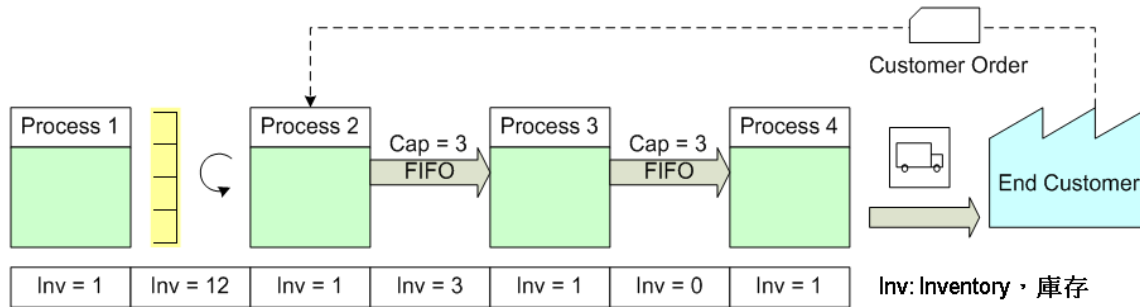
圖九

管子的直径比网球的直径稍大些。网球能顺利地通过管子，而无法改变一个个网球通过的顺序。实质上，没有『超车线』。同时管子的长度是固定的，所以在一个时间点，只有三个球能进入管子。

程序三 (process 3) 接着做哪一个呢？没有选择，就是 F 产品！因此程序三不需要独立的工作计划，其工作由拉式系统来指定。在图八的程序四 (process 4) 也是相同的道理。

如果程序二完成手上的工作，同时 FIFO 动线也满了，它要怎么办呢？程序二会停止工作！这个信号告诉程序二，目前它比系统的其他部分工作速度快。我们前面提过为什么，在这样的情况下不要程序二继续工作的道理。相同地，如果程序三要从 FIFO 动线取得下一个工作时，动在线没工作的话，怎么办呢？程序三停止工作！

这样的系统清楚呈现，在任何一个时间点上，那个程序是此系统最慢的程序。



圖十

图十呈现每个区块目前的库存情况。那个程序是目前最慢的呢？程序三目前是最慢的。称为系统目前的限制（constraint）。能轻松指出限制在那里，就在 FIFO 动线的上游与下游数量比率最大的地方。更简单来看，最慢的程序通常在最大的 WIP 堆积之后，我喜欢称之为『堆积管理（Management by Piles）』。

如果想要对系统的产出有立即的影响，应该将能量投入何处呢？需要专注于系统限制！如果使用拉式系统，监视观察系统以利决定限制所在之处，进而将资源投入该处，即能持续改善系统产出和整体的系统表现。

由于所有非限制在某些时间会没工作，所以也需要决定要这些资源在那些时间做什么事情。

在这个例子中，假如每个工作站限定一次只做一个产品或一个工作，而 FIFO 动线限定一次只能有三个产品，同时这个系统一天完成一个产品，那么一张只要一个产品的顾客订单需要多少前置时间（Lead Time, LT）才能完成？

$$\text{制造时间} = \text{WIP} / \text{比率} = (1+3+1+3+1) / (1/\text{每天}) = 9 \text{ 天}$$

$$[\text{Manufacturing LT} = \text{WIP} / \text{Rate} = (1+3+1+3+1)/(1 \text{ per day}) = 9 \text{ days}]$$

单一计划点究竟在哪里，它很有可能（通常乐意如此）是某些未开工的订单数量定期性地排队等待系统。我喜欢称为『On Deck』，就像棒球的 On-Deck box [下一个上场打球的选手等待的地方]。

On-Deck : 在一种拉式系统，订单等待被开始处理。（On-Deck – Orders waiting to be started in a Pull System）

假如有顾客订单含两个产品在等待，顾客的订单前置时间总共多久呢？

$$\text{整体前置时间} = \text{On-Deck 前置时间} + \text{制造前置时间} = (2/1) + (9) = 11 \text{ 天}$$

$$[\text{Total LT} = (\text{On-Deck LT}) + (\text{Manufacturing LT}) = (2/1) + (9) = 11 \text{ days}]$$

因为我们限制拉式系统中的库存数量，如此一来前置时间之预测就相当一致。这样对了解改善准时达交的表现，或为准时交付给顾客所需投入的努力很有帮助。

注意了，与顾客订单有关的制造前置时间从程序二开始。这个前置时间总是从单一计划点的位置开始。程序一前置间对末端顾客的前置时间没有影响，因为程序二可能需要的所有组件已到位，置于程序一与二之间的超市系统。

为什么 FIFO 动线在此情境比超市方式更有用呢？因为 FIFO 动线：

- 包含较少的库存（Contain less inventory）。
- 减低风险（Decrease risk）。
- 简化排程（Simplify scheduling）。
- 凸显目前的限制（Highlight the current Constraint）。
- 创造带动系统的流动节奏（Create a sense of Flow that drives the system）。

设限的先进先出动线（Capped FIFO Lanes）在下列情况下表现很好：

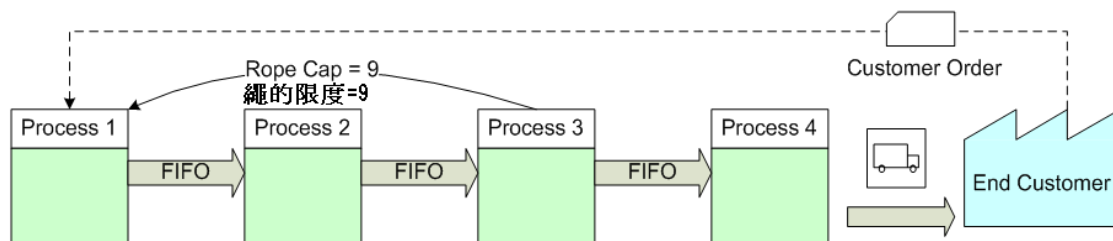
- 高产出数量（Output volume is high）。
- 对合适的零件类，其程序动线始终如一（Process routings are consistent for the applicable family of parts）。

实施设限的 FIFO 动线，必须决定下列的特性：

- 移动批量大小（Transfer batch sizes）。
- 每个程序之先进先出动线的限量（FIFO Lane Caps at each process）。
- 视觉控管以指出限量的状态（Visual controls to indicate cap status）。

鼓缓冲绳（Drum Buffer Rope）

鼓缓冲绳（Drum Buffer Rope, DBR）的是拉式系统的原型，为限制理论（Theory of Constraints, TOC）应用的一部份。与设限的 FIFO 动线系统有雷同的操作概念。



圖十一

在单一计划点与系统限制（资源）之间，有一个整体的最高限度。每次该限制完成一个工作，单一计划点则释出一个工作给系统。这个限度是『绳（Rope）』。由于该限制支配系统的步调，这是『鼓（Drum）』。设计最高限度以使系统限制总是有工作等着它去做，这是『缓冲（Buffer）』，必须确保系统限制（资源）总是忙碌。

注意了，FIFO 动线的下游从系统限制（资源）以下都没设限。这系统中不需要设限，因为下游程序都比该限制速度快。在这部分的动在线 WIP 应该绝不会累积。

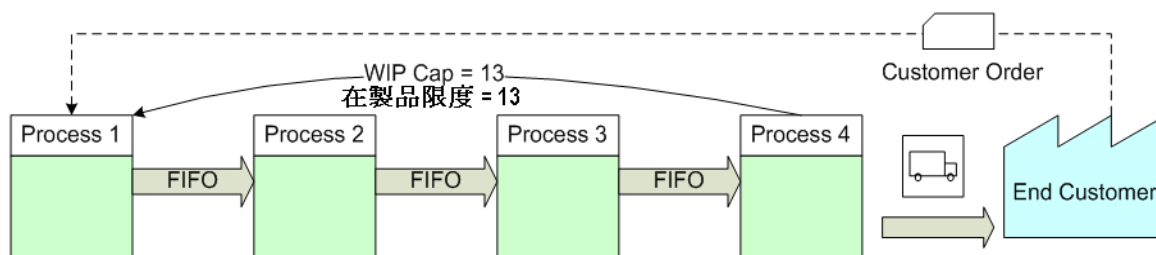
为了使用 DBR，系统限制必须位于相同的程序（或资源）。很多系统显然是这样的情形。实施 TOC 中，我们策略性地决定系统限制的所在，及实行政策以确定它留在原处。从财务的角度来看，当系统限制落在最昂贵的资财上，总是想要使投资报酬率优化。我常称此为『大机器（Big Machine）』。

股缓冲绳提供某些优于设限 FIFO 动线系统的功用：

- 在非系统限制（资源）上的干扰，变动等情况，较不可能使系统限制没事做，而局限了有效产出。这样的系统更能容忍程序变化造成的正常的起落现象及产品组合的改变。
- 只有一个程序（或资源）需要遵守相关的排程规则。单一计划点之程序不可能投入新订单，除非系统限制批准（能是一张简单的卡单（card），筹码（chip），容器(container)等等）。只要有工作，所有其他的程序就执行。在一种设限的 FIFO 动线系统，必须监控每个程序，并且每个程序必须遵守下游 FIFO 动线的最高限度。

在制品限度（WIP Cap）

在制品之最高限度的拉式系统与鼓缓冲绳系统相似，除了限度延伸到系统的最后一个程序，而非到系统限制为止。TOC 实践者经常称此法为『简易之鼓缓冲绳（Simplified Drum Buffer Rope, SDBR）』。我通常认为这种拉式作法最容易操作，与其他方法相比更有效能。



圖十二

在制品限度系统（WIP Cap system）提供一些优于 DBR 与设限 FIFO 动线系统的功用：

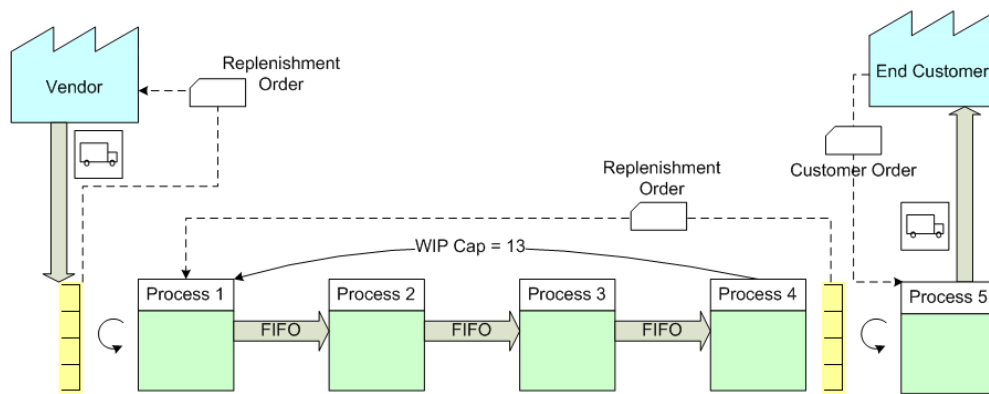
- 在非系统限制上的干扰，变动等，不可能使系统限制没事做，不使有效产出受限。

- 只有一个程序（或资源）需要遵守相关的排程规则。
- 不必要有一个固定的系统限制。
- 容易指出目前之系统限制。相较于设限的 FIFO 动线，此系统减少『不正确的解读』。
- 实施简单。事实上，经常开始的方式是，就以当前的情况设定 WIP 限度，然后用方法逐步降低。这会避免大部分受实施改善影响的人员提出借口的情况。

我认为对大多数进行中的应用构面而言，在制品限度的方式是最佳的拉式系统。非常有效用，容易准备也容易维护。假如你有兴趣了解细节，可参考我所书写的相关内容。

综合拉式系统（Combination Pull Systems）

大多数价值流包含两种或更多的拉式系统。一般而言，最少一种超市系统用于原物料及/或成品部分。WIP 通常受到设限 FIFO 动线系统，鼓缓冲绳，或 WIP 限度系统的控制。



圖十三

图十三呈现一种典型的综合拉式系统，其中组合许多我们提过的概念。这样的系统简易地排程与控管工作进度，从订购原料到交付成品。

III. 实施拉式系统(Implementing Pull Systems)

仅仅是限制库存需求的作法不足以启动拉式系统。我的经验是人们不会维持拉式的运作，除非建立一套正式的系统，不然难以落实拉式的运作方式。虽然相关的概念不难理解，但是却与一般的直觉背道而驰。在改变过程中遭遇困难或情况发生变化时，自然的反应倾向于走回头路，回到原来熟悉的作法。

同时，需要细心地设计拉式系统，以达最大效能。此外，只要数量、产品组合或其他关键变量有显著变化时，则需要调整拉式系统，或甚至重新设计。

要如何梳理当前情况，并开发及成功实施一套拉式系统呢？有很多的作法，最容易的地方是从价值流程图的解析（Value Stream Mapping）开始。

价值流程图解（Value Stream Mapping）

这个高层次的流程图解方式专注于指出无加值的时间（the Non-Value-Added time），产品在生产系统中停留，花费无加值结果的时间。由于在典型情况是前置时间的 95%以上是无加值时间，价值流程图解析方法能有效指出实施拉式系统，与缩短前置时间的机会。

一种典型的作法是选择一个产品或一种产品进行图解。如果选择一个种类的产品，这个种类的产品应该使用同类型的资源（机器，车间，员工等），或未来能使用同类型的资源。接着选择流程图的起点与终点，经常是从接到订单到送交产品为止。然而，在一开始实施拉式系统，经常观察价值流的一部份，例如一个部门。

建构价值流程图从起点到终点，都使用一套标准的标志说明，及视觉上的描述。注 - 本文中大多图形使用标准的价值流程图的标志。

一旦现况价值流图（the Current State Map），应用一系列的步骤来建构一张未来价值流图（a Future State Map），这张图描述一个近程的未来系统。未来价值流图上加有一些改善的缺口（Kaizen Bursts），这是要能从现况转成未来的重要改善之处。

价值流程图法的完整描述超出本文范围。The Lean Enterprise Institute (LEI)有相当详细的手册，名称是『Learning to See』。

实施拉式系统之额外备注：

1. **从学习开始** - 拉式系统是反直觉的，跟我们熟悉的想法不同。在选定的改善区块，初期你需要与公司的关键领导者及管理者的，开发使用该作法与利益的基本理解及共识。

2. **从小区块开始** - 选择能够建置一个超市/或有限制在制品的区块。仔细地盘算前置时间与绩效的改善方案。透过这个区块的改善过程进一步学习，并宣传成功案例。一旦这个区块的系统能自己运转，则实施另一个区块。
3. **保守地设定初始的库存限度（高）** - 有利于组织建立信心。随着时间，层层改善直到最终的目标。
4. **设立待命的支持团队** - 发生在系统中的问题，是人为的！这就是改善的机会。假如没有资源去协助系统中工作的人员，他们会回归到老方法。同时，假如没正确的实施，让改善成效显然可见，会导致错失拉式系统半数以上的利益。
5. **事先决定没工作时，员工要做什么** - 只有一个资源能是限制资源（the Constraint），而其他资源有时会没工作做。在这样的情况发生前，先想好到时员工该做什么事。多种职能训练与弹性派工方式就可减缓这个问题。

IV. 结语 (In Conclusion)

拉式排程系统 (Pull Scheduling Systems) 能大幅且快速地改善公司的绩效，能有利于促进长期的持续改善。然而，这些系统与本能的直觉相反，并未被大多数管理人与重要决策者充分理解。实施这些拉式系统并不困难，但是需要仔细规划实行的作法、培训及持续学习的毅力。

秉持持续改善的精神，希望能得到读者对本文的宝贵意见，作者希望能根据自己的实务经验与读者的回馈，持续修改本文的内容。

附录 A - 拉式系统比较表 (Appendix)

	超级市场补货 (Supermarket Replenishment)	设限的先进先出动线 (Capped FIFO Lanes)	鼓-缓冲-绳 (Drum Buffer Rope)	在制品限度 (WIP Cap)
提供选择 - 适合于当顾客之作业程序需要选择多种产品	是	不	不	不
相对的库存程度	高	中等	中等/低	低
过时库存的相对风险	中等	低	低	低
排程的难度	中等	低	低	低
支持单点排程	是	是	是	是
快速与正确指出当前限制的有效性	低	中等	中等	高
提升系统流畅度的创造力	低	高	高	高
非系统限制上的干扰会使限制没事做的可能性	高	高	低	中等
几个程序必须遵守排程规则	全部	全部	一个	一个
需要一个固定的限制	不	不	是	不
最适合的应用	有选择性及/或顾客的前置时间需要一个存放点	控制 WIP, 在大量重复性的制造情形	控制 WIP, 在知道定点的系统限制的情况	控制 WIP, 在多种多样的应用情况

专门术语与概念之词汇表 (Glossary of Terms and Concepts)

Terms	术语	概念	Concepts
Average Demand per period	一段期间的平均需求值	一段时间内的平均顾客需求值 (通常是一天或一周)。	Average customer demand per a period of time (usually a day or a week).
Buffer	缓冲	在一个限制资源前, 排队等待的工作。缓冲是有必要的, 以防止限制资源无事可做, 而降低有效产出。	Work waiting in queue in front of a Constraint. A buffer is necessary to keep the constraint from starving for work, which will reduce throughput.
Business processes are Dependent	业务程序的相依性质	改变一个程序会在某程度上, 影响某些或其他部份的程序。	A change in one process will tend to affect some or all of the other processes in some way.
Business processes are not additive	业务程序不是加总性质	改善一个程序通常不会正面地影响营业的整体表现及获利能力。事实上, 根据业务程序是相依性的原则, 可察觉到常常一个区块的改善, 会降低系统的整体表现。	An improvement to one process will usually not affect the overall performance and profitability of a business in a positive way. In fact, based on the <i>Business processes are Dependent</i> principle, what is perceived to be an improvement in one area will often reduce the overall performance of the system.
Capped FIFO Lanes	有限制的先进先出动线	一种拉式方法, 于工序间以先进先出的作业顺序, 来限制在制品的数量。	A Pull method that limits WIP between processes with a processing sequence of First-In-First-Out.
Constraint	限制	一个资源或一个政策限制一个系统的产出。	A resource or a policy that limits the output of a system.
Cost Formulas	成本公式	前置时间=成本=在制品/速率。这表示必须降低在制品, 以降低前置时间及成本。	Lead-Time = Cost = WIP / Rate This implies that you must reduce WIP to reduce lead-time and cost.
Customer Lead-time (LT)	顾客前置时间	从收到订单到送交产品的总时间。	The amount of time from order receipt until product delivery.
Customer Process	顾客程序	一个拉式系统的下流程序。	Downstream process in a Pull System.
Direction of	信息流的方向	在一个推式系统中, 在	In a Push System,

Information Flow		整个价值流的信息与产品流的方向相同。在一个拉式系统中，在整个价值流的信息与产品流的方向相反。	information and products flow in the same direction through the Value Stream. In a Pull System, information and products flow in opposite directions through the Value Stream.
Drum	鼓	在鼓缓冲绳的拉式系统中的限制程序，它提供系统中所有其他程序遵循的节奏或步调。	The Constraint process in a Drum Buffer Rope Pull System. Provides the cadence or pace which all other processes in the system follow.
Drum Buffer Rope	鼓缓冲绳	一种拉式方法，限制系统中之系统限制（鼓）与第一个程序间的在制品（绳）数量，并设计成总是有些在制品（缓冲）等在系统限制程序之前。	A Pull method that limits WIP (Rope) between the Constraint (Drum) and the first process in the system. Is designed so that there is always some WIP (Buffer) waiting before the Constraint process.
FIFO Lane	先进先出的动线	先进先出的作业顺序规则。	Processing sequence rule of First-In-First-Out.
Flow	流动	在价值流程上，持续向前移动的产品作业程序。	Processing state where products move continuously forward in the Value Stream.
Kanban card	广告牌卡	卡、片、容器或其他用具，用于一个拉式系统，以授权资源开工。	Card, chip, container, or other device used in a Pull System to authorize a resource to do work.
Lead-time reduction benefits	缩短前置时间的利益	<ol style="list-style-type: none"> 1. 增加销售 2. 改善质量 3. 降低营运成本提升产能及有效产出 4. 降低投资资产 5. 增加员工的满意度（士气） 	<ol style="list-style-type: none"> 6. Increased sales 7. Improved quality 8. Reduced operating costs 9. Increased capacity and throughput 10. Reduced invested assets 11. Increased employee satisfaction (morale)
Lean Cost Accounting	精实成本会计	成本会计、衡量与决策方法支撑及与精实生产一致，一般倾向缩短前置时间及改善现金流的决策。	Cost accounting, measurement, and decision methods that support and align with Lean Manufacturing. Will generally favor decisions that reduce lead-time and improve cash flow.

Little's Law	利特尔法则	前置时间=在制品/速率	Lead-Time = WIP / Rate
Manufacturing Lead-time (LT)	制造前置时间	从释出订单到制造，直到完成产品及可送交的总时间。	The amount of time from release of order to manufacturing until product is completed and ready for delivery.
On-Deck	预备区	一个拉式系统中等待开工的订单。	Orders waiting to be started in a Pull System.
Pull System	拉式系统	一个排程系统中，以某种方式限制库存数量。注：库存可以是原物料、在制品、成品等。	A scheduling system where inventory is limited in some way. Note – Inventory can be raw materials, WIP, finished goods, etc.
Push System	推式系统	一个排程系统中，库存数量并未被正式限制。	A scheduling system where inventory is not formally limited.
Rate	速率	每个时间单位中，完成的产品数目。	The number of products completed per unit of time.
Re-order point	再订购点	在超市补货系统中，只要是库存加上开工数量的总和，低于再订购点的数量，就发出一张新订单给供应程序。	In a <i>Supermarket Replenishment</i> system, whenever the inventory plus the open order quantity drops below the re-order point, a new order is generated for the Supplying Process.
Re-order quantity	再订购数量	拉式系统中产品的标准固定的再订购数量。	Standard fixed quantity of product to re-order in a pull system.
Rope	绳	在鼓缓冲绳的拉式系统中，限制在第一个程序与系统限制程序间的在制品数量。	The WIP limit in a Drum Buffer Rope Pull System between the first process and the Constraint.
Service Level	服务水位	乐见的补货时间比例，以维持超市货架至少有些产品，可用于顾客程序。	The desired percent of time that the Supermarket Shelf will have at least some product available for the Customer Process.
Single Schedule Point	单一排程点	在拉式系统中的单一程序，这是独立排程之处，其他所有程序则依拉式系统的运作排程。	Single process in a Pull System where the entire Pull System will be independently scheduled. All other processes are scheduled by the Pull System itself.

Single point schedule location	单一排程点的位置	整个拉式系统中，紧跟最后超市点之后的程序。	The process immediately after the last Supermarket in the entire Pull System.
Supermarket Replenishment	超市补货	拉式方法中，以超市链接顾客程序与供应程序，建立每一个部件的再订购点与再订购数目。	Pull method where a Supermarket connects a Customer Process with a Supplier Process. For each part number, a re-order point and a re-order quantity is established.
Supermarket Shelf	超市货架	在超市补货系统中，实体成品的存货位置。	The physical stocking location for completed product in a Supermarket Replenishment system.
Supplying Process	供应程序	拉式系统中的上游程序。	Upstream process in a Pull System.
Throughput	有效产出	系统在每个时间单位的完成工作的总数。一般而言，有效产出=（销售-直接物料）。	Amount of work completed by a system per unit of time. Typically, Throughput = (Sales – Direct Materials).
Traditional Cost Accounting	传统的成本会计	成本会计的方式试图将成本分摊到产品。这个方式，分摊经常费用，通常与直接人工成本成比例。	Cost Accounting approach that attempts to assign costs to products. In this approach, overhead is allocated, usually in proportion to direct labor.
Transfer batch	转移批量	容许被搬动或转移从供应程序到顾客程序的工作数量。	The amount of work that is allowed to collect up before being moved or transferred from a Supplier Process to a Customer Process.
Variability in Demand per period	每时期的需求变化性	期间性的需求数量变化，一般以标准偏差衡量之。	The amount of variation in demand from period to period. This is typically measured in Standard Deviations.
Visual Controls	视觉控制	员工能以视觉来控制程序，不需或只需很少的指令。典型的例子包含停车场、停车灯、空货架等之动线。	Controls that people can process visually with little or no instruction or direction. Typical examples include lines in parking lots, stop lights, empty shelves, etc.
WIP Cap	在制品限制	拉式方法中限制几个程序步骤间的在制品数量（不是限制每个程序的数量）。	Pull method where Work in Progress is limited across several processing steps (but not specifically limited at

			each step).
Work in Progress (WIP)	在制品	已开工，未完成的工作。	Work that has been started, but is not yet complete.