POLITECNICO
MILANO 1863

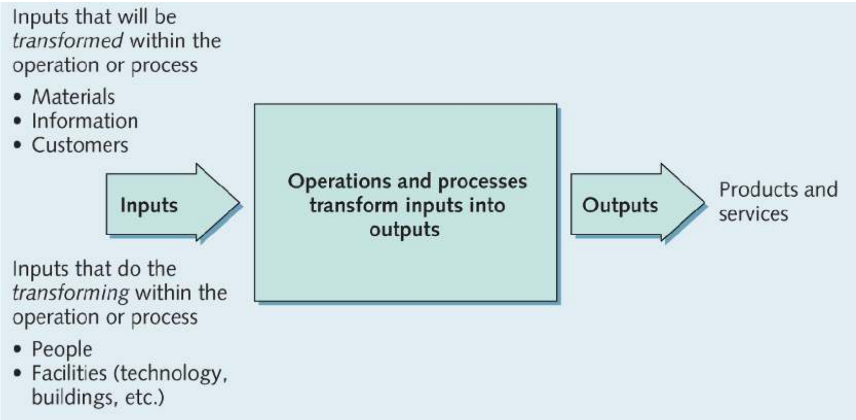
NOTES OF OPERATIONS MANAGEMENT

A.Y. 2021-2022

SUMMARY

1. INTRODUCTION

Operations and process management is about **how organizations produce goods and services**. The tasks, issues and decisions that are necessary to manage processes effectively, both within the operations function and in other parts of the business where effective process management is equally important. Every organization has an operations function because every organization produces some mixture of products and services. It is always concerned with managing the core purpose of the business – producing some mix of products and services. Processes are the component parts of operations. In fact, every part of any business is concerned with managing processes.



Operations management has expanded from treating only the core production processes in manufacturing organizations to include service organizations, non-core operations processes and processes in other functions such as marketing, finance and HRM. Operation mangers are concerned with each step in providing a service or a product. Consider that some activities, even if they are essential for the company, do not create any value (ex. Accounting, sales, etc.); operations are strictly related with the creation of value.

Why this course?

- Operations are central to the company's business and mankind's sustainability.
  Operations' relevance in a Company
  - Spending on material purchase can be 50-80 % of turnover/total cost
  - Spending on personnel can be 30-50 % of turnover/total cost (50-80 % in services)
  - The annual cost of the facilities (plants, information systems, etc.) can be 30-40 % of turnover/total cost
  - Investments are high and are binding for a long time, so making the right choices is crucial:
    - a site can cost tens of millions of € and take 2 years to be realized
    - develop distinctive competences even more
- The skills in the Operations area are critical and strongly impact the competitive advantage. Both in industrial companies and in the service ones.
- Operations contribute to Companies strategy
- Services are a special world, and a rapidly growing ones. Operations in services are at dawn: companies are struggling to find service-specific skills.
- There is a strong innovation taking place in the management of companies and organizations: LEAN.

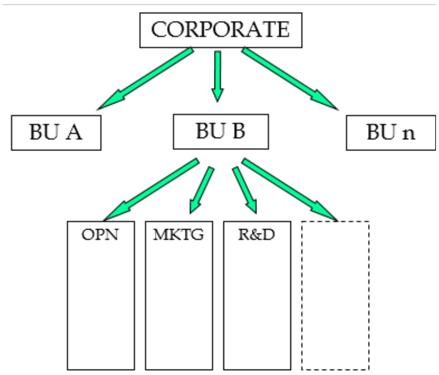
2. OPERATIONS STRATEGY

2.1. STRATEGY

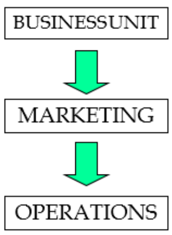
The success of a company is determined by the strategy it pursues in the market. With the business strategy, the company tries to gain a **competitive advantage** over the potential competitors in the market, and the survival of the company is directly related to the competitive advantage that the company has built over time.

2.2. STRATEGIC LEVELS

If we talk about strategy in general, we can have strategy that works at the **corporate level**. Imagine large multinational companies, for example Procter & Gamble or Nestlé, these are very large companies where the company encompasses many environments. Then there is the deployment of this corporate strategy at the lower levels, in the different **business units** of the company. Within a business unit, there is a business unit manager who manages the corporate strategy for a number of business units. So, he decides where he wants to compete in the market and how he wants to manage the different resources and functions within the business unit. Then there is the deployment at the most operation level, where customer contact is established.



In these views, it is a completely **hierarchical view**, where we have a brain in the holding that provides inputs to the lower levels, which then distribute those inputs and outputs and generate inputs to the still lower levels. So it's a top-down hierarchical view and that's the way traditional companies have been run.



This is the traditional view, which is perfectly coherent with the **top-down view**. There is the owner that has an idea. That idea is then passed on to the manager. The manager has to deliver results, so he has to generate revenue. First, he tells the sales-people and the marketers, "We have to sell this, we have to get these customers". So he first creates a plan with the sales-people and marketing, which then becomes a target for the operation function, which has to meet the requirements of sales.

2.3. RELEVANT FACTORS OUTSIDE THE COMPANY

This is the traditional view, but this started to change because of many reasons

- **Offer > Demand**: because the world is changing, what is changing? The supply has become much greater than the demand. Today, for 99% of the market, global capacity is much greater than global demand. And that is important because it is no longer true that every time you produce, the customer will buy it for sure. The market offers so many options that the customer can now choose which product to buy.
- **Customization**: If I have many options, then as a customer I start to want a product that is as close to my specific needs as possible. Then I start asking my suppliers to produce something that is suitable for my individual request. This is true for both the B2C customer and the B2B market.
- **Globalization**. Globalization is a cause of the first point. In this sense, it is the ability to capture the supplier worldwide, which means that we, as consumers, are at the base of the potential supplier, which is bigger. So, it means that the supply is bigger.
- **The speed of the innovation**. Basically, it is innovation. The faster the innovation, the newer opportunities. So, the greater the offer and therefore the greater the possibilities for the customer.

All of this makes competition between companies even more intense and underscores the importance of creating a competitive advantage for a company. This is much more important than in the past. The points we have just seen come from outside the company.

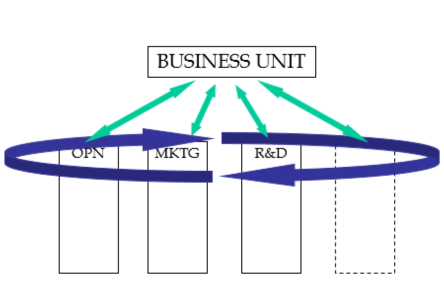
2.4. FACTORS INFLUENCING COMPANY'S RESOURCES

There are other factors that come from within the companies, from the operation side of the companies that will change the world and the opportunities that come from within the company itself.

- Cultural (educational level): the people who are in the company today are completely different from those we had in the past. The level of education of the people who go to work today is not comparable to what we had in the past. The people who go to work today are much more educated, they are much stronger with complex ideas that can make a real contribution even at the strategic level of the company. Nowadays, people in industrial companies often say on TV or on the radio that people are the most important resource. This is true, because there is a great potential, because people today are educated, they know a lot. So, they can really contribute to the success of the company.
- Another opportunity is ICT technology / Industry 4.0. Companies today have a great opportunity to use technology. Until 15 years ago, this was not the case. The Internet was not available to everyone because it was expensive. 3D printing was something like 'what are we talking about'. Today, there are a lot of technologies that are available to businesses at a low price. So, there is a technological opportunity on the company's side that they can take advantage of. It is not just given to me from the outside, but it is an input from the inside to develop new skills and new knowledge.
- Economical (wealth increase)
- Social (authority acknowledgment)

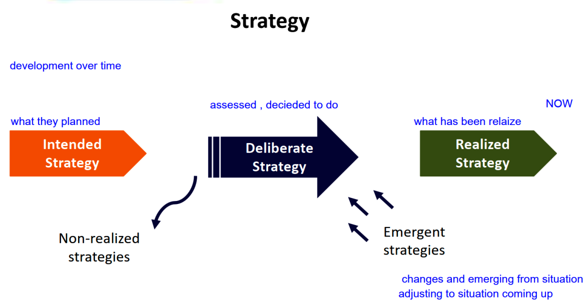
2.5. INNOVATIVE APPROACH

The world is changing and that leads to different requirements for companies. We cannot use a management model that was good 30 years ago. Therefore, the management model must change to follow the new demands and opportunities of the market. How is the managerial approach evolving? Instead of a rigid and
1 cumbersome hierarchical view, a more integrated view of the management approach is sought, where the different functions of the business unit complement each other. So, there is no longer the marketing department or the operations department, but there are also connections showed in the picture. So now when you set up the business unit, you consider the integration of the different functions. That means that when you set the strategy, you do not just go for those parts, you integrate both.



2 The second point is that it is no longer just top-down. The strategy, to be successful, has to be a **horizontally integrated strategy** where all functions consider opportunities and requirements at the same time. A strategy that also considers the opportunities that arise from **bottom-up** knowledge in the field.

Today, strategy must be a **continuous interaction between vision and mission**, board vision and feedback from the field. This innovative approach to create a strategy means that we end up with a strategy that is composed of a part of the strategy that we call the **deliberate strategy**, that is, the strategy that we develop at the board table, and the **emerging strategy**, i.e. the part of the strategy that comes from practice, from the opportunities that the different functions create in their daily work.



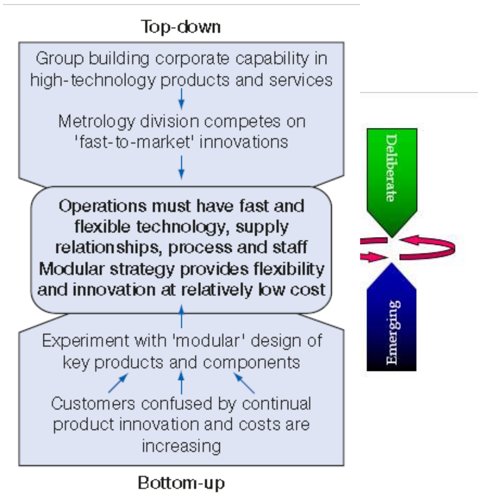
How does this affect operation strategy? It has an impact on the operation strategy, because first of all, the operation strategy has to have a meaningful level, which means that what has been decided by the business unit is not only **top-down**, but also **bottom-up**, because we have to influence the decision of the board based on the opportunities that we see. So even with the operation strategy, there is a part that is being implemented and a part that is evolving. So operations have to play a proactive role in defining the strategy.

2.6. STRATEGY

What is an operation strategy? It is a strategy that consists of the goal and the steps.

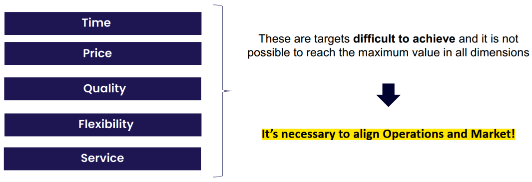
- The operation strategy refers to a **long-term vision** of the company. An example of an investment in an operation strategy is a new plant for the operations manager.
- Operation strategies refer to a **higher level of analysis** and to a **greater level of aggregation**. So, in the case of Barilla, we are not talking about the Penne Rigate item or the Parma market. But rather we are talking about the gluten-free family.
- Operations strategy uses higher **levels of abstraction**. In this simplified supply chain model, we do not go into the single workstation, but you need to abstract to work at these higher levels.

Example: Operation strategy is an integration of deliberate strategy and bottom-up strategy. A manufacturer of metrology instruments wants to increase the competitiveness of the company by improving time to market (he wants to be faster than the competition in developing new products). This is the point of view that requires a certain type of investment: R&D is expensive for this type of industry. What are the opportunities from below? Some experiments, some contacts with customers. Some functions have made experiments with modular design to innovate. Modular design is used to create sample innovations. They have done some experiments and it has worked well. This is an opportunity that influences operation strategy. How can we innovate? You can reduce time to market by using modular design innovation instead of direct product innovation. That is integration.



2.7. THE VOICE OF THE CUSTOMER

So how do we create an operation strategy? Operation strategy consists of setting a goal and planning the steps to achieve that goal. How do we set the goals for an operational strategy? We start by **understanding the needs of the customer** to satisfy them, because that is our goal. The goal comes from the voice of the customer. How do we practically implement the customer's needs? By **positioning**. Positioning is about am I cheap or expensive? Do I want to compete because I am cheaper? Do I want to compete because I offer better quality? Do I want to compete because I offer a better level of service? What do we give the customer in the end? We give him services.



What are the performances - **KPI**? TPQ FS

- **Time**: it can be seen in two main variables. The delivery speed and the punctuality. We can set a target for time performance or compare time performance with competitors.
  - Response time. The response time is the time that elapses between the customer's request and the delivery to the customer. For example, there is someone who chooses between Glovo and Deliveroo, depending on how long it takes to get to your home.
  - Punctuality: how well is the delivery date specified by the customer met. This is a typical performance that really affects the supplier. It is important to keep the delivery exactly on time. Either you deliver on time or you pay a penalty.

- **Price**: The price that for the operation perspective is the cost. The cost is usually not just the cost of purchasing, but the total costs (Total cost of ownership). We start a project with a company that manufactures coffee machines and wants to examine what the actual cost of the product is, i.e.,
  - The cost of production
  - The cost of distribution
  - The cost of using the product
  - The cost of maintaining the product
  - The cost of disposing of the product at the end of its life.

  This is a driver of competition. Who determines time performance in the company? Who is responsible for it? Operations.

match with customer needs

- **Quality**: there are two types of quality performance. The quality of design and the quality in terms of انطباق conformance. ability to deliver the system/product with the same performance

  Example: **Panda and a Ferrari**. The design quality is undoubtedly higher for the Ferrari. But what happens when we look at quality in terms of conformance, that is the promises made to the customer? Ferrari makes 10,000 cars a year, and the top speed is not the same for all cars. We have some cars with 300 kilometres per hour, other cars with 305 kilometres per hour. Maybe the Panda cars all have the same top speed of 170 kilometres per hour. So the conformance in this case is higher for the Panda than for the Ferrari.

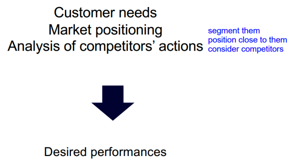
  Example: **McDonald's** is a company that has the quality of conformance as the basis for success. How does McDonald's compete? By price, of course. But a really important factor is quality of conformance. Because the performance of Big Macs is so high, that becomes a competitive advantage. Because wherever you go in the world, when you walk into a McDonald's, you know exactly what it is going to taste like. And that is really a competitive advantage over other suppliers. And where do they get that competitive advantage? Because the quality of conformance is very strict.

- **Flexibility**:

  new product with small effort,
  - Flexibility of the product, it means being able to offer a variety of products, we can provide the hamburger with meat, we can provide the hamburger with chicken.
  - Customization, how much you are able or how much you offer to customize the product you send to the customer. % of product you can customize
  - Variety of flexibility: it is similar to that of the product. 50 colors
  - Flexibility of the plan: It is about flexibility in delivery and the relationship you have with the customer. For example, if I have to choose between two doctors, one of whom offers an appointment two days a week in the morning, while the other gives you the option to write him a message and make the appointment on weekdays, that is the flexibility of the plan.

ability to change in customer order before delivering it. measure by frozen horizon (time before delivery date you can't change)

- **Service**: this is a sphere that includes many additional services that can be offered to the customer to increase the value of what is given to me. For example: you are an industrial manufacturer and you want to buy an industrial machine to produce paper rolls. You have a supplier that sells you only the machine, you can add the other supplier that sends you the machine, then if you want the maintenance service, then if you want the instructor to automatically update the software that you have installed on the machine.



2.8. OPERATIONS' OBJECTIVES

How do you translate customer needs into the goal for operations? They are translated using **KPIs**. We have KPIs that are required by the market, or target KPIs that we want to achieve. These KPIs are in trade-off in an equal resource condition, **if you drive one, you will lose the other**. It is very difficult to be the best in these conditions of equal resources. That is the first difficulty for the operations manager, because that means that because of the trade-off you have to make a decision based on the conflicting goals. It is too easy to say "ok, I want to be the best in all areas". You have to decide on the goal you want to achieve or the customers you want to serve. It is necessary to align Operations and Market.

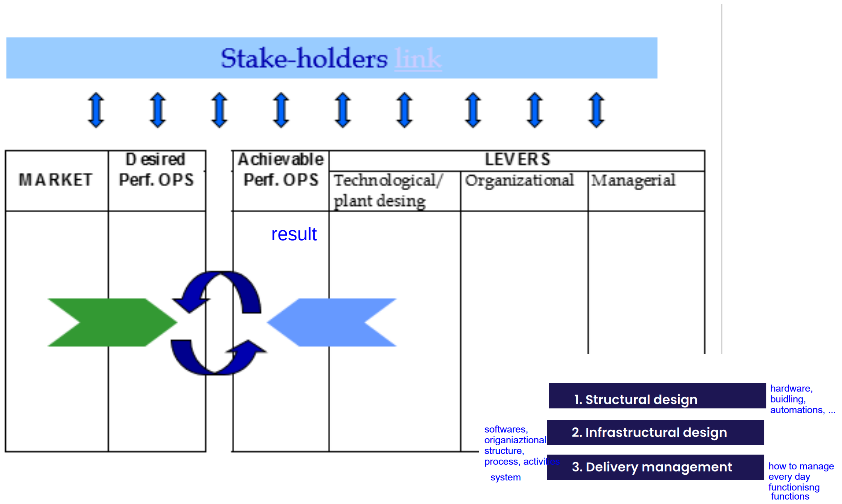
Starting with customer needs, i.e. market positioning, competitors and definition of performance. Then we have to define the second part of the operation strategy, which translated means which levers the operation manager has to activate or invest in the operation system

2.9. STRATEGIC LEVERS

What are the levers that we can activate to achieve our final goal?
This modelling is called the **reconciliation model**. The levers in the hands of the operations manager can be divided into three main categories اصلاح

- Structural design choices
- Infrastructural design choices
- Delivery management choices

Why is this model called a **reconciliation model**? Because this model helps the operations manager make decisions on operating levers that are consistent with the goals and objectives set by the business. Reconciliation because you need to align the levers (your decisions, your investments) with the market demand you want to satisfy.



2.9.1. Structural design choices (investment)

The **overall production capacity**, its allocation and location, i.e. how I set up the overall capacity.
Example: I have a demand of 10 cars per month and I am the operations manager. Next year we will sell 11 cars. It is up to the operations manager to decide how I set my capacity based on demand. I can set it at exactly (11), I can set it at a higher level than demand 15. What does it mean if demand is 11 and capacity is 12 or 15? These decisions 10, 12 or 15 have an impact on your performance. A capacity of 15 and a demand of 11 results in a saturation of 65%. What is the impact of low saturation on our performance? Higher costs than the other option. But we also have advantages because we have better time performance or better schedule flexibility performance. If I set production at 12, I have better cost performance, but I run the risk of losing market opportunities. Should I maintain a centralised plant or is it better to maintain plants in different countries and produce for the local market? Or do I specialise the plan for one type of family or in each plant a replicated production of the families. **Make or buy**. I have resources, do I manufacture all activities internally or do I outsource some activities. Which activities do I outsource? Do I push some activities or do I keep one activity in-house but outsource an entire product line? **Technological process and equipment**: what level of technology and automation do you want for your operating system? **Supply chain configuration**. Do I want to configure a few or single suppliers, or do I prefer to split supply among multiple suppliers? Do they want close collaboration?

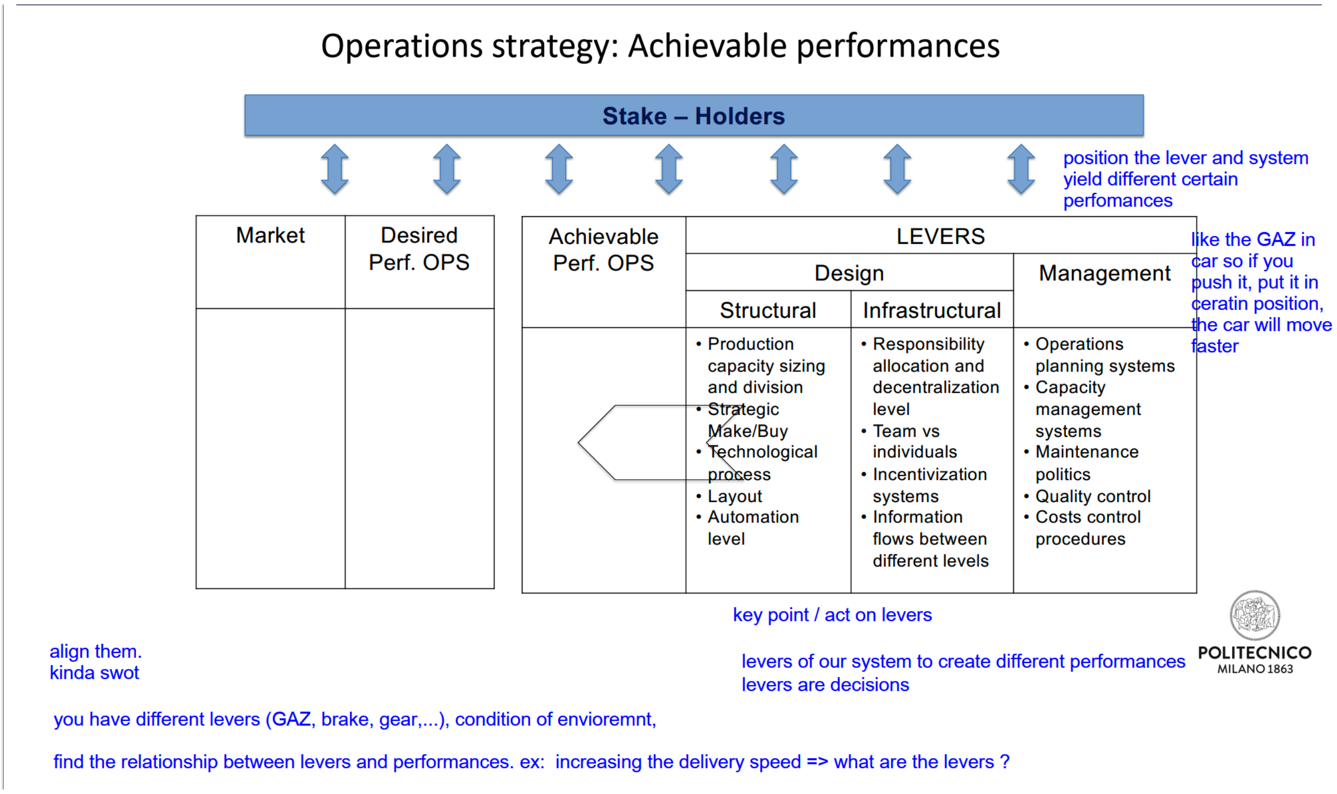
2.9.2. Infrastructural design choices (organizational)

How do I organise my resources? Do I divide them by process? I divide them by value stream. I put them in **teams** or make them work **individually**. This has an impact on performance. If you work in teams, you have better performance in terms of time and quality. But maybe you have higher costs because it is more difficult to saturate a single person when working in teams. Trade-offs! How do I set up my **incentive system**? How do I set up these bonuses? For example, if you make an incentive system for production, how many items you produce in the last year you will have a production manager who will push production as much as possible without caring what you are setting. This could be an opportunity, but also a risk.

2.9.3. Delivery management choices

These levers relate to the managerial sphere. How do you want to satisfy your customer? Do you want to satisfy them with a make-to-stock approach or with a make-to-order approach? What are the implications on performance? How do you coordinate with your supply chain partner? **Maintenance**, have you developed the maintenance capability internally and have a team to maintain your machines, or do you outsource the maintenance activity? Then establish a maintenance service agreement with your machine suppliers. What are the cost and time implications?

Continuous improvement systems. How many resources and how much time of your resources you devote to continuous improvement plans. What operational planning system do you use? Do you work for immediate release, do you adopt workload control logics? What are the priority rules you set, which customer do you prioritize? Do you have a priority? It is a FIFO logic. Is it a larger customer logic?



Capacity sizing, division and localization

Entry barriers
Face demand variability

Economies of scale
Volume flexibility
Minimum unit of increase/decrease

Work cost
Competences availability, services,..
Response time
Closeness to market

3. MODELS FOR OPERATIONS STRATEGY

Any operations strategy should reflect the intended market position of the business. Companies compete in different ways: some compete primarily on cost, others on the excellence of their products or services, others on high levels of customer service, and so on. The operations function must respond to this by providing the ability to perform in a manner that is appropriate for the intended market position. This is a market perspective on operations strategy. But operations strategy must do more than simply meet the short-term needs of the market (important though this is). The processes and resources within operations also need to be developed in the long term to provide the business with a set of competences or capabilities. If they are sophisticated and integrated, they can form the basis of the business's ability to offer unique and 'difficult to imitate' products and services to its customers. Operations strategy should reflect market requirements.

3.1. PERFORMANCES

A particularly useful way of determining the relative importance of competitive factors is to distinguish between **Order Winners (OW)** and **qualifiers (Q, QQ)**.

for short listing !? delivery and punctuality certificate,

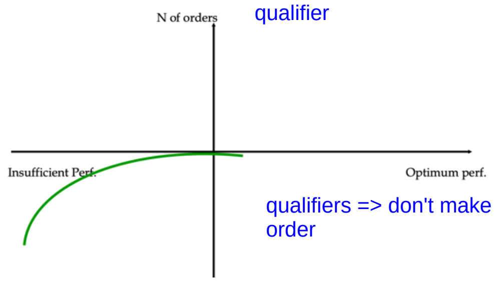
QQ: short list supplier, performances doesn't have relation with orders even if you improve it.
OW: we improve we get orders, we got lose order quicky if got worse

3.1.1. Qualifiers

Qualifiers may not be the main competitive determinants of success, but they are important in another way. They are those aspects of competitiveness in which the operation's performance must be above a particular level just to be considered by the customer. A performance below this 'qualifying' level may disqualify the operation from being considered by customers. But any further improvement above the qualifying level is unlikely to gain the company much competitive advantage. Qualifiers are those things that are generally expected by customers. Being great at them is unlikely to excite customers but being bad at them may disadvantage the competitive position of the operation. The graph aside is a representation of a performance which is a "qualifier".

If we assume that the graph origin is the starting point of the evaluation of our performance, we could notice how:

- If we worse our performance, we may be disqualified by customers from being on the "qualification list" – even if losing a bit is accepted.
- If we improve our performance, we will not have any advantages: no extra orders are allowed because the performance was qualifying but not strategic for our company



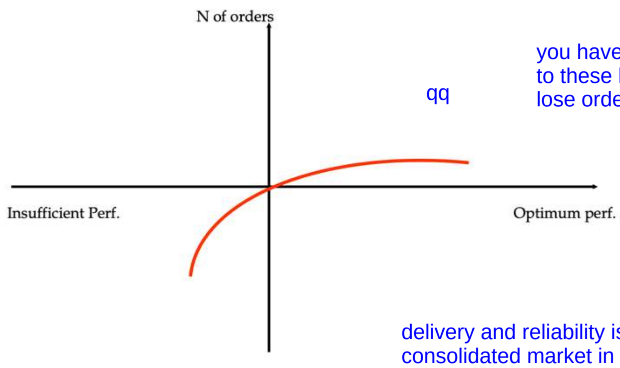
qualifier

qualifiers => don't make us have more order

The company needs qualified people to meet the requirements to be on the short list. The performance for these must still meet minimum level expectations so that resources are not wasted. This second graph on the left represents another kind of qualifier which is called "**order loser**".

In this case:

- If we worse the performance of these competitive factors we are not immediately qualifying on the "short list" of our customers
- On the other hand we are presently losing orders and in a short term period this lack could bring to a disqualification.
- This is the situation where the company is currently losing opportunities. Thus, it has to invest in those.



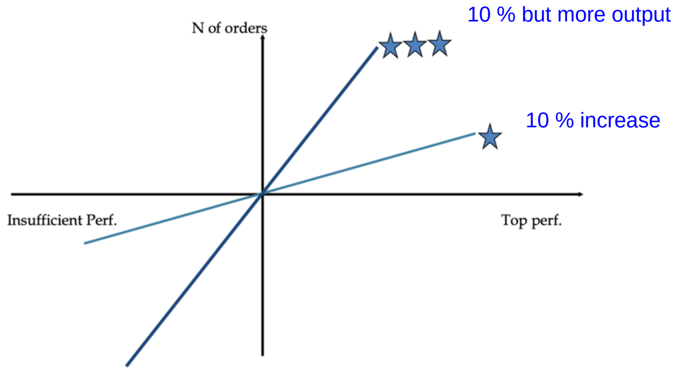
qq

you have to pay attention to these because you'll lose orders

delivery and reliability is order loser in consolidated market in hq case

3.1.2. Order Winners

Order winners are those things that contribute directly and significantly to winning business. They are considered by customers as key reasons for buying the product or service. Increased performance in an Order Winner will result in more business or improve the chances of getting more business. *Order Winners show a steady and significant increase in their contribution to competitiveness as the operation becomes better at supplying them*.



As the graph shows, higher ranking competitive factors would bring higher performance than lower ranking factors. Even though the ranking between the different performances usually shows minimal differences, we must be able to understand which one, with the use of the same amount of resources, is more important for the improvement of the company and therefore, which one has the steepest curve.

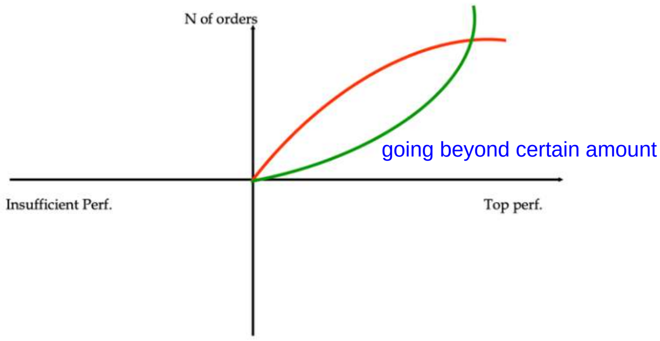
3.2. HOW TO DEFINE PERFORMANCES

In order to define performances, you have to consider a set of representative customers and a set of significant orders or any other type of contact customers had with the company. The first step is **how to collect this information**: you may use interviews or surveys in order to identify the previous order winners or the consumers' preferences. The second step, instead, consists of **comparing them with your internal performances** by considering the different behaviours of different customers and this collection must be done according to a unique method of evaluation.
Ranking them with stars could be one method, but it could end with the distribution of the same importance to all the competitive factors. **Distributing 100 points** could be another: this method would allow a deeper analysis of the evaluation. The scale is larger, the differences are more evident and all points must be summed up to 100.

What is the **process of identification, classification and weighting of different performance dimensions**? First, you have to understand which performances are important for the customers and you need to consider marketing data and orders analysis made by customers themselves. Once identified OW and Q, you have to:

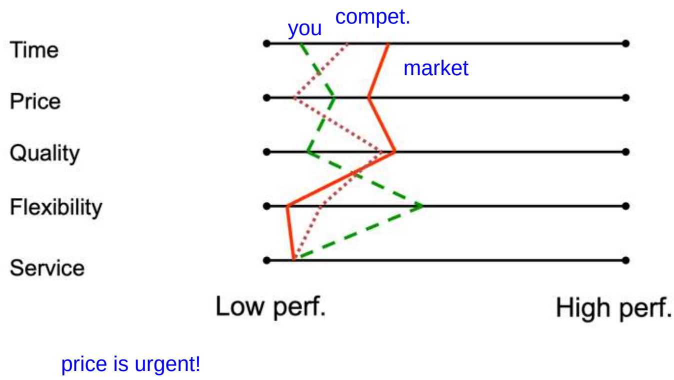
- **Identify eventual LOSERS among Q** and make a recovering plan
- **Assign priority to OW**, according to a ranking system (differentiation is important to give guideline)
- Consider different types and customers segments: market segmentation
    - searching and underlining differences
    - grouping homogeneous needs
    - adopting an operations point of view: marketing will probably ask you to improve each dimension due changes in market segmentation but it is asked here is to take an internal and operational viewpoint from the inside of the company.
- Review this process over time: over time Q could become OW and vice versa due to changes which could be:
    - market driven: the market changes and asks for new performances
    - company driven: the company has invested and has been transformed: it is now able to exploit those new resources by also focusing on them the attention of customers

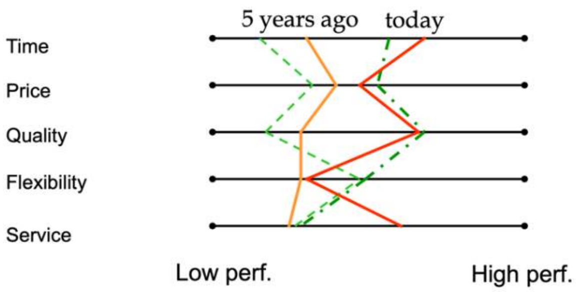
- Non-linearity: a linear dependence exists just when the change in the performance is small. In the real and complex world, the evolution of some competitive factors is never linear. The behaviours of those factors could diverge relevantly from expectances.



Mapping the performances could also help to have a better understanding of the market positioning of the company, once that the ranking of the different performances has been defined. In fact, it is always better to analyse the performances of the company in relation to the market and to the competitors. Thus, the analysis would be relative and effective.

For example, in the graph on the right, the green line represents the current performances of the company, the red dashed line represents the performances of the best competitor; while the red one represents what the market desires for each competitive factor. In this way it is clear which performances have to be improved in relation to the market wants and to the competitors (quality and time) and the other performances whose improvements are less urgent (price).





While mapping in relation to the market and to the competitors is quite easy to use; a more sophisticated approach is represented by the second graph on the right, where add the **contribution of "time"**. The "today" configuration is compared to a "past" configuration which is relevant for the comprehension of the present performances. In fact, we are able to understand the relative improvements in terms of time of both of us (dashed green line) and the market (orange and red lines).

For example, it emerges that the company has a good positioning in terms of "flexibility". In general, this characteristic is an important dimension because it involves the reduction of both costs (efficiency) and time (effectiveness) whenever a change happens. In this case, the graph shows how this characteristic is not relevant to the market, so that in 5 years' time it did not even increase. Moreover, we should question if the company had decided to keep it stable with the aim of showing new opportunities to the market and consequently did not succeed or if it had not acted to make anything change.
Therefore, the key principle is clear: if we intend to value the investment on a dimension which already has a marginal importance, we need to be able to make the market perceive this value, or, it struggles, we should reduce it.
Another example can be the one of "service": we observe how the company was not able to increase the performance of this dimension, meanwhile the market requests increased. Given the time horizon we refer to, this can be linked to a strategic decision of the top-managers who decided to invest in other fields.
Even in this case, we should question about the reason which is behind: it may be that the decision makers have some missing information about those characteristic or it could also happen that there was a missing

interest in increasing those performances because of the marginal importance given to the success of the company.
It is then clear how it is difficult to judge the strategic decision of a company. However, having a complete mapping of the performances could help managers to clarify and take the decision in relation to the environment, market and competitors, and to the evolution of this in time.
In conclusion, we could say that the drivers in choosing the **interventions** are, in a **priority list**:

- Importance: in terms of success and advantages; the higher the importance of a dimension, the higher the priority
- Distance: related to how distant the company is from the intended market; the far you are from market needs, the higher the priority
- Competitors: whenever the first two are the same, it is always better to react faster in those dimensions where the competitors are stronger in order to gain a fundamental competitive advantage; the higher the competitiveness, the far you are from the leader, the higher the priority
- Difficulties: in terms of costs and effort need to achieve the desired result. Sometimes it is not true that the hardest the project, the better the results; how much is the effort in organizational changes? It might not worth the case. Related to the case, a useful tool could be the PROJECTION SELECTION MATRIX (Pain - Gain).

which investment to do first!!

IMPACT

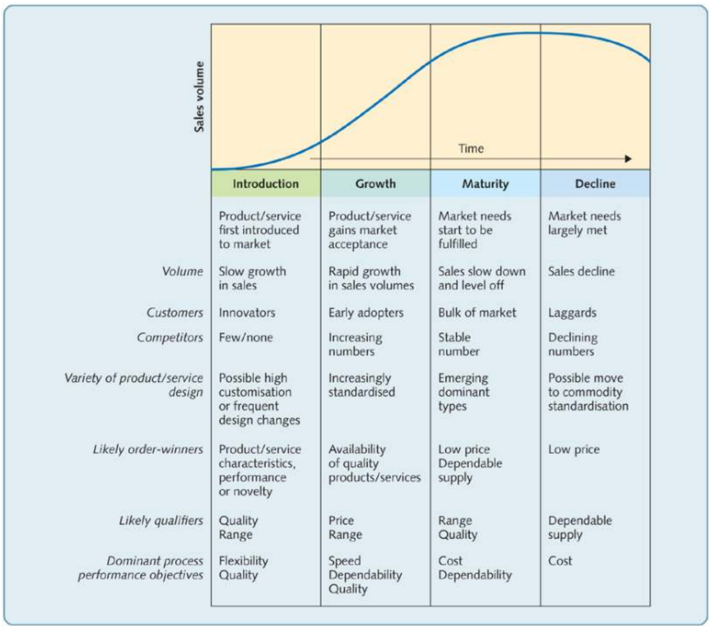
High		
Low		
	Low	High

EFFORT

3.3. LIFECYCLE S CURVE

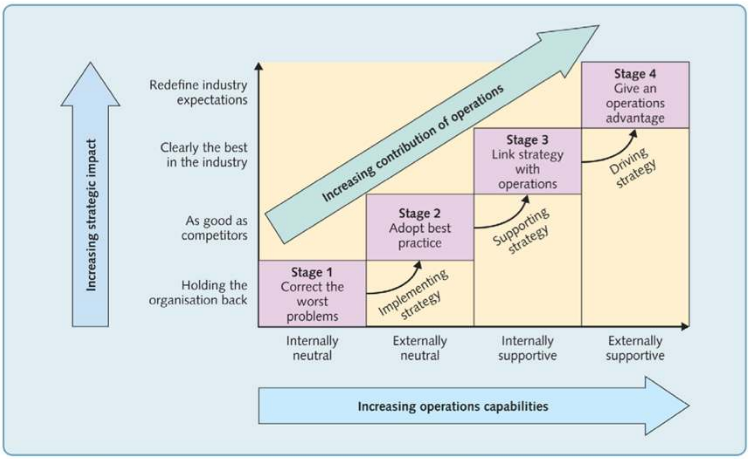
One way of generalizing the market requirements that operations need to fulfil is to link them to the life cycle of the products or services that the operation is producing. The exact form of product/service life cycles will vary, but generally they are shown as the sales volume passing through four stages – introduction, growth, maturity and decline.

- **Introduction stage**. When a product or service is first introduced, it is likely to be offering something new in terms of its design or performance. Few competitors will be offering the same product or service, and because the needs of customers are not perfectly understood, the design of the product or service could frequently change. Given the market uncertainty, the operations management of the company needs to develop the flexibility to cope with these changes and the quality to maintain product/service performance.
- **Growth stage**. As the volume of products or services grows, competitors start to develop their own products and services. In the growing market, standardized designs emerge. Standardization is helpful in that it allows the operation to supply the rapidly growing market. Keeping up with demand could prove to be the main operations preoccupation. Rapid and dependable response to demand will help to keep demand buoyant while ensuring that the company keeps its share of the market as competition starts to increase. Also, increasing competition means that quality levels must be maintained.
- **Maturity stage**. Eventually demand starts to level off. Some early competitors will have left the market and the industry will probably be dominated by a few larger companies. The designs of the products or services will be standardized and competition will probably emphasize price or value for money, although individual companies might try to prevent this by attempting to differentiate themselves in some way. So operations will be expected to get the costs down in order to maintain profits or to allow price cutting, or both. Because of this, cost and productivity issues, together with dependable supply, are likely to be the operation's main concerns.
- **Decline stage**. After time, sales will decline and competitors will start dropping out of the market. To the companies left there might be a residual market, but if capacity in the industry lags demand, the market will continue to be dominated by price competition. Operations objectives will therefore still be dominated by cost.



3.4. OPS STRATEGY MATURITY vs OPS STRATEGY EVOLUTION

The 'vision' for an operation is a clear statement of **how operations intend to contribute value for the business**. It is not a statement of what the operation wants to achieve (those are its objectives), but rather an idea of what it must become and what contribution it should make. A common approach to summarizing operations contribution is the Hayes and Wheelwright Four-Stage Model. The model traces the **progression of the operations function** from what is the largely negative role of 'stage 1' to becoming the central element of competitive strategy in excellent 'stage 4' operations.



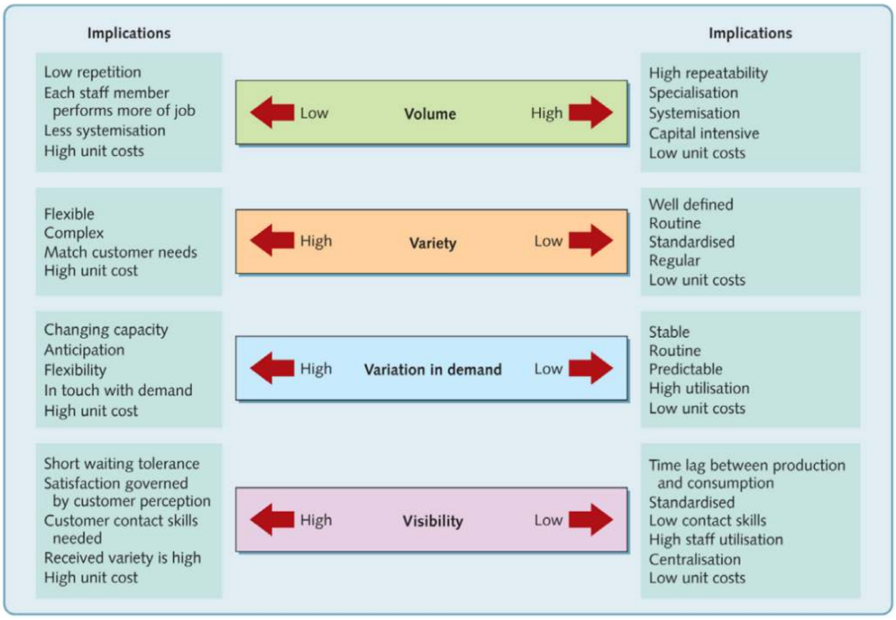
- Stage 1: **Internal neutrality**
  This is the very poorest level of contribution by the operations function. The other functions regard it as holding them back from competing effectively. The operations function is inward-looking and at best reactive, with very little positive to contribute towards competitive success. Its goal is to be ignored. At least then it is not holding the company back in any way. Certainly, the rest of the organization would not look to operations as the source of any originality, flair or competitive drive. Its vision is to be 'internally neutral', a position it attempts to achieve not by anything positive but by avoiding the bigger mistakes.

- Stage 2: **External neutrality**
  The first step of breaking out of stage 1 is for the operations function to begin comparing itself with similar companies or organizations in the outside market. This may not immediately take it to the 'first division' of companies in the market, but at least it is measuring itself against its competitors' performance and trying to be 'appropriate', by adopting 'best practice' from them. Its vision is to become 'up to speed' or 'externally neutral' with similar businesses in its industry by adopting 'best practice' ideas and norms of performance from others.

- Stage 3: Internally supportive
  Stage 3 operations have probably reached the 'first division' in their market. They may not be better than their competitors on every aspect of operations performance but they are broadly up with the best. Yet, the vision of stage 3 operations is to be clearly and unambiguously the very best in the market. They may try to achieve this by gaining a clear view of the company's competitive or strategic goals and developing 'appropriate' operations resources to excel in the areas in which the company needs to compete effectively. The operation is trying to be 'internally supportive' by providing a credible operations strategy.

- Stage 4: **Externally supportive**
  The vision for the operations function is to provide the foundation for competitive success. Operations looks to the long term. It forecasts likely changes in markets and supply, and, over time, it develops the operations- based capabilities that will be required to compete in future market conditions. The operations function is becoming central to strategy-making. Stage 4 operations are creative and proactive. They are innovative and capable of adaptation as markets change. Essentially, they are trying to be 'one step ahead' of competitors in the way that they create products and services and organize their operations.

3.5. A TYPOLOGY OF OPERATIONS



VOLUME

Processes with a high volume of output will have a high degree of repeatability, and because tasks are repeated frequently it often makes sense for staff to specialize in the tasks they perform. This allows the systemization of activities, where standard procedures may be codified and set down in a manual with instructions on how each part of the job should be performed. Also, because tasks are systemized and repeated, it is often worth developing specialized technology that gives higher processing efficiencies. By contrast, low-volume processes with less repetition cannot specialize to the same degree. Staff are likely to perform a wide range of tasks, and while this may be more rewarding, it is less open to systemization. Nor is it likely that efficient, high-throughput technology could be used. The implications of this are that high-volume processes have more opportunities to produce products or services at low unit cost.

VARIETY

Processes that produce a high variety of products and services must engage in a wide range of different activities, changing relatively frequently between each activity. They must also contain a wide range of skills and technology sufficiently 'general purpose' to cope with the range of activities and sufficiently flexible to change between them. A high level of variety may also imply a relatively wide range of inputs to the process and the additional complexity of matching customer requirements to appropriate products or services. So, high-variety processes are invariably more complex and costly than low-variety ones.

VARIATION

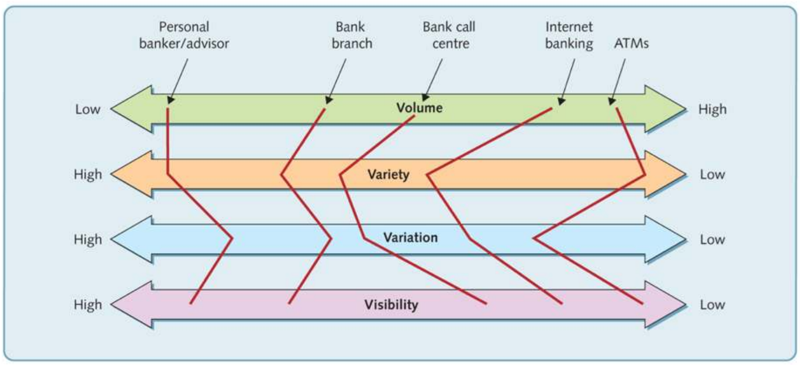
Processes are generally easier to manage when they only have to cope with predictably constant demand. Resources can be geared to a level that is just capable of meeting demand. All activities can be planned in advance. By contrast, when demand is variable and/or unpredictable, resources will have to be adjusted over time. Worse still, when demand is unpredictable, extra resources will have to be designed into the process to provide a 'capacity cushion' that can absorb unexpected demand. Because processes with lower variation do not need any extra safety capacity and can be planned in advance, they generally have lower costs than those with higher variation.

VISIBILITY

Process visibility is a slightly more difficult concept to envisage. It indicates how much of the processes are 'experienced' directly by customers, or how much the process is 'exposed' to its customers. Generally processes that act directly on customers (such as retail or health-care processes) will have more of their activities visible to their customers than those that act on materials and information. Low-visibility processes, if they communicate with their customers at all, do so using less immediate channels such as the telephone or the Internet. Much of the process can be more 'factory-like'. The time lag between customer request and response could be measured in days rather than the near-immediate response expected from high-visibility processes. This lag allows the activities in a low-visibility process to be performed when it is convenient to the operation, so achieving high utilization. Also, because the customer interface needs managing, staff in high-visibility processes need customer contact skills that shape the customer's perception of process performance. For all these reasons high-visibility processes tend to have higher costs than low-visibility processes. Many operations have both high- and low-visibility processes. This serves to emphasize the difference that the degree of visibility makes.

IMPLICATIONS

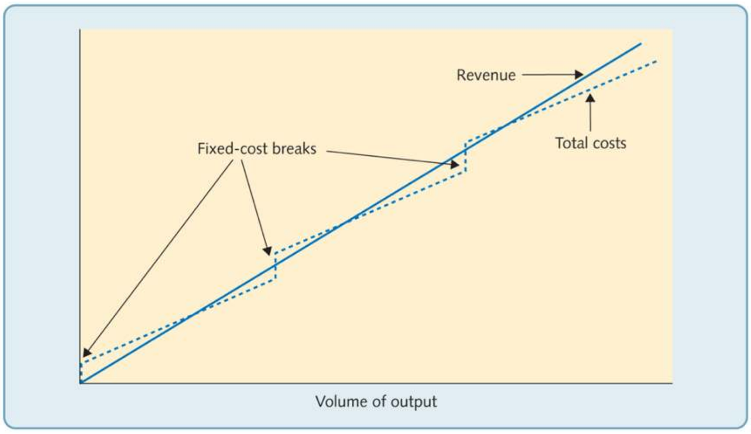
All four dimensions have implications for processing costs. Put simply, high volume, low variety, low variation and low visibility all help to keep processing costs down. Conversely, low volume, high variety, high variation and high customer contact generally carry some kind of cost penalty for the process.



This model helps in identifying those paths where we have a higher coherence between the 4 different characteristics. Moreover, it highlights the connection within the 4 and the impact that changing one would have on the changes of the others. This model also suggests the possible changes to improve performances and the example provides the analysis for some retail banking processes.

3.6. BREAK EVEN CURVE

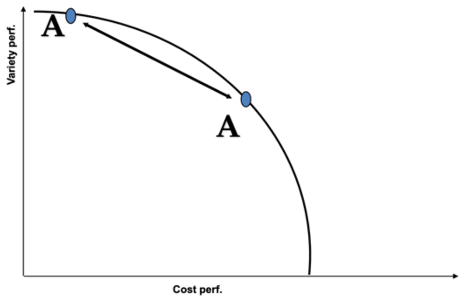
The **break-even** is used by managers in order to find out the break-even points. In the chart the curves of total costs and revenues are shown separately. The vertical axis indicates costs and revenues while the horizontal refers to the volume of output. While the revenue curve is illustrated as linear, due to a constant price, irrespective of the output; the total cost curve does not start from the origin and it shows some breaks because it is the sum of the fixed costs, which have to incur at certain levels of output, plus the variable costs, which grow linearly. Therefore, the break-even points can be found in the intersections of the two curves and it refers to the number of units the firm have to sell to earn enough revenues just to cover all the expenses of production, both fixed and variables. In the case, the firm does not earn any profit, nor does not incur any loss.

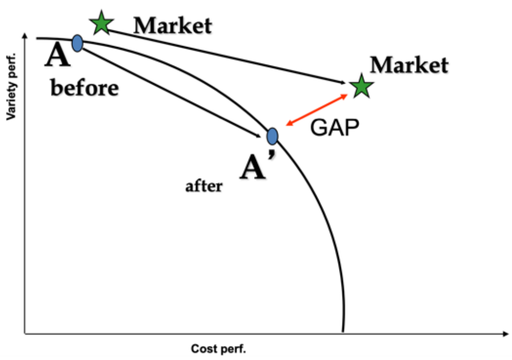


3.7. OPERATIONS STRATEGY SHOULD GUIDE THE TRADE-OFFS BETWEEN PERFORMANCE OBJECTIVES

An operations strategy should address the relative priority of operations performance objectives. To do this it must consider the possibility of improving its performance in one objective by sacrificing performance in another. So, for example, an operation might wish to improve its cost efficiencies by reducing the variety of products or services that it offers to its customers. Taken to its extreme, this 'trade-off' principle implies that improvement in one performance objective can only be gained at the expense of another, in terms of cost efficiency or product variety. There are two approaches to managing trade-offs result in two approaches to operations improvement.

The first emphasizes 'repositioning' performance objectives by trading-off improvements in some objectives for a reduction in performance in other (simple trade-off model).





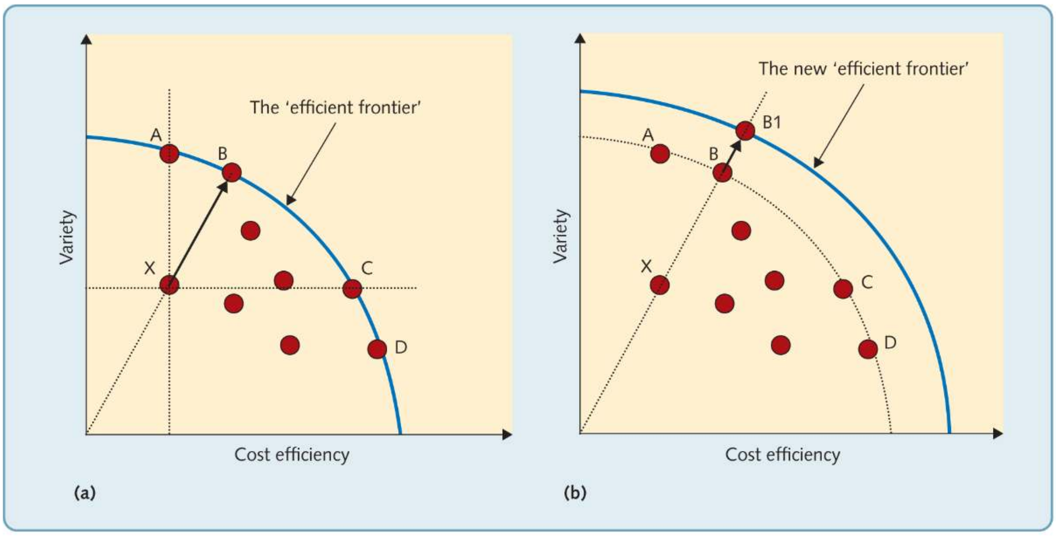
The other emphasizes increasing the 'effectiveness' of the operation by overcoming trade-offs so that improvements in one or more aspects of performance can be achieved without any reduction in the performance of others (trade-off model with the insert of the "efficiency" frontier). Most businesses at some time or other will adopt both approaches.
Moreover, it is very difficult to understand what to sacrifice at the expenses of the other one. Therefore, we have to be able to target the intended market in order to better understand which is the characteristics we should sacrifice in order to reach the minimum gap with respect to the evolution of the market.

3.8. TRADE-OFF AND THE EFFICIENT FRONTIER

The Figure shows the relative performance of several companies in the same industry in terms of their cost efficiency and the variety of products or services that they offer to their customers. Presumably all the operations would ideally like to be able to offer very high variety while still having very high levels of cost efficiency. However, the increased complexity that a high variety of product or service offerings brings will generally reduce the operation's ability to operate efficiently. Conversely, one way of improving cost efficiency is to severely limit the variety on offer to customers.

Operation X, however, has an inferior performance because operation A is able to offer higher variety at the same level of cost efficiency and operation C offers the same variety but with better cost efficiency. The convex line on which operations A, B, C and D lie is known as the 'efficient frontier'. They may choose to position themselves differently (presumably because of different market strategies) but they cannot be criticized for being ineffective. Of course they may choose to reposition themselves at some other point along the efficient frontier.

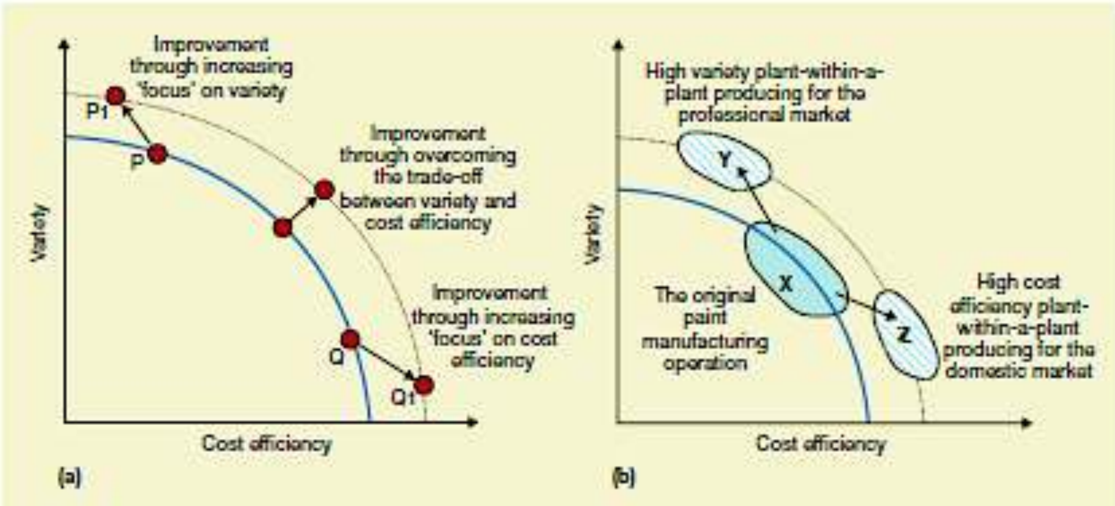
Operation X will generally have a strategy that emphasizes increasing its effectiveness before considering any repositioning. However, a strategy that emphasizes increasing effectiveness is not confined to those operations that are dominated, such as operation X. Those with a position on the efficient frontier will generally also want to improve their operations effectiveness by overcoming the trade-off that is implicit in the efficient frontier curve. This distinction between positioning on the efficient frontier and increasing operations effectiveness to reach the frontier is an important one. Any operations strategy must make clear the extent to which it is expecting the operation to reposition itself in terms of its performance objectives and the extent to which it is expecting the operation to improve its effectiveness.



3.9. IMPROVING OPERATIONS EFFECTIVENESS BY USING TRADE-OFFS

Improving the effectiveness of an operation by pushing out the efficient frontier requires different approaches depending on the original position of the operation on the frontier.

For both operations P and Q effectiveness is being improved through increasing the focus of the operation on one (or a very narrow set of) performance objective and accepting an even further reduction in other aspects of performance. The same principle of focus also applies to organizational units smaller than a whole operation. For example, individual processes may choose to position themselves on a highly focused set of performance objectives that matches the market requirements of their own customers. So, for example, a business that manufactures paint for interior decoration may serve two quite distinct markets. Some of its products are intended for domestic customers who are price sensitive but demand only a limited variety of colours and sizes. The other market is professional interior decorators who demand a very wide variety of colours and sizes but are less price sensitive. The business may choose to move from a position where all types of paint are made on the same processes (position X in Figure 2.10(b)) to one where it has two separate sets of processes (Y and Z): one that makes paint only for the domestic market and the other that makes paint only for the professional market. In effect, the business has segmented its operations processes to match the segmentation of the market. This is sometimes called the 'plant-within- a-plant' concept.



3.10. IMPROVING OPERATIONS EFFECTIVENESS BY OVERCOMING TRADE-OFF

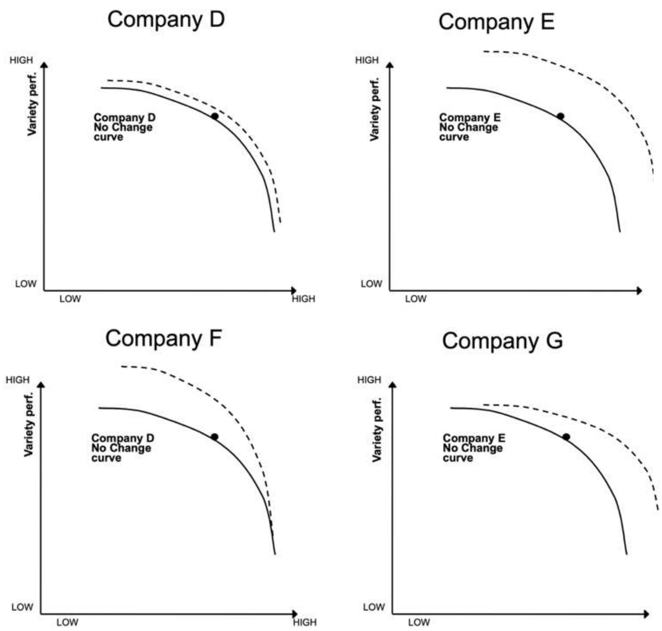
This concept of highly focused operations is not universally seen as appropriate. Many companies attempt to give 'the best of both worlds' to their customers. At one time, for example, a high-quality, reliable and error-free automobile was inevitably an expensive automobile. Now, with few exceptions, we expect even budget-priced automobiles to be reliable and almost free of any defects. Auto manufacturers found that not only could they reduce the number of defects on their vehicles without necessarily incurring extra costs, but they could actually reduce costs by reducing errors in manufacture. If auto manufacturers had adopted a purely focused-based approach to improvement over the years, we might now only be able to purchase either very cheap, low-quality automobiles or very expensive, high-quality automobiles. So a permanent expansion of the efficient frontier is best achieved by overcoming trade-offs through improvements in operations practice. Even trade-offs that seem to be inevitable can be reduced to some extent.

For example, one of the decisions that any supermarket manager has to make is how many checkout positions to open at any time. There seems to be a direct trade-off between staff utilization (and therefore cost) and customer waiting time (speed of service). Yet even the supermarket manager deciding how many checkouts to open can go some way to affecting the trade-off between customer waiting time and staff utilization. The manager might, for example, allocate a number of 'core' staff to operate the checkouts but also arrange for those other staff who are performing other jobs in the supermarket to be trained and 'on- call' should demand suddenly increase. Therefore, from an internal point of view, every company build its own frontier from a practical perspective.

But it is important to highlight that you cannot position yourself whenever you want, if you do not have the knowledge to build a company with the performances that overcome the frontier. From a theoretical perspective, the outer frontier represents the TECHNOLOGICAL, KNOWLEDGE and ORGANIZATIONAL FRONTIER that all companies operating in a sector yet did not overcome.

3.11. POTENTIAL CURVE

Moreover, even if two companies are at the same present situation, they are able to achieve two different positions thanks to their knowledge: for each company, the Potential Curve represents a curve it can easily manage to move closer to, thanks to its knowledge. In this case, E have potentially the knowledge to improve its performances much more than D.



ATTENTION: improvement does not have to be equal for both dimension
→ the new curve has not to be parallel to the previous one
(Maybe I know how to improve cost efficiency without impacting on variety as Company F, but I cannot do the vice versa; while Company G is able to achieve the exactly opposite position)

4. HQ CASE

WHAT HAPPENED 7 YEARS AGO?

HQ is an **injection moulding company** that was doing well until 10 years ago when HQ's biggest customer announced that it would phase out of the supplying components by HQ. So, it decided to take off all products supplied by HQ in a period of 2 years. In 2 years, the demand for HQ would decrease by 50%, leaving HQ with 50% of capacity available in the next two years. Against this background, HQ faced the emergence of small new competitors in its market.

WHICH OPTIONS:

The management had to make a decision: it could have changed the market strategy and targeted a new segment with high quality products (this is what the company decided to do). 10 years ago, the company was competing in the low-cost market, HQ selling products at low prices. Then the company decided to attack the new market segment and strive for high quality products.

- SAME MARKET, SAME PRODUCT, NEW CUSTOMERS
  We want to continue to compete with the same product by catching new customers or by exploring new markets. If you are so sure that you can sell the same product to more people, why have not you increased the volume in the past years? This means that it is not that easy. This is because the consolidated market of HQ is a market for standard products (not high quality), so most likely a very stable market with consolidated technology. It is a market where there are many and strong competitors and it is difficult to gain market share so easily.
  ⟹ Probably these options are not feasible and no sense

- DOWNSIZING:
  Downsizing: if you downsize a company or organization, you make it smaller by reducing the number of people working in it. We have to reconfigure our system for the new situation. Downsizing is very difficult, because this kind of company (industrial company) has a lot of fixed costs that are difficult to sell. A lot of old machines very difficult to sell. There are also a lot of people.
  ⟹ The point is that in our context it is not a good option.

- FOLLOW CUSTOMERS
  Following the customer with an innovation of the technology. It means changing technology completely, leading to few economies of learning, completely different processes and supply chains. It is about developing entirely new capabilities. We are switching from plastic to metal.
  ⟹ Maybe it is feasible, but only as an experiment (2-3%, not 50% of sales).

- NEW MARKET, SAME PRODUCT, NEW CUSTOMERS
  Enter in a premium market, with high quality products. It was a good choice, it appeared as the **best option**.

- VERTICAL INTEGRATION:
  Acquire the customer. Why don't we buy our customer? If we do, we can protect HQ demand and the customer will continue to buy the product. But vertical integration is difficult in this case for two reasons: first, we are in a market where the customers manufacture household appliances, so the customer is bigger than the supplier and acquisition is difficult. The second reason is that the customer operates in an environment that is not comparable to that of HQ. HQ only supplies some plastic components, which means that we are trying to merge with a company that has a completely different know-how. These two factors make vertical integration critical.
  ⟹ It is crucial and it is not possible in this sector.

- SAME PRODUCT, DIFFERENT MARKET.
  The current market only serves the UK market, so could look at geographical expansion. To implement this strategy, we need to understand which market deserves an opportunity, for example Asia or America. The product we want to sell is characterised by a highly standardised process, low quality and a low price. So if we consider a poorer country, it would be difficult to sell a cheap product in an already cheap market. Even if we consider selling in a richer country, a low-priced product would be more profitable, but here other factors have to be considered. Even if one could penetrate this market, transport costs, for example, would be higher, which would affect the low profit margins one would like to make on the products. The only reasonable markets could therefore be Germany and the Scandinavian countries, which are neighbouring and rich countries.

WAS IT A GOOD DECISION?

To operate in a new market, they changed something in their operating system, so what they did? Monitoring a period of 7 years. They did a huge amount of investment.

- Replaced machines, in particular they:
  - Sold old machines
  - Bought new machines

  The interesting fact is that the characteristics of the new machines where different; they had a **higher control**, a **higher pressure**, **less maintenance** and **lower cycle time** and **so higher capacity of the asset thanks** to new molding (from single impression to multi-impression ones). But pay attention, **the total capacity of the company remains the same** because if we look at the table below, they used to have 51 machines while the last year they had only 22 machines; so, they replaced the old machines with **new one faster** but at the same time the total number of the machines decreased, basically the total capacity remained more or less the same.

Machine group	Number of machines		
	Current year minus:		
	8	1	New
1	29	1	–
2	15	11	5
3	5	8	6
4	2	2	1
Total	$\boxed{51}$	$\boxed{22}$	12

  With the new machines it was introduced especially the multi-impression moulding equipment. With one cycle of machine we can have 4 items now, it needs **more accurate set up**. Another positiveness is that we can have an accurate colour matching. The design of the moulding is no ore in charge to the customers, but is in charge with HQ company, and this situation stressed the company → complain of the people of production. In the new market customers want the product as soon as possible.

Moreover, we can see through the **balance sheet** that:

- Revenues increased
- Profit increased

We should be happy, however there is a problem. The balance sheet shows the value from an economic point of view; If we go in detail, we see that revenues increased but happened also that there is an increase in the inventories. This is normal if you increase the volume of your production your stock will increase in order to make the operations working.

The problem is that we are operating in a **fashion market**, this kind of market has some characteristics as: "*what is cool this year maybe it won't be in the next year*". So, in most part of the cases what remains in your inventory after the season is something that does not give us gain anymore, we have seen an increase of the inventories but in order to make a company working, usually the level of inventories has not to be so high. We have already said that inventories are more or less 10% of the revenues. If we assume this and we look at the increase of the level of inventories, we can see that the value of the inventories is actually higher than the value of the one that I would expect. It means that every year we have some inventories that remain from the past season production, that has no value for the customer, and are considered for the company costs.

If we make the net between the net profit and the lost given by the cost of hold inventories, we will see a new evaluation of the profit completely different. We will have a decreased new profit computed by the accounting due to the cost of obsolescence given by the old stocks.

- $\boldsymbol{Physiological\ increment} = 10\% * sales$
- $\boldsymbol{Non\ physiological\ increment} = Inventory - Physiological\ increment$
- $\boldsymbol{Net\ profit\ corrected\ by\ obsolete\ stock} = \ Net\ profit\ before\ tax - non\ physiological\ increment$

Let us compute this and understand the corrected net profit as:

year	-7	-6	-5	-4	-3	-2	-1
sales	2552	2872	4212	4466	5810	5394	8021
Inventory	262	532	1029	1259	1559	2243	2567
Net profit before tax	146	185	274	362	564	708	1050
physiological inventory (10%)	255	287	421	447	581	539	802
non physiological increment	7	238	363	205	166	726	61
net profit corrected by obsolete stock	139	-53	-89	157	398	-18	989
profit compared with year -7	100%	-38%	-64%	113%	286%	-13%	710%
ROS*	5%	-2%	-2%	4%	7%	0%	12%

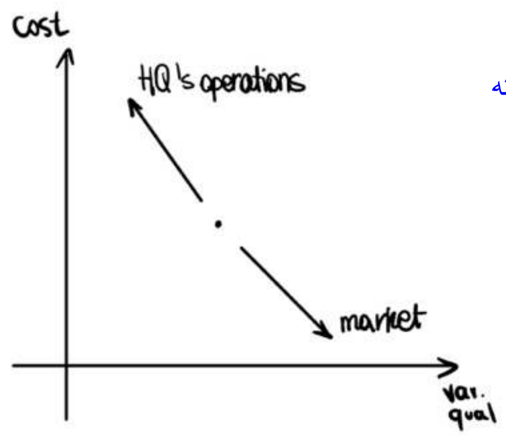
We have a completely different picture of the company; with this new picture we cannot say that we are not going well but we can say that there are events that we cannot control. These events make very high fluctuation in the performance, so make difficult to understand whether we are going in the correct direction or not. In conclusion we are not going well, that is because there is not alignment between what customers are demanding and what we are providing to the customers. This misalignment is due to the production, the think is that instead of having few products with high demand we have many products with low demand.
But how can we define whether this changeset in the structural configuration of the operation was a good choice or not? In order to evaluate this, we have to understand which is the impact on the performances that we provide to the customer and as we have seen the last time, we have the success of an operation strategy when we have the alignment between the performance provided by the operation and the performance demanded by the customers. Using the driver that we saw in the past lessons, what are the performances demanded by the customers?

PERFORMANCE	CONSOLIDATED CUSTOMERS	NEW CUSTOMERS
TIME:		
▪ Speed	-	***
▪ Reliability / Delivery	***	*
PRICE	***	-
QUALITY		
▪ Design	-	***
▪ Conformance	*	**
FLEXIBILITY		
▪ Product	-	-
▪ Range	-	*
▪ Plan	-	**
SERVICE	-	-

First thing to highlight is that it is impossible to compete perfectly on all the performances: you must give priorities to the performances you want to achieve. Secondly, serving different segments of customers we will have also different expectations on the performances that we are providing them. We need to understand the characteristics of the new market.

OLD MARKET	NEW MARKET
These customers are really concerned about performance. They compete mainly on cost. The customer is a distributor who gives a lot of visibility to production. So for the market, reliability and cost are most important, all other performance is secondary. Flexibility is unimportant, service too. And we add 1 point of quality conformity, why? Because of the HQ situation, because one of the points is that the market is developing and HQ is starting to have quality problems and the competition is offering better quality.	This new market is really focused on time speed (fast design, seasonality) and design, because customers want beautiful products. Reliability is not so relevant. Conformity becomes more important for the new customers. Flexibility range equal to 1 and flexibility plan equal to, because when we produce a new product, we need to scale up or down production faster.

Company is moving marketing in this direction but is moving operations in the opposite direction. HQ's operations are investing for a more efficiency.



اینو برعکس نوشته

Reconciliation model: visualizing tool to align the offering with the demand

LEVERS	IMPACT
Overall Capacity	Positive impact on: ▪ Costs High saturation keeps low costs. Negative impact on: ▪ Flexibility ▪ Time speed Because the surplus of capacity of low saturation led the company to react to unforeseen events, to react to the seasonal demand or peak of demand. If you have a surplus of demand, this leads you on being flexible on being reactive. If you have machines that are very high saturated, you cannot do it, you are going late for something.

Fractionation *"How they split the capacity?"*	Positive impact on: ▪ Costs Because you go for economies of scale. You have the total costs of the production capacity, which is lower, because two machines that produce 10 per machine are more expensive than one machine that produces 20 Negative impact on: ▪ Flexibility ▪ Time speed The bigger the machine, the more complex is it. Higher probability of breakdown. Which means a reduction of flexibility
Multi impression molding technology These technologies of multi molding have an impact on the fractioning. The use of multi-molding stressed the necessity of having bigger machines, because bigger machines have higher capacity of pression which is necessary for using the multimolding.	Positive impact on: ▪ Costs Multi molding led you to improve using economies of scales because with one pressure you do two items/three items per time. Negative impact on: ▪ Flexibility ▪ Time speed Multi molding makes the machine more complex (more complex to set up), which means a reduction of flexibility, which then directly means reduction in terms of time speed. You are less reacting.

Example of the impact of the fractioning of capacity
8h/shift
1h/setup
→ OEE= 7/8

# Of MACHINES	SETUP	# Of SHIFT	# Of products/machine	# Of PRODUCTS /day
51	0,5h	2	4	204
22	4	2	0,5	11

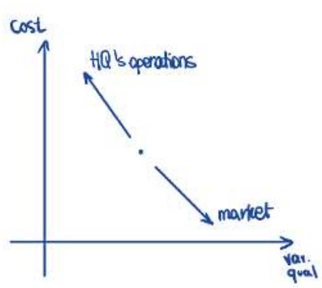
OEE is an indicator that measures the efficiency of the machines. It is a percentile indicator where you evaluate how much we are producing finished products compared to the hours the machine is on. In this case, 7/8 is equal to 85%, which is very good! So for the same volume, the number of items delivered to the customer, we have a very different capacity in terms of flexibility. If we fraction a lot the capacity, we have a lot of small machines that are easy (or quick) to set up and we can produce 204 different variants per day. On the other hand, if I have 22 big machines, the set-ups are very complex. It takes me four hours to set up. So when you set up a product, you have to let it produce for a long time. Because you cannot keep switching from one product to another. Otherwise, you lose four hours every time. This is another effect of fractional capacity.

Multi-moulding and big machine stress an organizational lever. How did they organise the management of the machines? The knowledge and skills for repairing the machine or for setting up the machine were with a few people and the rest with a few people. So they had an organisation and kept the organisation going with people who were really specialised. That is the lever for competencies and responsibilities. You have an organisation where specialisation has created certain people for certain roles in the company who are able to operate the machine and control the machine. This high degree of specialisation is a lever. It is a decision of the operations manager how you structure your operating system.

LEVER	IMPACT
Competencies	Positive impact on: ▪ **Costs** The organization has few expensive people with the competencies and many cheap employees that just have to run the machine ▪ **Quality of conformance** Due to the fact that you have a high level of specialization. Negative impact on: ▪ **Flexibility**
Responsibility Responsibility is directly linked with the competencies. In a system where you have few people with responsibility, you have competencies that are concentrated with few people with a responsibility	Positive impact on: ▪ **Costs** ▪ **Quality of conformance** Negative impact on: ▪ **Flexibility** ▪ **Time Response** Because there are the managers who become the bottleneck of the system. In fact, people are no longer able to stay in their traditional position because market demand has changed the needs within the company. So that the person who used to be able to set up the machine is now involved with the r&d department and no person is able to do the installation in the field anymore. So these machines stop or they do not have people to do this setup. They have made a consistent choice in the setting of skills or responsibilities. The choices between the two are consistent, but they are not aligned with the business strategy

To sum-up

Why did they have a bad performance? Because they push the operating system in one direction while the market wants it in the other. These misalignments lead to bad results. If they align it, they should get better performance. Should they have kept the old machines without investing in new ones? No, because they definitely had to invest because the old machines were no longer able to meet the new demands of the customers. The market has evolved, regardless of the customer segment. New competitors compete on the level of quality. The initial low level of conformance quality became a problem, an order loser. This growing problem was the signal that investment was needed. If they did not invest, they would definitely lose the market because competition was increasing and someone was stealing their market share. The point is that if they wanted to prioritise the new market, they should have made other investments. On the other hand, if they had wanted to pursue the old market, they should have invested as they did.

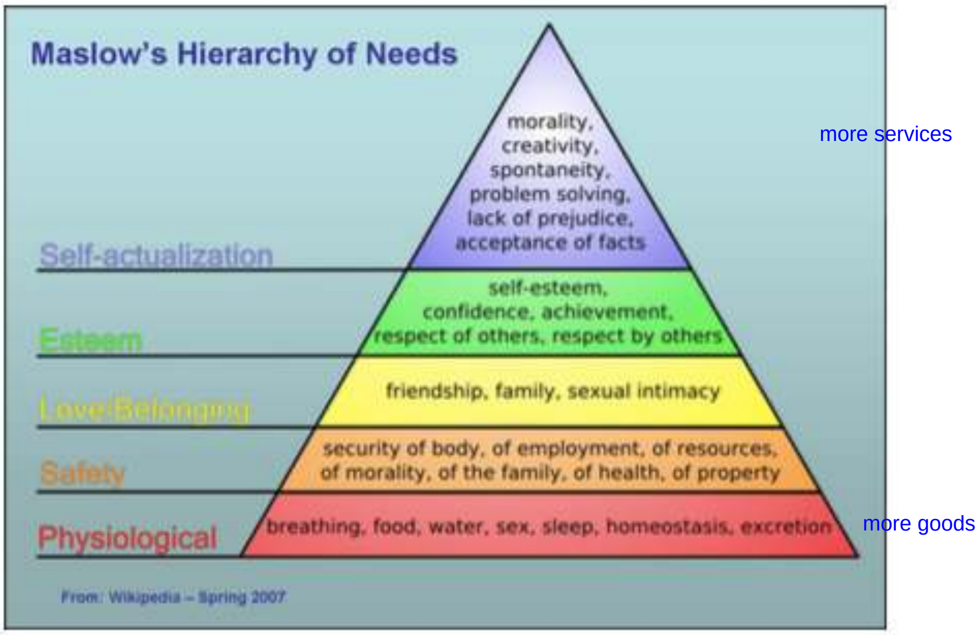


5. INTRODUCTION TO SERVICE

The service sector is becoming increasingly more important in terms of GDP and the number of employees. This growth trend is due to several factors:

- Socio-economic evolution:
  - Evolution of needs
  - Welfare: people income increase enabled them to buy more. They had enough income to satisfy secondary needs
  - Entertainment (e.g. people earn more money so they could do other things)
- Technological progress ICT: ICT allows services to improve distribution and communication: e.g. Social Networks
- Globalization: transport, tourism, services for industrial companies (without these ones neither service will exist so US started a re-manufacturing process)
- Outsourcing: service provider for industrial companies. It is only a balancing between service and manufacturing rate

Satisfying different needs: MASLOW PYRAMID. The higher the levels the more intangible the needs are:

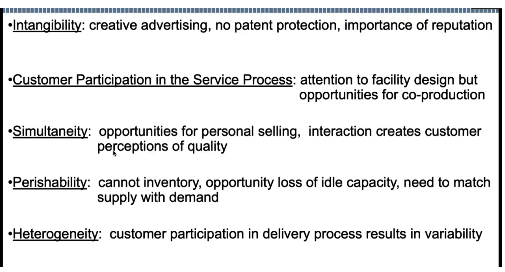


Customers have different level of needs. At the basis there are the basic needs necessary to survive (physiological needs). Once satisfied them, humans start exploring other needs that are at higher level. New companies that are growing are continuously pushing the services to the highest level of the pyramid. The improvement of service pushed them to the higher levels (self-actualization). This means an innovation also of the processes that we have to design in order to satisfy the new needs. So the design of the operations become really significant. This complexity comes more than the traditional companies because they are more complex. Take the example of the manufacturing company that produces manufacturing items.

5.1. WHAT IS THE DIFFERENCE BETWEEN PRODUCTS AND SERVICES?

DISTINCTIVE CHARACTERISTICS IMPLICATIONS	IMPLICATIONS
INTANGIBILITY	Creative communication, reputation, ease of generation and customization, difficulties in assessing the quality ex ante, importance of trust and experience. The more this becomes relevant, the more operations should consider it and change its traditional approach. The service is intangible and the manufacturing item in tangible.
CUSTOMER PARTECIPATION	Attention to facility design, opportunities for co-production → manufacturing process becomes a potential lever to sell: the more the customer is inside the delivery system, the more operations need to change. There is a set of processes that permits the customer experience. The customer is an act of the place, is part of our processes and this create variability. Longer customers interaction.
SIMULTANEITY (production/consumption)	There are no stocks, direct relationship with the customer, control over the process and not on the product, close interaction, front-office key → service quality depends on customer perception. The customer exploits the service at the same time they buy it.
PERISHABILITY	Unmet demand is lost, cyclicality and seasonality of demand with peaks and valleys also very high → problem of saturation → political alignment, need to match supply with demand side. With service you cannot make stocks. The external variability of the demand is giving a lot of impact, and if you cannot stock items, it is a problem, the capacity must be perfectly aligned with the demand otherwise we lose money.
HETEROGENEITY	Customer participation in delivery process results in → manufacturing company are trying to reduce variability because it would be better for them (standardization) but services processes cannot reduce variability so operations have to accommodate customers' needs. Service quality is variable because it depends on many factors.

What is the impact on operations? Stress the **tangible** point. It has a really strong impact. Quality in the manufactured system is under control, in the service this is not so easy because the service is not tangible, so the meaning is that is not measurable. And you cannot well define it. The measures are not representing the service, you can understand only a part of the quality.



EXAMPLE: Uber: how can you evaluate the service? How can you measure if 15 min are too much? How can you measure the good driving of the driver? They are feature that are difficult to measure. Measure the price, how fast it responses to reach you. Different cases. If you are late for the airplane, you hope that the cars reach you in the less time possible. If you are out of the disco and you want to stay with a girl, you hope that the car reaches you in the highest time possible. Different situation that leads to difficulties in evaluating a service.

EXAMPLE: how can you measure a restaurant?

Tangible is the explanation for which web likes trip advisor etc.., references are used but not so much.
The intangibility creates a lot of difficulties in measures. The intangibility is directly linked to the experience.
When we design a service, we keep the attention on the design on the experience delivered to the customers.
The simultaneity is directly linked with the experience. You have just that moment to evaluate the service. You have to design very well that moment.
The external variability is giving a lot of impact in the manufacturing industry. In the services industry we are not able to stock services. Or we deliver the service, or we lose the money.

Customer is an element that you cannot control. It is not a trained customer. The stability of the processes is much lower than the stability of processes in the manufacturing industry. Instability of processes comes from the participation of the customers. The customer is an active player of the processes, for this reason he creates variability. Services industries are very affected by a big internal variability.
Variability is the nightmare of Operations.

Service is touched by a big variability respect of the manufacturing company.

*Example: **Airplane company***
Example of Perishability: if an airplane has 50% of free seats, the airplane company loses the money. It cannot storage the seats for another flight.

5.2. SERVICE PROCESSES

The process is one key element to define a service delivery system. It is classified:

- Visible vs. not visible to the customers: the level of interaction of the customer is a feature that can be described
- Volume vs variety
- Processing: people, objects, information

Processes have many differences which could be the base for competitive advantage
Operations is a front office resource.

5.3. OPERATIONS SYSTEM CHARACTERISTICS:

1. Interaction with the customer
2. Volume to handle vs variety offered
3. Variability (of demand, of capacity) and uncertainty

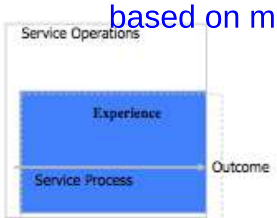
5.3.1. INTERACTION WITH CUSTOMERS

The customer receives an **outcome and an experience** and its experience is defined by personalization, speed, flexibility and personal access, privacy, courtesy and competence.
Depending on the relative importance of outcomes and experience, we can distinguish generally two different settings of operations.

	EXPERIENCE	OUTCOME	EXAMPLE
FRONT-OFFICE focus operations	Higher importance	Taken from granted Assumed	Cinema Entertainment Park
BACK-OFFICE focus operations	Lower importance	Higher importance	Postal service

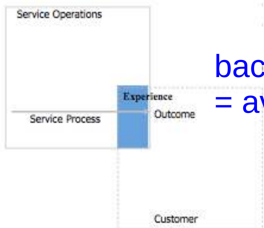
FRONT OFFICE: The outcome is taking for granted, and the experience is making the different that create the competitive advantage. Therefore, the outcome is still important but not as much as feelings along the process. They build their competitive advantage on the interaction between the customers and the front office, must develop the capability of managing the customers, absorbing the variability coming from the requests of the customer and the ability of build up the experience for the customers. The outcome is the most tangible part, while the experience is the less tangible part.

front office: capacity based on max demand



EXAMPLE: Education: you perceive the value when you come to class; Cinema or entertainment park: there are services where the physical outcome is not existent, for example going to the Cinema gives out just an experience: the outcome is the most tangible part, while the experience is the less tangible part.

BACK OFFICE: these can be addressed in the same way as manufacturing companies because the experience is not as important as in the front office and they are **still focused on the outcome**. The competition here is about the best outcome.



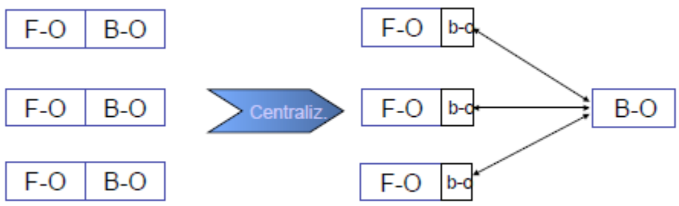
back: capacity = avg demand

EXAMPLE: Postal Service's experience is not relevant: you want the service to be completed in a small amount of time.

Different industries are classified from the output and the experience and therefore it is better to improve them all. The level of front office or back office depends on the sector. The level of front / back office opens to the operations manager different type of efficiency.

NB: CENTRALIZATION

In general, to be effective you have to set FRONT OFFICE capacity at the pick of the demand otherwise you will create a queue in the system. However, if you decouple BACK OFFICE activities, you can better absorb demand. If you are able to move the activities from the front office to the back office you are able to concentrate the back office activities in one central. This is a lever.



Pros:

- No interruptions (efficiency)
- Fast decrease along the experience curve
- Specialization of activities (thanks to volume)
- Lower manpower cost
- Possibility to follow personal attitudes
- Less volume variability
- Thanks to high volume:
  - Economies of scale
  - Automation, with advanteges in terms of cost and quality

Cons:

- Possible GAP with the Front-Office
- Activities overlapping
- Longer lead times
- Greater rigidity
- Inability to keep a product base approach, therefore switching to an activity based approach

* Higher efficiency → you are reducing the variability and so the saturation of the resources;
**Rigidity → because we must assure that in the back office arrives a standard request.

5.3.2. VOLUME to handle vs VARIETY offered

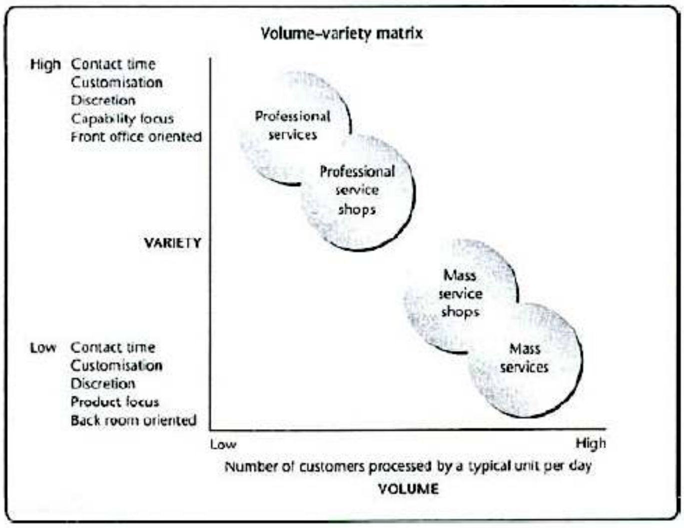
VOLUME – VARIETY MATRIX

It follows the same logic as for the manufacturing world:

1. **Processes for low-range and high volumes**, it is an equivalent "line" services: although it is a much less common logic, adopted by all those companies that offer services characterized by stages out of sequence, they have to always play the same activities which can be separated from stocks, with the products that follow a strict predefined sequence. A classic example should be the postal office: once the customer has taken a letter has travelled to the door, the operator will charge the object that will be stored in an appropriate stock, then will be taken, pursuing a series of intermediate steps before reaching its destination). They are mass services.

2. **Processes for high-variety, low-volume unit**, it is an equivalent: what was an aggregation by technology and departments, in a service will result in a business combination for skills, and most of companies are structured in this way (it's "classic" is the hospital: it is being organized with the radiology department, to orthopaedics to ophthalmology ... and the customer who enters the system can follow different paths on the basis of what are their needs and skills required). They are professional services.

The following **graph** does not refer to the characterization of the processes of service of a company as a whole, but rather it **classifies a single service process or a group of resources that work together and are dedicated to the execution of a series of activities**:



1. **Volumes**, it means to refer to units processed, or to the customers served by this group of resources in the unit of time;

2. **Variety**, it means to refer to the number of different units processed (or different services offered), or different requirements met by the resource group in the unit of time;

It can be firstly **observed that**:

- **Increasing varieties**, the *time of contact with the customer will grow* and the company will be able to *personalize* the service and the *degree of discretion of the front office* to decide what to do; the skills of its staff are a key element for the company and most of the value perceived by the customer is given by the front office activities;
- **The lower the variety**, the more it *reduces the time of contact* with the customer, the degree of *personalization* of the service and the *discretionary* decision-making of those involved in the *front office*. In this case, the focus is primarily on the product or service and on the output data to the client, and management attention will firstly give priority to the management and organizational activities of back office, which generate the output.

5.3.2.1. MASS SERVICE

These service processes are characterized by a **high number of transaction per saving** (or units processed): examples are the toll booth, the operator of the call center and the teller of the post office. The process is very **standardized**: considering the case of the motorway (the procedure is always the same: the user hands in the ticket, the amount is shown, money are collected, it is given the eventual rest, the bar is raised). It is also clear that there is a **low/short interaction** with the customer, which makes the management focuses primarily on back office activities.

Productivity is a key performance: we must maximize the value of output produced by each serving (must serve a lot: a typical measured performance at the call center is "the number of calls for operator per hour"). But we must also **pay attention to compliance/conformance** (because you are selling a service standard, and the customer goes to get that precise standard, well-defined features, it will be necessary to ensure that the 'output meets those expectations').

About the level of productivity, process standardization and output conformity, an interesting example to be mentioned is the service provided by *McDonald*: at any point of sale, any time it is, and whatever the geographical area, procedures require that the burger must always have the same look, the same shape, the same flavour. There is an almost obsessive attention: for each of the ingredients, there is a single source for geographical area (a single supplier to ensure consistency, it is also possible that this will serve more areas). There is a manual that explains exactly what the procedures are to follow to cook each product and implemented a system of auditing extremely precise in order to verify that, indeed, all activities are carried out according to predefined rules, thus the standardization is high.
This opens up a large area for the use of computerization: if, in fact, in manufacturing processes, high-volume, standardized process means introduction of extensive automation, worldwide service concept is jointly mutual, with the introduction of information systems able to maximize productivity and also compliance.

For example, applications that are finding more and more spread, are so-called **system of "trouble shooting"** (born in the field of banking help desk). When the customer has a problem, and phones the call center, any operator, regardless of his skills, might be able to resolve this problem using this application, by asking a series of questions to the customer (and having as input his answers), the monitor shows what should be the most appropriate procedure to resolve it. A system of this type is very expensive, as it must be designed and dedicated ad hoc on a service provided by the company; however, it is justified precisely by the high volumes that the system has to face: it can be made available to a number of different operators, increasing computerization and productivity, so it is quite advantageous.

Finally, from all these characteristics, it is evident how the skills are **mainly incorporated within the system, and the size of innovation concerns mainly the process**. The procedures are also embedded in the system, and not only them, but also the competences to complete a certain task.
Note, in fact, that defined procedures (like the case of the mass service) is just a way to incorporate knowledge into the system instead of resources: that speech was already made, alleging that, carefully checking the input, and the process, using predefined rules and standards, enables to achieve therefore a consistent output. The procedures also allow you to obtain the efficiency of the system because the skills are incorporated into the service delivery system and, to perform tasks, it will be sufficient to take those rules that are defined, no major training efforts. Always with reference to McDonald: anyone can cook one of its products, it is sufficient to follow the default rules; this, of course, allows you to leverage **low cost of staff training**.

Innovation is in the process, the way of reaching a simple output. The internal performance is the productivity. The competitive advantage is given by competences that are embodied in the internal processes.

5.3.2.2. PROFESSIONAL SERVICES

The service processes are the opposite of the above, as characterized by poorly defined processes and lower transaction volumes for each serving (or units processed). They are examples of **law firms, specialized clinics and private consultants and psychotherapists.**

The main characteristics are:

- **Low volume of transactions for server/unit**: think about the number of patients who visit a psycho-therapist studio.
- **Strong interaction** with the client and, above all, the fact that the goal is to **provide solutions**; while, in the mass services, the focus is first and foremost the final output and the back office (when you go to McDonald, the goal is to get a sandwich based on what you expected). When you buy a professional service, the customer buys firstly the opportunity to have a solution. Assuming a private consultant, because it is expected to identify the solution to a particular problem; those who go by the lawyer because it is expected that he will be able to determine the best scenario for resolve legal problems.
- The problem is **not standardized** and it is not so easy to solve. **No rigidity**. Not so structured.
- **Product innovation**: precisely because they sell solutions, each time you will be able to obtain different outputs, so the characteristic feature lies in the final result obtained.
- Finally, skills and abilities of the people are the critical asset of a system of this type: a professional company sells its powers and its ability to create solutions, and not a standard process and repetitive for which you can define the procedures in advance.
- Attention to provide solutions, no matter what: the final result is unknown, but the expert will find a solution. What is sold is the capability to find a good solution to the problem.
  EXAMPLE: Doctor, consultancy. You will deliver a project in 2-3 weeks. It is not about delivering a McChicken. Every customer has his own solution. McChicken is always the same product. If you want to innovate you have to find new solutions for customers.

5.3.2.3. PROFESSIONAL SERVICE SHOPS

Professional service shops are the development of the professional services as DIMENSION INCREASES. This represents the evolution of the type of "professional services": with increasing volume of transactions, the objective becomes to try to **increase efficiency** (because of the higher volumes) but without penalizing too much the level of customization. A fairly example could be: think of three different private consultants who, one day, decided to set up together a consulting firm; this is just a transition from "professional services" to "professional service shops". In a situation like this, the aim of making efficiency without reducing too much customization, it means finding a compromise between the following two extreme situations; assuming that you present a client to show its problems and require a solution, the company could:

- **Identify a similar case**, look at what has been the approach adopted in that situation, provide exactly the same solution at the new customer
- Commit their resources on delivering a **highly focused solution** specialized to solve the specific problem of the customer

In the first case, it comes to an execution of activities too pushed towards the mass services (recycling of a solution already identified). In the second, however, is accomplished much customization on the market that It is not able to guarantee the same margins that would give a professional service (where it would be justified this approach, but the context is another, and this, not to waste, would impose too high a price for individual counselling, which would exit the market).
The goal will be to find the right compromise between reticence and customization based on what are the characteristics of your target market.

A second fundamental factor is that of the **sharing of knowledge**, which should be incorporated into the system, for example, thanks to knowledge management systems, able to integrate and make available the experience and knowledge of the various subject. With reference to the example of the consulting firm, once decided to join, the three resources will have to collaborate in order to share their cultural heritage and their knowledge, so that the company's overall result was the integration of the three individual and not the mere sum "in silos" of the experience of the three parties (but each jealously protects).

As a result of these early features, there is the need to develop “a style of the house” because the subjects are no longer working privately but are part of the same company. Although the activities still leave margins of freedom, there should be a common approach to operations, marketing, purchases...; resources must not act in a way inconsistent and uncoordinated but, while respecting freedom of certain decision-making, must develop ways to act as a common factor.
Here, the **autonomy decreases**: the individual will not be totally free to act according to his will but should at least follow the general guidelines (he can no longer make the price considered most appropriate to the situation but will have to stick to a certain range; he will not offer customers everything they want, but it will be limited).

Arise, therefore, semi-professional figures, no longer highly specialized on a particular activity, but expert on a number of them. Remember that the objective is always to achieve efficiency: no longer have a role for the activity X, one for the task Y, one for the activity Z and one for the activities W, but a figure that could occupy the tasks of X and Y, and another that could occupy of tasks W and Z. Consulting companies changed the organizational structure in this direction: the companies hired people experts in communication and not in the specific field. In most cases the managing directors already know what they should modify in the market and ask consultants for some help in understanding if this choice is correct. Therefore, consultants do not have to arrive with new proposals but just have to evaluate them.

Finally, it is possible **to develop figures specifically dedicated to treating the customer relationship**. Who interacts directly with clients does not always have the sensitivity to deal with it so you can consider establishing an ad hoc figures that manage that interface?

5.3.2.4. MASS SERVICE SHOPS

They represent the development of the mass services to **a greater range of services offered:** in essence, when a mass service realizes to operate properly and provide benefits that meet the customer satisfaction, you may try to increase penetration of the customer by offering additional services. Examples are the **supplier of computers** which also offers activity of data transfer from the old equipment; or a supplier of computers which also offers the supply of printers and fax machines; or a computer supplier offering advice on the design of the corporate information network.
The focus of this class is to emphasize that it is the customer that drives and require the provider to expand its range of services. It is clear that the more services are offered, the more you have difficulty in following procedures, adopting the logic of standardization and, more generally, referring to all those principles for the mass service. **Increase the difficulties in terms of compliance monitoring (or, on other words, the strict control over the process)**: in case a problem occurs, it is more difficult to understand if this problem stems from the fact that indeed something is not working or because there are different products.
In expand its offering, the company’s service cannot be separated from increasing the degree of discretion front office in order to allow employees to self-manage different situations without affecting the compliance of the process (but training employees more). They also have to be able to understand the customer’s need: the company must be able to provide the right service among those in the predefined range.
While, in the mass service, the logic of customer approach was "this is the range that I propose, is up to you to choose according to your needs", in the mass service shops, the client communicates what are its characteristics, its need, its prospects and expectations, and they are subsequently offered the possible solutions, among those available.

POSSIBLE USES

1. **Verification of internal consistency**: depending on the configuration (mass service, professional service, mass service shop, professional service shop), you have to go and check that, actually, the internal levers are structured in a manner consistent with this configuration (so, for example, if it is logical "mass service", we must verify that procedures have been defined, it was designed a system of compliance monitoring, and so on)

2. **Verification of external consistency**: it is always the same principle of alignment. We have to understand if the way it is structured is appropriate and aligned with the concept sold on the market, or if the market looks at the company with different eyes. Secondly, you have to make sure that, given the configuration, this is actually consistent with what the targeted market wants. You have to apply this logic in a dynamic manner (going to ask if there have been developments in the market, with respect to the moment in which the system was configured, which could make necessary changes)

3. **Compare different units/businesses:** different levels of units or businesses may need different levels of control

4. **Management of change**: if you want to move from a point to another, you have to understand what the interventions are to be done, induced consequences, and ensure consistency between the operations change and the message to give to market.

To be consistent with the new market, operations will have to modify in terms of:

- More flexible procedures and/or increasing the number of procedures defined (to be able to handle possible exceptions)
- Introduction of new skills and roles (as it offers more services, you must also have someone who knows such services)
- Increase in the degree of discretion of the front office: increasing the time of interaction with the customer, it needs to develop an attitude to relationship with the latter (the servant had to enjoy interacting with the customer)
- More generally, to make investments in training, review tasks: operations must figure out how to orient the customer, it must be able to understand their needs and I lead him to the offer he would be more interested in
- Accept this greater degree of delegation to the front office
- In general, a different structure and processes approach

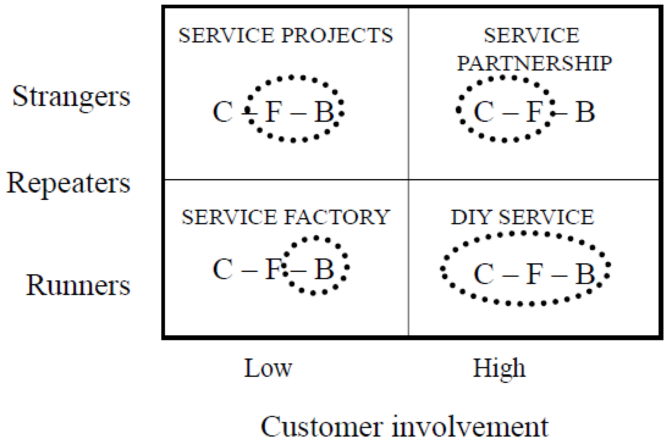
To better **understand the variability** of the process, the variety degree that the process needs to manage, we can divide the service requests:

- **Runners**: applications that always need the same operations / activities. Often predictable, and in large volumes. Requested that we made a lot of times. Opportunity for automation and process review. *Example: oil change, McDonald's classical cheeseburger*. 90% of the cases.
- **Repeaters**: requests that refer to activities known but grouped differently. Unpredictable and medium/low volumes. Events are planned but rare. Everything is defined and we are able to satisfy. *Example: potatoes with the sausage of Crispy McBacon.*
- **Strangers**: requests that require the creation of new activities. Often unpredictable. Are the exceptions: events are NOT provided. You can prepare it but not totally in advance. *Example: going to McDonald's and ask for a totally new burger.*

Service	Runners	Repeaters	Strangers
Car Service	Oil change Tyre replacement	Gear box repair	Product recall Electronic malfunctioning
Airline	Check-in In flight service Maintenance	Overbooking	Hijacking
Hospital	Routine surgery Operating book Rehabilitation	Post-surgery complication	New surgical technology

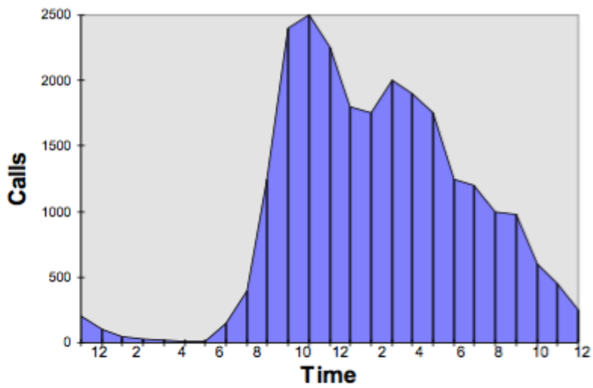
A mass service would have 100% of runners, a professional service would have the majority of strangers events.

WHERE TO FOCUS: The customer involvement depends on the occurrence of Strangers, Repeaters and Runners.



- SERVICE PROJECTS: it is like the consulting firm example. The customer interacts with the FO but then the key point is the interaction between front and back office. It is important to analyse the results of the partnership together. For example: market research.
- SERVICE PARTNERSHIP: the larger part of the activities is done with the customer (rehabilitation in health care center). The solution is worked out differently with each customer. An example for this kind is an operation consulting company. The consultants go on place to change the company together with the employees. For example: consultancy, the doctor.
- SERVICE FACTORY: the competitive advantage is built in the back office (postal service), the attention is mostly on the back office: efficiency and speed are the most important characteristics of the firm. For example: post office.
- DIY SERVICE (do it yourself): you have to balance the activity either in F/B office and with the customer. The process could turn to be unfeasible if the customer is not able to manage it (do it yourself service). The company is successful is somebody else (the customer) is completing the task. The company has to develop the ability to communicate to the customers the right directions. For example: entertainment industries, cinemas, service restaurant.

5.3.3. VARIABILITY (of demand, of capacity) and UNCERTAINTY



- **VARIABILITY** is the gap between the actual value and the average value
- **UNCERTAINTY** is the gap between the actual value and the expected one: how easy is to forecast?
  We may have a situation which is characterized by low variability and high uncertainty.
  On the other way round uncertainty is variability that cannot be explained.

How to manage variability?
Variability not explained caused uncertainty. The company has to transform the uncertainty in variability.
COROLLARY: uncertainty can be reduced by explaining it, thus reducing it to variability (which is much easier to manage because it depends on you).

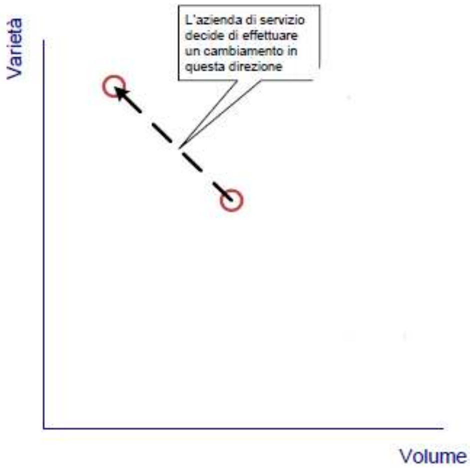
There are 2 types of variability:

- ***determined by the company*** that can control or remove it
- ***outer variability*** that could be limited by the company

*EXAMPLE: **Pareto rule**: 20% of the products/services generates 80% of variability. If it is not possible to remove variability, isolate it. It was figured that in the hospital 1 hour was not saturated: we may saturate it better by trying to explain the reason of the uncertainty. Maybe a surgeon is faster than another one, therefore it could be **controlled by the company**. The techniques may be analysed, the average data of the time of the service would be measured and in the end the uncertainty would be understood, but the variability still would be high. In the other case, the uncertainty from **external events** should be limited. For example, this can be done by inserting a queue in the system and this would allow to split the system into two subsystems in order to isolate the variability. A smaller system would also include a smaller volume and lower interactions and so it is easier to understand the course of events and therefore the variability decreases.*

Therefore, the better approach is **to handle the variability at the front office**. The reason is that the front office could be trained to interact with the customers and to understand their needs. This process may take more time but it could better understand the customers' expectations and bring the customers to one of the back offices which result to be more suitable for them. **The greater the variability, the greater the front office's competences and discretion needed**: in this case the front office loses control of the system but understands people better and is able to switch the to one of the services or does not accept the customer at all.

ATTENTION: as regard **professional service** configuration there is a certain degree of customer variability you do not want to eliminate because it is related to the *presence of the customer within the service delivery system*. On the other hand, **mass service** configuration aims at *reducing this variability* to the minimum value, because of the lower interaction with customers.
In a SERVICE SYSTEM variability must be managed at the front office: the greater the variability, the greater the front office's competences and discretion needed. Variability can't affect BO activities, so you need competent employees in the FO capable to absorb variability related to customers and outside causes. In the FO you have to listen to each customers' request and find which service will be more suitable for it.



A system based on command and control is very inefficient in managing variability. This refers to system based on procedures that do not allow variability inside the process. You can no longer divided the procedures in steps to be assigned to workers who will not have the power to make any change.

6. SERVICE CONCEPT

6.1. LUNCHING A NEW SERVICE

- **Restaurant**: Our idea is about a restaurant which offers dishes' customization. It is mainly for meat lovers because it provides a wide visual choice of different cuts of meat of different parts of different animals
- **Supermarket**

The service concept aims at aligning the System towards the inside and the outside. Aligning the goal is crucial because the process is not totally controlled. It reduces the gap between what the customer expects and what he really received.
The service concept focuses on the difference between:

- How **the company** would like the service **to be perceived** form clients, employees and stakeholders
- How **clients, employees and stakeholders see** the service

The Service Concept is made of:

1. **Organizing idea**: it regards the ***essence of the service brought to, or used by customer***
   It is a wonderful alignment tool for employees and for customers, a guideline for all choices and a continuous reinforce of the competitive differentiation of the service offered. It is linked to an improvement in customer perception. Operating translation of the vision of the company. It is a statement for the priority on the choice of the employees of the company giving the position of the company in the market.
2. **The service provided**: it is related to:
   - the ***processes*** to deliver the system
   - the ***output***

   The ***Service Delivery System*** is used to identify the key characteristics of the output. It is the environment where the customer is included.
3. **The service received**: it is different from the service provided because it regards:
   - the ***outcome,*** received by the customer.
   - the ***experience***, we have all the feature that we set before, the no tangible part.

   e.g. Food and calories

This is the service from the customer perspective: what is the advantage for customers? There is no difference between the output and the outcome, only for the perspective. The **output** is *what the company deliver to its clients*; while what the customer receives, so the outcome, is different: it is a *mixture with his individual perception*. Operations can design a process related to a certain output, not outcome. The difference between the output and outcome is: they are basically the same thing looked at two different point of view; output is what we think o do for the customer and the outcome is the perception.

E.g. Output: music → outcome: noise or sensations
Output: car problem solving → outcome: the car still does not perform well

Designing the ***experience*** is important according to these factors:

- ***customization*** of the process
- response ***speed*** (of the delivery system)
- employees' ***flexibility***: home banking systems are characterized by a low flexibility while other services can allow different behaviours of customers
- ***intimacy*** with the customer: colloquial or formal language
- ***accessibility*** of the personnel: ease to find information to ask to the personnel
- ***perception*** of being valuable: through information detection
- ***curtesy*** and ***competence***: fundamental to the right perception of the company

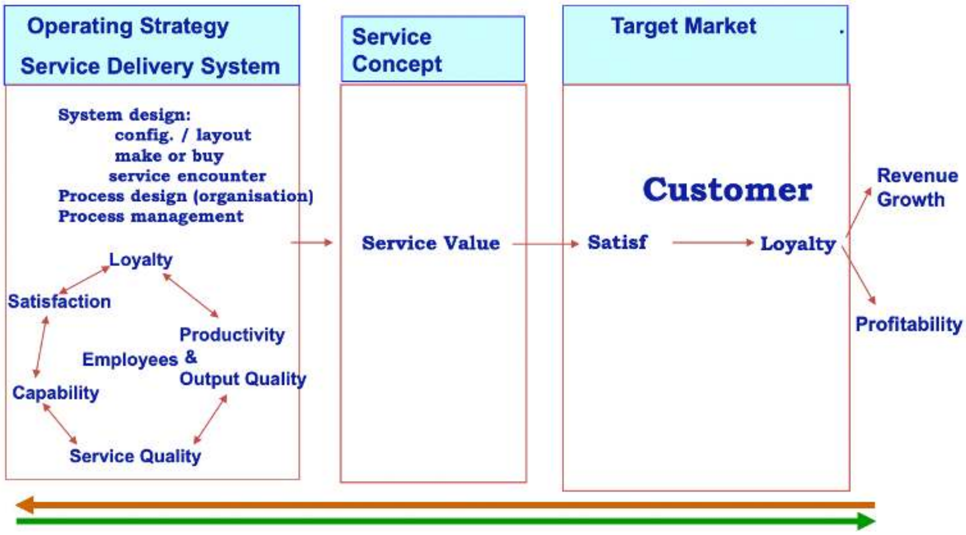
As roller coasters' designers should be able to design the sequence of feelings and sensations that the customers have to experience, service companies also have to deliver to the customers a wide range of opportunities which define their personal experience.

Complete the proposal and add the description of the business from the customer perspective

Example: Bus: outcome → the moving from one point to another; the output → for the service company is to have design a certain route that is longer and the bus that moves from the initial step and goes to the final step. In the journey it has many stops that include the customer's one.

NB: We design operation for a certain output in order to satisfy a certain outcome.

6.2. THE SERVICE PROFIT CHAIN



This scheme is used to *align operations with the market*, with an additional part regarding the service profit chain. The investigation of the more profitable companied highlights the presence of ***customer loyalty***, which is linked to the ***satisfaction*** of the customers in the *target market*. In fact, customer loyalty is the only attribute which enables companies to gain revenues and profitability.
The *service concept* regards the value proposition: the ***service value*** which consists in what you are selling to a specific market segment. It is based on the perception, if there is a misalignment: Or I invest in making him perceive or I disinvest in my product.
Moreover, at the level of the *operating strategy*, apart from the system design, companies should also consider the importance of ***employees***. These are in direct contact with the customers, they do not only provide the service, but they sell it. Successful companies have loyal employees, which are realized by their job. Investments in training may attract employees, make them satisfied and more capable. And this is a profitable loop which helps both companies and employees, providing better productivity and output quality, but also better service quality.
Many measures of customer retention and customer loyalty are focused on past analysis more than finding elements to ensure customers' future loyalty and surprisingly, there is more attention on finding new customers than keeping those already engaged, as if customers were inalienable.
A customer that is satisfied it is easier to supply, also because his friends and so reputation.

6.2.1. CUSTOMER VALUE EQUATION

Customer's value is related to outcome and experience and can be compared with the price they actually buy the products and the acquisition costs in order to get to the wanted product (therefore the total acquisition cost).

$$\textit{customer value} = \textit{outcome} + \textit{experience price} + \textit{acquisition cost}$$

Customer value: is the comparison between the value perceived by the customer and what he spent to enjoy it.

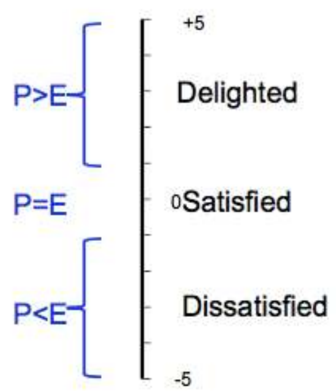
Acquisition costs: all the additional costs that are necessary for the customer in order to enjoy the service. They are relevant to be analysed because often the acquisition costs are higher than the main business itself (for example: the total cost of Disneyland minus the entrance is much higher than the entrance cost itself). It is important to analyse the total cost that is related to the whole business.

Most of times, the real business is not the business itself but what there is around it.
Customer's loyalty: it is an indicator of the satisfaction of the customer, and so of revenues, growth and profitability. Many measures of customer retention and customer loyalty are focused on **past analysis** and not on new elements that will ensure customers' future loyalty; switch costs are usually low.

Companies must switch to the customer's perspective:

- **Southwest Airlines** is able to **understand customers' needs** thanks to their employees: the people working with customers. Postponing flights timetables to the time which is most suitable with workers' clients. The information technology is making this easier and possible and on a continuous base.
- **Shouldice Hospital** is, on the other hand, **selecting the customers** by asserting that they are the wrong supplier for certain kind of disease.
- **US Army developed focused packages** of insurance to dramatically reduce the number of checks of documents. This would reduce costs and time spent in doing the checking. It was used because of the higher mobility the experimented, but also to provide a better service for customers, people followed by an insurance company who already knows them well and is in contact with them. This sample checking is affordable because of the army education of the people.
- **Progressive corporation** is one of the largest providers of car insurances in the United States. It **developed focused service packages** of insurances addressed to people who get caught by being drunk by understanding the characteristics affecting the risk of this specific situation. For example, people with a child appear to be more responsible than other customers, therefore are lower risk and are asked lower price insurances. Meanwhile higher risk situations have to pay more. And this management and understanding of the operations gets higher profitability for the company.
- **Managing expectations** is also fundamental. As the Western world's markets are not growing, it is important to meet the expectations of the customers.

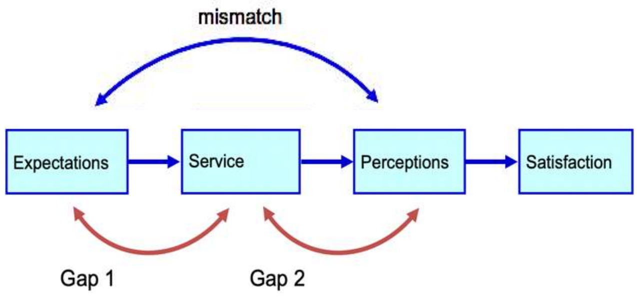
The ***Prisoners Dilemma***: what is the best solution?
The prisoner's dilemma is a paradox in decision analysis in which two individuals acting in their own self-interest pursue a course of action that does not result in the ideal outcome. The typical prisoner's dilemma is set up in such a way that ***both parties choose to protect themselves at the expense of the other participant***. As a result of following a purely logical thought process, both participants find themselves in a worse state than if they had cooperated with each other in the decision-making process.
Mobile phones telecommunication providers had a loss in customers which was really high (25%) and a corresponding gain in new customers which was almost equal. But this system is much more expensive for those companies who have to provide the newcomers, cheaper contracts.
In the past years, the fight between phone provider was to steal the customer to competitor, instead of making the already acquired customers happier, loyal and make them buy more. It would be easier and cheaper to focus on the existing customers.
In last years they tried to change this system: they tried to make the existing customers happier by adding phone minutes to their offer, which is absolutely costless for them.



There is an area where customers are **satisfied** and perception is aligned with expectations.
If perception is lower than expectations, customers are **dissatisfied**.
If perception is higher, customers are **delighted**.
What is the expectation? It is coming from the reputation and the total costs: the higher is the cost, the higher is the expectation. The same is valid for marketing that is linked with the reputation → the more the more.

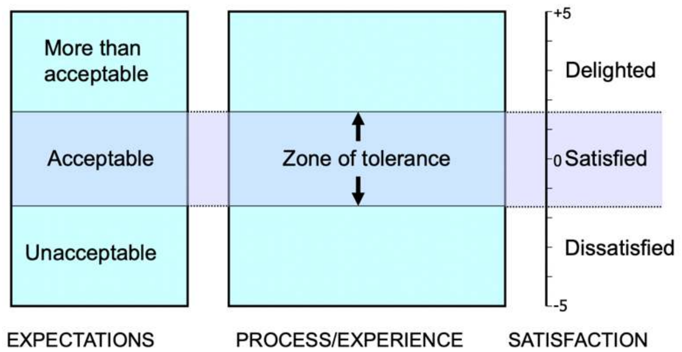
There are two possible mismatches: you must manage the 2 gaps.

- *Gap 1*: promising more to the potential customers may create a gap between expectations and the service which is actually provided. This is bad for customers but easy for companies.
- *Gap 2*: the customers are not perceiving the value of the service they have been provided. The work is not on the service, but on the perception. E.g. comparing your service with the one of the others.

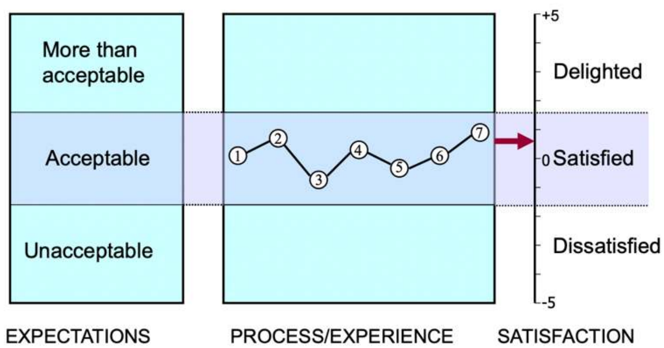


EXAMPLE: You are serving the best wine in the city (service), but if the customer is not able to feel the difference between your wines and the other there is a gap. What are you providing is not covering the expectations? You should work and improve your different communication to the customers.
It is important to improve these points.

Therefore, a ***zone of tolerance*** exists where expectations are aligned with perception which is acceptable for companies and on the other hand make customers satisfied.



This, companies should try to create a ***system of checks*** that will help to reach the intended objective, because in the end if the customers are dissatisfied, they are difficult to change.



7. ONTARIO PACKAGING

INTRODUCTION

The management of the company took into account the investment proposal of the replacement of the roller (purchased 12 years ago), with a new, more modern process, which would provide the same capacity. At a cost of **2.5 million€,** the investment would ensure savings in labour costs, consumption of materials and other elements, such as to provide a payback of 4.5 years at current production volumes. However, this solution does not respect the constraint imposed by the group leader of a pay-back time of 4 years. The work team responsible for evaluating the investment has been exploring the possibility of **increasing the volume of sales by 40%** going to serve a new market segment. With the increased volume of production, you could better saturate the machine, and bring down the payback in 3.5 years. The team proposed this new hypothesis to the parent company, which asked for **confirmation that the pay-back did not exceed four years before giving its approval.**

The new process would have had the same setup times, but a rolling speed double so a **single machine could replace the current process**, which was formed by two machines. Considering an increase in production volumes of 40%, the **new process would be saturated at 80-85%** (before 60%).

The need for change is essential and urgent: the marketing has in fact recognized as the current **machines will no longer be able to meet 100% new customer demands** (it is a **question of quality** of specifications: marketing has highlighted the fact that, customers are evolving their needs and the old process will no longer be able to satisfy them).

MARKET REQUIREMENTS

Given the previous description of the Ontario Packing, in the table we would evaluate which are the most suitable requirements in order to have the best benefits from the operation system and which of them instead, constitute the present performances of the company.

→ **Reconciliation Model** is the alignment between Market requests and company deliver.

<u>Customer analysis:</u> What are the performances requested by our traditional customers? High quality and Services (time, reliability).

PERFORMANCE		CLASSIFICATION	NOTE
TIME	Speed	OW	It is an OW as it is a reason for the customer's choice of supplier. Creates competitive advantage
	Reliability	OW	Similar to the previous
QUALITY	Design	Q (order user: bad values will decrease the number of users)	The premium price comes from the service, so it does not provide new orders, but now we are facing quality issues (we are in a lower level than the threshold, i.e. qualifier performance)
	Conformance	Q	
COST		Q	We are operating in high quality market, it is taken for granted
FLEXIBILITY	Plan	OW	It is a premium service; it gives a competitive advantage
	Product	Q	
SERVICE	No info		

Here there is a description of the reasons why the previous market performances were classified as Order Winners, Qualifiers or Order Losers:

- <u>Product flexibility</u> is certainly one of the most important performances, as well as the strengths of Ontario: **it operates on demand**, which means that is the client to say what he wants, to provide design, while the company takes charge of the creation of packaging in the manner and desired characteristics. This means an amplitude of almost infinite range, extremely high flexibility of the product.
- The <u>speed</u> is a fundamental parameter (in terms of "drivers" of the number of orders received by the company): when the customer requires a product, you need to be as quickly as possible, tending to lots of small dimensions.

- <u>Plan flexibility</u>: as customers may change their requirements on the order until a few days prior to its implementation, it is necessary to tend to zero the time horizon frozen; It must be in a position to change whenever the customer requests it.
- The **cost** represents a “qualifier”, since it has at least to remain within an acceptable range.
- The **quality conformance** is a "qualifier"
- The **quality of specifications** is also a "qualifier", but is borderline in becoming "losers": as highlighted in the introductory phase, in fact, currently the specifications meet the minimum requirements of the customer, but the situation should evolve in a negative way in cases where it is decided not to invest and not to change (the customers are becoming more demanding in terms of accuracy specifications, the number of sheets of rolling, number of processed materials ... and, this process, in a bit will not be able to meet them)

LEVERS	IMPACT
HIGH SATURATION (TECHNOLOGICAL CHOICE)	Positive impact on: ▪ **Lower costs** Negative impact on: ▪ **Flexibility,** not directly. ▪ **Time speed**
FRACTIONING DECREASE (TECHNOLOGICAL CHOICE)	
SETUP INCREASE (TECHNOLOGICAL CHOICE) Affected by: • Fractioning	Negative impact on: ▪ **Flexibility,** not directly. ▪ **Time speed** I can group all similar product together in order to reduce the impact. Suppose to have: ABCABCAB I schedule as: AAABBBCC → only 3 setups instead of 8. This behaviour is affected by the time of setup. The longer the setup time the bigger the batches that I schedule.
LESS PEOPLE (ORGANIZATIONAL CHOICHE) Affected by: • Setup increase	Positive impact on: ▪ **Lower costs**
BATCH SIZE Affected by: • Setup increase • High saturation	Negative impact on: ▪ **Flexibility,** not directly. ▪ **Time speed** Batching is the factor that leads to a reduction in time speed. Suppose to have: ABCABCAB I schedule as: AAABBBCC with one batch per day. The first C order is produced now after 7/8 days instead of 3. The flexibility of making changes is changed negatively

<u>OSS</u>: If the **setup time becomes too heavy**, we try to reschedule orders to minimise the number of setups and consequently their costs the batch size increases, which is the most important factor in deciding the frequency of a certain production. **Batching** influences the company's resting time.

- If the **batch size is smaller**, we are able to respond in fewer days, we are shorter, more responsible. We have better reaction leverage even if we need a change in the plan.

- If the **batch size is higher**, changing the plan (e.g. for an urgent order) will have a bigger impact on the reliability/speed of the company.

WILL THE INVESTMENT BE APPROVED?

No, because it will not work. On the contrary, it will worsen the performance of the company's Order Winners, the ones that give competitive advantages. At the same time, we cannot avoid the investment, because if we do nothing, we will still have quality problems.

When we calculate the payback time the error is:
Considering the scenario with payback time equal to 6.5 years, under current conditions with the 2 old machines, with the new machines we have a cost that is low. The new machines give us a gain compared to the old situation. This gain is actually our margin of the new machine over the old machines. This margin is the thing that increases our profit, the higher the margin the higher the profit rate. But this is not enough because they have not considered that the revenue with the new machine is stable, this is a necessary investment, if we do nothing, we will lose revenue and market share and the payback time usually does not consider this. The slope of the increase in profit is the sum of: the fact that the new machine is working better than the current machine, and the fact that with the new machine we are able to keep the revenue as it is, and therefore guarantees us more revenue. If we consider both impacts, the payback time is no longer 6.5 years but will be shorter. We do not know if it will be shorter than 4.5 years, but we are sure that it will be shorter.

What happens if Ontario Packing does nothing?

There might be issues in QUALITY. Quality is an Order Loser performance → we are losing market share → we are losing revenues. Under current conditions, it would not be able to achieve it (serve four and a half to pay back the investment), so it has decided to search for a new market segment to be served. In this way it would have saturated the free capacity of the machinery, thus improving the evaluation indicator.
→ 4,5 years as Payback Time is impossible with +40% of revenues. Let us try with 6,5 years.

MESSAGES OF THE CASE

Evaluating the scenario, what are the effects of actual customer perception?
1. It is about how it is going to change performance to deliver to the customer;
2. The levers influence each other and are strongly correlated;
3. The evaluation could only be done on a number, so only in the payback time. A decision cannot be made based only on such short information

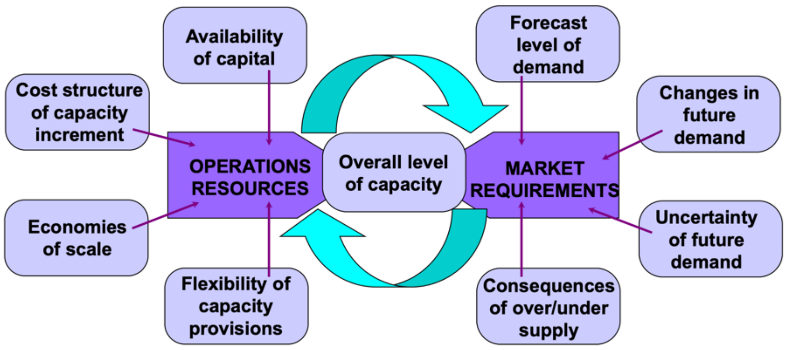
8. STRATEGIC CAPACITY MANAGEMENT

The decision on **how much capacity installed** within the system and how to segment it is not a "static" choice but needs to be revised (generally upward) at different times. With strategic we mean the **medium-long term**. In this sense, it is necessary to define different key points:

- **Timing of the change**, or when you need to change;
- **Magnitude of change**, or how many units you need to change;
- **Attention to transitory phase**, how do you manage the transitory phase.

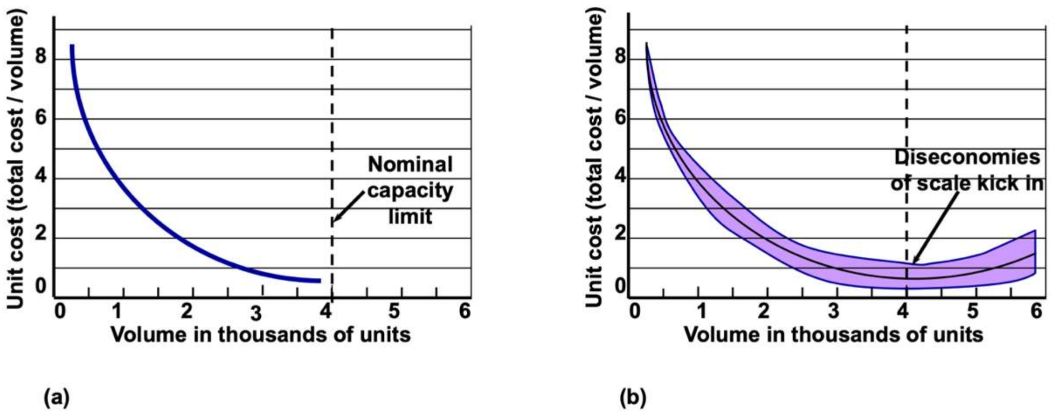
To understand when it is appropriate to change, we must, first of all, evaluate:

- The **lead time of the change**, e.g. the time required to effectively obtain the expansion of the production capacity;
- Interior **flexibility to change**;
- The influence of possible **economies of scale**: if they have a very important significance, will need to make larger units, otherwise we would limit extensions to "*incremental*";
- The **expected demand**, making predictions;
- The **level of uncertainty** associated with these predictions (which will always and in any case wrong). The higher the uncertainty, the riskier the investment;
- The **behaviour of competitors**;
- Customer behaviour (in particular, in terms of level of service required).



Some factors influencing the overall level of capacity

We may have elements influencing capacity coming from the MARKET (EXTERNAL ELEMENTS) of related to the RESOURCES owned and their management (INTERNAL MANAGEMENT).
There is an area which is under the unit cost which is referred to RISK



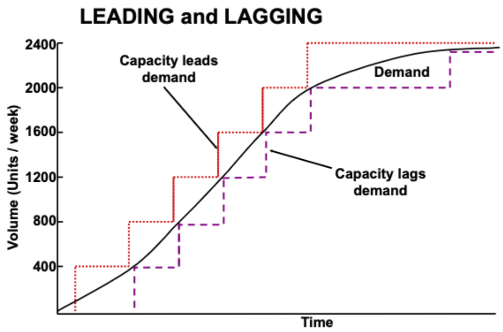
Unit cost curve

8.1. TIMING OF CHANGE

8.1.1. Leading and Lagging Strategies

Given a certain level of current production capacity, and a demand forecast for the future, it is possible to process two different strategies for capacity adjustment:

- **LAGGING**: a *reactive* strategy, to **pursue the demand**: the capacity is expanded at the time when it is expected there will be a sufficient level of demand for saturate it;
- **LEADING**: a *proactive* one, to **anticipate the demand**: capacity is expanded at a time when the current value is saturated by market demands (basically, you go to anticipate possible increases in demand)



The best strategy does not exist, because each strategy can be contextualized and has pros and cons

The obvious disadvantages of a strategy of anticipation (**LEADING**) are:

- **Greater inefficiency** (: you may have higher impact from overestimating demand and therefore low-capacity utilization);
- **Greater financial exposure** (= higher production unit cost or outbound cash flow: in relative terms compared to another situation: you make an upfront investment of an increase only expected, but not yet reported).

In contrast, however the advantages are:

- You can **always leverage to meet the needs of its customers**, being able to guarantee him a high level of service;
- If there be a higher than expected demand, there would be **extra capacity** (*spare capacity*) to be able to meet the pick (in the case in which one would follow, however, it would come to a maximum saturation, based on the forecast, not being in the condition to acquire new orders);
- In case you had problems with new plants (for example, failures or stopped), **the impact on customers is lower** (which can be observed looking at the graph: since a delay in the production, means to move the curve capacity more towards the right, we see immediately that, in the case in which moves the one that anticipates the demand, the area of demands in delay has a certain value; in the case in which moves the one that chases the application, the requests within delay extend along a much larger area);
- Finally, we have **delivery times shorter** and **more reliable** (as in the case of Ontario Packaging, with unsaturation plant) with extra capacity, you can deliver in times more content and, in case of problems, you have the ability to recover precisely because it has free capacity.

It is evident that they are **two opposite strategies** (LAGGING), advantages and disadvantages of a policy to "chase" the question is **exactly the dual of the previous one**.

The advantages are:

- **High plant utilization**
- **Low production cost**
- **Lower impact from overestimating demand**

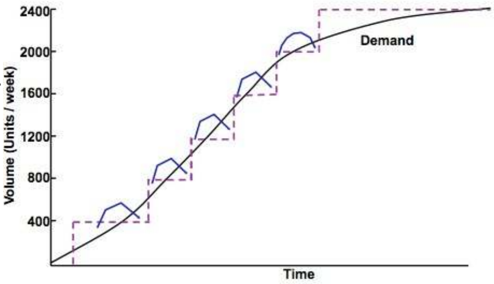
The disadvantages are:

- **Longer response rate**, due to the high saturation. The system is rigid, there are rigid plans to follow.
- **Lower delivery reliability**, the rigidity means that if there is an unforeseen event (failure) this one has a huge impact on the system;
- **Higher impact from underestimating demand**, the company loses the opportunity.

In general, it can be said that where the market rewards a particularly low time of delivery, a high level of service, and where the price is not a key variable, it is preferable to adopt a strategy aimed at the advance of the application (LEADING). On the contrary, where the market looks first at the price, and require efficiency, with marginal attention to the level of service and delivery, then you can choose to "*chase*" (LAGGING STRATEGY).

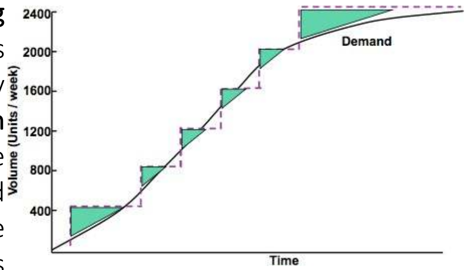
8.1.2. Smoothing

A first **intermediate situation** is one that sees frequent increases in production capacity. In a case of this type, you can use the **created stocks in moments of unsaturation** with the first implant, to **cover the possible capacity deficit** while making the extension (it is obvious that, if such increases is more sporadic, cannot be stock up now for a future, too far away, where the expansion will occur).



8.1.3. Filling products

A second interesting alternative is to use the products fillers. The **filling products** are goods or services, not directly related with the company's core business, but it can make or deliver as they possess the technology and skills needed. **These products are used to saturate the production capacity not currently used**. The logic is as follows: although these products do not fall within the main market that is going to serve, and although we are aware that you would not be competitive on the market going to compete with the whole cost structure, this segment is used instrumentally to repay only the company's overall fixed costs (usually high), making it sufficient that give a positive contribution margin. If those products are not delivered, there will not be a significant difference in the market. For example, if you are a company with high fixed costs and low variable costs, if you have spare capacity, the additional value added to the saturated capacity is higher. You can also sell these products at a price which is higher than the variable costs only, and therefore, which is lower than competitors, you will gain a competitive advantage. Saturating capacity would provide product fillers. If I have spare time, I can do it. It is something that I do not have in my core business, but it is still something that I can do.

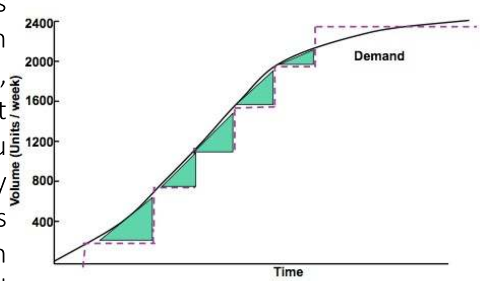


Using filling products is like modifying demand:

- It is possible to influence demand by increasing it: it is another way to have additional demand to different industries by the same technology
- Not all products are the same: it is possible to stop the production of filling products at any time, but not the one of main products.
- It is useful to identify which product sacrifice in case capacity is not enough.

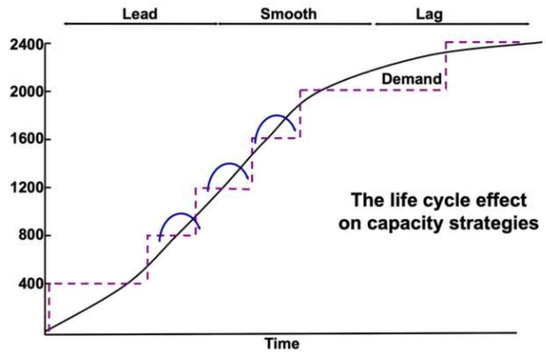
8.1.4. Outsourcing

Another possibility is to decide to **outsource**, which, however, is less obvious than you might think; to turn to outside suppliers you must, in fact, have done upstream the whole process of preparation, research, evaluation and selection of the most suitable supplier. Also, so that it can be achieved the effectiveness and efficiency in the report, you need to put the supplier in ideal conditions, preparing the technology (if necessary) and providing all necessary information to the case. It is a middle one strategy that is closest to the lagging strategy: it is when you use an external supplier to fill the customer's demand in the period when production capacity is not sufficient. So what part of demand is best outsourced?



8.2. THE LIFE CYCLE EFFECT ON CAPACITY STRATEGIES

The common behaviour: usually the company adopts the leading strategy: if the company believes in the product, it adopts it to be aggressive in the market and capture market share by creating the brand and the name, when the product is maturing the company usually changes the strategy and goes from smoothing, in the final stages, completed maturity, it usually uses a lagging strategy to minimize investment and has capacity that is not useful. Strategic capacity choices have to be coherent with the phase.

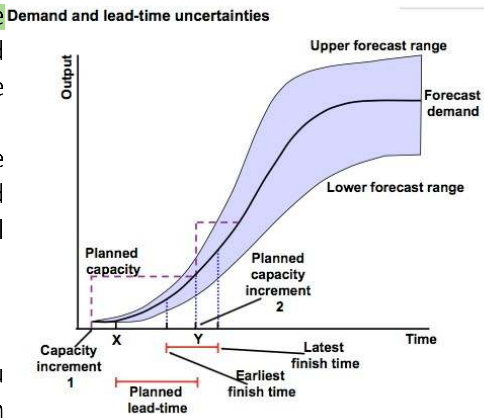


Information area is not certain

A factor which cannot be neglected when we talk about "timing" in **uncertainty**. We have two main sources of uncertainty:

8.2.1. Errors is DEMAND FORECAST

As all previous considerations were made on the base of the trend of the demand estimated, we need to be able to understand what the demand could be. In these terms the most effective solution is to give a confidence interval on the expected value.



Moreover, there is a difference between the planned lead time and the actual one. The lead time refers to the difference between the latest and the earliest finish time. If the size of the change should be bigger, the level of uncertainty has a big role.

8.2.2. Errors in planned Lead-Time and Capacity increment

The uncertainty in demand forecasting is not the only aleatory factor; you must also keep in mind the uncertainty related to the time when such capacity will be available. In fact, it also has to include buffers for both foreseen and unforeseen events. The actual value could be shorter or longer, and obviously the best choice is to be shorter than expectancies.

8.3. SCENARIO ANALYSIS

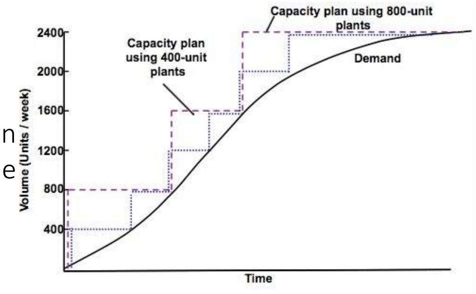
As the two main variables are the time of uncertainty about the availability of capacity and the uncertainty on the demand, it is then possible to build scenarios, and conduct sensitivity analysis (which, given a set of variables, evaluate the impact on the final result of a variation of a certain percentage of each of them). The **scenario analysis** provides us the picture of what could happen: given all the combinations between the forecasted demand and the planned lead-time, I have to understand which scenario will be more likely and which of them will bring the best results.

Moreover they have to be carried out to:

- **Anticipate possible constraints/opportunities**: on the company from a whole series of future events; once associated with each of them the value of probability, they will neglect those of low- impact, and attention will be paid to those of high-impact and high-probability, finding solutions to anticipate possible constraints and following problems.
- **Prepare countermeasure**: note however that it has a cost: you cannot be sure of everything, but we must make choices; you cannot protect themselves from everything, but you have to decide what risks face, on which be more exposal and which need to be prevented in advance.
- **Limit the impact of unforeseeable events**
- Prepare a fast, **more efficient** and **more effective response**
- The **drawback** is that they obviously have a cost → *no risk no gain*

All the decisions have certain level of risk, even if you decide not to invest has a certain level of risk. Decisions are evaluated according to:

- Value (€)
- Robustness: (Risk and Variability)



Capacity plans for meeting demand using either 800 or 400 unit capacity plants

SIZE OF INCREASE

The decision to be made on the management strategy is the size of expansion (or reduction) of this capacity; the more the growth of such dimension, the more the problems described are amplified.

For smaller steps, the risk is lower than higher steps:

- You fill the additional capacity in shorter period time
- The economies of scale under saturation are higher for higher demand
- The investments are higher for higher demand → higher costs

So, you have to consider:

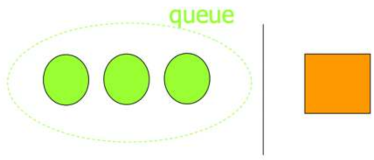
- The influence of **economies of scale**: if they have a very important significance, will need to make larger units, otherwise we would limit extensions to "incremental";
- The **variability of demand**: smaller units, giving greater flexibility in volume (it is easier to open and close a unit smaller than a large unit, because it takes less to saturate it), and even greater flexibility in the management of resources (it is easier to dispose of a small unit rather than a large one which, in general, implies a series of fixed costs that are a cause of inefficiency);
- **Financing availability**: though often the convenience is that to buy a single large power and industrial units, in reality this does not happen because you do not have the capital to do so. It would be appropriate to try in every way to find these resources, to achieve what would be the optimum configuration for your operations, not settle for suboptimal. It is therefore clear that, by this account, the financial aspect and the aspect of operations are often closely intertwined: it is to figure out how to reconcile both needs.
- **Over/underutilization costs**: use filling products in under saturation
- **Outsourcing possibility**

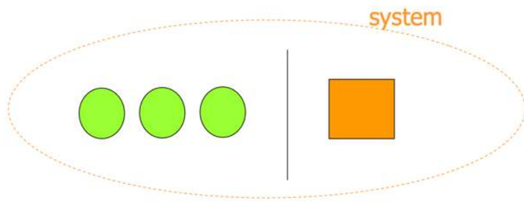
9. QUEUE MANAGEMENT

QUEUE THEORY 1

The **queue management** is another way in which companies can manage their **internal capacity**. Through queue management, companies can understand **criticalities** and where they are concentrated in order to provide a better service. Modelling through queuing theory supports companies in order take better operations decisions for example, they can manage how many row materials they need to have by understand better their system.

It studies the **waiting times of customers/products**: we are able to understand where our customers within our system are. We want to reduce at minimum the time that our customers are spending in our queue, so we need to manage our system in order to avoid them.





The line where there are people waiting is **queue**.
The queue + server = **system**.

9.1. WHAT IS A QUEUING SYSTEM FORMED OF?

A queuing system is formed by one or more **customers** waiting to be served by one ormore **servers**:
Customers' examples:

- People waiting at the cashier's desk for a bank operation;
- Pieces waiting to be processed by a lathe in a job-shop;
- Container waiting to be loaded on chassis.

Corresponding servers:

- Cashiers;
- Job-shop's lathes;
- Overhead traveling cranes' array for the loading.

Example:
Bank historical data show that on Monday there are on average 12 customers/hour at he only counter; the counter operator is able to serve 10 customers/hour on average. Is a queue going to form? In this case a queue is going to form, our services are overloaded.
Mu>lambda

Example 2:
Bank historical data show that on Monday between 10 and 11 am there are on average 10 customers/hour (1 customer every 6 minutes on average) at the only counter. The counter operator is able to serve 12 customers/hour on average (1 customer every 5 minutes on average).

Is a queue going to form? In this case it depends because we do not know when they are going to arrive, so we need to know in advance that it is not taken for granted that every customer is coming every 6 minutes, so the distribution can be all the time different, it might happen that the queue is going to form.
Mu<lambda

In a perfect stable context, if mu<lambda, there will never be a queue because there is not variability. Actually, there is variability, so there may form a queue.

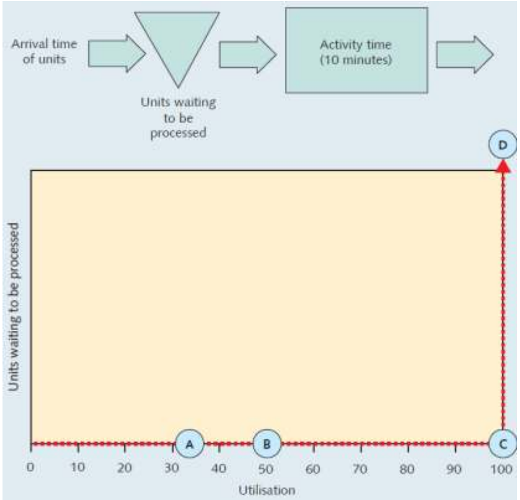
Queues form due to a not perfect balance between demand rate and service rate, it might happen that we cannot manage the customers, this is affected by two main balances:

- **Structural imbalances**: it characterizes a system whenever the service rate is always lower than arrival rate, so we know that due to our low number of servers available, we are not going to deploy the server in the right way

.

- Incidental imbalances: which are characterized by:
    - Variability: difference between the actual and the average value
    - Uncertainty: it is linked by the probability, estimated through statistics.

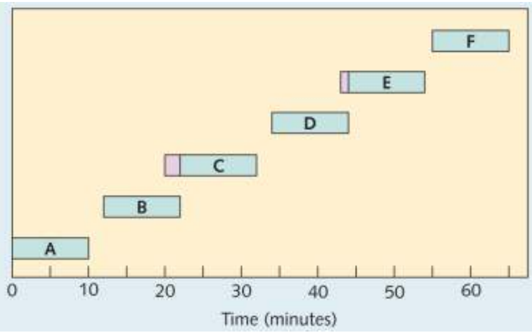
Example: **Structural imbalance** (case of no variability with μ = 10 min)
A: customer arrives every 30 min → 33.3% of saturation
B: customer arrives every 20 min → 50% f saturation
C: customer arrives every 10 min → 100% of saturation
D: customer arrives every <10 min → more than 100%



In the Y we have units waiting to be processed, X is utilization. Here (point A and B) we are not going to face problems because the arrival rate is lower than the service rate, but then if we reach 100% of the saturation, from C we jump directly to D. If service rate is lower than arrival rate, there will be infinite queue. In order to prevent this situation, we need to reduce our utilization rate (limit oursaturation).

Example: **incidental imbalance** (case of variability)
Each activity takes around 10 minutes. At the beginning with unit A we have no problem, the unit B arrives 2 minutes after, we spent 10 minutes for theactivity. The problem is that at time 20, we were there managing the activity for B so C has to wait for 2 minutes, and it is going to start its activity at 22 and it will finish at 32. D arrivesat 34 and the server starts immediately. We would have time to manage them, but we do not know when they arrive. Unfortunately, someone is going to wait for our service.

Unit	Arrival time	Start of activity	End of activity	Wait time
A	0	0	10	0
B	12	12	22	0
C	20	22	32	2
D	34	34	44	0
E	43	44	54	1
F	55	55	65	0



Usually, the tendency is to consider average values only (the one related to the variability). But there is always a variability that needs to be considered and managed because it provokes incidental imbalance, we need to understand how to manage the variability or how to manage the time of being in line. Forecast and historical data help, but they are not the truth.

The structural imbalance of a queuing system is that the average pace of customers' arrival into the system is greater than the average of the server's serving rate; in this sense, queues are a result, a symptom. Queues are not the problem, but a symptom that something is going wrong, they reflect the performances of the system.

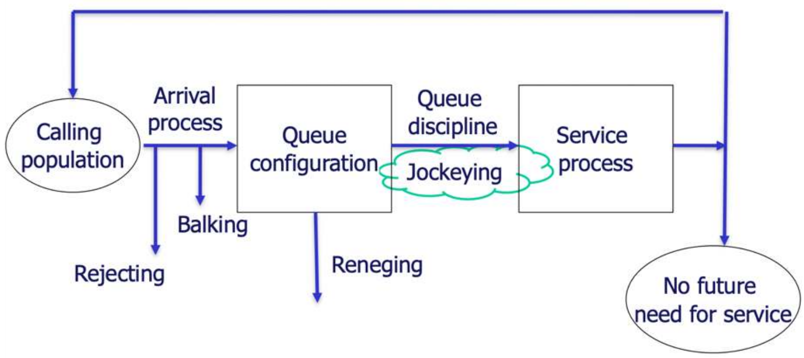
Example: waiting line versus stock

Characteristics	Stocks in a productive system	Waiting lines in services
Costs	Capital opportuniy cost	Time opportunity cost
Space	Warehouse	Waiting area
Quality	Infefficiencies index	Negative Impression
Decouplement	Allow to decouple the productive process phases	Allow job division and specialization
Utilization level	WIPs allow to keep machines always busy	Waiting customers keep the servers busy
Coordination	There's no need of a detailed scheduling	Allow to not precisely balance demand and offer

Regarding quality, the waiting causes a negative impression in customers' mind because they are losing time. Regarding coordination, having people in the queue allows to deploy the service but we cannot precisely balance demand and offer.

WHY queues are particularly critical in the **service companies**? Because pieces waiting do not create problems, while customers (specific of service companies), which are human being, create more problems related to psychological behaviours. If customers are not satisfied, there will be a negative word of mouth.

9.2. SYSTEM MODELLING



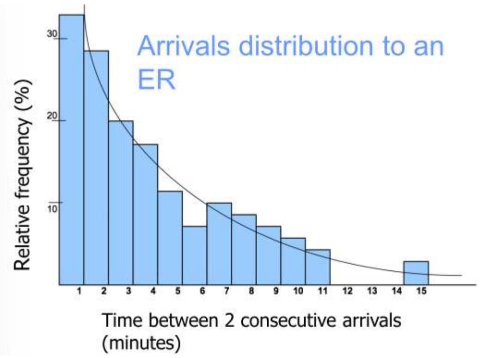
- **Calling population**: all the people that arrive to have our service, when they arrive they could behave in different ways.
- **Rejecting**: we decided to accept certain type of customer so we reject those who do not correspond to the one we wanted (for example, we do not accept people who did not take an appointment). If a piece arrives to a machine but it is not expected to be worked with that machine, it is rejected. It refers both the production and the service systems.
- **Balking**: customer does not want to wait so he decides to go away. After the arrival process, with the balking and rejecting people, we arrive to the queue configuration. It refers only to the service system (customers).
- **Queue configuration**: Where the queue actually forms.
- **Reneging**: When they wait for, they are annoyed so they decide to go away (it happens during the queue, you enter the queue but after 10 minutes in the queue you decide to go away).
- **Service process**: Moment in which we deliver the service. It this moment, depending on the queue that was formed before, people will decide whether to come back or not.
- **Jockeying**: when there is a queue less full you can switch from one to another: it is a way to manipulate the system to increase performances

9.3. POPULATION:

The customers' population is the input source, it can be:

- **Infinite**: if the number of clients in the system does not affect the demand pace for the service made by new customers; in this case people cannot be stocked, we do not know how many people are going to arrive.
- **Finite**: if the potential number of new customers for the system is significantly affected by the number of customers already in the system; we cannot manage an infinite population but just a certain maximum number of people.

The arrival process describes how customers show up. It is described by the distribution of interval times, which represents the time intervals occurring between two consecutive arrivals. Generally, data are given by collecting the actual arrivals time. Different empirical studies show that very often the distribution of inter-arrival times is well described by a negative exponential distribution.



The negative exponential distribution has a continuous probability density function as: $f(t) = \lambda * e^{-\lambda t}$

λ = arrival rate → average number of arrival per time unit

t = interarrival time → Time that intervals between 2 consecutive arrivals

Mean = $\frac{1}{\lambda}$

Variance = $\frac{1}{\lambda^2}$

The cumulative probability function, which describes the probability that the time between 2 consecutive arrivals is t or less than t, is: $F(t) = 1 - e^{-\lambda t}$

The mean and the variance characterize the negative exponential distribution.

EXAMPLE: The average inter-arrival time between 2 patients to the doctor is 2.4 minutes. Considering that a patient is just arrived, what is the probability that another will arrive inthe next 5 minutes?

$\lambda = 1/2.4 = 0.4167$ arrivals/minute

$F(5) = 1 - e^{-\lambda t} = 1 - e^{-0{,}416*5} = 1 - 0.124 = 0.876$

9.4. THE POISSON DISTRIBUTION:

The **Poisson distribution** is in a one-to-one relationship with the exponential distribution. When the inter-arrivals time are exponential, the number of events N(t) that takes place ina given time t is a Poisson process. In this case the focus not on the time but on the number of events, which are something that is discrete and not continuous.

The Poisson distribution gives the probability of n arrivals during an interval time t.

The Poisson distribution has a discrete probability function as:

$$f(n) = P(N(T) = n) = \frac{(\lambda t)^n * e^{-\lambda t}}{n!}, \quad n = 0,1,2,3,\ldots$$

- λ = arrival rate: average number of arrival per time unit. t = number of interest time periods (usually t=1).
- n = number of arrivals (0,1,2,3,...).
- Mean= t*λ
- Variance = t* λ

Example:
At a telephone exchange phone, calls arrive at a pace of 12 calls per hour. Knowing that the number of calls can be shaped with a Poisson distribution, answer the following questions.

What is the probability that 10 calls will arrive in the next hour?

λ=12calls/h n=10 calls t=1

$$P(10 \text{ in } 1 \text{ h}) = \frac{(\lambda t)^n e^{-\lambda t}}{n!} = \frac{(12*1)^{10} * e^{-12*1}}{10!} = 0.148$$

What is the probability that there will not be any call in the next 5 minutes?

λ=12 calls/h n=0 calls t=5 minutes

$$5\min * \frac{1}{60\frac{\min}{h}} = 0.083h$$

$$P(0 \text{ in} 0.083 \text{ h}) = \frac{(\lambda t)^n e^{-\lambda t}}{n!} = \frac{(12*0.083)^0 * e^{-12*0.083}}{0!} = 0.3679$$

What is the probability to have more than 2 calls in the next 10 minutes?

$$P(N>2)=1-P(N \leq 2)=1-[P(N=0)+P(N=1)+P(N=2)]$$

$$10\min * \frac{1}{60\frac{\min}{h}} = 0.167h$$

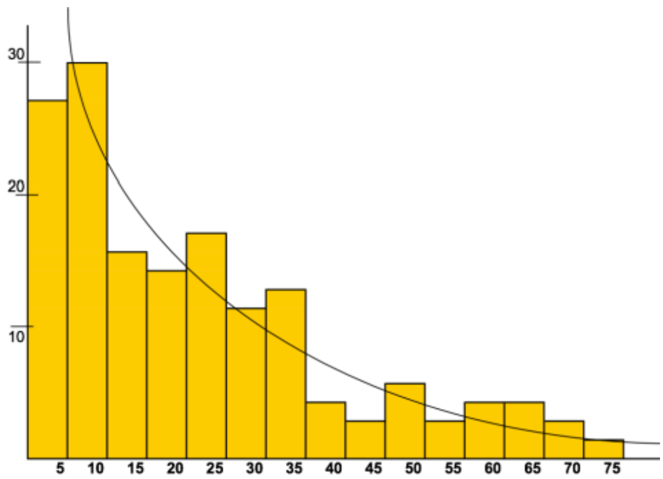
$$P(0 \text{ in } 0.167h) = \frac{(12*0.167)^0 * e^{-12*0.167}}{0!} = 0.135$$

P(1 in 0.167h)= 0.27013 P(2 in 0.167h)= 0.27067

P(N>2)=0.3242

9.5. THE SERVICE PROCESS:

The **service process** describes how each server delivers the service, in particular it definesits duration. It is defined in terms of service times distributions (also in this case we have a negative exponential distribution). Usually, data are given by collecting the actual service times. The service times distribution is often well represented by an exponential distribution.



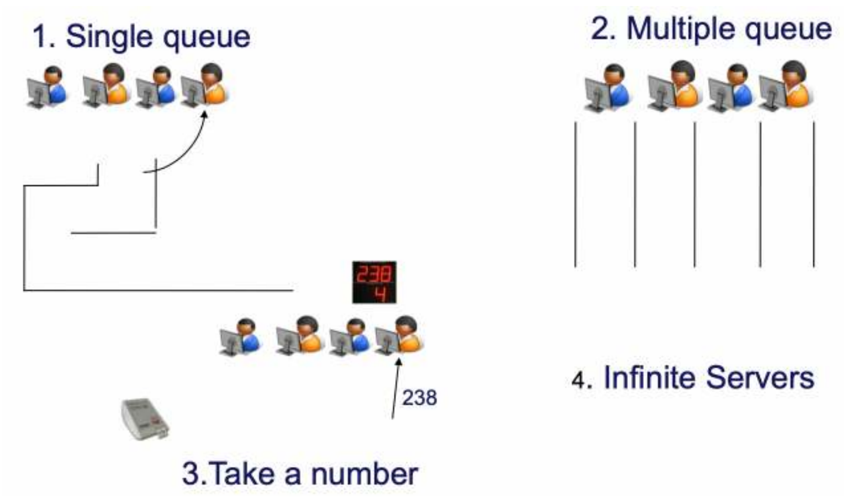
9.6. QUEUE CONFIGURATION:

- **Single queue**: Only one line, there are different servers but the line to reach them is just one.
  - Pros: FCFS; minimize the average waiting time, resulting in less frequent reneging actions; avoiding of the jockeying problem.
  - Cons: There can be the balking because the customer is scared

- **Multiple queue**: the different server can be different in terms of time → ex: in the majority of supermarkets each server has its own queue, so you decide in which queue you want to stay.
  - Pros: diversifying the work, possibility to choose the best server; allows jockeying.
  - Cons: the average waiting time is higher than single configuration

- **Take a number**: the same concept of single queue, you take a ticket and wait your number. In addition to that, you know how many people are before you.
  - Pros: FCFS; possibility to forecast the waiting time by measuring the number of people before me; no anxiety for the choice of server; the customer can relax while waiting for his turn.
  - Cons: It could scare the customer if he/she does not have anything else to do.

- **Infinite servers**: for example, the online chatbots, amazon → a lot of users simultaneously; it is often self-service, you are your own server.



The queue configuration should change according to the goals that we want to achieve:

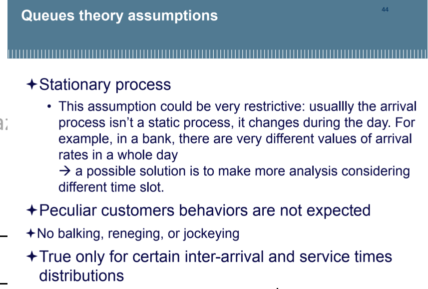
- High customization → it is better to opt for multiple servers to diversify the processes
- Service time high variability → it is better to opt for single queue to reduce it

Queue capacity:

- **Limited queue**: number of customers in queue + served is limited; if the capacity is overfilled, the new customers that arrive will be rejected.
- **Unlimited queue**: the number of customers in the queue plus the ones who era being served is potentially unlimited

Queue discipline: the rule according to which the customers are served, often related on the queue configuration. There are two types:

- **Static**: does not change over time or if the conditions change
- **Dynamic**: the discipline can change over time and/or at change of the system conditions



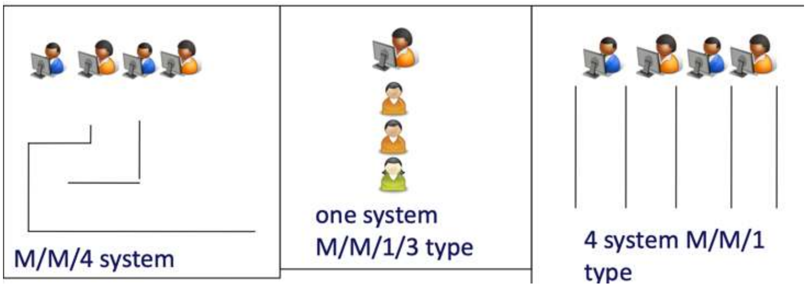
What is the best method used for queue systems analysis?

METHOD	PROS	CONS
Deterministic analysis	Simple and intuitive	Does not consider transitory Does not consider arrivals and service time variability Queue is ON-OFF
Queueing theory	Considers inter-arrivals and service times We can calculate significant variables	Does not consider transitory Very complex
Simulation	Considers transitory Very flexible	Time consuming Skills needed to be designed

9.7. KENDALL'S CODIFICATION

We can classify our system by using a certain codification that enables us to have someinformation, the so-called ***Kendall's codification***, which is characterized by letters:

- A = Identify customers arrival distribution;
- B = Identifies service times distribution;
- C = Identify number of servers;
- K = Identifies queue capacity (buffer by default is endless) → how many customers are inline.
- M = Identifies the population dimensions;
- Z = identifies service discipline (FIFO)

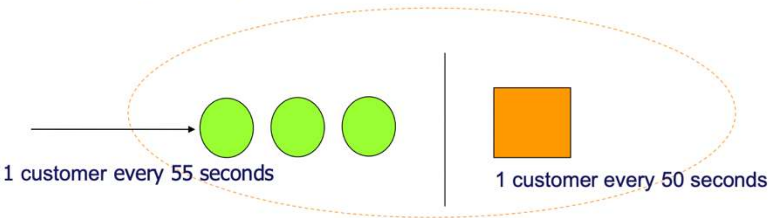


Example:

Pierre Cahon, barman of one of the most popular discotheques of St. Catherine Street, Montreal, serves cocktails at a pace of 1 every 50 seconds (1 customer every 50 seconds). Recently, during an evening when the club was particularly full, every 55 seconds someone was at the bar asking for a drink.

We need to shape the system: How to do that?

As we can see, on the left there is queue and right servers. Queue: 1 customer every 55 seconds
Server: 1 customer every 50 seconds



Which are the relevant parameters?

- λ = Arrival rate (customer/time unit)
- μ = Service rate (customer/time unit): How many customers can we manage?
- n = Number of customers in the system (waiting in line + serving e.g. at the cashier)
- L_q = Average number of customers in line
- W_q= Average waiting time in line by the customer
- L_S = Average number of customers in the system
- W_S = Average waiting time in the system by the customer
- P_n = Probability that there are n customers in the system
- P_o = It indicates the probability not to have any customer in the system.

With these parameters I must find the system utilization rate as:

System utilization rate: $\frac{\lambda}{\mu * s}$ → It is the coefficient that describes the utilization degree of the system, where:

- S = number of servers
- λ = arrivals rate
- μ = service rate

$\rho = \frac{\lambda}{\mu}$ → probability of being waiting in line

In the Pierre Cahon's example case:

- Assuming that Pierre is able to serve customers following a FCFS (first customer is the first one served) method, how much would you expect to wait before you will have your drink in hand (Ws)?
- How many people would you expect to find in line on average?
- What is the probability that three or more people will be in the system?
- What is the probability that you can be served immediately without none in line?
- What is Pierre's saturation level? (How busy is he?)

M/M/1 Queueing System:

P_n=probability that there are n customers in the system

$$P_0 = 1 - \rho \qquad P(n \geq k) = \rho^k \qquad P_n = P_0 \rho^n$$

L_s=average number of customers in the system: $L_s = \frac{\lambda}{\mu - \lambda}$

L_q=average number of customers in line: $L_q = \frac{\rho\lambda}{\mu - \lambda}$

W_s=average time waiting in the system by the customer: $W_s = \frac{1}{\mu - \lambda}$

W_q=average time waiting in line by the customer: $W_q = \frac{\rho}{\mu - \lambda}$

$$W_q + \frac{1}{\mu} = \frac{\rho}{\mu - \lambda} + \frac{1}{\mu} = \frac{\lambda}{\mu(\mu - \lambda)} + \frac{\mu - \lambda}{\mu(\mu - \lambda)} = \frac{\mu}{\mu(\mu - \lambda)} - \frac{1}{\mu - \lambda}$$

LITTLE'S LAW → $L = \lambda \cdot W$

$L_s = \lambda \cdot W_s$

$L_q = \lambda \cdot W_q$

$$W_s = W_q + t_{processing} = W_q + \frac{1}{\mu} = \frac{\rho}{\mu - \lambda} + \frac{1}{\mu} = \frac{\lambda/\mu}{\mu - \lambda} + \frac{1}{\mu}$$

Ws: the time spent in the queue and the time to be served → Ws = Wq + tprocessing
If $\lambda > \mu$: structurally imbalanced

Solution:

λ = 60 / 55 = 1.09 customers/minute
μ = 60 / 50 = 1.2 customers/minute

a) Assuming that Pierre is able to serve customers following a FCFS method, how much would you expect to wait before you'll have your drink in hand (Ws)?

$Ws = [(1 / (\mu - \lambda)] = 9.09$ minutes

b) How many people would you expect to find in line on average?

$Lq = [(\rho * \lambda) * (1 / (\mu - \lambda)] = 9.0$ customers

c) What is the probability that three or more people will be in the system?

$P (n \geq 3) = (\lambda / \mu)^3 = 0.75$

d) What is the probability that you can be served immediately without none in line?

$P_0 = 1 - (\lambda / \mu) = 0.092$

e) What is Pierre's saturation level? (how busy is he?)

$\lambda / \mu = 0.908$

QUEUE THEORY 2

Example 2:

A consultant is searching for a way to understand the exact number of counters to leave open in a new branch bank, in order to be able to offer a good service to the customer. How many counters should be left open in the rush hour if you want to keep the waiting time (Wq) in line below 15 minutes? Each counter will have its own customers' queue in front of itself.

The consultant, from previous surveys, knows that during the rush hour there are on average 100 customers per hour and that each server at the counter takes on average 5 minutes to be able to serve each customer (arrival and service distribution are negative exponential). MM1

Objective: Understand how many servers we need to have in order to manage our customer, we want to have a waiting line below 15 minutes (during rush hours).

DATA:

λ	100 customer/hour
μ	12 customer/hour
λ/μ	8,33

$$\frac{\lambda}{c} = \frac{100}{c} \qquad \frac{100}{c} < \mu$$

$$\frac{100}{\mu \cdot c} < 1$$

$$\rightarrow \quad c \geq 9$$

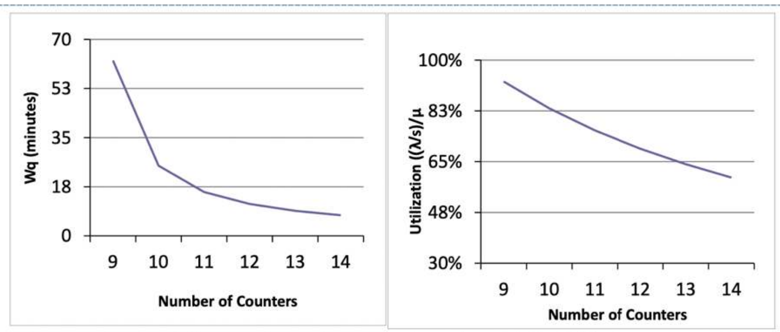
The λ of the single queue is given by the λ of the system, divided by the number of servers.

Which preliminary indications can we give to the consultant based on these data?
We need to have a minimum number of servers which must be greater than 8,33.
This means at least 9 servers.

	Number of counters					
	9	10	11	12	13	14
(λ/s)/μ	92,59%	83,33%	75,76%	69,44%	64,10%	59,5%
Wq (min)	62,5	25	15,625	11,36	8,93	7,35
Ws (min)	67,5	30	20,625	16,36	13,93	12,35

11,36 min < 15 min

As it is shown on the table, we need to look at the number of counters, and we see that we achieve only 12 customers.

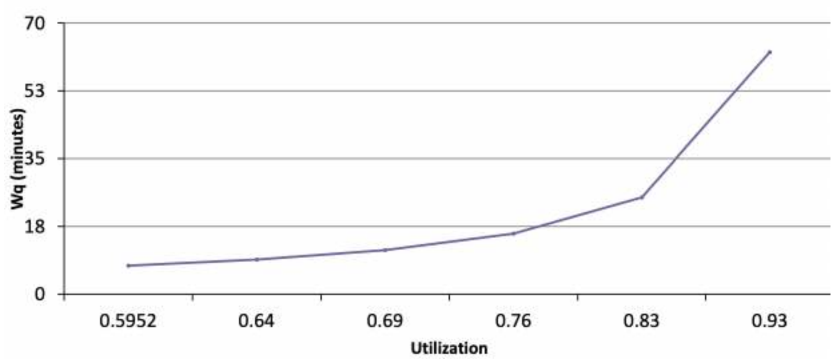


Increasing by 1 the number of servers from 9 to 10 (+11%) decreases the throughput time of 60%!!

This picture is showing that, in fact, the reduction of the waiting time is not linear at all. In reality, if we increase the number of counter of just 1 more, we reduce the waiting time not in a linear way. This is also showing the opposite: If I increase my saturation, I will increase my waiting time by eliminating one server that will have a really strong impact on the system.

In other words, the **trend is not linear**:

Waiting time vs. saturation:



The waiting time does not grow in a linear way with the increase of the system saturation time.
Even with a system utilization rate rather low (around 60%) there will be waiting times.

The queue theory is addressed because of:

- System sizing and designing (ex. consultant who needs to understand how manycounters at least have to be left open);
- Evaluation of costs and gains to improve the service (ex. barman who tries to understand how to make his customers more satisfied):
    - Queue length (Long lines suggest a poor customer service and cause congestion and dissatisfaction);
    - Average waiting time in line (Long waits are linked to a poor service);
    - Average waiting time in the system (it can suggest problems with customers,servers' efficiency or capacity);
    - System utilization rate (Check costs without unacceptable reductions in services)

9.8. TWO ACHIEVABLE GOALS WHEN SIZING A SYSTEM:

1. SERVICE QUALITY PRINCIPLES

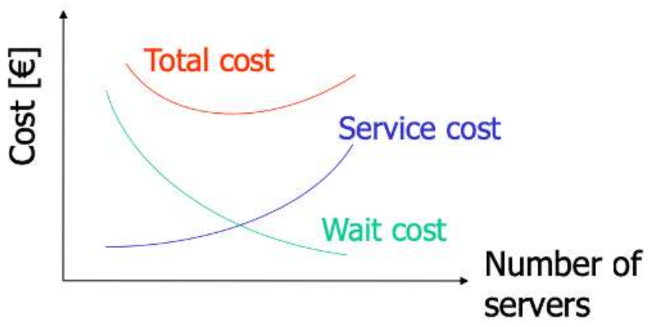
We need to look at the service quality principles, so we can manage our capacity by setting a quality level.

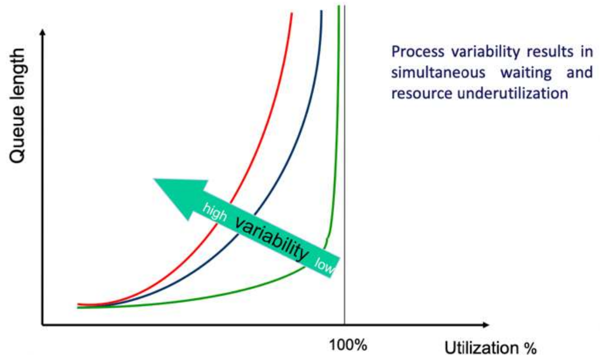
a. Average waiting time of the customer (banks, restaurants):
   $E(W^*)<\beta$: (beta is a certain value expressed by us)
   Example: sizing with the goal that the expected value of the time that the customer has to wait before sitting at a table is less than 10 minutes.

b. Maximum waiting time (public services):
   $P(W^*>w) \leq \alpha$
   The probability that the maximum waiting time is greater than a certain value (W)must be lower than a certain value (alpha). We cannot exceed the waiting time.
   Example: sizing with the goal of waiting that the probability that the client has to wait more than 5 minutes before receiving the service is less than 5%.

2. COST PRINCIPLES:

On the other side we also have cost principles, so we know that, for sure, the service is an opportunity for our image and our brand, so we try to satisfy costumers as maximum as possible. The problem is that this creates cost, so we need to find the right trade-off; because we need to estimate the service cost and on the other side the cost of making customer waiting in line (the waiting cost); the last one increases when the waiting rate is too high.





We also have to think about the variability: If I know that it is the distribution of our customers is always the same (so they arrive always at the same time), the variability is low and we can decide to have a higher utilization rate.

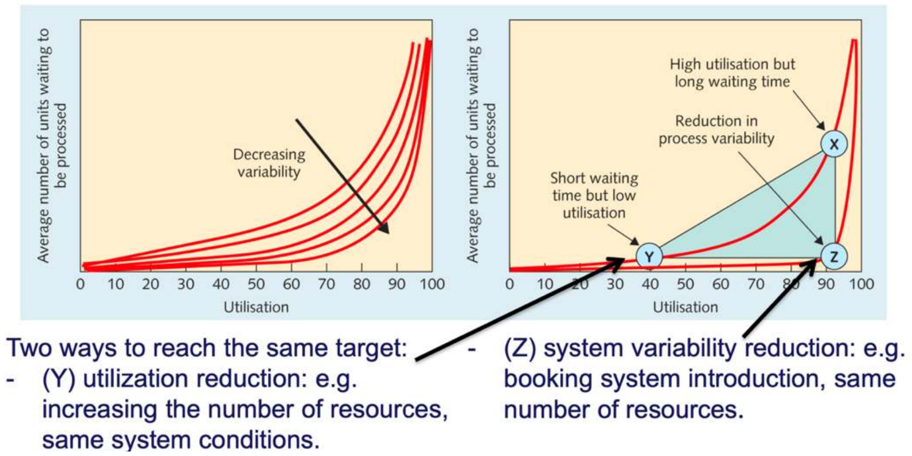
On the contrary, if the variability is too high, we need to change our capacity and add at least one server.

The higher is the variability, the lower must be the utilization rate and so the higher must be the server in our service system.

The graph is showing on the Y asses the average unit waiting to be processed and on the X asses we have the utilization.

In order to move from the point X (where the waiting time is too much) we can act in two ways: in order to reduce the average number of unit waiting to be processed we can move towards Y or towards Z.

In case of Y I reduce my utilization, in case of Z I reduce the variability.



To sum up: We can either reduce the utilization (by adding new servers) in order to deploy 0 waiting time, either reduce the variability (e.g. introducing booking systems).

9.9. MANAGERS MODIFY QUEUEING SYSTEM IN ORDER TO GET BETTER PERFORMANCES

In order to manage the queue, the manager can use a set of levers:

1. **Levers and countermeasures-offer side**: In this case we have a service and we can change the configuration by adding or removing some resources, we could also decrease the service time (technology utilization or training), we could increase theflexibility of resources or decrease the service time variability.

2. **Levers and countermeasures-demand side**: In this case we can act on the demand side, for example we could decrease the uncertainty level (by putting the possibilityof booking or improving forecast), we could also decrease the variance (by putting incentives for non-rush hour moments).

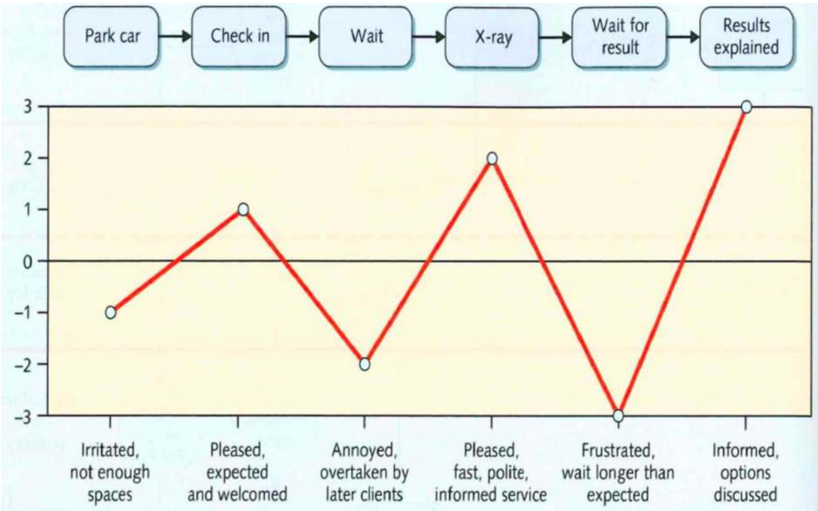
3. **Levers and countermeasures for soft aspects/related to the psychology of waiting**:In this case it depends on the customer expectation; on one side we could increase the satisfaction:

satisfaction= perception – expectation

The satisfaction level is important because when the customer is there, he might have a negative behaviour. If for the customer is the first time there, u need to give the right impression, if thefirst time you fail, you will have more problems in the future, so u need to make them come back.

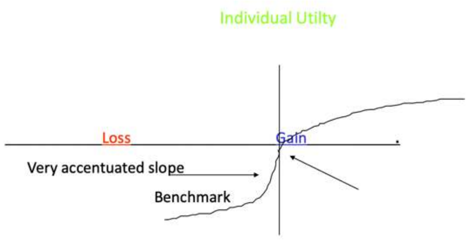
9.10. PSYCHOLOGY OF WAITING:

Measuring the customer satisfaction is important to understand where the critical point of the system is. This is a customer experience map of a visit for an X-ray investigation.



The **customer does not find immediately the parking**, then he goes to check in and the servers are quite good, so he is pleased, then between the check in and the X ray has to wait so he is annoyed, after everyone is polite and he is pleased. Even though we know how to manage the time where the customer is with us, we also need to think when the customer is not with us.

The **prospect theory** is explaining the idea that what is a gain for us has less value than a loss. You remember and you are much more influenced by failures, so this is what we have to think when we need to deliver a service.



Psychology of waiting should be managed and you need to think about countermeasures: for example, the magazines. The more the service is of value or the more the customer senses its utility, the more he/she is willing to wait. Waiting alone seems longer than waiting in a group

QUEUE MANAGEMENT

EXERCISE 1

In the first morning hours customers arrive at the post office where there is a *single take a number queue* (1 single queue but service delivered by different types of servers), at an average pace of 54 customers/hour (Poisson) while each server can manage to complete a service in an average time of 4 minutes (negative exponential)

DATA:

- Poissonian distribution, λ = 54 customer/hour,
- Service rate equal to 4 minutes (described by a negative exponential distribution).

QUESTIONS:

1. If there are 6 servers at the counter: what is the average number of customers in the system, the average waiting time in line and in the system? **(Ls, Wq, Ws).**
2. Determine the smallest number of counters that need to be opened to keep the average time in the system lower than 10 minutes (**Ws < 10min**).
3. If an employee costs 30 €/h and customer's waiting time stands for a cost of 20 €/h for customer, what would be the optimal number of servers according to purely economic considerations?

1. DATA:

We need to map our system and understand how it is composed of:
It is shaped as an M/M/6, it is a Marcovian distribution, 6 is the number of servers.
Situation can be shaped as a M/M/6 system, where:
λ= 54 customers/hour.
μ= 1 customer every 4 minutes = (1*60)/4 = 15 customers/hour.

<u>SOLUTION:</u>

ρ = λ/μ = 54/15 = 3.6 (arrival rate/service rate) → We do not consider S because we are not in a multiple queue, we have a single queue.

In order to estimate Lq we have a **table**, I need to combine the ρ with the Lq in order to find the number of servers; as it is shown in the table we find that c= 6.

$$L_s = L_q + \frac{\lambda}{\mu} = L_q + \rho$$

$$L_q = \left[\frac{\lambda}{\mu}; c\right]$$

Lq = 0,295 (average number of customers in line)

Ls = Lq + ρ = 0,295 + 3,6 = 3,895 customers (average number of customers in the system)

Wq = Lq / λ = 0,00546 hours → around 20 seconds (19,67 seconds)

Ws = Lq / λ + 1 / μ = 0,072 hours → around 260 seconds (259,67 seconds) → I find Ws in this way because we are in M/M/c

Lq results of model M/M/c

λ/μ	c=1	c=2	c=3	c=4	c=5	c=6	c=7	c=8
0,15	0,026	0,001						
0,20	0,050	0,002						
0,25	0,083	0,004						
0,30	0,129	0,007						
0,35	0,188	0,011						
0,40	0,267	0,017						
0,45	0,368	0,024	0,002					
0,50	0,500	0,033	0,003					
0,55	0,672	0,045	0,004					
0,60	0,900	0,059	0,006					
0,65	1,207	0,077	0,008					
0,70	1,633	0,098	0,011					
0,75	2,250	0,123	0,015					
0,80	3,200	0,152	0,019					
0,85	4,817	0,187	0,024	0,003				
0,90	8,100	0,229	0,030	0,004				
0,95	18,050	0,277	0,037	0,005				
1,0		0,333	0,045	0,007				
1,1		0,477	0,066	0,011				
1,2		0,675	0,094	0,016	0,003			
1,3		0,951	0,130	0,023	0,004			
1,4		1,345	0,177	0,032	0,006			
1,5		1,929	0,237	0,045	0,009			
1,6		2,844	0,313	0,060	0,012			
1,7		4,426	0,409	0,080	0,017			
1,8		7,674	0,532	0,105	0,023			
1,9		17,587	0,688	0,136	0,030	0,007		
2,0			0,889	0,174	0,040	0,009		
2,1			1,149	0,220	0,052	0,012		
2,2			1,491	0,277	0,066	0,016		
2,3			1,951	0,346	0,084	0,021		
2,4			2,589	0,431	0,105	0,027	0,007	
2,5			3,511	0,533	0,130	0,034	0,009	
2,6			4,933	0,658	0,161	0,043	0,011	
2,7			7,354	0,811	0,198	0,053	0,014	
2,8			12,273	1,000	0,241	0,066	0,018	
2,9			27,193	1,234	0,293	0,081	0,023	
3,0				1,528	0,354	0,099	0,028	0,008
3,1				1,902	0,427	0,120	0,035	0,010
3,2				2,386	0,513	0,145	0,043	0,012
3,3				3,027	0,615	0,174	0,052	0,015
3,4				3,906	0,737	0,209	0,063	0,019
3,5				5,165	0,882	0,248	0,076	0,023
3,6				7,090	1,055	0,295	0,091	0,028

DATA:
Situation can be shaped as a M/M/6 system, where:
λ= 54 customers/hour
μ= 1 customer every 4 minutes =15 customers/hour.

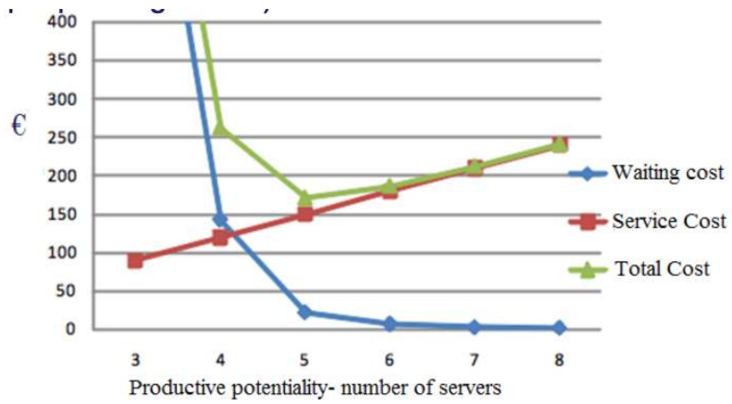
Goal:
Ws < 10 min having the smallest number of open counters, in order to calculate it we need to understand how many servers we need.

SOLUTION:
First, we need to think about how we should estimate Ws, that is still related to Lq, in fact, Lq in linked to the number of servers:
Ws= (Lq / λ) + (1 / μ) (when c varies, so does Lq!)
We do not know Lq, but we know the arrival rate that is 54 customers/hour and we know the service rate, which is 15 customers/hour, so we know that we must have an Lq lower than 5.4:
10 min / (60min/hour) > Lq / 54 + 1 / 15

Lq < 5,4 with ρ=3,6 → looking at the table, it results that c ≥ 5 server (because for c=4, Lq)

2. DATA:
Calculate the optimal number of servers that minimizes service cost and wait cost. In this case we need to balance our cost and identify the minimum cost, so we need to manage on one side the number of servers that we need and on the other side the problems related to the service delivery to the customers.

SOLUTION:
Server cost: 30 €/h*c
Wait cost: 20 €/h*Lq (with Lq depending on c)
So we need to estimate c in order to have the minimum cost. The equation in the graph depends on c.
So: the total cost is given by theminimum of the two curves, which is: ***minimum 30*c + 20*Lq***.

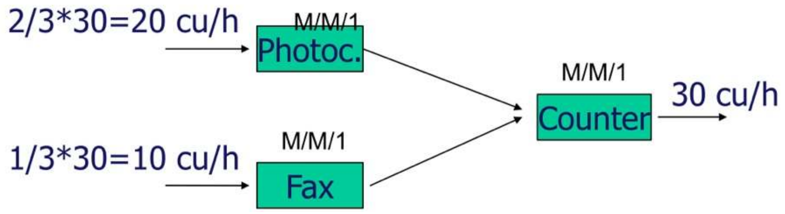


EXERCISE 2

At a photocopier shop, customers usually enter to make copies or send a fax. Once they have finished, they go to the cash counter to pay. The customers' arrival rate is 30 cu/h. Of those, 2/3 enter to make copies while 1/3 to send fax. The arrival rates are distributed like a Poissonian. On average a customer takes 2 minutes at the photocopier; 4 minutes at the fax; 1 minute to pay.

Supposing that service rates are distributed like a negative exponential:

1. What is the probability that the shop is empty?
2. What is the system average throughput time?
3. How many people on average are in line at the cash counter?
4. How many customers on average are at the shop?
5. What is the probability to have less than 6 people in line at the photocopier?
6. What is the probability to have 1 customer at the shop?



1. What is the probability that the shop is empty?

System shaping:

	Photocopier	fax	Cash counter
λ	30 cu/h * 2/3 = 20 **cu/h**	30 cu/h * 1/3= 10 **cu/h**	20 cu/h + 10 cu/h = 30 **cu/h**
μ	1/2 cu/min * 60 min/h = 30 **cu/h**	1/4 cu/min* 60 min/h = 15 **cu/h**	1 cu/min * 60 min/h = 60 **cu/h**

So, the probability that the shop is empty:

$$Probability = \boldsymbol{P} = \boldsymbol{P_{0photocopier}} * \boldsymbol{P_{0fax}} * \boldsymbol{P_{0cashcounter}}$$

$$\boldsymbol{P_0 = 1 - \rho = 1 - \frac{\lambda}{\mu}}$$

Photocopier	1 - 20/30=0.333333
Fax	1 - 10/15=0.333333
Cash counter	1 - 30/60=0.5

→ Probability = 5,5 %

2. Average system throughput time (Ws):

- Photocopier throughput time Ws = 1/(μ - λ) = 1/(30-20) = 1/10 hours = 6 minutes
- Fax throughput time Ws = 1/(μ - λ) = 1/(15-10) = 1/5 hours = 12 minutes
- Cash counter throughput time Ws = 1/(μ - λ) = 1/(60-30) = 1/30 hours = 2 minutes

System expected throughput time: weighted sum of all types of customers' throughput times (There are two types of customers):

- Photocopier customers → cash counter 20 / 30 customers (2/3)
- Fax customers → cash counter 10 / 30 customers (1/3)

20/30 * (6 min + 2 min) + 10/30 * (12 min + 2 min) = 10 minutes (this is the **Average Waiting Time**)
→ total amount of throughput time.

3. Average people in line at the cash counter:

Lq= ρ*λ / (μ - λ) = (30/60 * 30)/ (60-30) = 0,5 customers

4. Average number of customers at the shop:

We need to consider the average number of customers at the photocopier plus the average number of customers at the fax machine plus the average number of customers at the cash counter.
Ls shop = Ls photo + Ls fax + Ls cash counter Ls= λ / (μ - λ)
So, for each single system we have to estimate the Ls:

	Photocopier	fax	cashcounter	total
Ls	20/(30-20)= **2**	10/(15-10)= **2**	30/(60-30)= **1**	2+2+1= **5**

Total Ls: 5 -> average number of customers at the shop, because it was computed with thesum of the average number of customers in each system.

5. Probability to have less than 6 customers in line at the photocopier:

Probability to have less than k customers in the system where the system is line + service; because of this reason, we need to add the 6 customers in line for the photocopier plus the 1 that is at the server. We have the specific formula to calculate this probability:

$$\boldsymbol{P(n < k) = 1 - P(n \geq k) = 1 - \rho^k}$$

Probability to have less than 6 customers in the system:
OSS: [IN THE SYSTEM = LINE + IN SERVICE]

$$\boldsymbol{P(n < 7) = 1 - P(n \geq 7) = 1 - \rho^7}$$

[6 customers in line + 1 in service]

$$\boldsymbol{P(n < 7) = 1 - \rho^7 = 1 - \frac{20}{30}^7 = 94,15\%}$$

The probability to have less than 6 customers in line is 94.15%. 94.15% of time there are less than 6 people waiting in the system.

6. What is the probability to have 1 customer at the shop?

This means that we must have at least one customer in each system:
Or 1 customer at the photocopier, or 1 in the fax or 1 in the cash counter.

$$\boldsymbol{P = P_{1photo} * P_{0fax} * P_{0cashcounter} + P_{0photo} * P_{1fax} * P_{0cashcounter} + P_{0photo} * P_{0fax} * P_{1cashcounter}}$$

$$\boldsymbol{P_1 = P_0 * \rho^n}$$

$$\boldsymbol{P_0 = 1 - \rho = 1 - \frac{\lambda}{\mu}}$$

→ Probability is 10.18%

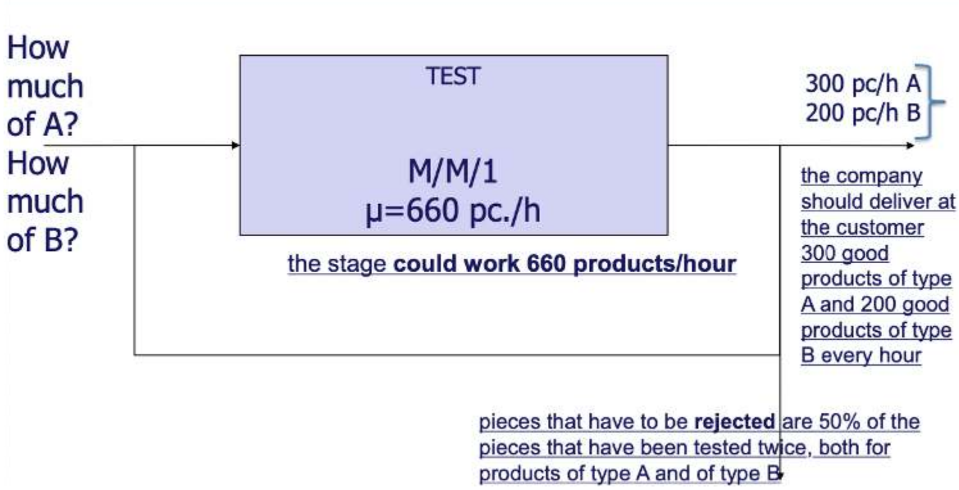
	Photocopier	fax	cashcounter
P_1	0.2222222	0.2222222	0,25
P_0	0.333333	0.333333	0.5

Some messages from queuing system shaping:

- A complex system is the composition of more elemental sub-systems that interact among themselves (we need to map the system and identifying every step of the process);
- Very often in a complex system transit different type of customers (we need to know in advance every types of customers of one own system);
- The system expected throughput time is a representation of an "**average**" customer and not of the specific one!
- The throughput time varies according to the variation of the workload in the systems and at the variation of the input.

EXERCISE 3

The test phase of a stage is a stage where enter, without different priority logics, product A and B that form the productive selection of the company. In particular, the testing machine is not of the latest generation. Indeed, every once in a while, it could signal as nonstandard even products that actually are good. For this reason, in case nonstandard products are signalled, they are tested once more. On average, 15% of products A and 15% of products B that the machine tests have to be tested again. The pieces that have to be rejected are 50% of the pieces that have been tested twice, both for products of type A and of type B. The rest of the products instead are good and can be delivered to the customer. In case the stage could work 660 products/hour and in case the company should deliver at the customer 300 good products of type A and 200 good products of type B every hour, how much material is it necessary to be put in the test phase every hour? Shape the system reporting all the significant parameters.



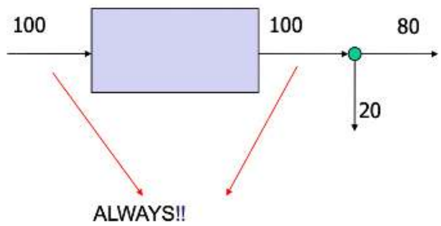
DATA:

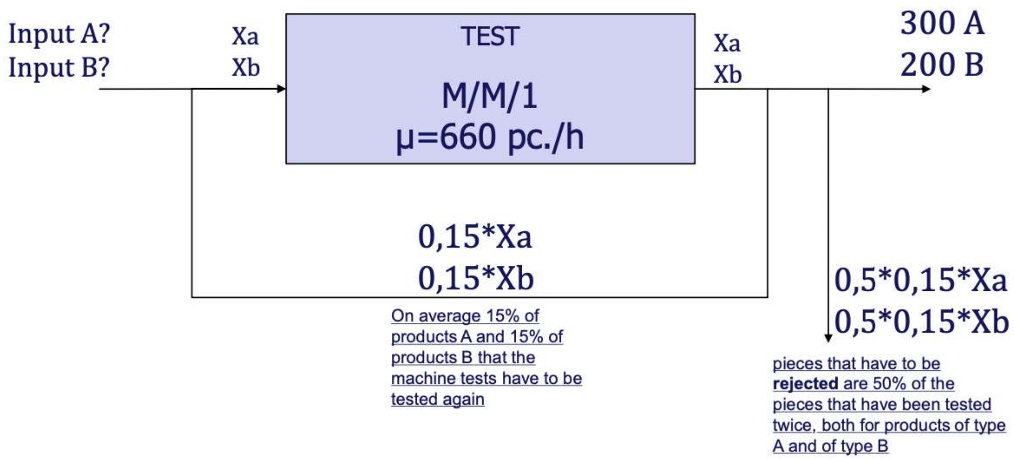
We must deliver a certain amount of products to our customer. We need to deliver 300 products of A per hour and 200 of products of B per hours. Our machine has a service rate of 660 pieces/hour and we know that sometimes some of the products have to be reworked and some other must be rejected.
We need to estimate how much of raw material for A and B we need; to do so, we must look at the NODES. The nodes are the point in which there are flows of materials that enter and exit the system and the sum of all the flows that enter and exit must be always equal. This means that what enters must always be equal to what exit in the system. Node balancing is fundamental because we base our calculations on this theory. Here we have a system and as they enter the node, the sum of what enters must be equal to what exists.

Coming back to our system, we must estimate the input of A and input of B: The testing machine is composed of 1 server (M/M/1), the service rate is 660 pc/hour and part of X and Xb has to be worked and tested again (15% of them), moreover we know that pieces that have been rejected are the 50% of the pieces that have been tested twice. So, 15% of the products must be tested again and, in addition, those pieces that have been tested twice, the 50% in discarded.

Node Balancing:
Everything that enters in the node is equal to everything that comes out from it.
Just like Energy, nothing is created or destroyed!!

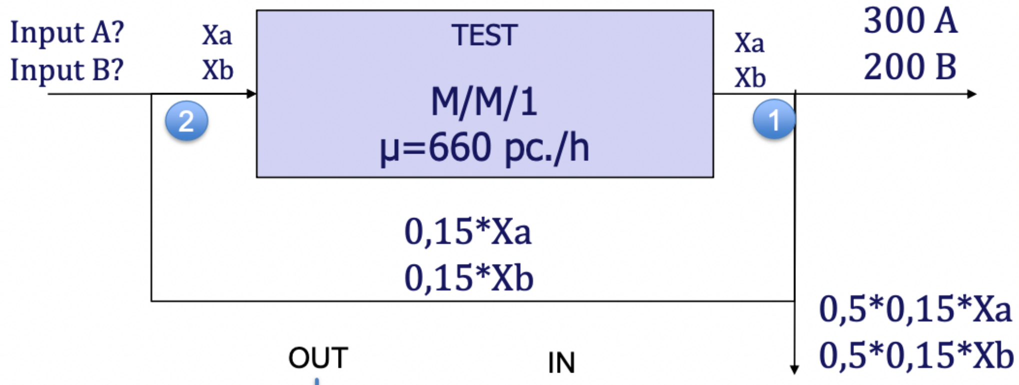
The flow in the system (line + server)
Everything that enters in the system is equal to everything that comes out. In the system there is not any flow loss. (The service process works on customers. With neither creating nor destroying any!)





SOLUTION:
In order to estimate the input A and input B (Xa and Xb) we need to do the node balancing: every time we have more than 1 product, we must focus at one single products per time and after the other.

NODE 1:



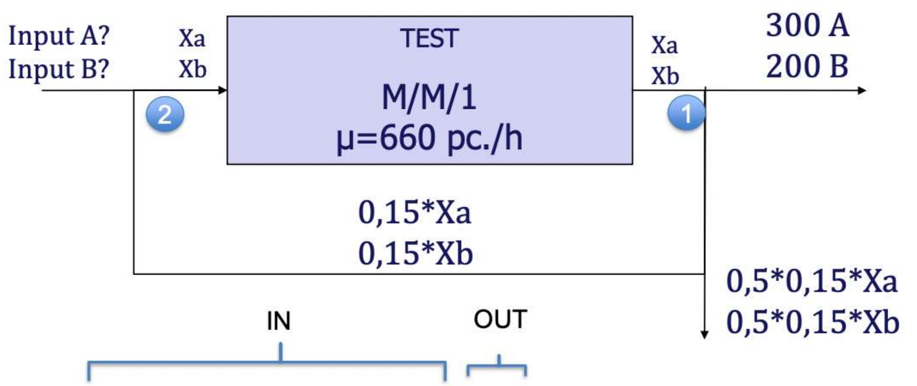
1

A) $\underbrace{300+0{,}15Xa+0{,}5*0{,}15Xa}_{\text{OUT}}=\underbrace{Xa}_{\text{IN}} \rightarrow Xa=387{,}09\ pc/h$

B) $200+0{,}15Xb+0{,}5*0{,}15Xb=Xb \rightarrow Xb=258{,}06\ pc/h$

NODE 2:

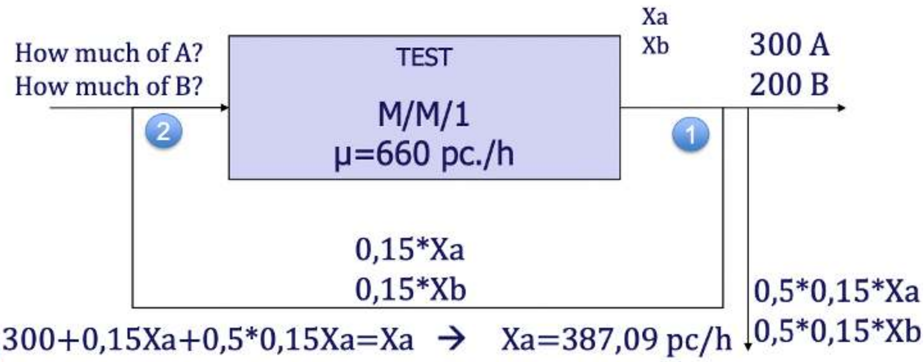
It is the one that is calculating my input.



2 Input A: input A + 0,15Xa=Xa → input A = 329,02 pc/h

Input B: input B + 0,15Xb=Xb → input B = 219,35 pc/h

TO SUM UP:



1 300+0,15Xa+0,5*0,15Xa=Xa → Xa=387,09 pc/h
200+0,15Xb+0,5*0,15Xb=Xb → Xb=258,06 pc/h

2 Input A: input A + 0,15Xa=Xa → input A = 329,02 pc/h
Input B: input B + 0,15Xb=Xb → input B = 219,35 pc/h

QUEUE THEORY 3

EXERCISE 1: RINASCIMENTO CLINIC

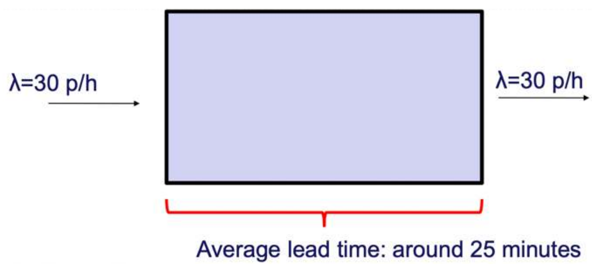
"Rinascimento" clinic is a small health care facility, located in Austin, Texas.
The medical examinations take place without an appointment and the flow is described as follows: a patient enters the clinic and goes to reception, where the receptionist, according to the patient's problem, sends her to Dr. Carter, a general practitioner, or sends her to Dr. Romano, a specialized physician. In some cases, the patient could be seen by both Dr. Carter and Dr. Romano.
The collected data are depicted in the following table. Arrival rate and service rate follow a negative exponential distribution.

Variable	Average
Patients arrival rate	30 p/h
Patients that are sent to Dr. Carter (fraction of the total number entering in the system)	2/3
Patients that are seen by Dr. Carter and then by Dr. Romano (percentage of the number of patients that are sent to Dr. Carter)	0.15
Receptionist service rate	40 p/h
Dr. Carter service rate	30 p/h
Dr. Romano service rate	15 p/h

1. The priority rule is FCFS. You are required to calculate:
    a. The average waiting time in each queue.
    b. The average waiting time in the system for each patient typology.
    c. The expected LT (Lead time of a generic patient).
    d. The average time of inactivity (expressed in minutes each hour) for the receptionist, Dr. Carter, and Dr. Romano.
2. Rinascimento clinic is considering implementing a priority system to manage in a different way the visits held by Dr. Romano. The idea is that the patients going to Dr. Romano after being seen by Dr. Carter, have a non-pre-emptive priority on the patients that arrive directly to Dr. Romano. Which are, in your point of view, the main effects introducing this priority rule has on the LT (for the 3 typologies and the generic patient)?

The average waiting time in each queue.

In this system, of course, none is discarded, 30 patients are going in, 30 patients are going out.

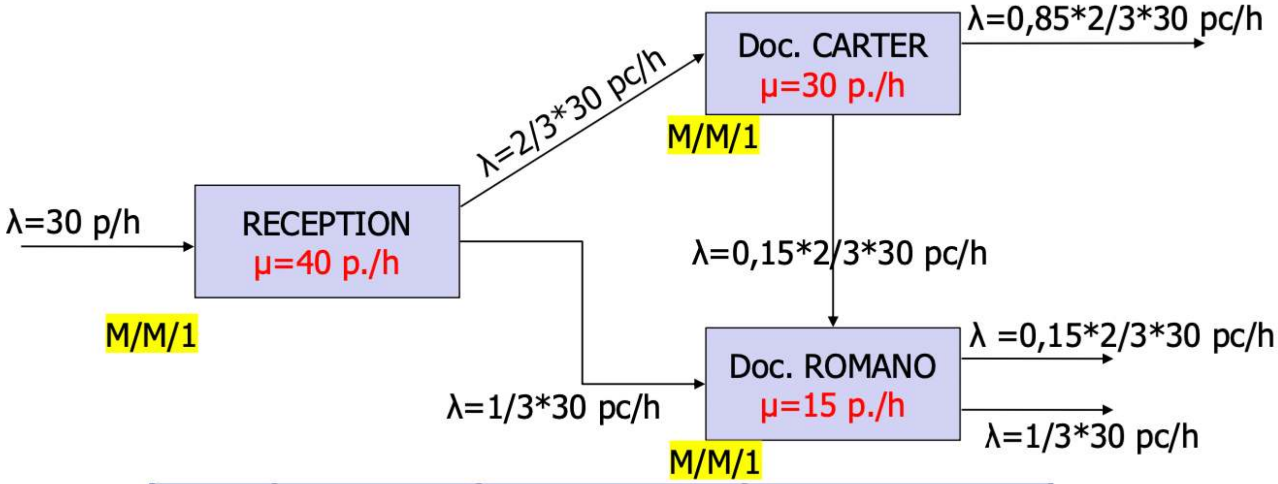


The **system average throughput time** is around 25 minutes. But the patients complain of a very variable time and that the reason behind it is not clear. Time varies from less than 5 minutes to more than 45!

Customers' complains risk damaging the center image and losing customers! So, for this reason, a specific analysis of the process is necessary to comprehend the causes of variability!

- Each process is characterized by a certain patients' arrival rate and a certain service capacity. Both parameters influence queue length and throughput times! It is important **know the system**
- The system is formed by different entities (different type of patients) who have different behaviours and paths. You have to **identify different typology / families**!

Mapping the system



	Reception	Doctor Carter	Doctor Romano
λ	30 p/h	20 p/h	10+3 =13 p/h
μ	40 p/h	30 p/h	15 p/h

The reception is composed by one single server (M/M/1); 2/3 of 30 patients will go to Dr. carter, but we also know that the 15% of this will be also sent to Dr. Romano, so:

- λcarter = 2/3 * 30 pc/h with M/M/1 system, then 15% of customers are sent by Dr. Roman, thus the λ is the sum of those patients that arrive directly from the reception and those that arrive from Dr. Carter.
- λRomano= 1,15*2/3*30.

Let us estimate the **Wq**, **Ws** and the **inactivity time** of each sub-system:

Wq= (λ/ μ)/(μ- λ)	**Ws=1/(μ- λ)**	**inact.t.=(1- ρ)*60 min/h**

	Reception	Doc. Carter	Doc. Romano
Wq	(30/40)/(40-30)*60= **4,5 min**	(20/30)/(30-20)*60= **4 min**	(13/15)/(15-13)*60= **26 min**
Ws	1/(40-30)*60= **6 min**	1/(30-20)*60= **6 min**	1/(15-13)*60 = **30 min**
Inactivity time	(1-(30/40))*60= **15 min**	(1-(20/30))*60= **20 min**	(1-(13/15))*60= **8 min**

Ws is given by the sum of Wq and service rate.
The inactivity time is in some same the opposite of the saturation, so the calculation is the opposite.
Even though there is a waiting time there is also inactivity time, and this is due to the variability and uncertainty, and in order to estimate the average waiting time in line for each of the three typologies we must estimate what types of behaviours they have within our system:
We have three types of customers/types of path in the system:

- Reception -> Dr. carter
- Reception -> Dr. Carter -> Dr. Romano
- Reception -> Doctor Romano

	Reception (Wq)	Doc. Carter (Wq)	Doc. Romano (Wq)	Wq
Type 1	4,5 min	4 min	-	4,5+ 4 = 8,5 min.
Type 2	4,5 min	4 min	26 min	4,5+4+26 = 34,5 min.
Type 3	4,5 min	-	26 min	4,5 + 26 = 30,5 min.

This is showing that, in order to estimate the average throughput time, we must understand the quantity of these patients, which is called the OCCURANCE ESTIMATION. We need to do occurrence estimation:

	Occurrences
Type 1	((2/3)*30) - (0,15*(2/3)*30))/ 30 = 0.5666
Type 2	(0,15*(2/3)*30) / 30 = 0.1
Type 3	((1/3)*30) / 30 = 0.3334

- TYPE 1: All the customers who enter the system minus the ones sent to Dr. Romano; this sum must be divided by the input in order to give us the percentage of the patients who enter the system, go to the reception and then just go to Dr. carter.
- TYPE 2: Sum of all the patients who entered in the reception, went to Dr. Carter and also to Dr. Romano. The sum will be divided by 30 in order to calculate the percentage.
- TYPE 3: Those that enter the reception and went to Cr. Romano.

Obviously, the sum of all these values must give 1.

Once we did this, we can calculate the expected throughput time of all the system.
We need to sum the Ws of each type of customers, that is then multiplied by the relative occurrence.

	Reception (Ws)	Carter (Ws)	Romano (Ws)	Ws	Occurence	Expected time
Type 1	6 min.	6 min.	-	**12 min.**	**0.5666**	12*0,5666 +42*0,1+3 6*0,3334=
Type 2	6 min.	6 min.	30 min.	**42 min.**	**0.1**	
Type 3	6 min.	-	30 min.	**36 min.**	**0.3334**	
EXPECTED THROUGHPUT TIME					**1**	**23 min.**

We have to multiply each Ws for its occurrence, we find the expected time for each type and summing them together we find the total throughput time of the system which is 23 minutes.

- The system expected throughput time refers to an "average product". It does not have to be taken as an absolute benchmark;
- It is crucial to know the system and to map it: which phases? Which products? Which problems?
- For each type of product, the variability is much lower: it is explained a part of the initial uncertainty;
- If it is possible to assign the type of product at the beginning of the process, the uncertainty decreases and it is possible to better align performances to expectations, therefore it is possible to start effective countermeasures.

For this reason in order to better provide the service could introduce the **priority rule**, which is the idea of giving priority to a certain type of customers (e.g. pregnant women).

M/M/1 system with assigned priority to a customers' class

It is a system formed by just one server who is able to serve two classes of customers (1 and 2) that have a different service admission

Hp: customers of class 1 have a service admission priority higher than customers of class 2

In this case we have one server and diverse types of priorities given to our customers.

The following cases will be considered:

- Pre-emptive priority
- Non pre-emptive priority

We might have two types of priorities:

1) **Pre-emptive priority**: In case a customer of class 1 enters the system, while the customer of class 2 is served, must receive the service by stopping the delivery service of customer 2.
   If a customer of class 1 enters the system while a customer of class 2 is being served, the service to the customer of class 2 is interrupted in order to give immediate precedence to the customer of class 1.

The following formulas allow us to calculate the system throughput time of the two classes of customers:

- $E(S_1) = \frac{1/\mu}{1-\rho_1}$
- $E(S_2) = \frac{1/\mu}{(1-\rho_1)*(1-\rho_1-\rho_2)}$

Example of pre-emptive priority:

$\lambda_1 = arrival\ rate\ of\ type\ 1\ customers = 0{,}2$

$\lambda_2 = arrival\ rate\ of\ type\ 2\ customers = 0{,}6$

$\mu = service\ rate = 1$

$E(S) = \frac{1}{1-0{,}8} = 5$

In case that customer 1 has absolute priority:

$\boldsymbol{E(S_1)} = \frac{1}{1-0{,}2} = 1{,}25 \qquad \boldsymbol{E(S_2)} = \frac{1}{(1-0{,}2)*(1-0{,}2-0{,}6)} = 6{,}25$

2) **Non-Pre-emptive priority**: If a customer of class 1 enters the system while a customer of class 2 is being served, the service to the customer of class 2 is terminated and then the customer of class 1 is taken care of.

Following formulas allow us to calculate the system throughput time of the two classes of customers

- $E(S_1) = \frac{(1+\rho_2)/\mu}{1-\rho_1}$
- $E(S_2) = \frac{(1-\rho_1(1-\rho_1-\rho_2))/\mu}{(1-\rho_1)*(1-\rho_1-\rho_2)}$

In case that customer 1 does not have absolute priority on customer 2

$\boldsymbol{E(S_1)} = \frac{1+0{,}6}{1-0{,}2} = 2 \qquad \boldsymbol{E(S_2)} = \frac{1-0{,}2(1-0{,}8)}{(1-0{,}2)*(1-0{,}2-0{,}6)} = 6$

2) Main effects of the non-pre-emptive priority:

Class with higher priority: customers arriving from Dr. Carter (S1)

Class with lower priority: customers arriving from the reception (S2)

- $\rho_1 = \frac{\lambda_1}{\mu} = \frac{3}{15} = 0{,}2$
- $\rho_2 = \frac{\lambda_2}{\mu} = \frac{10}{15} = 0{,}667$
- $E(S_1) = \frac{(1+\rho_2)/\mu}{1-\rho_1} = \frac{(1{,}667)/15}{1-0{,}2} = 0{,}1389h = 8{,}33\ minutes$
- $E(S_2) = \frac{(1-\rho_1(1-\rho_1-\rho_2))/\mu}{(1-\rho_1)*(1-\rho_1-\rho_2)} = 0{,}6083h = 36{,}5\ minutes$

	Reception (Ws)	Carter (Ws)	Romano (Ws)	Ws	OCCURRENCE	EXPECTED TIME
Type 1	6 min.	6 min.	-	**12 min.**	0.5666	
Type 2	6 min.	6 min.	8,33 min. (priority 1)	**20,33 min.**	0.1	
Type 3	6 min.	-	36,5 min. (priority 2)	**42,5 min.**	0.3334	
EXPECTED THROUGHPUT TIME					**1**	**23 min.**

The expected throughput time is the same as before because we did not change anything: We have the same quantity of resources and the same types of patients, the only difference is in the type of service we are going to provide to people.

The alteration of the priority logic does not change the system expected throughput time; however, in some cases it is essential to increase the customer's satisfaction degree and to improve customer service.

EXERCISE 2 – STERILISATION CENTRE

The two surgery blocks (S.B.) of the hospital "A.O. Curiamo Tutti" do not have a cleaning instruments zone. The surgery instruments kits are directly sent to the sterilization center, once used by different specialized departments.

To improve the service level offered to "A.O Curiamo Tutti", Eng. Smith is in charge of studying the sterilization system during the rush hours when all the 12 operating theatres of the hospital work simultaneously.

The specialized departments are as follows:

- Orthopaedics
- General Surgery
- Gynaecology
- Ophthalmology

The sterilization center is located on the same floor of S.B. At the sterilization center entrance, a nurse receives all the surgery instruments kits. She is in charge of recognizing and categorizing all the instruments kits, and of sending them to the right dedicated area.

In each area (in particular, there are 4 areas dedicated to the specialized departments: Orthopaedics, General Surgery, Gynaecology, and Ophthalmology) there are specialized nurses that are in charge of the cleaning phase. The instrument kits are allocated to each nurse according to the single queue configuration. Once the cleaning phase ended, the instrument kits are sent to 2 specialized operators that are in charge of preparing the material before putting that onto the sterilization's machines. The two operators work in parallel, according to the multiple queue configuration.

Eng. Smith collected the following data:

- During rush hours, two diverse types of surgery kits arrive, some classified as "urgent" and some others as "not urgent". The arrival rate of the urgent kit is 5 kits/hour, and they have pre-emptive priority on the non-urgent kits. In particular, non-urgent kits have an arrival rate of 10 kits/hour. The nurse can recognize and categorize one kit every 2 minutes. Only in this stage, there is a distinction between urgent kits and non-urgent kits; after this stage, they are characterized by the same priority.
- The 30% of the whole amount of kit in entrance go to the Orthopaedics cleaning area, the 20% of the whole amount of kit in entrance go to the General Surgery cleaning area, the 30% of the whole amount of kit in entrance go to the Gynaecology cleaning area, the 20% of the whole amount of kit in entrance go to the Ophthalmology cleaning area.
- The instrument kits that arrive at the Ophthalmology cleaning area are categorized as "normal"(40% on the total amount in the entrance at the Ophthalmology cleaning area) or "special" (60% on the total amount in the entrance at the Ophthalmology cleaning area). The special ones have a non- pre-emptive priority over the normal ones.

Another data, collected by Eng. Smith, is that 15% of the instruments kit cleaned in the Orthopaedics cleaning area, once cleaned in this area, are sent to the General Surgery cleaning area, where will be cleaned another time.

Operators	**Service time/operator**
Orthopedics (3 nurses)	2 kits every 30 minutes
General Surgery (2 nurses)	1 kit every 20 minutes
Gynecology (2 nurses)	1 kit every 12 minutes
Ophthalmology (1 nurse)	1 kit every 12 minutes
Categorization (1 nurse)	5 kits every 10 minutes
Preparing phase (2 Operators)	2 kits every 10 minutes

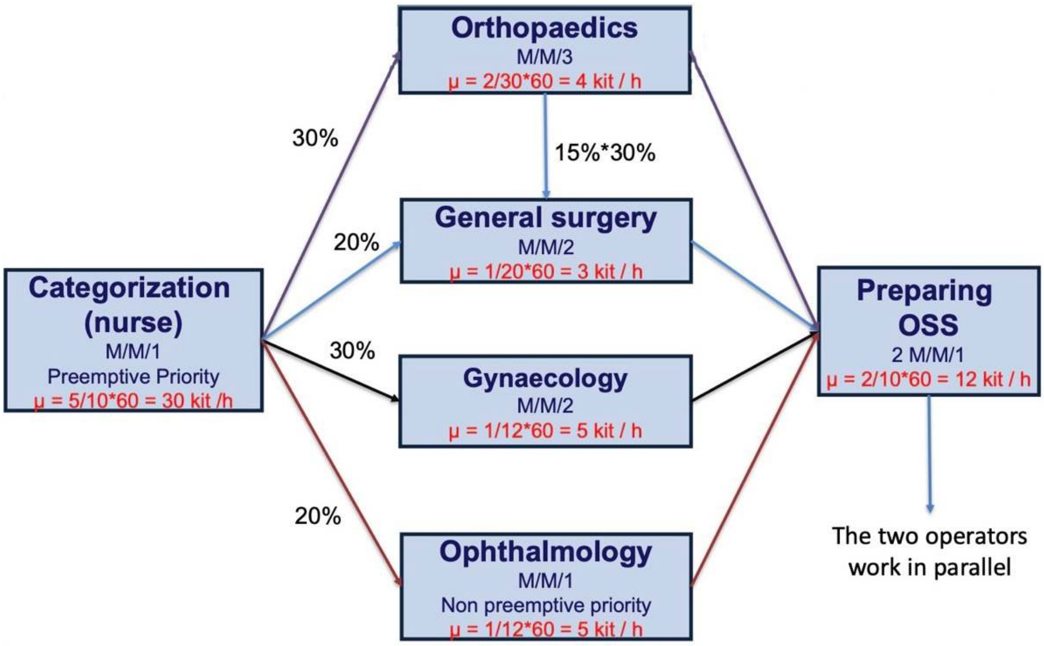
The service rate and the number of operators involved in the flow are depicted in the following table.

Service rates and arrival rates are distributed according to a negative exponential. Queues are managed with a FCFS rule.

You are required to:

a) Map the system in a detailed and precise way, putting all the important parameters.
b) Calculate the average time between an instruments kit entering the sterilization center, until it is ready to be put into the sterilization machines.
c) Identify which kind of improvements you would like to propose to improve the whole system. (Examples: decreasing cost without increasing queues, reducing queues without increasing resources, etc.). You are required to motivate each proposed improvement.

a) Map the system in a detailed and precise way, putting all the important parameters.



In order to calculate all the parameters, we need to start from the first stage and from the text:

	Catergorization (nurse) (set/h)	Ophthalmology (set/h)	Orthopaedics (set/h)	Gynaecology (set/h)	General surgery (set/h)	Preparing (OSS) (set/h)
λ	Urg.= 5 Not urg.=10 Total = 15	20%*15 = 3 Spec.=60%*3 = 1,8 Norm.=40%*3= 1,2	30%*15 = 4.5	30%*15 = 4.5	20%*15 + 15%*30%*15 = 3.675	15/2 (half demand to each server!)
μ	30	5	4	5	3	12
	M/M/1	M/M/1	M/M/3	M/M/2	M/M/2	2 M/M/1

Single queue configuration

Multiple queue configuration (work in parallel)

λ arrival rate
μ serice rate of a single operator

When they enter in the ophthalmology they are divided in special and normal (60% and 40%); in order to calculate the arrival rate, we need to take into consideration the fact that the 20% of the 15 kits are sent there (so 3 kits in the ophthalmology) and we need to multiply them according to the percentage of the "special" and the percentage of the "normal":

- Special= 60%*3=1,8
- Normal=40%*3=1,2

We do the same calculation in order to find the arrival rate for all the other departments; for the general surgery we also have to calculate the kits that are first sent to the orthopaedics :

20%*15+15%*30%*15=3,675.

For the preparing arrival rate: overall amount of kits divided by two, because each server will see only part of the queue and on average (since they are 2 servers) they will see only half of the queue.
The service rate was simply the conversion from the table that was done in minutes into hours.
In the table there are all the information but then we must estimate the Ws that can be estimated only if we know what type of customer and what type of products we take.

In fact, A total of 12 typologies of "customers" (12 kinds of products):

- 2 kinds of kits (urgent / not urgent):
  - Nurse urgent kit
  - Nurse not urgent kit

- 6 different paths:
  - Nurse => Ophthalmology (special)=> OSS
  - Nurse => Ophthalmology (normal)=> OSS
  - Nurse => Orthopaedics => OSS
  - Nurse => Orthopaedics => surgery OSS
  - Nurse => surgery => OSS
  - Nurse => Gynaecology => OSS
  -

So we have twelve types of customers and once we have mapped the types of products in our system, we must understand what the frequency and the occurrences is, in order to find the average waiting time of the system because we multiply each path for the occurrence and find the overall throughput time.

b) Calculate the average time between an instruments kit entering the sterilization center, until it is ready to be put into the sterilization machines.

Calculate the average time between an instruments kit entering the sterilization center, until it is ready to be put into the sterilization machines.

	nurse		Orthop.	General surgery	Gyneac.	Opthal.		OSS	Frequency
	Urgent	Not urgent				Special	Normal		
1	X		X					X	8.5%
2	X		X	X				X	1.5%
3	X			X				X	6.67%
4	X				X			X	10.0%
5	X					X		X	4.0%
6	X						X	X	2.67%
7		X	X					X	17.0%
8		X	X	X				X	3.0%
9		X		X				X	13.33%
10		X			X			X	20.0%
11		X				X		X	8.0%
12		X					X	X	5.33%
								Total	1

How can we compute the occurrences?

URGENT

1. Urgent- Othop – OSS → 5/15*(30%-15%*30%) = 8,5%
   First is the parentage (5/15) that corresponds to the urgent sets, this is then multiplied by those products sent to the orthopaedics minus the one sent also to the general surgery.

2. Urgent –Orthop –Gen Surgery –OSS → 5/15 * (15%*30%) = 1,5%
   The second path is characterized by those products that once they are sent to the orthopaedics, they are also sent to the general surgery, so this are the urgent (5/15) multiplied by the 15% of the 30% (those sent to the general surgery, once sent to the orthopaedics).

3. Urgent- Gen Surgery –OSS → 5/15*(20%) = 6,67%
   Urgent that are sent directly to the general surgery.

4. Urgent- Gyneac-OSS → 5/15% * (30%) = 10%
   Urgent that are sent to the gynaecology and then to the preparation.

5. Urgent – Ophtal Spec –OSS → 5/15*(20%*60%) = 4%
   Those sent to the ophthalmology, but in this case, the distinction is not only on the urgent, but on the ophthalmology, we have normal or special kits, so there are different types of products.

6. Urgent – Ophtal Spec –OSS → 5/15*(20%*40)=2,67%
   Urgent, ophthalmology and special.

NON-URGENT:

The same must be done for the non-urgent, we have exactly the same values but we need to multiply them for 10/15:

7. Non Urgent- Othop –OSS = 10/15* (30%-15%*30%) = 17%
8. Non Urgent –Orthop –Gen Surgery –OSS= 10/15 * (15%*30%) = 3%
9. Non Urgent- Gen Surgery –OSS= 10/15*(20%) = 13,33%
10. Non Urgent- Gyneac-OSS = 10/15*(30%) = 20%
11. Non Urgent – Ophtal Spec –OSS = 10/15*(20%*60%) = 8%
12. Non Urgent – Ophtal Spec –OSS = 10/15*(20%*40%) = 5,33%

Once we found the occurrences, we can now calculate the throughput time:

Throughput time - nurse

Ws with pre-emptive priority (M/M/1)

- $E(S_1) = \frac{1/\mu}{1-\rho_1}$
- $E(S_2) = \frac{1/\mu}{(1-\rho_1)*(1-\rho_1-\rho_2)}$

Nurse (preemptive priority)	Urgent kit	ρ_1 = 5/30= 0,17	$E(S_1)$= 2,4 min
	Not urgent kit	ρ_2 = 10/30 = 0,333	$E(S_2)$ = 4,8 min

Throughput time - ophthalmology

Ws with not pre-emptive priority (M/M/1)

- $E(S_1) = \frac{(1+\rho_2)/\mu}{1-\rho_1}$
- $E(S_2) = \frac{(1-\rho_1(1-\rho_1-\rho_2))/\mu}{(1-\rho_1)*(1-\rho_1-\rho_2)}$

In both of the case the Ws of the ophthalmology is extremely high, this does not impact on the whole system

Ophthalmology (NOT preemptive)	Special kit	ρ_1 = 1,8 / 5 = 0,36	$E(S_1)$ = 23,25 min
	Normal kit	ρ_2 = 1,2 / 5 = 0,24	$E(S_2)$ = 40,125 min

it is a way of improving or boosting a certain type of customers.

Throughput time - orthopaedics

In this case we have an MM3 so c=3, so we need to use the table:

Queue system, M/M/3 type ρ = 4,5 / 4 = 1,125

Lq (1,125; c=3) = 0,073 look at the table!

Ws= Lq/λ + 1/μ = 0,073/4,5 + 1/4 = 15,97 min

We need to do the interpolation because we cannot find the right value on the table:

(y – 0,066)/ 0,094 – 0,066 =

= (1,125 – 1,1)/ (1,2 – 1,1)=> **Y=Lq=0,073**

Lq results of model M/M/c

λ/μ	c=1	c=2	c=3	c=4	c=5
0,15	0,026	0,001			
0,20	0,050	0,002			
0,25	0,083	0,004			
0,30	0,129	0,007			
0,35	0,188	0,011			
0,40	0,267	0,017			
0,45	0,368	0,024	0,002		
0,50	0,500	0,033	0,003		
0,55	0,672	0,045	0,004		
0,60	0,900	0,059	0,006		
0,65	1,207	0,077	0,008		
0,70	1,633	0,098	0,011		
0,75	2,250	0,123	0,015		
0,80	3,200	0,152	0,019		
0,85	4,817	0,187	0,024	0,003	
0,90	8,100	0,229	0,030	0,004	
0,95	18,050	0,277	0,037	0,005	
1,0		0,333	0,045	0,007	
1,1		0,477	0,066	0,011	
1,2		0,675	0,094	0,016	0,003
1,3		0,951	0,130	0,023	0,004
1,4		1,345	0,177	0,032	0,006
1,5		1,929	0,237	0,045	0,009

Throughput time – General surgery:

Queue system, M/M/2 type ρ = 3,675 / 3 = 1,225

Lq (1,225; c=2) = 0,744 (Interpolation as before) Ws= Lq/λ + 1/μ = 0,744/3,675 + 1/3 = 32,15 min.

Throughput time - gynaecology:

Queue system, M/M/2 type ρ = 4,5 / 5 = 0,9

Lq (0,9; c=2) = 0,229

Ws= Lq/λ + 1/μ = 0,229/4,5 + 1/5 = 15,053 min

Throughput time – OSS:

2 queue systems, M/M/1 type for each queue: λ = λOSS / 2 = 7,5

Ws= 1/(μ - λ) = 1/(12 – 7,5) = 13,33 min

Now, we need to find the mean **Ws for the whole system**:

	Nurse	Orthop.	General surgery	Gyneac.	Opthal.	OSS	Occurrence	Weig. time
1	2,4	15,97	-	-	-	13,33	0,085	2,69
2	2,4	15,97	32,15	-	-	13,33	0,015	0,96
3	2,4	-	32,15	-	-	13,33	0,0667	3,19
4	2,4	-	-	15,05	-	13,33	0,1	3,08
5	2,4	-	-	-	23,25	13,33	0,04	1,56
6	2,4	-	-	-	40,125	13,33	0,0267	1,49
7	4,8	15,97	-	-	-	13,33	0,17	5,80
8	4,8	15,97	32,15	-	-	13,33	0,03	1,99
9	4,8		32,15	-	-	13,33	0,1333	6,70
10	4,8	-	-	15,05	-	13,33	0,2	6,64
11	4,8	-	-	-	23,25	13,33	0,08	3,31
12	4,8	-	-	-	40,125	13,33	0,0533	3,10
							TOTAL (min)	40,51

We need to evaluate the Ws of each system and, knowing the path, we can sum the Ws characterizing of every single system characterizing each path, and then, to compute the overall throughput time of the system we need to compute the average value by having the occurrences.

c) Identify which kind of improvements you would like to propose to improve the whole system. (Examples: decreasing cost without increasing queues, reducing queues without increasing resources, etc.). You are required to motivate each proposed improvement.

We could have different solutions:

1. Have a single queue system for OSS (40% Ws reduction for OSS stage).
2. Understand the main causes of second working for kits processed in orthopaedics.
    a. Ad hoc process or tools to avoid second working.
    b. More training for orthopaedics' operators.
3. Improve service process in orthopaedics: 2 kits in 25 minutes instead of 30 minutes. Ws is still about 16 minutes (16,1 minutes in front of 15,93 minutes) but it could move one operator to other areas.
4. …

QUEUE THEORY 4:

EXERCISE 1 - QUEUEING THEORY DESPACITO:

A company has a central document-copying service where employees of this company go occasionally to make copies. Arrivals are assumed to follow the Poisson probability distribution, with a mean rate of 15 per hour. Also, service times are assumed to follow the Poisson distribution. With the present copying equipment, the average service time is 3 minutes. A new machine is available that will have a mean service time of 2 minutes. The average wage of the employees who bring the documents to be copied is €8 an hour. Employees need to wait at the machine while it is copying, and all employees are fully saturated during their work.

1: If the machine can be rented for €2 per hour more than the old machine, should the company rent the new machine? Consider lost productive time of employees as the time they spent waiting in the system.

2: The area around the copy machine can accommodate only three employees. What percentage of time will this area be inadequate for waiting employees when using just the old copy machine?

3: Due to the extremely low prices, the company thinks about purchasing many small-scale copy machines. Assume the small-scale copy machines are set up with a first come first serve queue policy in a single queue system, and that each small-scale copy machine has an average service rate of 5 employees per hour. Determine the smallest number of copy machines the company needs to provide to keep the average waiting time in the system for the employees lower than 15 minutes.

Solution:

$\boldsymbol{\lambda}$= 15 $\frac{\textbf{customer/employees}}{\boldsymbol{h}}$

$\boldsymbol{\mu}$1 = 60/3 = 20 with 3 = average service time → case AS-IS

$\boldsymbol{\mu}$1 = 60/2 = 30 with 2 = average service time → case TO-BE

Wage of employees = 8 €/h

Question 1: **Δ**cost for renting = 2 €/h

c1 < c2 → DON'T RENT

c2 < c1 → RENT

M/M/1 with Ws = $\frac{\mathbf{1}}{\boldsymbol{\mu}-\boldsymbol{\lambda}}$ → Ws1 = $\frac{\mathbf{1}}{\mathbf{20-15}}$ = 0.2 h = 12 min, Ws2 = $\frac{\mathbf{1}}{\mathbf{30-15}}$ = 0.06 h = 4 min

Case 1: 8 €/h * 15 employees * 0.2 h/employees = 24 €/h

Case 2: 8 €/h * 15 employees * 0.06 h/employees + 2 €/h = 10 €/h

Yes, the company should rent the new machine.

Question 2:

P(n≥4) = ρk = $(\frac{\lambda}{\mu})^4$= $(\frac{15}{20})^4$ = 0.3164 * 60 min/h = 18.98 min

Question 3:

Ws target ≤ 15 min → M/M/C with C?

λ= 15 employees and μ = 5 employees → $\rho = 3$

Ws = $\frac{Lq}{\lambda} + \frac{1}{\boldsymbol{\mu}} \leq$ 15 min

Lq ≤ 0.75 → Lq (ρ; c) ≤ 0.75 → c = 5

EXERCISE 2: MECCANICA SPA

MECCANICA Spa is a little manufacturing company. The customers ask the company a rate of 100 pieces per hour. The selection is composed by 3 types of products: type A (50% of the demand), type B (30% of the demand) and type C (20% of the demand).

The production system is simple, it is formed by:

1) Milling (2 milling machines. Service rate: 90 pieces per hour).
2) Drilling (3 drilling machines. Service rate: 75 pieces per hour).
3) Testing (1 testing machine). The service rate of the operator dedicated to the testing machine is 160 pieces per hour.
4) Packaging (1 packaging line with one operator able to package 130 pieces per hour). The assumption is that all the service times are described by negative exponential distribution. The flow is the same for each type of products:

MILLING => DRILLING => TESTING => PACKAGING

In particular, at the milling stage, both the machines are able to work all types of products and the company decided to dedicate one machine to type A products and the other machine to products of type B and type C. In the drilling stage there are three identical machines, each one dedicated to work on one type of product. The testing machine reveals that sometimes type B and type C products have to be worked again because of an error during the drilling process. These products have to be processed again by the drilling machine and the re-working time is the same of the first time.

On average, 20% of type B and 10% of type C products that are introduced in the system have to be worked twice.

About the twice worked pieces, 85% are good and 15% are rejected and trashed.

At the packaging stage a product of type C has a priority: when a server is free type C product is next to be processed.

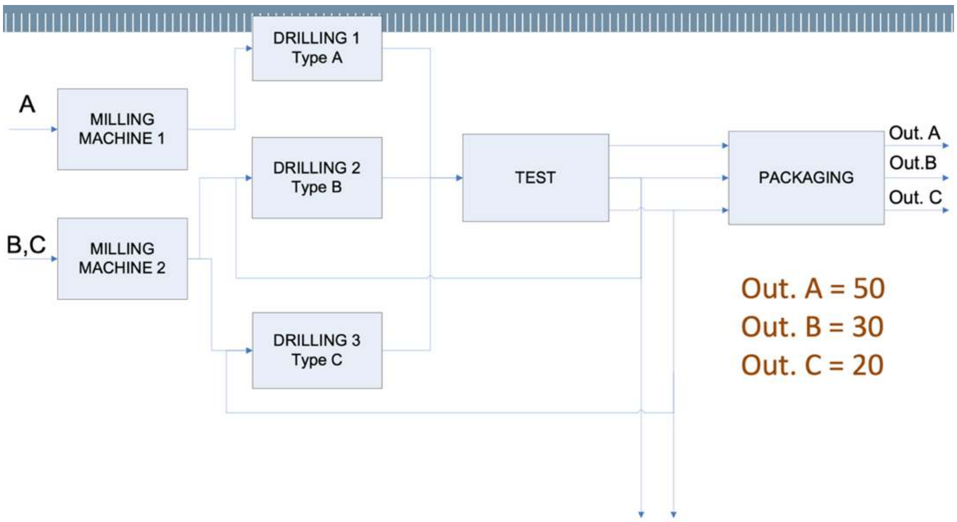
Shape Meccanica Spa. System, define relevant parameters (arrival rates, service rates,queue type, inactivity time of resources) and map all the flows and paths of products.

Solution:

First of all, we know that we have to deliver 100 good pieces each hour, divided as:

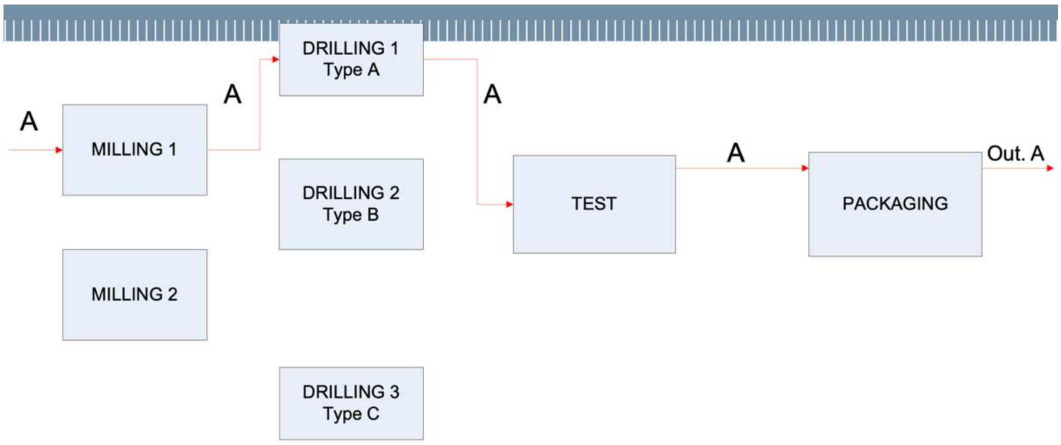
- 50 types A;
- 30 types B;
- 20 types C.

The following graph is showing the system modelled:



We have product A that enters the drilling machine that is dedicated to this product type, then it is sent to the testing and packaging machine, that are for all the types of products. Type B and C are both worked on one milling machine (the second one), then they are sent to each type of drilling machine according to the product (product B in drilling machine B and product C in drilling machine C); they are both sent to the testing machine and some of products B are reworked, others are discarded and the same happens for product C.

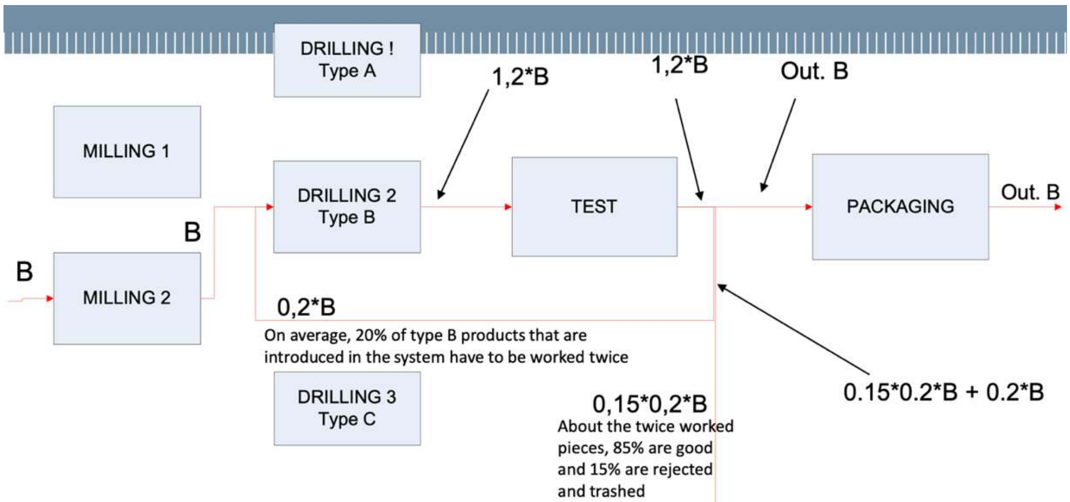
We now need to estimate the arrival rate of the ***product A***



$$A = Out.\ A = 50\ pc./h$$

We know that for product A nothing happens (nothing is scrapped or re-worked), so what enters is exactly what goes out from the system.
So A = out. A = 50 pc/h

Arrival rate for ***product B***:

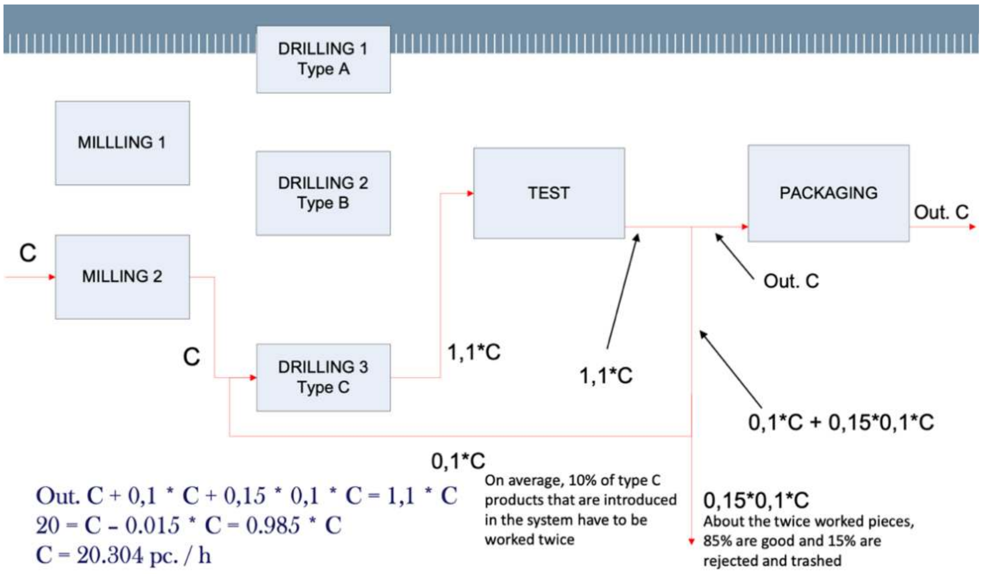


$$Out.\ B + (0{,}2 * B + 0{,}2 * 0{,}15 * B) = 1.2 * B$$
$$30 = 0.97 * B$$
$$B = 30.93\ pc./h$$

Product B enters in the milling machine, to the drilling machine, to the testing machine and then it is re-worked.

On average 20% of B that are introduced in the system have to be re-worked so 0,2*B has to be re-worked. We know that between drilling and testing machine we have 1,2 of input B (because it is the sum of input B + 0,2*B) so 1,2 *B.
After the testing machine, the product can be also scrapped and this happens to the 15% ofthe products that are tested twice.
So before the packaging we have an output that is equal to the output B. So the output B is:
Out.B + (0,2*B + 0,2*0,15*B) = 1,2*B30 = 0.97*B
B = 30.93 pc/h

Arrival rate ***for product C:***



The product C follows the same path of B but the 10% of C are re-worked and then the 15% of the product worked twice are discarded (15*10*C), so:
Out.C+0,1*C + 0,15*0,1*C=1,1*C20=C-0,015*C= 0.985*C → C= 20.304 pc/h.

Now we have all the values and we can put these information in the table:

	Milling1 (pdt. A)	Milling 2 (pdt B,C)	Drilling pdt. A	Drilling pdt. B	Drilling pdt. C	Testing	Packaging
λ (pc./h)	50	51,23	50	37,12	22,33	109,45	100
μ (pc./h)	90	90	75	75	75	160	130
Queue	M/M/1	M/M/1	M/M/1	M/M/1	M/M/1	M/M/1	M/M/1
Priority	-	-	-	-	-	-	Non Pre
Idle time (min/h)	26,67	25,85	20	30,30	42,136	18,96	13,85

Idle time = 1- λ /μ

Each machine has an arrival rate that could be of one single typology or could be characterized by more than one product type (for instance, milling machine 2 is characterized by both product B and C). When we perform the calculation we must compute every calculation for EACH product type. Everything was characterized by M/M/1, the drilling for instance is 3M/M/1, but in reality they are different because they work in a single product and in the test is not written that itis a multiple queue.
The idle time (1-ρ*60).

Key message: mapping phase is fundamental to understand the system and its inefficiency.

10. SHOULDICE CASE

WHAT IS THE VALUE PROPOSITION OF SHOULDICE HOSPITAL?

- Excellent work
- Take care of the patient
- Quality of the service
- Compared to other clinics they provide the service at a lower price
- Intimacy between patient-patient and patient-staff
- Everything in one visit
- Fast recover
- Higher salary with low working time
- Specialized in a specific segment, on a specific target
- Auto diagnosis
- Self-recovering
- Similar people in the room → Shouldice methodology
- Few type of customer → standard

The value proposition is not here.

Service concept

1. **Organizing idea** → value proposition, the mission, a sentence that should synthetize how do we approach. You have to abstract from the operative area.
2. **Service provided**: it has to be linked to the service received, the outcome and the experience that we want to provide to the customers. We must activate levers.
3. **Service received**: divided in two main parts, to be defined as deployment of organizing idea:
   a. **Outcome** → ex: cheap, quality (conformance, design), time ()
   b. **Experience** → which kind, which should be the characteristics?

		SHOULDICE CASE
ORGANIZING IDEA	What we define as the **value proposition** or mission. How do we approach?	**Solving the hernia** (starting point) How do they want to provide this? How do they want to differentiate? Shouldice methodologies, standards. The time that you devote to solving the hernia must be short, nice (you do not have the feeling to be in hospital but you are on vacation), accessible to everybody (money speaking). When we make investment we make it to stress this point. The slogan "***Shouldice clinic as a vacation club***", in order to give the idea of a place of peace and relaxation, where the customer goes to exit healthier and let the staff take care of him.
SERVICE PROVIDED	What is Shouldice Hospital output? In order to have Intimacy or Confidence what are the outputs? Outputs are the result of operative steps → Levers	<u>Output</u>: **Social Interaction** How to create Social Interaction in order to create intimacy? What are the levers? ▪ Double-Room, Single dinner, U shape configuration[1] (*Structural*) ▪ Put together Patient-Staff, Similar people in the room (*Managerial*)

[1] It makes the inter-relationships and the exchange of views almost "natural" every time there is a problem or a need, just because the different rooms are virtually "facing" one another

		Low recurrence How to have low recurrence rate in order to create a high-quality service for the customers? What are the levers? ▪ Few types of customers (*Managerial*) ▪ High specialization, Shouldice procedure[2], (*Organizational)* **Self-recovery system** You provide a rehabilitation process that is based on self. What are the levers? ▪ Special Stairs, carpets & moquette[3] (*Structural*) ▪ Auto-diagnosis, Few nurses (*Organizational)*
SERVICE RECEIVED	Which kind of service we want to offer to customers. ▪ Outcome ▪ Experience These two points has to be deployed from the organizing idea	Which should be the characteristic of the experience offered by Shouldice Hospital? ▪ **Intimacy:** Patients with Patients and Patients with staff ▪ **Confidence** ▪ **Being in a SPA** Which should be the characteristic of the outcome offered by Shouldice Hospital? ▪ **Cheap** ▪ **High quality services:** in terms of conformance, but even quality of "design", for example the time that you need to stay at the clinic is shorter respect to other clinics/hospitals

In order to have a certain output we must activate levers: if you want to manage it
you must be able to correctly define the goal and the operative steps to do in order to achieve the final goal. This is an example of the difficulty of the company to distinguish what is a lever, what is a performance, how you translate performance in business success.

Their performances in terms of:

- Speed → bad because they have a very long queue
- Reliability → well
- Flexibility → very bad

This is strictly correlated to the strategy and the value proposition they built up. In particular, what are the choices that we put in place in order to make the customer receiving these points?

[2] They serve only a particular type of hernia's problem

[3] They create an environment where the recovery for patients is easier. This solves the trade-off between the need to immediately resume the mobility of the recently operated patient and the danger that implies the post-operation movements.

Structural decisions:

- Double rooms
- U shape, to make easier the work among the doctors
- Single dining room, they all eat in a single room
- Special stairs: the steps of the stairs are not normal, they are smaller stairs, in this way it is easier for the patience to start to work
- Carpets, moquette: in this way it is more difficult for the patient to sly, they created an environment where it is easier for the people to start rehabilitation independently

Organizing decisions:

- Specialization/standardization: we do not cure every hernia, we cure only certain type of hernia → Shouldice methodologists
- Shouldice procedures: they work for a low recurrency rate in trade-off with flexibility, we do not cure everything because it would be more difficult for us to respect the quality rate, so we exclude some type of customers
- Auto diagnosis:
- Few nurses: compared to the public center (1:15 compared to 1:4), lined to the self-recovery systems and also let the relatives to stay in the clinic, in the same room of the patient, so I have less nurses to take care of the patient.

Managerial decision:

- The timing of lunch and dinner is at the same time together P-S
- Put together similar people in the room, with similar habits → this is a trade-off: if you have a lot of people, you have a queue of people that can better match together with the habits, if you have a short queue you have not so many people and can be more difficult to match people with same interests
- Few types of customer: you must have a good health otherwise we do not accept you.

All the levers are coherent with the organizing idea and the value proposition, everything is aligned.

HOW IS THE HOSPITAL GOING TO REACT TO A POTENTIAL MARKET INCREASE?

Company wants to **increase the business**, but how?

- Working on Saturday
  Changing the working shift from 5 to 6 days, the capacity would increase by 20%. This logic, however, tested for a certain period, has not given the desired results. Basically the staff is not happy about this decision. On the first hand, not all surgeons were willing to accept it, because it contradicted what was promised initially (little work in the weekend, more free time to devote to the family). On the other side, the result was the creation of two sub-groups within the medical staff:
  - Those who had agreed to work on Saturday were paid a bonus and felt "more important";
  - Those who did not accept this condition, felt a bit like the "*black sheep*".

  This scenario is contrary to the principle of cohesion and "*big family*" of the company´s philosophy, they were killing the key points that had made the business successful: *homogeneous group, no difference, collaboration, high salaries, clean working time*. The solution was soon abandoned especially for disadvantages in terms of group cohesion.

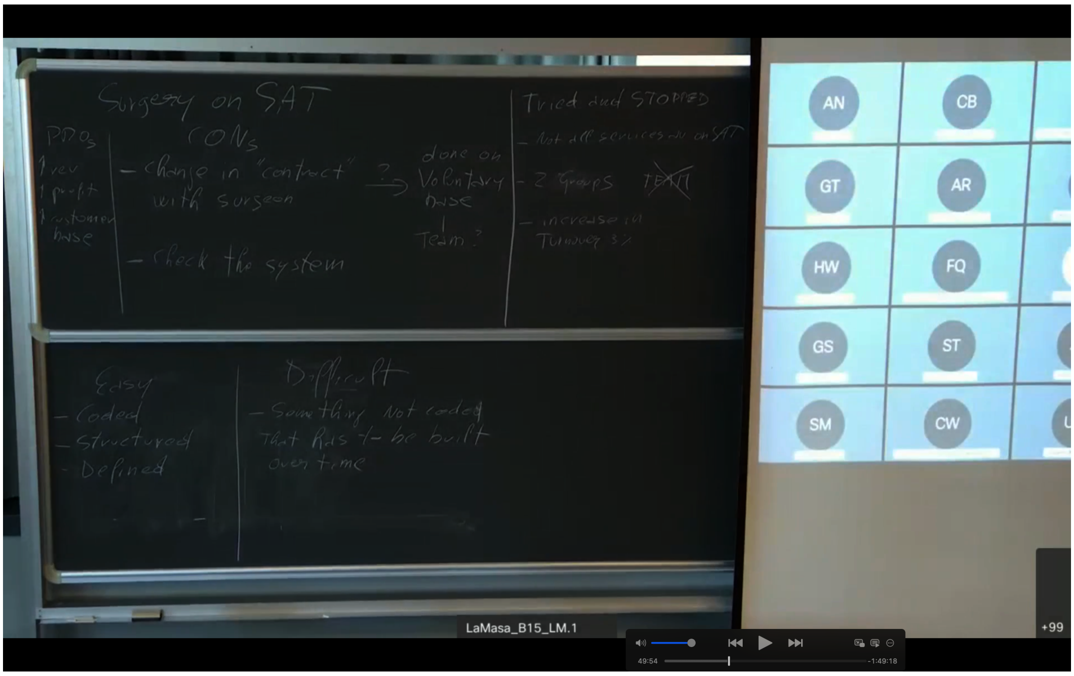
- Opening a new centre in the US
  Opening a second clinic is an alternative that could be viable even in economic and management terms, but not from the strategic and operations perspective.
  It will be easier to replicate the elements "hardware" of the current structural (patient rooms, the structure of the operating blocks, stairs ...). However, **it would be much more difficult** (if not impossible) **to do so with "software"** (infrastructural). The cohesion (staff family), the relationship with

patients and confidence in the procedure, are all elements that, only the current headquarters, has been able to build over time, relying on past history and experience. Moreover, the **procedure Shouldice** and the **process that accompanies the customer through the system**, for the current clinic, are now established mechanisms and **"order winners"**. By opening a new office, and hiring new surgeons and nurses, all these elements will be absent for a certain period. The key point is that the system that has been created inside the clinic, has characteristics that do not derive from the mere application of procedures, but such procedures, they are the result of an evolution which, if not retraced, it will never be able to give the same results.
There is the risk of **losing the control**. The more behaviours and activities are coded and codeable (standardization towards mass service), the easier it is to replicate it. On the contrary, if a system is characterized, mostly, with professional elements and not coded, that has developed over time, it will be very difficult to implement an expansion, because it is **difficult to replicate**. You have to replicate the entire system. It means that you have to replicate all soft aspects (all characteristics of the system built up in long-time) in order to replicate the success.
→ For all these reasons, **the alternative opening of a new clinic is not viable**: the most reasonable is to develop a long-term plan that will lead to implement a series of improvements leading to grow slowly, ensuring, however, the maintenance of the leading position in the field worldwide.

CONCLUSIONS

- **Service concept model** is a useful tool in order to align business strategy-operations.
- **Soft aspects** for an operation system there are many aspects that cannot be replicated just buying objects. What gives the competitive advantage is the knowledge and how do you exploit it, the experience and how far you are from your competitor on the ability of using this technology.

You can look at the Shouldice case as a **fantastic example of alignment between operations and market**. It is also necessary to point out that, this alignment is not one way, but is **bidirectional**: it is not just the company that chooses a configuration according the needs of the customer's segment that wants it to serve (implementing plant Technological (structural/hardware), organizational(infrastructural/software) and managerial elements that are best suited in making them evolve over time), but also the fact that, based on these choices, the company is able to understand what the market is more likely to turn to.



11. YIELD MANAGEMENT

11.1. CONTEXT

The fundamental responsibility of operations management is to **provide the capability to satisfy current and future demand**. Trade-offs between customer service and cost. How can we cope with mismatches between demand and the capability of the service delivery system? How can we manage variability in our system?
The shape of the demand is **variable**, while the shape of the capacity is **flat**. You can adjust (make it more coherent) the shape of the capacity following the shape of the demand. Booking system is a technique through which we can make the shape of the demand flatter. Making it more under control.
In the supermarket, to follow an increase of demand the supermarket can put more casher (*worker flexibility*) to make the shape of the capacity more variable. You can move operators from one point to another one.

Yield management is in the middle, that increase the control of the demand and exploit capacity and flexibility.

11.2. DEFINITION

Yield Management is a variable pricing strategy based on understanding, anticipating and influencing consumer behaviour in order to maximize revenues or profits form a fixed, perishable resource (such as airline seats or hotel room reservations or advertising inventory).

Yield management exploits the information on customers' behaviour that Operations get while delivering the service, so it improves the competitiveness of the company.
It therefore **helps aligning demand with capacity**.

11.3. HISTORY

The beginning of this managerial approach started in 1978, when with the "*Airline Deregulation Act*" prices were liberalized in the aerial field. This brought to the rise with **new companies with lower prices** which had two effects:

- Traditional companies were in danger thanks to the new competitors who were offering very aggressive prices.
- The aerial fields widened, including the customers who before were not willing to pay higher prices and prefer car or train transportation.

New competitors that work steeling market share.

Later, during the 80s, airline companies started to have an increasing computational power thanks to the larger amount of data they had in hand and, looking at the information, they understood how to improve and to influence the participation of customers by:

- Setting the **highest price to those customers who are willing to pay** them and **lower prices to the others**
- Selling at **different prices to different customers**
- **Saturating planes**
  → Not every ticket has the same price

In fact, if the profit cannot grow because of the **competitive pressure** and the company is bound to hi**gh fixed costs**, the costs are reduced by selling to more people.

Idea: saturate seats' capacity offering at lower cost some of the unoccupied seats
Moreover, it was necessary to determine seats' optimal allocation between economy and business classes **to avoid the cannibalization effect** (having customers willing to pay higher prices who buy low cost seats). To proceed with service differentiation, in order to reduce cost underestimation or overestimation, information system and allocation model are fundamental to keep track and determine rates.

We can say that **Yield Management** acts in an integrated way on demand and capacity in order to "*sell the right capacity, to the right customer, at the right time and the right price*" so to **maximize profits**.

Moreover, **application fields are constantly increasing** and the main industries that adopt Y.M. are:

- Transport: airline companies, naval transportation, railway companies, rent a car company
- Entertainment: tour operators, cruise boats, golf course, Movie theatre, Advertising
- Hotels, Restaurants

11.4. IDEAL CHARACTERISTICS FOR YIELD MANAGEMENT

- Uncertain demand
- Fluctuating demand
- Low marginal sales cost and high fixed costs
- Fixed capacity: it is very difficult to change the capacity in shorter.
- Same capacity to different market segment (same service to different customers at different prices. The service/price could be exactly the same or slight differences are made at the last stages of the production process)
- Ability to segment markets
- Perishable inventories
- Product booked/sold in advance
- High capacity change cost

Yield management mainly uses **2 different tools**:

1. Capacity allocation
    a) Price policies: setting the price
    b) Demand forecast
    c) Protection policies (e.g. set the protection level): how many seats are reserved to the full price
    d) customers.
2. Overbooking
   For example, selling more seats than the capacity is possible, knowing that a percentage of the customers will not attend the flight

Capacity allocation and market allocation
Customers are divided into segments. In the example, customers are divided according to their willingness to spend, while the number of the seats available is 110.

Group	Price to purchase	Number of customers
C1	$ 230	20
C2	$ 180	30
C3	$ 150	40
C4	$ 120	50
C5	$ 40	50

Linear model
The problem is in the **allocation of each group to the capacity to maximize the margins**. Decision variables to explain the number of tickets allocated to each class.
The objective formula represents the *maximization of margins* for unit sold. The **first is a constraint** reflecting the capacity of each class and states that the number of units reserved to customers of class j has to be lower or equal to the number of people of class j. While the **second constraint** reflects the overall capacity and represents that the amount of units considering all classes has to be equal or lower than the overall capacity.

X_j= Number of tickets allocated to class j
M_j= Unit margin for a sale to class j
N_j= Number of persons of class j
C= Overall capacity

$$\max \sum_j M_j x_j$$

Subject to

$$x_j \le N_j \quad \forall j$$

$$\sum_j x_j \le C$$

For this example:
X_1=20
X_2=30
X_3=40
X_4=20
X_5=0

What is missing?
Time: people are willing to trade time for money. If you buy in advance, you pay less, but sometimes this is not possible if dealing with people

The other huge problem lies in the fact that **this data is not immediately available**, so it needs to be forecasted. Through this forecasting, knowing the willingness of the customers of buying full price products will enable managers to protect the capacity for each class of customers. In fact, not all tickets can be sold in advance: otherwise, there are no tickets which ca be sold at higher prices.

Capacity allocation and demand forecasting

To forecast the demand, we can exploit some software starting from the data that is available. In order to find patterns, clusters of customers, you need to have a huge collection of customer's data. This data allows to have information about personal consumption, but also about the way people consume it.

Data provide a lot of information because of:

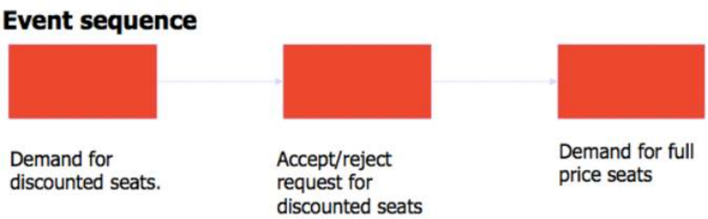
- High detailed level (origin, destination, rates, day of the week, timetable)
- Quantities to keep track:
  - Demand per each category/rate/period
  - Demand elasticity

Forecasts difficulties stands in:

- Seasonality
- Trend
- Special events
- Competitors actions
- Small variations in a variable can have large effects on others

Protection level sizing

x_1, x_2 = demand quantities for the *full* and the *discounted* tickets.
The objective is to determine the quantity of the capacity dedicated to full price paying customers



$$S_1 = Protection\ level$$
$$A_2 = C - S_1 \quad Capacity\ allocation, low\ cost$$

11.5. THE YIELD MANAGEMENT GAME

In the simplification of the game, we only have two price differentiation:

- Discounted ticket
- Full ticket

From the expiration date, only full price tickets can be sold. The problem is when the demand for the discounted price is much higher than capacity. There, you must decide how many rooms you have to protect, not to sell, at a discounted price, even if you cannot know how many people would accept to buy a higher price. Moreover, not all who have booked will actually show up (No-Show)

→ The **goal** is the maximization of the revenues of the hotel

Decisions:

1. Define how much **capacity you want to protect** (PL for full customers)
2. Define how much **overbooking** you want to leave
3. Define a guideline to decide which requests you will accept and which one you will refuse

11.6. THE YIELD MANAGEMENT GAME

- Number of rooms = 100
- Walking cost = 500€
- Average NO-SHOW = 20%
- Standard Deviation = 3

DISCOUNTED DEMAND:

- Demand well exceeding capacity
- Average price = 300€

FULL PRICE DEMAND:

- Demand (purchase) follows a normal distribution with average = 43 and standard deviation = 6
- Average price = 1100

For each request, you have to accept it all or refuse it all.

Arrivals of customers paying a discounted price

A. 20 Representatives at 275 €/customer
B. 30 Professors at 200 €/customer
C. 15 MBAs at 150 €/customer
D. 5 People of a Family at 250 €/customer
E. 12 People attending a convention at 200 €/customer
F. 16 People attending a conference at 500 €/customer
G. 35 Sport reporters at 350 €/customer
H. 10 Basket team at 250 €/customer
I. 8 High school band at 400 €/customer
J. 9 Government officers at 350 €/customer
K. 10 Marketing managers at 200 €/customer
L. 8 Golf players at 400 €/customer
M. 11 Nurses at 250 €/customer
N. 20 Members of a political party at 450 €/customer
O. 5 Vacationers Club med at 375 €/customer
P. 8 Patients of Salus clinic at 300 €/customer
Q. 12 PhD students at 150 €/customer
R. 6 Finance managers at 220 €/customer
S. 4 People of a family at 220 €/customer

Arrivals of customers paying a full price

1. 3 People on a business trip at 875 €/customer
2. 1 CEO Yamaha at 1400 €/customer
3. 4 Managers IBM at 850 €/customer
4. 2 Managers Honda at 1300 €/customer
5. 8 Ambassadors at 900 €/customer
6. 8 Travel agents at 850 €/customer
7. 10 Japanese people on holiday at 900 €/customer
8. 8 Medical doctors attending a conference at 790 €/customer
9. 2 McKinsey consultants at 1400 €/customer
10. 1 Operations manager at 1500 €/customer
11. 1 CEO of an automotive company at 800 €/customer
12. 4 Traders at 1300 €/customer
13. 1 Operations manager's assistant at 1400 €/customer

Arrivals with No-Show effect

CUSTOMER PAYING A DISCOUNTED PRICE				
BOOK ID	# BOOKED	CUSTOMERS	PURCHASE	REVENUE
A	20	Representatives	16	4400
B	30	Professors	25	5000
C	15	MBAs	12	1800
D	5	People of a Family	4	1000
E	12	People attending a convention	9	1800
F	16	People attending a conference	13	6500
G	35	Sport reporters	30	10500
H	10	Basket team	8	2000
I	8	High school band	7	2800
J	9	Government officers	8	2800
K	10	Marketing managers	9	1800
L	8	Golf players	6	2400
M	11	Nurses	8	2000
N	20	Members of a political party	15	6750
O	5	Vacationers Club Med	5	1875
P	8	Patients of Salus clinic	6	1800
Q	12	PhD students	10	1500
R	6	Finance managers	4	880
S	4	People of a family	3	660

CUSTOMERS PAYING A FULL PRICE				
BOOK ID	# BOOKED	CUSTOMERS	PURCHASE	REVENUE
1	3	People on a business trip	2	1750
2	1	CEO Yamaha	1	1400
3	4	Managers IBM	3	2550
4	2	Managers Honda	1	1300
5	8	Ambassadors	6	5400
6	8	Travel agents	7	5950
7	10	Japanese people on holiday	10	9000
8	8	Medical doctors attending a conference	7	5530
9	2	McKinsey consultants	2	2800
10	1	Operations managers	1	1500
11	1	CEO of an automotive company	1	800
12	4	Traders	3	3900
13	1	Operations manager’s assistant	0	0

$PROFIT_i = REV_D + REV_F - (OVB+PEN)$

ID of Solution	PROFIT	SAT	OVB
1			

Protection level definition

$S_1 = Protection\ level$

$A_2 = C - S_1\ Capacity\ allocation,\ low\ cost$

The protection level definition can be happened according to two different methods:

- The **Marginal Analysis**
- The **Heuristic Expected Marginal Seat Revenue** (EMSR)

11.7. MARGINAL ANALYSIS

It is based on the cost of overestimate or underestimate the protection level S_1

x_1 = demand for the full ticket

Two costs related to the risk:

c_u = **cost to underestimate** the demand of a full price unit

- This is equal to the additional margin we would have gained if we would have sold the same unit at a full price instead of at a discounted price.
- Is the demand of full that you cannot satisfy. It is a risk because the full is more profitable than the discounted.

c_o = **cost to overestimate** the demand of a full price unit

- This is equal to margin we would have gained by selling the same unit at a discounted price, instead of not having sold it at all.
- We lose the opportunity of a discounted, obtaining an empty room. This is the risk.

As the problem regards a trade-off between the two costs, we need to balance the underestimate expectation cost and the overestimate expectation cost.

$$\boxed{\text{P(under.) *Underest cost} \geq \text{P(over.) *Overestimate cost.}}$$

$$P(X_1 \geq S_1) * C_u \geq P(X_1 < S_1) * C_o$$

$$[1 - P(X_1 < S_1)] * C_u \geq P(X_1 < S_1) * C_o$$

$$P(X_1 < S_1) \leq \frac{C_u}{C_u + C_o}$$

In our problem,

C_u = full price – discounted price = 1100 – 300 = 800

C_o = discounted price – empty room = 300 – 0

(Hp: variable costs are negligible)

P (X1<S1) ≤ 800/1100 = 0,727

→ ϕa = 0,6

S1 = PL = $\mu + \phi a * \sigma = 43 + 0,6 * 6 = 47$

Φ_a	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0.1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0.2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0.3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0.4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0.5	[illegible]	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0.6	0,72575	,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0.7	[illegible]	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78524
0.8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0.9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1.0	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1.1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1.2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1.3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91308	0,91466	0,91621	0,91774
1.4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1.5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1.6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1.7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1.8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1.9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2.0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2.1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2.2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2.3	0,98928	0,98956	0,98983	0,99010	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2.4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2.5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2.6	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2.7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2.8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2.9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861

A few messages:

- It would be possible to saturate the hotel selling all rooms at a reduces rate. In this way, however, the profit would not be maximized. Should therefore bear a risk against a potentially superior margin.
- The uncertainty and variability must be managed: overbooking and risk on the fixing of the protection level
- Having information on the past history is essential to a set of yield management system, as well as having good demand forecasts. The information system is central to the success of this management approach
- The ability to effectively segment the different types of customers and know in advance is a crucial point, which the operations are to exploit
- It is important to pay a lot of attention to choose of prices.

11.8. HEURISTIC EXPECTED MARGINAL SEAT REVENUE (EMSR)

Fix the level of protection for each of the classes / rates

n # available rates

f_i unit revenue associated with rate i

$$f_1 \geq f_2 \geq \ldots f_n$$

μ_i Average demand for rate i

σ_i^2 Variance in demand for rate i

ϑ_i Level of protection for class i and more expensive classes

D_i Available demand to pay rate i, or more expensive

BASIC CONCEPT

Expected Value of the Fair = Fair * P(fair arrives)

Expected Value of the **D** = P_d * **1**

Expected Value of the **F** = P_f * **P(X > ϑ)** ϑ = Protection Level

Example

P_d = 100€

P_f = 200€

P(X > ϑ) = 0,2

Expected Value (Full) = 200*0,2 = 40€

Little Wood Rule *We protect the seat until the expected value for the discounted is higher or equal than the expected value for the full*

$$f_d >= f_f * P(X > \vartheta_f)$$

$$Mf * P(X_1 \geq S_1) \geq Md$$

$$\Updownarrow$$

$$\overline{f}_i * P(D_i \geq \vartheta_i) = f_{i+1}$$

Where…

$$\overline{f}_i = \frac{\sum_{j=1}^{i} \mu_j f_j}{\sum_{j=1}^{i} \mu_j}$$ Average rate of class i and more expensive

$D_i \sim$ N $\quad \overline{\mu}(i) = \sum_{j=1}^{i} \mu_j \quad \overline{\sigma}^2(i) = \sum_{j=1}^{i} \sigma_j^2$

$$\overline{f}_i * [1 - P(D_i < \vartheta_i)] = f_{i+1}$$

$$P(D_i < \vartheta_i) = 1 - \frac{f_{i+1}}{\overline{f}_i}$$

$$F(z_\alpha) = 1 - \frac{f_{i+1}}{\overline{f}_i}$$

$$\vartheta_i = \overline{\mu}(i) + z_\alpha \overline{\sigma}(i)$$

EXAMPLE:

Determine the protection level for classes 1 and 2

Class	Rate	Mean	Variance
1	€ 100	30	50
2	€ 80	30	80
3	€ 40	50	120

Average weighted rates and aggregated Averages and Variances

	$\overline{f_i}$	$\mu(i)$	$\sigma^2(i)$
1	€ 100	30	50
2	€ 90	60	130
3	€ 67,30	110	250

Protection level for class 1

$$F(z_a) = 1 - \frac{80}{100} = 0.20 \Rightarrow z_a = -0.84$$

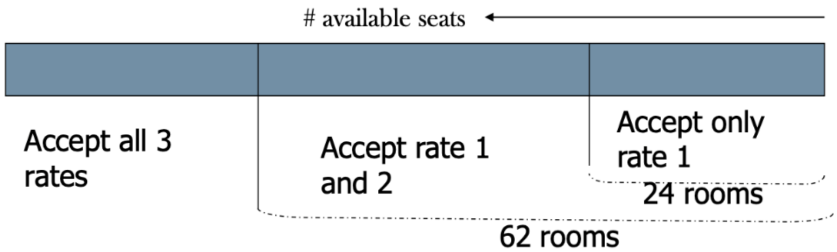
$$\Rightarrow J_1 = 30 - 0.84 * \sqrt{50} = 24.06 \approx 24 rooms$$

Protection level for classes 1 and 2 together

$$F(z_a) = 1 - \frac{40}{90} = 0.5556 \Rightarrow z_a = 0.14$$

$$\Rightarrow J_2 = 60 + 0.14 * \sqrt{130} = 61.6 \approx 62 rooms$$

There's no protection level for the most economic rate



12. YIELD MANAGEMENT - EXERCISES

12.1. EXERCISE 1 – FUTURA SPA

Public administration decided to celebrate renewal of the Giglio theatre with a gala party. The organization of the event is delegated to an important public relation Company named Futura. Futura is in charge of guaranteeing maximum profits allocating the 500 available seats of the theatre in the best way, through yield management theory. The price agreed to sell the ticket is equal to 1000 €/ticket (full price). However, because the probability to sell all tickets at this price is very low, Futura decided to introduce another price. Historical data showed that probability to sell less than 350 tickets is equal to 90%. Demand for tickets (at 1000 €) is distributed as a normal distribution, average is equal to 200. The second price introduced is 320 €/ticket (discounted price). Tickets related to this price must be sold in advance. The manager of Futura knows that if they decided to sell at 320 €/tickets, they would sell all the tickets (500) without any problem. However, they decided to sell tickets at two different prices in order to obtain a higher margin.

To the customer interested in buying tickets at full price, several services are offered:

- Buffet provided by a catering Company, paid in advance (30 days before the event) by Futura. The cost is 15 €/ person.
- Flyer containing the Theatre History. The Flyer is offered by an external sponsor. The cost is 8 €/flyer.
- Program of the evening printed on high quality paper. Cost is 40 €/program.

To the customer that purchases tickets at discounted price less services are offered:

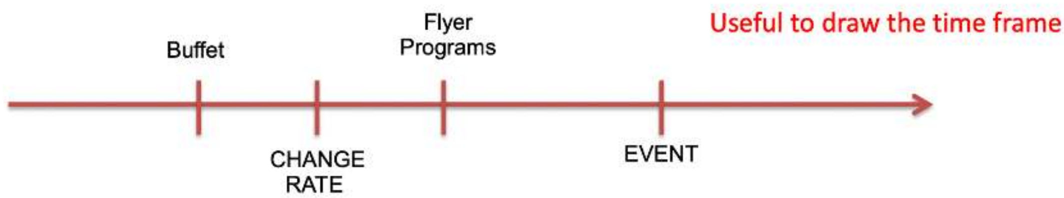
- Flyer containing the Theatre History. The Flyer is offered by the public administration. The cost is 8 €/ flyer.
- Program of the evening printed on low quality paper, each costs 25€.

To print flyer and program, 7 days are needed. Tickets at discounted price are sold until 20 days before the event. After that date it is only possible to purchase tickets at full price.

Capacity= 500 seats
Demand follows a Normal Distribution N(200, ð)
P(X<=350) = 90% → σ = ?
Pf=1000, Pd=320

Costs:

- Buffet = 15 €/p → FULL
- Flyer = 8 €/p → FULL, offered by sponsor
- Program HQ = 40 €/p → FULL
- Flyer = 8 €/p → DISCOUNTED, offered
- Program LQ = 25 €/p → DISCOUNTED



Questions

1. You are required to calculate protection level for full price tickets.
2. For every event, it happens that some people, those who have not purchased tickets before and are not minded paying full price, try to purchase ticket directly at the entrance. 15 minutes before the beginning of the events, managers of Futura check how many tickets have not been sold. So, they decide to sell the **remaining tickets at the price of 150 €**. Data shows that usually no-show related to people that have purchased tickets at discounted price is equal to 10% (The **ticket is not refundable under any circumstances**). You are required to comment on the impact of the introduction of this new price on the protection level.

First thing to do: resume of the costs and timeline.

We always need to specify to which kind of price costs are related to and if the company sustains these costs before or after the event (preventive or final cost).
In a chronological perspective, the protection level (**PL DECISION**) should be defined before all the events and costs of the system, when we are not sure about anything.
Then there is the **EVENT**, which divides the time in PRE-EVENT and POST-EVENT: in this case all costs are sustained by the company before the event.
Sometimes we could have post event costs like the cost of the buffet for only those who participate to the event.

- **Buffet**: FULL, PRE. We pay without knowing the effective number of full price tickets (we will know the number only the day of the event). The actual and effective demand for full price ticket is important.
- **Flyer**: both FULL and DISCOUNTED, PRE. It is provided by another external company, so it not a cost sustained by Futura, we do not consider it.
- **Program**: both FULL and DISCOUNTED, PRE.

1) Protection level definition

Co = company paid for h.q. program and buffet for a seat that remains empty. Moreover, it lost the opportunity of selling that place at a discounted price (with related costs).

→ The day of the event I discover that there are some empty places. Their cost is the overestimation cost, that we should suppose in advance, before the event. Overestimation cost is the lost related to the income that would have come from of the selling with the discounted price (320) + pre costs related to full price tickets already sustained by Futura (costs are positive and savings negative) (40+15) – costs not sustained for discounted tickets not sold (25).

Co = Full + Discounted = (Program HQ + Buffet) + (Discounted price ticket – Program LQ)

Co = (40 + 15) + (320-25) = 350

$$\overbrace{P(X_1 \geq S_1)}^{\text{Probability to have underestimated the demand of full price customers}} * C_u \geq \overbrace{P(X_1 < S_1)}^{\text{Probability to have overestimated the demand of full price customers}} * C_o$$

$$[1 - P(X_1 < S_1)] * C_u \geq P(X_1 < S_1) * C_o$$

$$P(X_1 < S_1) \leq \frac{C_u}{C_u + C_o} \longrightarrow \alpha$$

$$pl = \mu + z_\alpha * \sigma$$

Legend
X1 Full price customer demand
S1 (pl) Protection Level
Cu Cost of Underestimation
Co Cost of Overestimation

Cu = company lost the opportunity to sell a ticket at full price (not sustaining also related service costs) but gained the rate of discounted ticket (and paid for its service costs)

→ The day of the event the seats are full and I think that I could have sold more full price tickets, but I did not, because I sold discounted tickets. Underestimation cost is the lost related to the income that would have come from the selling with the full price mitigated by the selling of discounted tickets (1000-320) + costs not sustained for the full tickets that I would have sold (-15-40).

Cu = Full – Discounted = (Full price ticket – Buffet – Program HQ) - (Discounted price ticket – Program LQ)

Cu = (1000 – 15 – 40) – (320 - 25) = 650

Cu = 650
Co = 350

$$P(X_1 \geq S_1)*C_u \geq P(X_1 < S_1)*C_o$$

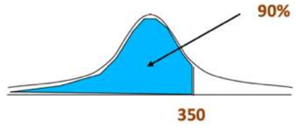
$$[1-P(X_1 < S_1)]*C_u \geq P(X_1 < S_1)*C_o$$

$$P(X_1 < S_1) \leq \frac{C_u}{C_u + C_o}$$

$$P(X_1 < S_1) \leq \frac{650}{650+350} = 0{,}65$$

Use the Normal Distribution Table to estimate S1

Normal distribution, mean = 200
P(X<=350) = 90% → σ = ?

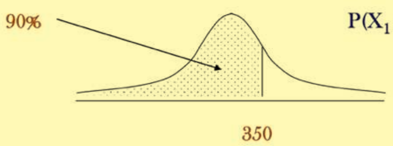


How much is the standard deviation?
Conversion to a standard normal distribution N(0,1):

$$P(X_1 \leq K) = P\left(\frac{X_1-\mu}{\sigma} \leq \frac{K-\mu}{\sigma}\right)$$

First of all, we need to calculate the **standard** deviation of the distribution σ using the formula. This is needed for the protection level formula. Then we calculate **the value of Zα** for the formula of the protection level. We always round up the value of the protection level. At the end we can also define the number of tickets that we sell with the discounted price.

•Probability of demand lower or equal to 350:



Conversion to a standard normal distribution N(0,1):

$$P\left(\frac{X_1-\mu}{\sigma} \leq \frac{350-200}{\sigma}\right) = 0{,}9$$

F(Zα)=0,9
Zα=1,28

$$\frac{350-200}{\sigma} = 1{,}28 \rightarrow \sigma = \frac{150}{1{,}28} = 117{,}2$$

F(Zα)= 0,65
Zα = 0,39

Full price tickets Protection level:

$$S_1 = 200 + 0{,}39*117{,}2 = 245{,}7$$

Company has to reserve 246 seats for Full price customers

From the Normal Distribution Table

How many tickets at discounted price could company sell?
(Capacity – Protection Level)
500 – 246 = 254

2) No-show phenomenon of discounted tickets (10% of probability that who bought discounted ticket does not turn-up). Introduction of last-minute ticket, price = 150 €.

We could have empty seats due to:

- Overestimation
- No-show

Tickets are not refundable under any circumstance.

<u>Last minute tickets</u>: no further service offered than the ticket itself (any service, any cost)

→ The day of the event empty seats and I sell these seats at a very low price to some last minute participants.
Last minute tickets apply for both empty seats and no-show phenomena!
NO SHOW CAN BE ON BOTH CATEGORIES! In this case is just for discounted price category.

Co = company paid for h.q. program and buffet for a seat that remains empty. Moreover, it lost the opportunity of selling that place at a discounted price (with related costs) but sold it at last-minute price.
→ In an overestimation situation, with empty seats, we should subtract the income that would have come from of the selling with the last minute price (last minute ticket would have all been sold because of the very low price, like discounted one) (150). I could have also sold the empty seats to discounted participants and because of the fact that no show is only for the discounted one (10% discounted), I could have sold last minute tickets also to the discounted participants that do no-show (10%*150).

Futura sells empty seats due to:
- Mistake of overestimation: last minute Discounted customers no-shows (10% of probability that who bought discounted ticket does not turn-up)
- Remaining empty seats because not sold to the full price customers

Co = Full + Discounted
Co = 40 + 15 - 150 + (320 - 25 + 0,10*150) = 215

Cu = company lost the opportunity to sell a ticket at full price (without sustain related service costs) but gained the rate of discounted ticket (and paid for its service costs).

Cu = Full – Discounted
Cu = 1000 – 15 – 40 – (320 – 25 – 15 + 0,10*150) = 635

Co = 40 + 15 + (320 - 25 + 0,10*150) - 150 = 215

Cu =(1000 - 15 - 40) - (320 - 25 + 0,10*150) = 635

When company sells a discounted ticket in 10% of the cases it has an additional revenue of 150 € by last minute ticket

Cu = 635
Co = 215

$$P(X_1 \geq S_1) * C_u \geq P(X_1 < S_1) * C_o$$
$$[1 - P(X_1 < S_1)] * C_u \geq P(X_1 < S_1) * C_o$$
$$P(X_1 < S_1) \leq \frac{C_u}{C_u + C_o}$$
$$P(X_1 < S_1) \leq \frac{635}{635 + 215} = 0,747$$

Use the Normal Distribution Table to estimate S1

F(Zα)= 0,747
Zα = 0,67

Full price tickets Protection level:

$$S_1 = 200 + 0,67 * 117,2 = 278,5$$

Company has to reserve 279 seats

From the Normal Distribution Table

How many tickets at discounted fare could company sell?
500 - 279 = 221

In this way the protection level for full price tickets is higher, so better for the company!
Costs that are not differential for PL decision are:
- Cost paid by an external source
- Fixed costs
- Costs to be paid in advance (PRE) for both (together) Full and Discounted fares: for example, a buffet to be paid for all tickets.

These costs will never appear in Co and Cu!

12.2. EXERCISE 2

The organizers of Operations management world conference are excited for the opportunity to organize the best event in the operation management field. So, they want to obtain the maximum profit from the event. They have to pay the rental rate of the conference room, that is able to host 1000 participants, and it costs 80 €/participant. They decide to propose two different prices: one discounted and one full.
With the discounted price (600 €/participant) they are sure to sell out all the tickets, but they want to earn more from the event. So, they decided to sell tickets at discounted price in advance, and to also sell tickets at full price (1200 €/participant). Historical data shown that with the price of 1200 €/ participant it would be possible to sell less than 400 tickets with an 80% probability. Demand for tickets (at 1200 €) is distributed as a normal distribution, average = 250.
They decide to print the proceedings for all the participants (both full and discounted price). That cost 50 €/participant. They have to communicate the numbers of proceedings copies 20 days before the event.
The organizers, also decided to offer a coffee break to all the participants (both full and discounted price), and they have to pay the catering Company 90 days before the event (10 €/participant).
Moreover, they decided to offer a gala dinner to the participants that purchase ticket at full price. They agreed with the restaurant the price of 50 €/participant, and that cost is paid directly by the organizers to the restaurant after the dinner. Historical data shown that on average 70% of the participants (full price) will attend the dinner.
Sales are organized as follow:

- From 80 days before the event until 30 days before the event: tickets at discounted price.
- From 30 days before the event until 1 day before the event: tickets at full price.

Questions

1. You are required to calculate protection level for full price tickets.
2. How the protection level will be modified if they decided to introduce a last minute price of 300 €/participant, to sell during the event day.
3. Calculate the break-even point of the two different scenarios.

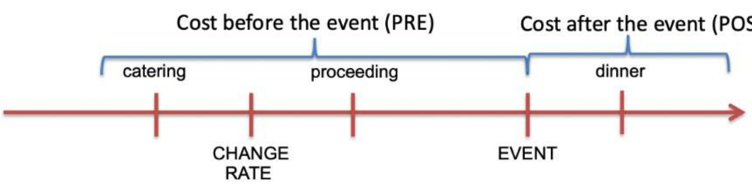
1. You are required to calculate protection level for full price tickets.

Capacity= 1000 seats
Demand follows a Normal Distribution N(200, ∂)
P(X<=400) = 80% → σ = ?
Pf=1200, Pd=600

Costs:

- Rental of the room = 80 €/p → BOTH (PRE)
- Proceedings = 50 €/p → BOTH (PRE)
- Catering = 10 €/p → BOTH, (PRE)
- Dinner = 50 €/p → FULL (POST, only for the70%)

These costs are not differential!



Co = company lost the opportunity to sell discounted tickets. The costs of the two classes are not differential. The only cost indeed refers to the opportunity cost

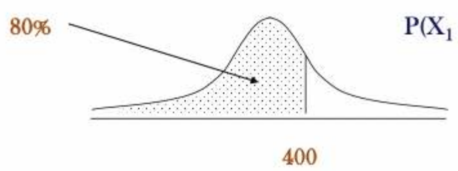
$$Co = \overbrace{80+ 50 + 10}^{\text{FULL}} + \overbrace{600 - 80 - 50 - 10}^{\text{DISCOUNTED}} =600$$

Cu = company lost the opportunity to sell a ticket at full price (without sustaining related service costs) but gained the rate of discounted ticket. The costs are differential.

$$Cu = \underbrace{(1200 - 0{,}7*50)}_{\boxed{\text{FULL}}} - \underbrace{(600)}_{\boxed{\text{DISCOUNTED}}} = 565$$

N.B the cost post the event for full price customers must be considered. Indeed, this cost is really differential because the company does not give to full price customers the possibility to purchase the tickets, thus it does not spend money for their services. Moreover, in case they would have purchased the ticket, only the 70% would have participated to the dinner

• Probability of demand lower or equal to 400:



$P(X_1 \leq 400) = 0{,}8$

Conversion to a standard normal distribution N(0,1):

$$P\left(\frac{X_1 - \mu}{\sigma} \leq \frac{400 - 250}{\sigma}\right) = 0{,}8$$

F(Zα)=0,8
Zα=0,84 → From the Normal Distribution Table

$$\frac{400 - 250}{\sigma} = 0{,}84 \rightarrow \sigma = \frac{150}{0{,}84} = 176{,}57$$

$$P(X_1 < S_1) \leq \frac{Cu}{Co + Cu} = \frac{565}{600 + 565} = 0{,}485$$

F(Zα)= 0,485 → It is lower than 0,5
Use the complementary value of 0,485 in the table
1-0,485= 0,515 → Zα = -0,04

Full price tickets Protection level:

$$S_1 = 250 - 0{,}04 * 176{,}57 = 242{,}93 >> 243 \text{ seats}$$

2. How the protection level will be modified if they decided to introduce a last-minute price of 300€/participant, to sell during the event day.

Introduction of last minute ticket for not sold tickets, price = 300 €
Co = opportunity cost to sell a discounted ticket but at the same time, company sells a last minute ticket for the unsold.

Co = 600 – 300 = 300

300 = LM tickets for empty seats of full price customers

Cu = company lost the opportunity to sell a ticket at full price (without sustaining related service costs) but gained the rate of discounted ticket

Cu = (1200 – 0.7*50) – 600 = 565

N.B. The Cu remains the same since there are no differences in respect to the first situation

$$P(X_1 < S_1) \le \frac{Cu}{Co+Cu} = \frac{565}{300+565} = 0{,}653$$

F(Zα)= 0,653
Zα = 0,39

Full price tickets Protection level:

$$S_1 = 250 + 0{,}39 * 176{,}57 = 318{,}86 >> 319 \text{ seats}$$

3. Calculate the break-even point of the two different scenarios.

Break-even point → **Revenue = Costs**

1)Rental rate = 80000 €
Standard situation
No-shows

2)Rental rate = 66000 €
Standard situation
No-shows

CASE 1 - STANDARD SITUATION

$$\text{Revenue} = X*1200 + (1000 - 243)*600 = 1200*X + 454200$$

Real demand of full price customers | Full customers price ticket | Capacity | Protection Level | Discounted customers price ticket

$$\text{Costs} = 1000*50 + 1000*10 + 0{,}7*50*X + 80000 = 35*X + 140000$$

Capacity | Proceedings | Catering | Dinner | Real demand of full price customers | Rental Rate

Rental Rate = 80000 euros

Break-even point → Revenue = Costs

$$\underbrace{1200*X + 454200}_{\text{Revenues}} = \underbrace{35*X + 140000}_{\text{Costs}} \rightarrow X = -1904{,}24$$

→ X= negative → company is sure to reach the breakeven point

CASE 1 - NO-SHOW PHENOMENA

$$\text{Revenues} = X*1200 + (1000 - 319)*600 + \overbrace{(319\text{-}X)}^{\text{Empty seats}}*300 = 900*X + 504300$$

Real demand of full price customers | Full customers price ticket | Capacity | Protection Level with No-phenomena | Discounted customers price ticket | Last minute price

$$\text{Costs} = 1000*50 + 1000*10 + 0{,}7*50*X + 80000 = 35*X + 140000$$

Capacity | Proceedings | Catering | Dinner for the 70% of full customers | Real demand of full price customers | Rental Fee

Rental Rate = 80000 euros

Break-even point → Revenues = Costs

$$\overbrace{900*X + 504300}^{\text{Revenues}} = \overbrace{35*X + 140000}^{\text{Costs}} \rightarrow X = -421{,}156$$

→ X= negative → company is sure to reach the breakeven point

CASE 2 – STANDARD SITUATION

Break-even point with the new rental rate 600000 €

$$\text{Revenue} = X*1200 + (1000 - 243)*600 = 1200*X + 454200$$

$$\text{Costs} = 1000*50 + 1000*10 + 0{,}7*50*X + 600000 = 35*X + 660000$$

R=C

$$1200*X + 454200 = 35*X + 660000$$

→ X= 177 company has to sell 177 full price tickets to reach the break-even point

CASE 2 – NO SHOW PHENOMENA

Break-even point with the new rental rate 600000 € and No-Shows

$$\text{Revenue} = X*1200 + (1000 - 319)*600 + (319\text{-}X)*300 = 900*X + 504300$$

$$\text{Costs} = 1000*50 + 1000*10 + 0{,}7*50*X + 600000 = 35*X + 660000$$

R=C

$$900*X + 504300 = 35*X + 660000$$

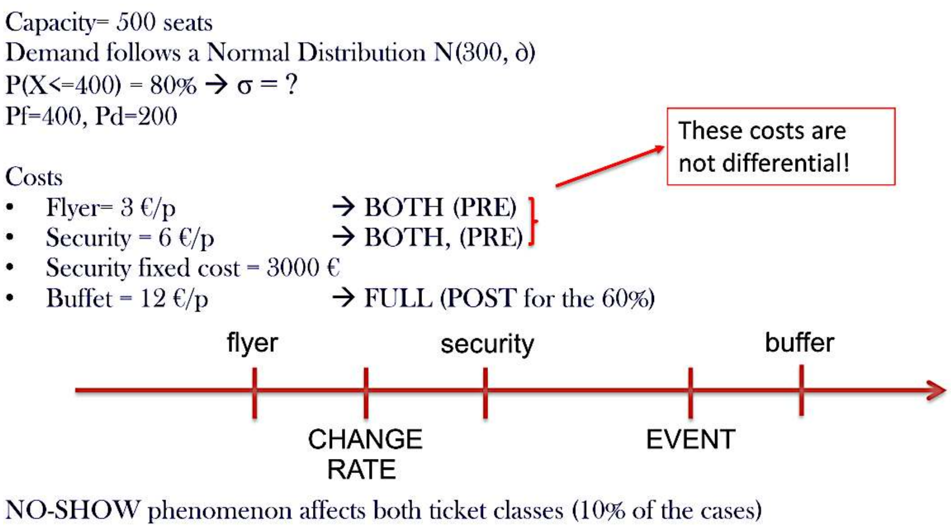
→ X= 180 full price tickets

12.3. EXERCISE 3

The Sales Manager of a Company decides to apply Yield Management for an important sport event. If the customers will book in advance (40 days before the event), they can buy tickets at a special price of 200 €. The manager, based on historical data, knows that if he decides to sell all the tickets at this special price, he will not have any problem to sell all the 500 tickets. But he knows also that the yield management strategy can help to reach more profit from the event. So, he decides to introduce another price, 400 €, from 40 days before the event till the day before the event. The demand for tickets sold at 400 € is distributed as a normal distribution with average 300 and probability of 80% to sell less or equal 400 tickets.
Both the typologies (discounted and full) are characterized by no show phenomena. Data shows that usually no-show related to people that have purchased tickets both at discounted and full price is equal to 10% (The ticket is not refundable under any circumstances).
The manager decides to print a flyer to promote the event. He will print the flyer 2 months and a half before the event. The cost for each flyer is 3 €.
Moreover, he has to hire a security service for the day of the event. He has to communicate to the security service Company (and to pay) the right number of stewards that he needs, 35 days before the event. The cost for this service is calculated as 6 €/ticket. Also, there is a fixed cost of 3000€ for the stipulation of the contract. He decides to offer a buffet after the event, just to the full price customers. It will cost 12 €/ person, and it will be paid after the event. The historical data show that only the 60 % of the full price tickets sold will attend the buffet after the event.

Questions:

1. You are required to calculate the protection level for the full price.
2. The manager promised to his Company to reach at lead a net profit of 45000. The rent for the sport arena is 50000 €. Do you think that the manager will reach this net profit with a probability of 95%?
3. He decides to introduce a last-minute price (175€) to sell just the day of the event. How does the protection level change? Comment on it.



1. You are required to calculate the protection level for the full price.

Co = opportunity cost to sell a discounted ticket (the other costs are not differential)

$$Co = (400 - 200) = 200$$

Cu = company lost the opportunity to sell a ticket at full price (without sustaining related service costs) but gained the rate of discounted ticket.

$$Cu = (400 - 0,6 * 0,9 * 12) - 200 = 193,52$$

0,6 → because only the 60% of the full price customers takes part to the buffet

0,9 → because 10% of no-shows of full price customers → only the 90% will participate and will have the buffet

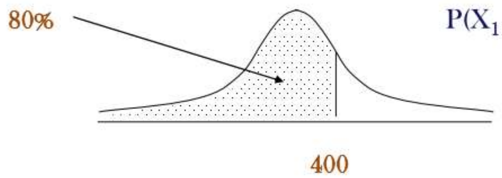
•Probability of demand lower or equal to 400:

F(Zα)= 0,80 → Zα = 0,84 → Use the linear interpolation

$$\frac{y - y1}{y2 - y1} = \frac{x - x1}{x2 - x1}$$

$$\frac{y - 0{,}84}{0{,}85 - 0{,}84} = \frac{0{,}80 - 0{,}79955}{0{,}80234 - 0{,}79955}$$

Y=0,842=0,84



$P(X_1 \leq 400) = 0{,}8$

Conversion to a standard normal distribution N(0,1):

$$P(\frac{X_1 - \mu}{\sigma} \leq \frac{400 - 300}{\sigma}) = 0{,}8$$

F(Zα)=0,8
Zα=0,84

$$\frac{400 - 300}{\sigma} = 0{,}84 \rightarrow \sigma = \frac{100}{0{,}84} = 119{,}05$$

Φ_α	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0.1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0.2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0.3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0.4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0.5	0,69146	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0.6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0.7	0,75804	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78524
0.8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0.9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1.0	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1.1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1.2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1.3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91308	0,91466	0,91621	0,91774
1.4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1.5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1.6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1.7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1.8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1.9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670

$$P(X_1 < S_1) \leq \frac{Cu}{Co + Cu} = \frac{193{,}52}{200 + 193{,}52} = 0{,}4918$$

F(Zα)= 0,4918 → 0,4918< 0,5

Use the complementary value of 0,4918 in the table
1-0,4918= 0,5082 → Zα = -0,02

Full price tickets Protection level:

$$S_1 = 300 - 0{,}02 * 119{,}05 = 297{,}6 >> 298 seats$$

2. Breakeven point

Revenue – Costs = Target Profit

Rent of arena = 50 000€

Profit target = 45 000€

$$Revenues = X * 400 + (500 - 298) * 200 = 400 * X + 40400$$

$$Costs \;=\; 500 * 3 + 500 * 6 + 0{,}6 * 0{,}9 * 12 * X + 3000 + 50000$$
$$Costs = 6{,}48 * X + 57500$$

$$Target\ Profit = 45000 \text{ (from data)}$$

$$400 * X + 40400 - (6{,}48 * X + 57500) = 45000$$
$$X = 157{,}8 \rightarrow 158\ (Minimum\ number\ of\ full\ price\ customers\ ticket)$$

$$P(X_1 \geq 158) = P\left(\frac{X_1 - \mu}{\sigma} \geq \frac{158 - \mu}{\sigma}\right) = P\left(Z_1 \geq \frac{158 - 300}{119{,}05}\right) =$$
$$P(Z_1 \geq -1{,}193) = 1 - P(Z_1 \leq -1{,}193) = 1 - 0{,}117 =$$
$$= 0{,}883$$

The probability to reach profit target is lower than 95%

3. Last minute ticket

Introduction of last-minute ticket, price = 175 €

Co = opportunity cost to sell a discounted ticket (and additional last minute revenue in 10% of cases) but at the same time, company sells a last minute ticket for the not sold tickets

$$Co = (200 + 0{,}1 * 175) - 175 = 42{,}5$$

LM tickets for no-shows of discounted price customers that happens for 10%

LM tickets for empty seats of full price customers

Cu = company loses the opportunity to sell a ticket at full price (without sustaining related service costs) but gained the rate of discounted ticket (for both tickets there is an additional last minute revenue from no-show phenomenon)

$$Cu = (400 - 0{,}6 * 0{,}9 * 12 + 0{,}1 * 175) - (200 + 0{,}1 * 175) = 193{,}52$$

LM tickets for no-shows of full price customers that happens for 10%

LM tickets for no-shows of discounted price customers that happens for 10%

$$P(X_1 < S_1) \leq \frac{Cu}{Co + Cu} = \frac{193{,}52}{42{,}5 + 193{,}52} = 0{,}819$$

$F(Z\alpha) = 0{,}819$

$Z\alpha = 0{,}92$ → From the Normal Distribution Table

Full price tickets Protection level:

$$S_1 = 300 + 0{,}92 * 119{,}05 = 409{,}5 >> 410 seats$$

Φ_α	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0.1	*0,53983*	*0,54380*	*0,54776*	*0,55172*	*0,55567*	*0,55962*	*0,56356*	*0,56749*	*0,57142*	*0,57535*
0.2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0.3	*0,61791*	*0,62172*	*0,62552*	*0,62930*	*0,63307*	*0,63683*	*0,64058*	*0,64431*	*0,64803*	*0,65173*
0.4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0.5	*0,69146*	*0,69497*	*0,69847*	*0,70194*	*0,70540*	*0,70884*	*0,71226*	*0,71566*	*0,71904*	*0,72240*
0.6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0.7	*0,75804*	*0,76115*	*0,76424*	*0,76730*	*0,77035*	*0,77337*	*0,77637*	*0,77935*	*0,78230*	*0,78524*
0.8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0.9	*0,81594*	*0,81859*	*0,82121*	*0,82381*	*0,82639*	*0,82894*	*0,83147*	*0,83398*	*0,83646*	*0,83891*
1.0	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214

Revenue – Costs = Profit

Rent of arena = 50 000€

Profit target = 45 000€

$$Revenues = X * 400 + (500 - 410) * 200 + 0{,}1 * [X + (500 - 410)] * 175 + (410 - X) * 175$$

$$Revenues = 242{,}5 * X + 91325$$

$$Costs = 500 * 3 + 500 * 6 + 0{,}6 * 0{,}9 * 12 * X + 3000 + 50000$$

$$Costs = 6{,}48 * X + 57500$$

$$R - C = Pr$$

$$242{,}5 * X + 91325 - (6{,}48 * X + 57500) = 45000$$

$$X = 47{,}3 \rightarrow 48$$

$$P(X_1 \geq 48) = P(\frac{X_1 - \mu}{\sigma} \geq \frac{48 - \mu}{\sigma}) = P(Z_1 \geq \frac{48 - 300}{119{,}05}) =$$

$$P(Z_1 \geq -2{,}12) = 1 - P(Z_1 \leq -2{,}12) = 1 - 0{,}017 =$$

$$= 0{,}983$$

The probability to reach profit target by introducing last minute ticket is greater than 95%!!

Some examples: How to shape the following systems?

1. The tickets sold for a concert are more than the available places. The day of the concert, some people do not show up and thus some places remain empty. Map the situation and try to think about what you would do.

2. Alitalia have sold all the flight tickets and thus the crew have prepared the meal for all the passengers who have purchased the tickets. Moreover, they have purchased some gadgets for those people flying in the first class, but these gadgets can be reused for the next flight in case of no show-up of some passenger.

12.4. EXERCISE 4 – Marcus Automotive

Marcus, a famous British automaker, is going to propose to the market a new super-car model, limited edition. The company decided to produce only 250 cars of this model.

Marcus, to give exclusivity to its new model, made a contract with the well-known English circuit of "Brands Hatch" flat-rate fee of 700 € / service package, to give the opportunity to the prospective customer to access, on certain dates, the circuit. This means that, when Marcus requires the access to the circuit, he will pay the discounted rate of 700 €. Marcus established that the margin associated to this service is 200 € / package (customers will pay to Marcus € 900 to access the circuit).

Moreover, because Marcus participates to FIA "Le Mans Series" championship, the organizer reserved to Marcus the right to buy the tickets for the opening race in Monza at privileged price of € 60 each. Marcus intends to take the opportunity, because he is sure to sell all the tickets at € 100 each. Finally, the company set to give the car keys in a precious box of brier-root, with the symbol of Marcus, prepared in advance for all 250 cars.

Sales function decided to sell the car and all the services (access to "Brands Hatch" circuit and ticket for the "Le Mans Series" opening) at the discounted price of 150,000 € to customers that book the vehicle ahead of its official launch at the London Motor Show. The production cost of the vehicle, regardless of whether it is sold at discounted or at full price, is 70,000 €.

The price of bookings that will be made after this event will be € 170,000, services not included. Customers will have the opportunity to separately buy the services. It is estimated that the services will be purchased by 70% of this customer base.

It is also estimated that the demand for cars at full price, following the presentation at the show, can be described by a normal distribution with mean 150 and variance 15. After 3 months, the effect of the London Motor Show is finished and unsold cars have to be sold through another channel, which requires an additional cost of 30,000 € / car. The selling price and the opportunity to get additional services remain unaffected.

The Marcus costs structure is specified in the table below.

London Motor Show participation	10.000 €
Stand preparation	50.000 €
Staff (hostess, steward)	15.000 €
precious box of brier-root	100€/box
Opening “Le mans series” ticket	60 €/ticket
Access to "Brands Hatch" circuit	700 €/access
Advertising banner	200 €/banner
registration “Le Mans Series” championship	20.000 €
gadget e merchandising for events	5.000 €

Questions

1. Is it correct set a protection level for after London motor show booking? How many cars? → protection level
2. How much does the protection level change if the distribution of after London motor show demand is described as a normal distribution with mean 150 and probability of sell a number of cars between 110 and 190 equals to 76%?

1. Protection level

Marginal analysis:

$$P(X_1 \geq S_1) * C_u \geq P(X_1 < S_1) * C_o$$

$$[1 - P(X_1 < S_1)] * C_u \geq P(X_1 < S_1) * C_o$$

$$P(X_1 < S_1) \leq \frac{C_u}{C_u + C_o}$$

Co = cost related to have dedicated too many cars to after London exhibition sales, company was not able to sell all cars and has to pay additional costs to sell them through other channels

Cu = cost related to have allocated too cars to discounted rate, so company loses the opportunity to sell cars at a higher price after the London exhibition.

•Co = 150.000 – 60 – 700 + 30.000 – 170.000 – 0,7*[(100-60)+(900-700)] = <u>9.072 €</u>

•Cu = 170.000 + 0,7*[(100-60)+(900–700)] + 60 + 700 – 150.000 = <u>20.928 €</u>

$$P(X_1 \geq S_1) * C_u \geq P(X_1 < S_1) * C_o$$
$$[1 - P(X_1 < S_1)] * C_u \geq P(X_1 < S_1) * C_o$$
$$P(X_1 < S_1) \leq \frac{C_u}{C_u + C_o}$$

$$P(X_1 < S_1) \leq \frac{20928€}{20928€ + 9072€} = 0{,}6976$$

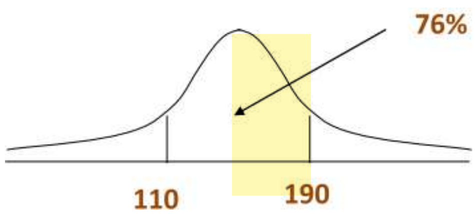
F(Zα)= 0,6976
Zα = 0,52

Full rate cars protection level: $S_1 = 150 + 0{,}52 * \sqrt{15} = 152$

152 cars allocated to after London motor show sales.
98 cars to sell before London motor show at discounted rate.

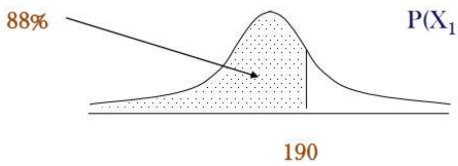
2. Δprotection level

After-exhibition demand is described by a normal distribution with mean 150 and
probability of sell a number of cars between 110 and 190 equals to 76%.



How much is the standard deviation?

The probability that the demand of cars after London exhibition is lower than 190 is equals to:



$P(X_1 \leq 190) = (1\text{-}0{,}76)/2 + 0{,}76 = 0{,}88$

$$P\left(\frac{X_1 - 150}{\sigma} \leq \frac{190 - 150}{\sigma}\right) = 0{,}88$$

F(Zα)=0,88
Zα=1,175

$$\frac{190 - 150}{\sigma} = 1{,}175 \rightarrow \sigma = \frac{40}{1{,}175} = 34{,}04$$

New protection level: $S_1 = 150 + 0{,}52 * 34{,}04 = 168$

168 cars to sell at full price after London exhibition.
Only 82 cars to sell before the event at discounted rate.

12.5. EXERCISE 5 – Astoria-Hotels

The Astoria-Hotels is a chain of luxury hotels.

The Engineer Bianchi, Head of Operations at the Villa Reale Hotel in Milan, has noted that in some days many rooms of the hotel remained empty. It is strange because the hotel receives so many reservation requests that it partially refuses. That is why he asked the engineer Cantu to reconsider the optimal level of overbooking.

Customers of the Villa Reale Hotel can book by phone or internet and at the time of booking they have to pay entire booking price. This amount is not refundable under any circumstances.

There is a single rate of 145 € per room per night, the total number of rooms of the Villa Reale Hotel is 120 and Engineer Cantu estimates that the No-show rate is about 10%.

The cost structure of the Villa Reale Hotel is shown in Figure 1

staff	1.500€/day
breakfast**	15€/meal
newspapers***	3€/room
Bed sheets and towel washing****	5€/set

**breakfast is included in the booking rate (145€). There is not the possibility to have lunch or dinner.

***It is included in the booking rate.

****The set is cleaned only if the room is used. Washing is given in out-sourcing.

FIGURE 1

When a customer that bought a reservation arrives at the Villa Reale Hotel and she finds it full the operators at the reception try to find an alternative solution. In 60% of cases it is possible to find a free room in another hotel, Futura Hotel, belonging to Astoria- Hotels chain, the remaining 40% of cases receptionist should contact to other hotels.

For customers who go to Futura Hotel it is paid the taxi to the hotel (average travel cost of 15 €) and a dinner at the Golf Club Milano restaurant that costs 70 €. In addition, the Villa Reale Hotel has to pay 25 €/day to Futura Hotel.

For customers who go to a hotel not belonging to the Astoria-Hotels chain, Villa Reale Hotel pays the taxi to the hotel (but in this case the average travel rate is 20 €) and offers the same dinner whose cost is 70 €. Usually these hotels belong to same class of Astoria-Hotels. The price of accommodation in these hotels is directly paid by Villa Reale Hotel and it is on average 120 €.

In any case the sum of 145 € for the reservation is not refunded to the customer.

Questions:

1) Compute the optimal overbooking level (number of reservations that should be accepted by the Villa Reale Hotel).
2) How does the overbooking level change considering the cost of loss of image because of the hotel does not guarantee the room that the customer has booked? (Estimated loss of image cost 250 €)

Attached 1

Binomial distribution parameters n and p, n = hotel capacity (120 rooms) e p = probability of success (in this case 0,9 because of 10% of no-show).

Q=120 rooms p=0,9

#acc.book	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145
P(X<=Q)	1	1	1	1	1	1	0,99	0,98	0,95	0,91	0,85	0,77	0,68	0,58	0,47	0,37	0,28	0,21	0,15	0,1	0,07	0,04	0,03	0,02	0,01	0

The number of reservations to accept is related to the probability resulted by marginal analysis.

1) How many reservations should Villa Reale hotel accept?

Marginal analysis adapted to **overbooking**

- **Cu**= cost of underestimating No-show customers
- **Co**= cost of overestimating No-show customers

$$C_u P(Ovb \le NS) \ge C_o * P(Ovb > NS)$$

=>

$$C_u P(Ovb \le NS) \ge C_o [1 - P(Ovb \le NS)]$$

=>

$$P(Ovb \le NS) \ge \frac{C_o}{C_u + C_o} \quad \Leftrightarrow \quad P(X \le C) \ge \frac{C_o}{C_u + C_o}$$

Cu→ a room remains empty

- **Cu** = 145 € - 15€ - 3€ - 5€ = **122€**

Co= company has to assure a room to the customer that booked a room

- **Co** = 0,6*(15€+70€+25€) + 0,4*(20€+70€+120€) – 145€ = **5 €**

$$P(X \le C) \ge \frac{C_o}{C_u + C_o} = 5 / (122+5) = 0{,}04$$

#	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145
P(X<=Q)	1	1	1	1	1	1	0,99	0,98	0,95	0,91	0,85	0,77	0,68	0,58	0,47	0,37	0,28	0,21	0,15	0,1	0,07	0,04	0,03	0,02	0,01	0

→ **Villa Reale Hotel will accept 141 reservation requests**

2) How does the overbooking level change considering the cost of loss of image because of the hotel does not guarantee the room that the customer has booked? (Estimated loss of image cost 250 €)

$$P(X \le C) \ge \frac{C_o}{C_u + C_o} = 255 / (122+255) = 0{,}68$$

#	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145
P(X<=Q)	1	1	1	1	1	1	0,99	0,98	0,95	0,91	0,85	0,77	0,68	0,58	0,47	0,37	0,28	0,21	0,15	0,1	0,07	0,04	0,03	0,02	0,01	0

→ **Villa Reale Hotel will accept 132 reservation requests**

12.6. EXERCISE 6 – Easy-Fly

The Easy - Fly is an airline company that is introducing new air routes to the United States. In particular the management is evaluating the variable rate for the route from Milan to New York. The aim of the company is maximizing the profit with 3 different classes of tickets. The plane can host 500 people. The seats on the plane are undifferentiated, so the class of customer cannot be recognized once a customer has bought the ticket.

Rate	price	From	To
Easy	500 €	Booking start	90 days before the flight
Premium	700 €	89 days before the flight	20 days before the flight
Executive	1000 €	19 days before the flight	Flight day

The available data indicates that the different demands are described by a normal distribution with the parameters in the following table.

	average	variance
Executive Rate	100	70
Premium Rate	250	50
Easy Rate	400	100

The company's management is evaluating to include in the flight ticket sale some services. The company's management does not intend to provide a service of prestige because it understood that the customer perception of the air service has changed over time, customers prefer price over service. For this reason, to keep costs to a minimum, it has decided to include in the service only a blanket and a small pillow. For this service, the company has stipulated a flat contract with a company that provides the washing service for blankets and pillows.
Easy-Fly pays an amount of 2 € / set for expected use of pillows and blankets at the beginning of year for all the quantity you expect to order during the year. The contract is already paid. In addition to this service, but already very common among airline service offers, the management gives to customers the chance to eat lunch for a fee on the plane. For this reason, just before the departure of the flight, once passengers and luggage boarded, airline staff orders to the catering company about 300 meals. This estimate is made based on the experience of the hostesses and no-one has ever studied how many meals are really consumed. The average revenue for this service is 35 € / meal. All revenues belong to the catering company. The airline company does not have earnings from this sale. However, it collaborates on meal demand forecasting because the lunch service must be provided to the passenger.

Questions:

1) How many seats do you allocate to each rate?
2) Analyse and show the impact on the protection level of a decrease in the Easy rate (other things being equal, assuming that a reduction in the Easy rate does not mean a change of rate for any values of the mean and variance).

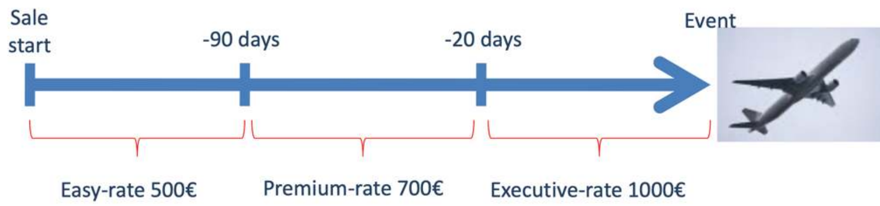
Optional question:

A colleague of the chief operating officer is considering the possibility of using the 3 rates in a different way. In particular, she thinks to sell the Easy-tickets only shortly before the start (last minute ticket), if she realizes that he cannot sell all the seats allocated at the Executive-rate:

Rate	Price	from	to
Premium	700 €	Booking start	15 days before the flight
Executive	1000 €	14 days before the flight	2 days before the flight
Easy	500 €	1 days before the flight	Flight day

With this pricing system he would be sure to sell all the seats without any problem at a price less than or equal to 700 €. The distribution of the demand for places at 1000 € in this circumstance is described by a normal distribution, with mean 100 and 70% of probability that the value of the demand stays in the range between 40 seats and 160 seats. How many seats do you allocate to the Executive rate in this circumstance?

To sum up:
The offer is composed by 3 different rates → Heuristic EMSR



HEURISTIC FRAMEWORK

1) Calculate protection Level for class 1, θ1 (highest class)
2) Calculate protection Level for class 1+2, θ2 (highest class + next lower class)
3) ....calculate the protection level for class 1+2+3+..+n, θn ....
4) The number of seats reserved to the lowest class (the cheapest) isn't determined by protection level formula. It depends on the available capacity.

$$Mf * P(X_1 \geq S_1) \geq Md$$

$$\Updownarrow$$

$$\overline{f_i} * P(D_i \geq \vartheta_i) = f_{i+1}$$

$$\overline{f_i} = \frac{\sum_{j=1}^{i} \mu_j f_j}{\sum_{j=1}^{i} \mu_j}$$

average revenue of i-rate and «more expensive than i» rates

$D_i \sim N$ $\qquad \overline{\mu}(i) = \sum_{j=1}^{i} \mu_j$ $\qquad \overline{\sigma}^2(i) = \sum_{j=1}^{i} \sigma_j^2$

$$\overline{f_i} * [1 - P(D_i < \vartheta_i)] = f_{i+1}$$

$$P(D_i < \vartheta_i) = 1 - \frac{f_{i+1}}{\overline{f_i}}$$

$$F(z_\alpha) = 1 - \frac{f_{i+1}}{\overline{f_i}}$$

$$\vartheta_i = \overline{\mu}(i) + z_\alpha \overline{\sigma}(i)$$

Protection Level for class 1, θ1 (highest class)

Executive rate
Price: 1000 €
Average demand: 100
Variance: 70

$$P(D_i < \vartheta_i) = 1 - \frac{f_{i+1}}{\overline{f_i}} \Rightarrow P(D_i < \vartheta_i) = 1 - 0,7$$

Zα = - 0,52

class 1 Protection Level (Executive rate):

$$100 - 0,53\sqrt{70} = 96 \text{ seats}$$

Protection level θ2

Executive + Premium rates
Weighted average price: 785,71 €
Average demand: 350
Variance: 120

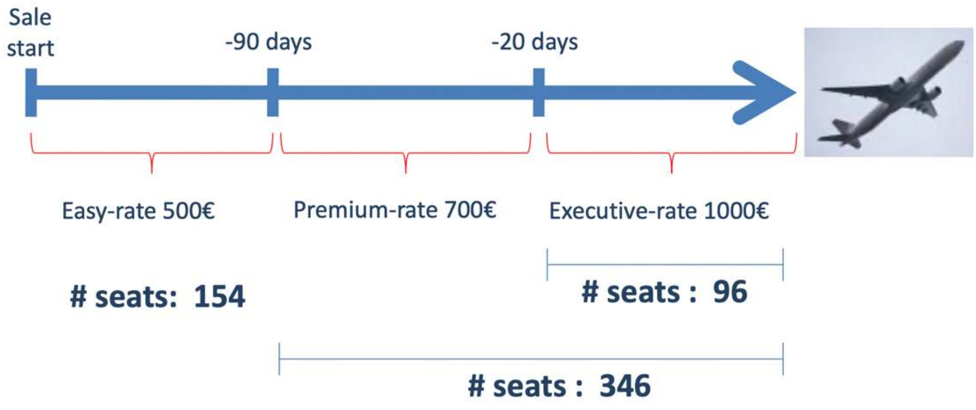
$$P(D_i < \vartheta_i) = 1 - \frac{f_{i+1}}{\overline{f_i}} \quad \Rightarrow \quad P(D_i < \vartheta_i) = 1 - 0,636$$

Zα = - 0,35

Classe 2 protection Level (Executive + Premium rates):

$$350 - 0,35\sqrt{120} = 346 \text{ seats}$$

1) Protection Level



2) Reduction Easy-rate

Assumptions: no changes on average and variance of demands' distributions

$$\overline{f_i} * P(D_i \geq \vartheta_i) = f_{i+1}$$

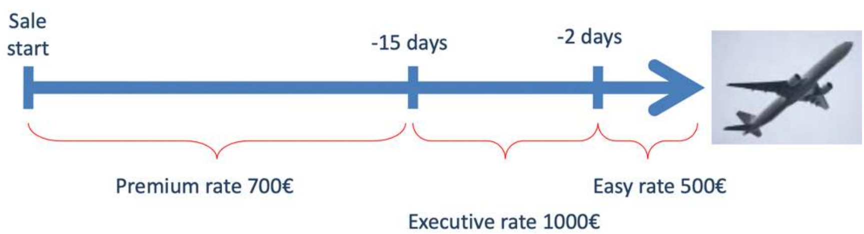
If Easy rate is decreased, what happens to protection levels?

Revenue by Easy ticket is lower.
Expected cost of θ2, protection level 2 (Exec. + Prem.), growth is lower.
θ2 Protection level 2 (Exec. + Prem.) increases.

- Protection level 1, only for Executive tickets, does not change
- Protection level 2, only for Premium tickets, changes (it increases)

3) Last Minute

Company modifies its sales strategy



In this case Easy rate becomes last minute ticket.
It will be used only if company does not sell all Executive tickets
Company has only two main fares: Premium and Executive
Strategic use of Easy rate to saturate airplane capacity

Define protection level for Executive rate

→ In case that all dedicated Executive rate seats weren't occupied, company could sell remaining seats at Easy rate.

It decreases risk, overestimated cost is lower!

Marginal Analysis:

$$P(X_1 \geq S_1) * C_u \geq P(X_1 < S_1) * C_o$$

$$[1 - P(X_1 < S_1)] * C_u \geq P(X_1 < S_1) * C_o$$

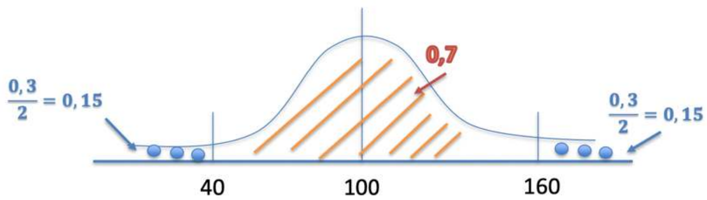
$$P(X_1 < S_1) \leq \frac{C_u}{C_u + C_o}$$

Underestimated cost→lost margin because company sold at discounted price (Premium rate) a place that could be sold at full price (Executive rate)

Cu = 1000 – 700 = 300 €

Overestimated cost→lost margin: a seat reserved to executive rate is not sold at full price because the demand was lower than protection level and it is sold at a last minute rate (Easy) instead of discounted rate (Premium)

Co = 700 – 500 = 200 €



P(X≤160)=85%
Zα=1,04
Dev. std. = (160-100)/1,04
Dev. Std = 57,69

Protection Level

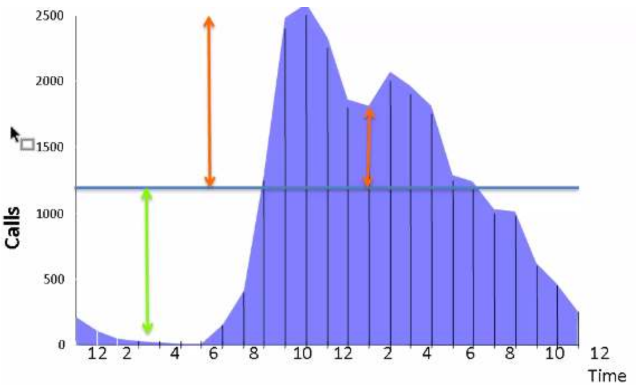
$$P(X_1 < S_1) \leq \frac{C_u}{C_u + C_o} = \frac{300}{300 + 200} = 0,6$$

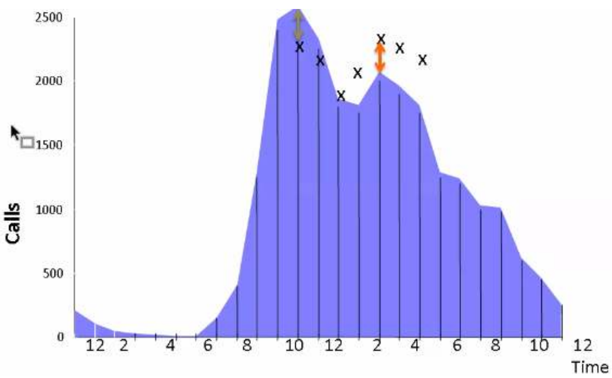
Zα=0,25 -> from the normal distribution table

S_1= 100 + 0,25*57,69 = 114 seats

13. COPYING WITH VARIABILITY AND UNCERTAINTY:

The first graph is showing us the **variability** (green arrow and the red arrow), which is the difference between the average value and the actual value. Higher is the variability higher will be the capacity that we need to put in place. A way to measure variability is with the standard deviation.
Example: sometimes we have shortages, sometimes we have increase of the demand, resulting in stockout or too much stock
The second graph refers to **uncertainty** i.e., the difference between the actual value of the demand and the forecasted one.





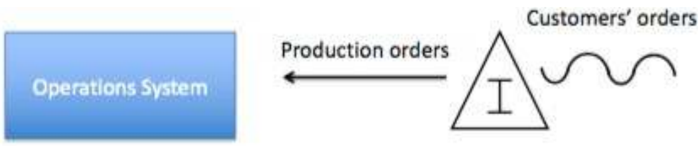
The two phenomena are not really connected, since there is not any kind of relation or proportion, so it is possible to have a high variability and low uncertainty, or high uncertainty and low variability.
How can we cope with these situations?

13.1 MANAGING VARIABILITY

Variability can be addressed (reduced) using 3 strategies:

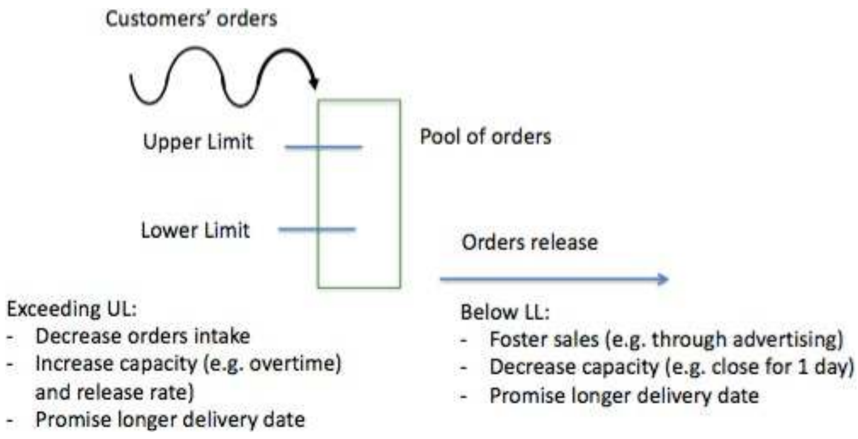
1. **Decoupling demand and capacity (Buffering):** putting a buffer in the middle and decoupling the demand from the market. We separate the offering and the demand. Two techniques:

   a. **Decoupling capacity**: with this approach we put a buffer to make the inventory responds to the customers' order. Anytime the warehouse is going below a certain level, production orders are issued to the production system. In this case the buffer is a PHYSICAL BUFFER.



   b. **Pre-shop pool:** In this case we have customers' orders, we collect the orders and once the orders are received, we put them in a pool that is accumulating them. Then, when we reach a specific number of orders, we release them in an operation system to complete and deliver them. We collect several orders by choosing the proper length of time to create the correct compromise in terms of delivery and terms of mitigation. The average demand rate in the long period has to be the same as the average release rate.
   N.B.: This is not a way to respond to lower capacity. This idea can be implemented in manufacturing companies but also in service industries.

c. **System regulation:** In this case, we still work with the pool but the length in which we collect orders is not pre-determined, but it depends on how many orders we have. If we produce too less, we are not able to respond (lower limit) but also if we get a higher demand is inefficient as well (upper limit).



How to respond when we get out of the comfort zone (out of these limits)?
If we are exceeding the upper limit, we can reduce the orders or we could deliver just a part of the orders, we could go with some temporary increase of the capacity, or we can postpone the delivery date. In case we have a lower limit, we are not able to saturate our capacity, in this case we can increase the demand by increasing the advertising or we can increase the capacity (instead of 6 days we work 5 days). Another way is to wait for collecting more orders in order to reach a position where we are over the upper limit in order not to sustain higher cost for the production.

In those two cases (pre shop pool and system regulation), the buffer is not physical,because it is related to time (the buffer can relate to a physical aspect or to a non- physical, like time).

2. **Managing capacity:** In order to manage the capacity, we can use the capacity management by extending the capacity in order to absorb variability. Critical aspects for capacity management:
   - **Flexibility**: By adding capacity you are flexible (in time) because the time spent in which you are able to increase or decrease your capacity. Flexibility is also measured by the minimum size of change.
   - **The minimum size of the change**: Higher (the cost) is the set-up, the minimum size of change is higher.
   - **Minimum time the change lasts**: when you are changing the capacity you are not able to produce therefore u lose and then u should be able to recover, it is a matter of compromise.
   - **How much in advance the demand is known**: If we are able to have a forecast in the longer period we can prepare for the change, but if we are not able it's difficult, because our change time is lower than the cycle of our change, this means that we could be too late.

   How can we cape with these issues?
   - **Increase customer participation**: if we increase the part of the activity that is performed by the customer by its own, we are increasing the capacity. It is a very cheap capacity, but it is not easy to create, because the customer is not prepared and they are heterogeneous, so they increase the whole system variability.
     Example: Ikea founded its success on this, designing products that are assembled by the customers themselves
   - **Capacity sharing**: making part of my capacity available to other companies in period of low demand

- **Employees cross-training**: if people are able to do different tasks, they will be more flexible and can move from an activity to another when necessary.
- **Part-time employees**: they work only when it is necessary, feasible only if the demand is enough predictable
- **Daily shifts scheduling**: we start from the demand forecasting and we understand the shape of the capacity that we need and calculate the minimum need for each day of the week. We identify the different types of possible shifts according to national contracts and schedule them to minimize the total working time (through Integer Linear Programming) and finally assign people to shifts.

3. **Managing demand**: In order to manage the demand, we use the demand management through the development of complementary services, through the segmentation of the demand and price incentives.
   - **Demand segmentation**: you can create clusters of customers with homogeneous behavior, so supports the company on understanding customers' behavior and in this way we reduce the variability.
     Example: orthopedic clinic, we have doctors and customers that complain about excessive workload and excessive waiting. Average daily visits: 200. Variability: 150. To reduce the variability, we create two clusters of customers: planned visits and urgent visits. The resource is the same, but the needs are different. We move from a flat distribution of planned visits, concentrating it in the end of the week, leaving free the week for urgent visits. The result is that the total demand is flatter.
   - **Price incentives**: we change the prices according to the demand of the specific period, in order to make it flatter. For example, we reduce the price when the demand is lower.
   - **Promote demand in low season**: looking for new services to offer.
   - **Development of complementary services:** to do more and better businesses, to exploit synergies and reduce the effects of seasonality.
   - **Development of reservation system**: there are advantages for both parties. The customer has a reserved seat, and the company has an early sale.

To sum up, management of variability can be managed med by 3 strategies: decoupling the demand, managing the capacity and managing the demand.

13.2 MANAGING UNCERTAINTY:

We can use 3 strategies:

- **Additional decoupling between demand and capacity**: The additional decouple attack the cause of uncertainty and find an explanation for the deviations from the forecasted value; the additional decoupling of demand and capacity transforms uncertainty into variability.
- **Capacity buffer**: always available (for example the doctor who receives a call he is always available)
- **Manage delays in deliver**: Way to absorb the buffer time.

These strategies are in the short term, in case of long term, uncertainty is never given and we should work in a way to reduce uncertainty by (e.g.) trying to explain, to understand and reduce the level of uncertainty letting it become variability

14. SHARING ECONOMY

Sharing Economy is "a non-reciprocal pro-social behaviour".
Benkler (2004)

"Sharing Economy is the act and process of distributing what is ours to others for their use, as well as the act and process of receiving something from others for our use".
Belk (2007):

The traditional paradigm consisted in companies that owned the resources and the customers bought the products.

In the 90s' we moved from this paradigm, with the asset that is given to the customer for a certain period of time, without the change of ownership.

After the financial crisis of 2007-2008 the sharing economy arose, resulting in a change in both customers and service providers:

- A **change in the customers**: People experience financial difficulties since then, they would re- evaluate their consumption patterns and the value of the ownership. Customers behaviour.
- A **change in service providers**: Individuals who lost their jobs needed to find new job opportunities and companies such as Uber offered an attractive alternative

Generally speaking, there was **change in the value system**, since the value is perceived differently and no more depending on the ownership of the asset. This was also made feasible thanks to the **technologic evolution** and the spread of **internet**.

They **link customers in need** of a service **with the suppliers** of that service

- typically simple, single-stage services.
- Temporary access / non – ownership

They **operate through digital platforms** such as a website or a mobile app

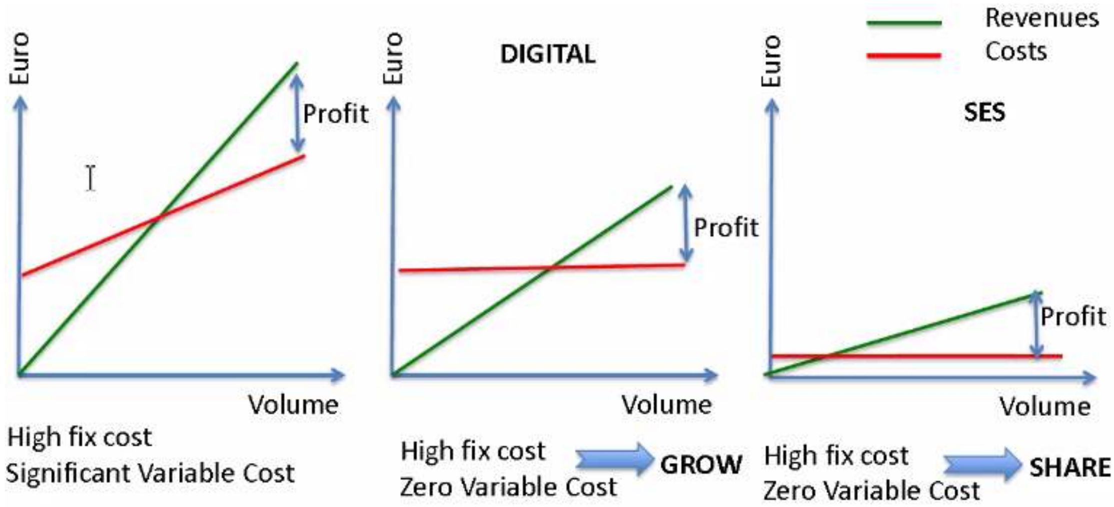
- Online platform, intermediary role
- Easy accessibility of both of service and platform

They offer services **based on idle physical assets** and/or labour.

- Large flexibility in managing their capacity.
- Idle Capacity (under- utilized resources)
- Capability of operating at near zero marginal cost

They are **based on social network** (not the app)

- Peer-to-peer connection
- Trust and network-based activity
- Sustainability (no additional resources)
- The service suppliers are typically amateurs



- **Traditional paradigm**: there are fixed and variable cost. For each product that we produce there are additional variable cost. If we want to be profitable revenues have to cover both fixed and variable cost.
- **Digital advent**: The cost structure of digital service (e.g. software), is composed by fixed cost (initial development of the digital service), the scaling up of the digital service has no cost for the company.
- **SES:** It cut off the initial investment. They put the threshold of the risk at a very low level compared to the digital advent.

In the digital and sharing economy, the variable costs are very low because for Microsoft there is a very low difference in costs for selling one licence or one million, so the slope is quite flat. The difference is that in the sharing economy the fixed costs are very lower than in the case of digital economy, since we do not own the assets.
Example: Airbnb vs Hotel
The initial amount of investment of a hotel is different from the Airbnb one. Because the investment in the assets of the Airbnb apartment has been already done. I share my assets to exploit the moment in which the asset is idle. I transform an idle asset into a shared asset.

14.1 SHARING ECONOMY DEFINITION

Definitions:

"We define access-based consumption as transactions that may be market mediated in which no transfer of ownership takes place"
Bardhi & Eckhardt, (2016)

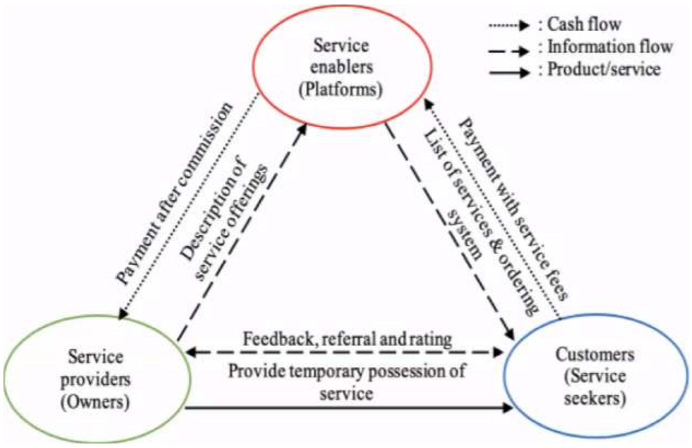
"The sharing economy is an economic platform based on principles and models of economic, social and environmental impact".
Frenken & Schor, (2017)

14.2 SHARING ECONOMY CHARACTERISTICS

What are the characteristics of sharing companies?

- They meet a customer's need with a supplier of the solution for that specific problem
- They operate based on a platform that meets potential customers and potential providers. The success is based on the strength of the soil network, which means peer-to-peer connection, word of mouth, references. Service suppliers are typically amatory
- SES characteristics:
  - **Network effects**: the more providers there are, the more convenient is the service for the customers and also the quality of the service is better. Vice versa, the higher is the number of customers, the more convenient is for the service providers.

- **Economies of scale**: It takes a large investment to develop a digital platform, but once the platform is ready, it can be made available to additional customers without incurring any additional cost.
- **Switching cost**: it is very accessible, easy both to enter and exit for both customers and suppliers. There is problem related to the higher difficulty to retain customers

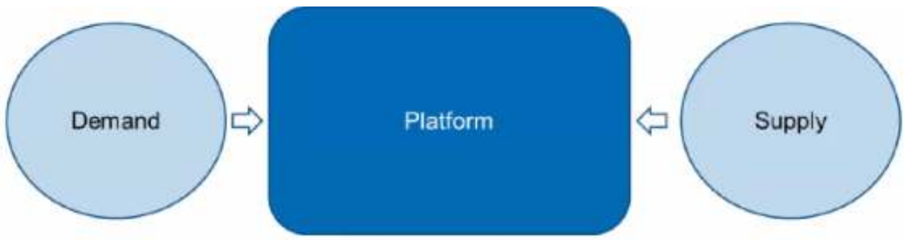


We still have the **customers** and the “**single” supplier** (the person who owns the apartment). Moreover, there is the **service enabler** (Airbnb platform). In many cases the service provider and the customer do not communicate. The customer communicates with the service enabler.

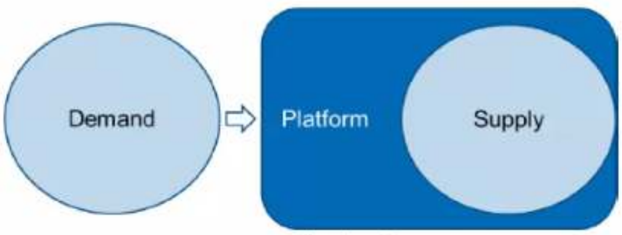
14.3 SHARING ECONOMY CONFIGURATION

We can have different configurations:

- **Peer-to-Peer**: there is a clear distinction between suppliers and customers, with the platform that behaves only as an intermediary.



- **Business-to-peer**: in this case there is not a net distinction between the suppliers, which are integrated within the platform, which collects all the offering.



Complex services mean complex operations that support the service. The challenges in this kind of companies are different from manufacturing or traditional service companies. Different configurations:

- **ON-demand rental network**: The platform is the owner of the assets. *Enjoy* owns the cars. Characterized by no advance booking, resources are spatially distributed and resources rented from one location can be returned to another.

- **P2P resource sharing** (ex. Airbnb): This configuration is characterized by many buyers and sellers, the supply side is not distinct from the buy side, supply stimulates demand and vice versa and A resource unit can sustain the consumption needs from more than one consumer
- **On demand service platform** (ex. Uber): the platform is not the owner of the assets. The supplier is the owner of the assert. Uber does not own the cars. The driver ("single" supplier) buys the car. This configuration is characterized by the capacity that affects demand and vice versa, capacity that can be controlled only indirectly via wages and prices and both capacity and demand vary temporally and spatially very fast.

14.4 CLOUD COMPUTING AND CLOUD MANUFACTURING

It is a trend that consists in companies that make part of their capacity available on the platform. There are some criticalities:

- Intellectual property: it is easier to be copied
- Environment: relationships usually are not just one shot in the traditional paradigm. Now it is changing since there is more competition.

The impact of this trend on the B2B market is:
For simple services, which survives on big volume, the traditional paradigm is changing of high volume and low variety. In fact, traditionally the company must be clearly positioned within the market (high volumes/low variability or low volumes/high variability). The new paradigm exploits the idle capacity of the suppliers, to deliver to the final customer a variety of production/customization, keeping in any case a high volume, because the inefficiencies of high variety are no more concentrated in a unique production system, but distributed on many prod systems. So the impacts are lower thanks to the exploitation of idle capacity. It commutes the request of the market in a new frontier, which is a huge opportunity for small and innovative companies, because a small company can enter in a market in an easier way (lower barriers to entry), being able to compete also in the high segment of market, which is maybe occupied by other big players. The platform collects the variability of the small different firms that allow to compete against the big players.

Advantages:

- Agility
- Follow technology/demand
- Scalability
- High quality of services
- Continuity of the service
- Capex becomes opex

15. SYSTEM PHYSICS

It was introduced by some physicians, using some quantitative approaches and applying some physics theories to develop a very structured approach for lean management. We start considering basic concepts.

15.1 PRODUCTION CAPACITY IN TERMS OF VOLUME AND MIX

Cycle time = the time a resource takes to produce a certain output. It is possible to measure it in different ways and to focus on different levels (single resource, department, company).
At single resource level, the fundamental element is theoretic capacity (e.g. Unit/hour).
What is the capacity of the system in theoretic situations? Theoretic capacity refers to the conditions when everything goes well.
Eg. Unit cycle time (C/T): 6 min, theoretic capacity ($\frac{1}{C/T}$): 10 unit/hour
To determine the actual capacity, it is necessary to also take into account other factors.

15.2 AVAILABILITY

A resource may be not available to produce. Part of working time during which the resource is available for working is named Availability. (Its symbol is A). Availability is never at 100%. Causes for unavailability can be:

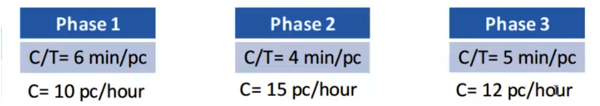
- Breakdowns
- Interruptions (for problems, of for calls)

Actual Capacity = Ct * A Actual capacity (C) = C_t * A

For example, if availability is 80%, a resource with a theoretic capacity of 10 units/hour, in the mid/long term will produce only 8 units/hour. It is the effective time that you have excluding break downs and other drops of other organizational issues. It is a % of the theoretical capacity.

15.3 SERIAL SYSTEM

What is the capacity of the system?



We start considering this simple system, where each phase has a different C/T, that is a measure of its capacity. In a system like this, we do not have any indication about stocks. We should consider what is the worst C/T, that is the worst rhythm that affect the whole production capacity of the system.
→ In this case it is 6 min/pc. It is as strong at its weakest station, that is the longer one.

C= 10 pc/hour C= 15 pc/hour C= 12 pc/hour

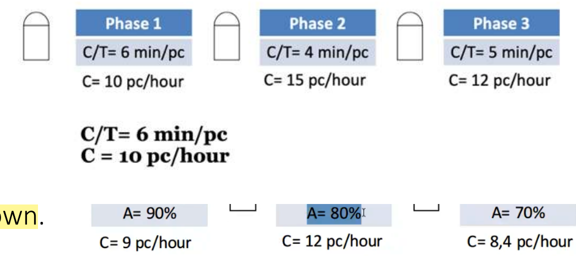
C/T= 6 min/pc
C = 10pc/hour

Capacity is limited by the slowest phase C/T = 6 (medium-term viewpoint) C = 10 units/hour.

Categories of system:

1. Decoupled serial system

In this case we have stocks between each phase. Each phase can work at its own rhythm without any problem. By respecting the pace of their own phases, stocks are going to be accumulated between phases.

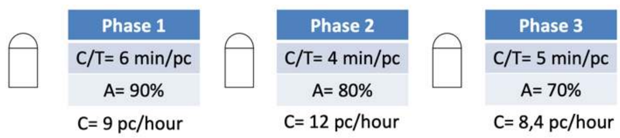


So stocks are used to make each resource work alone, on its own.

- System capacity is 10 u/h.
- Queues between each phase do not balance out different C/T NB: Availability can change system capacity and the bottleneck.

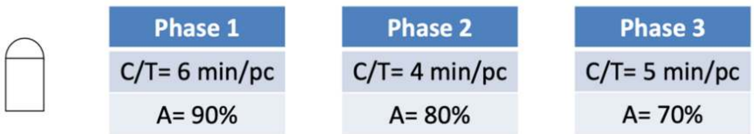
You are trying to protect by having a certain buffer upstream. These buffers are served as decoupling the phases. Stations are separated by the buffers.
The capacity of the system in this case is the same. The presence of a buffer does not impact on the capacity, because the buffer works on the **short term**, but at the end of the day so on the **medium-long term**, the capacity is the same at it is still determined by the weakest. You can treat each phase as its own system.

Actual capacity (C) = C_t * A



Actual capacity (C) = C_1 * A = 8.4

2. Coupled serial system (line) We don't have stuck, they're dependant No Queues



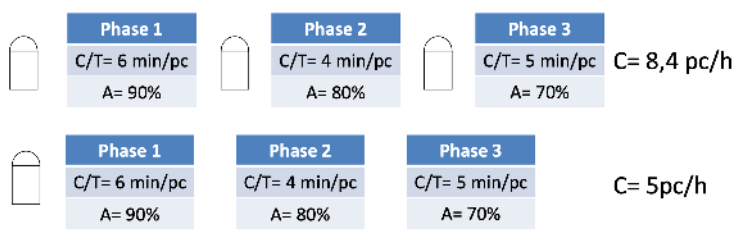
Phases work together, respecting the same schedule, so they are no more independent as in the previous case. Each time a phase stops for any reason, all the production is interrupted.

We have different availabilities and C/T. Each time that one machine stops also the others stops, so we have a joint probability.
Now no longer are queue and buffers → Queues disappear.
The availability of the whole system is given by 90% * 80% * 70% and C/T is defined by the slowest C/T (bottleneck) between the different phases, which is the one that affect the work of the other phases, in this case it is phase 1.

- Queues disappear
- System availability is the product of the single availabilities A = A1*A2*A3 = 0,5

System C/T is the largest C/T. Capacity is therefore 60 * 0,5/6 = 5 u/h
Why this system? What are its advantages?
There are no stocks: a stock is a queue, because stocks are waiting to be processed. If we do not have stocks, we do not have queues inside the system and waiting time, so maybe the throughput time is lower.

15.4 DECOUPLED VS COUPLED SERIAL SYSTEM



Production changes from 8,4 pc/h to 5 pc/h

It seems like queues/stocks bring an improvement, but it's not true:
They are just an increase of waste

Advantages: no stock between the phases (no queues, no waiting time, -possible- lower throughput time).

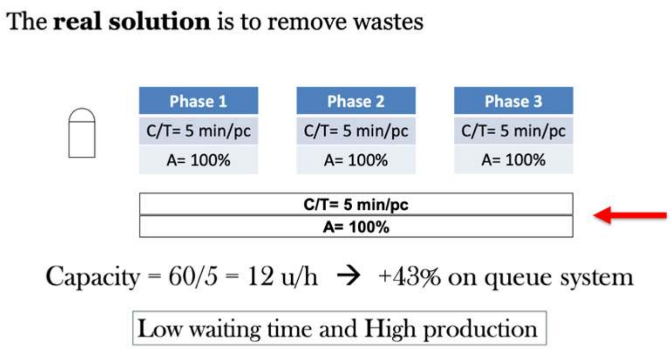
If you increase the availability and you will remove wasted, and so inefficiencies, by removing buffers, you acted on reducing waste in time, money. If you act on the efficiency will be increasing the actual capacity and here we have low waiting time and high production.

Produce more in a shorter amount of time
[ABB from 5 days to 0.9 days]

Adding queues
If we add queues to the previous system, we choose the worst one as the bottleneck of the system (the worst C/T) and the production capacity of the decoupled system becomes 8.4 u/h, so it is higher than 5 u/h.

- It seems like queues/stocks bring an improvement, but it is not true: they are just an increase of waste!
- Eliminating stocks, we can significantly decrease the waiting time, being more efficient.

3. Ideal coupled serial system



With this system we can reach a capacity of 12 u/h, more than any previous system. Going towards an ideal system, the coupled system becomes more effective and the same happens for the decoupled system, reaching both 12 u/h (there are no differences between the ideal coupled and ideal decoupled systems).
If we balance the C/T (5 for all phases), we improve the availability (100% all phases), we eliminate the activity related to inefficiencies and we improve all the parameters of the system, the capacity of the system will tend to ideal, both for coupled and decoupled systems. There will be less waiting time and lower throughput time.
→ Decoupled system is better for throughput time, but to adopt it we have to work a lot on the inefficiencies.

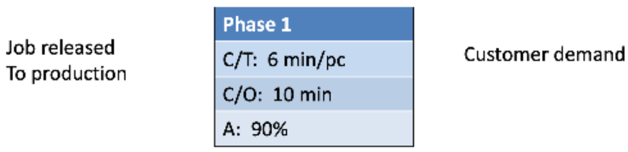
15.5 SET-UPS AND SYSTEM FLEXIBILITY

Set-up is not included in availability, but it is considered as an inefficiency of the system, because it is a non-value added time. We have a system with these characteristics:

- The system works 8 hours/day and daily demand is 60 units (constant in volume).
- The system has a range variety of 10 products, and they are all requested every day
- Is there any problem in delivering to order the requested quantities with a delivery time of 1 day?
  Tp: Time to produce (only to process)
  Tp = Average demand (pc / day) * (C/T) / A
  Tsu: Time to Set-up
  Tsu = Number of set-ups (set-up/day) * C/O
  Ta: Available time (daily)

change over time

Example 1:
The system works 8 hours/day and daily demand is 60 products (constant in volume).
The system has a range variety of 10 products (variants), and they are all requested every day.
Is the company able to deliver the requested quantities within 1 day?



They are cycle time and the change over time. I take 10 min to change from ....
EPE = every part every day

- **Tp** = Time to produce (only to process)
- **Tp** = $\frac{\text{average demand } (\frac{\text{pc}}{\text{day}}) * C/T}{A}$
- **Tsu** = Time to setup
- **Tsu** = number of set-up per day $(\frac{\text{pc}}{\text{day}}) * C/O$
- **Ta** = available time (daily)

The **feasibility analysis** is the **Tp + Tsu ≤ Ta**

Tp + Tsu = 60pc/day *6 min/pc /0,9 + 10 setup/day*10 min/setup = 500 minutes → 10 setups because of 10 variants

Ta= 480 minutes

$$Tp+Tsu>Ta$$

[Feasibility condition]

The company cannot satisfy market demand in make to order, thus it must intervene on the production process in order not to lose profits! What can the company do?

15.6 BATCHING

The company does not produce the variety requested one day, but with a wider window of the demand. In order to have more time to produce, companies usually batch production of each single product.

Example: Doing set up each day, if we have to produce 60 units for 5 days a week, we will produce 300 units weekly and if we have 10 set ups per day, in a week there are 50 set up. The quantity to produce remains the same (60 → 300), but if I consider a weekly demand, the time to set up is lower. The same can be done for 1 month.

Definitions of the 2 types of batching:

- **Quantity Batching** defines the exact (sometimes minimum) quantity to produce every time a product is to be produced.
- **Time Batching** defines the times a product is produced over a defined time horizon (e.g. once/week)

If Time batching is used, and each product is realized every two days (instead of being produced daily, so 5 different products/day instead of 10):

5 set up every day

Tp + Tsu: 6/0,9*60 + 5*10 = 450 minutes

Ta = 480 minutes

In order to have more time to produce, companies usually **batch production** of each single product. That implies working with decoupled systems:

- Advance of unrequested production
- Having inter-operational/finished products stocks
- Increase of system response times
- Increase of throughput time variability, system congestion (need of space) and coordination needs (between product types)

Batching creates problems

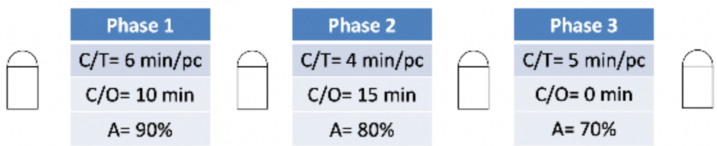
- There are stocks (need for space, costs increase...)
- Production in advance of demand not yet requested!
- Risk to keep stocks of not requested units! (For example, when demand has variability)
- Flexibility decreases (if today code F is requested and it is not in stock nor scheduled for today's production, it could be produced only adding a rush order that creates inefficiency)
- Increase coordination and supervision needs in production process

Reducing the setup time: Batching is THE way companies adopt, even if it entails some disadvantages and it is NOT FOR FREE!
It is a lot better to pursue the reduction of set up times!! (Reducing set up time from 10 to 8 minutes, it is possible to produce every day each product in the requested quantity!)

Example 2: minimum batching
The system works 8 hours/day and daily demand is 60 products (constant in volume).
• The following stocks decoupled the different phases.
• The system has a range variety of 10 products, and they are all requested every day. Each machine produces 10 different products (with setup at every product change).

Decoupled system. The minimum batching is a measure of the flexibility of the system (KPI). It is calculated for each production unit, which are decoupled. The final objective is to improve the production reducing the quantity of the batch, so producing more frequently with a lower quantity.

What is the minimum batching for each of the 3 phases?



To calculate the minimum batching, it is necessary to understand how many setups can be done in each phase every day.

Phase 1:
In order to calculate the minimum batching, it is necessary to understand how many sets up can be done in each phase every day → How many set-ups may we do per day?
6/0,9*60 (min/hour) + X*10 = 480 min → **8** setup per day so I can produce only 8 variant per day.
The **min batch quantity** for each product is: how many products for variant?
Total demand/number of set-up per day → 60/8 = **7,5** pc/batch
(I cannot produce batches with a lower quantity, it is the limit of the system)

Phase 2:
How many set-ups may we do per day?
4/0,8*60 + X*15 = 480 min → **12** set-up per day. (That is, if orders are grouped per day, there is no need to produce more than the ordered quantity). The concept of minimum batching can be applied, but it is not important, because in this phase the company is able to produce the quantity and variety asked by clients.
The **min batch quantity** for each product is:
Total demand/number of set-up per day → 60/12 = **5** pc/batch (total demand/pieces for batches)
NB: if orders are grouped per day, there is no need to produce more than the ordered quantity

Phase 3:
There are no constraints on the batch, because there is no set up (C/O = 0). I can work with batches of 1 unit (A,B,C,D,F and so on repeating).

<u>Serial coupled system</u>.
The system works 8 hours/day and daily demand is 60 products (constant in volume).
• The following stocks decoupled the different phases.
• The system has a range variety of 10 products, and they are all requested every day. Each machine produces 10 different products (with setup at every product change).
If one phase stops for setup, so do the others!

It is therefore important to understand how much time the line is stopped to make setups: total time of stop for the line is the total time of stop for set-ups.

In the case that for each product change corresponds a set-up at each phase, when a phase does the set-up, so do the others.
The time for a set-up is 15 minutes (longest set up time, if set-ups are performed in parallel. It is the sum of set-up time if setups are performed sequentially. E.g. by one single person)

What is the minimum batching for each of the 3 phases?

Phase 1	Phase 2	Phase 3
C/T= 6 min/pc	C/T= 4 min/pc	C/T= 5 min/pc
C/O= 10 min	C/O= 15 min	C/O= 0 min
A= 90%	A= 100%	A= 95%

What does it mean that the system is coupled?
If one phase stops for setup, so do the others!
It is therefore important to understand how much time the line is stopped to make set-ups:
total time of stop for the line is the total time of stop for set-ups.

C/T= 6 min/pc
C/O= 15 min*
A= 85,5%

*longest set up time, if setups are performed in parallel. It is the sum of setup time if setups are performed sequentially. E.g. by one single person

How many set-ups may we do per day?
6/0,855*60 + X*15 = 480 min → **4** set-up per day
The **min batch quantity** for each product is:
Total demand/number of set-up per day → 60/4 = **15** pc/batch → **minimum average batch.**

Example 3: minimum batching

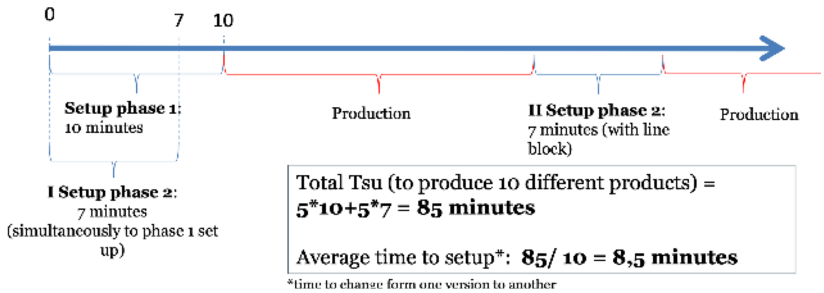
The system works 8 hours/day and daily demand is 60 products (constant in volume).
• The system is coupled (no stocks between stages)
• The system has a range variety of 10 products, and they are all requested every day.
• There are 5 different product types at phase 1.
• At phase 2 there is a further differentiation leading to 10 different product types.
What is the minimum batching for each of the 3 phases?

Phase 1	Phase 2	Phase 3
C/T= 6 min/pc	C/T= 4 min/pc	C/T= 5 min/pc
C/O= 10 min	C/O= 7 min	C/O= 0 min
A= 90%	A= 100%	A= 95%

<u>Coupled system.</u> Each time I have to do a set-up in a phase, also the other phase does set-up with its rhythm (there is a parallel set-up) and who finishes before waits for the other. Then we move to production and after some time, there is another set-up needed for phase 2 (7 min). The cycle continues. We have to do this for each product, so 5 times for phase 1 and 10 times for phase 2. Each time I have to do a set-up in one phase, the other phases wait.

At system level:

- 5 setups at phase 1
- 10 setups at phase 2
- How much is the stop time for the line because of sets up?



C/T= 6 min/pc
C/O= 8,5 min
A= 85,5%

System:
How many set-ups may we do per day?
6/0,855*60 (min/hour) + X*8,5 = 480 min → **7** set-up per day
The **minimum batch quantity** for each product is:
Total demand/number of setup per day → 60/7 = **8,6** pc/batch*
*It corresponds to the average demand of 1.43 days for each type of finished products

Setup and system flexibility

- Production change requires set-ups.
- The longer the set-up time is, the less frequent the production changes will be.
- The longer the set-up time is, the greater the risk of producing something that is not requested.
- Set-ups have impacts on performances, increase costs (stocks, necessity of coordination…), increase throughput time variability and decrease system flexibility.

Production capacity:

- Theoretic production capacity indicates the number of products realized in the time unit that the resource is able to produce under optimal conditions.
- Availability takes into consideration resource's stops that decreases the theoretic production capacity.
- Setups decrease the time available for production and system flexibility. Batching is used to increase single resource's efficiency but it creates problems at a system level.

Decoupled VS coupled serial system:

- A system decoupled by stocks allows to make production phases more independent, but it is a system that requires stocks to work, thus showing that is a system with problems.
- A serial coupled system links the production phases. This highlights problems and allows to see them and address them.
- The objective is to remove problems (remove the causes of the problems), so that the system improves.

16. SYSTEM PHYSICS – EXERCISES

16.1 EXERCISE 1 – WASHING SPA

Washing Spa is a manufacturing company producing household appliances. Its production system is composed by 3 production stations decoupled by stocks, whose main characteristics are highlighted in the table below.

	Station 1	Station 2	Station 3
Cycle time	22 sec/pc	25 sec/pc	24 sec/pc
Setup time	10 min	5 min	5 min
Availability	80%	100%	90%

The daily demand is equal to 2000 units. The facility is open 220 days per year, three shifts per day of seven hours each. The quality control at the end of the line intercepts defective units, which are equal to 5% of the processed ones. Defective units are scrapped (discarded).

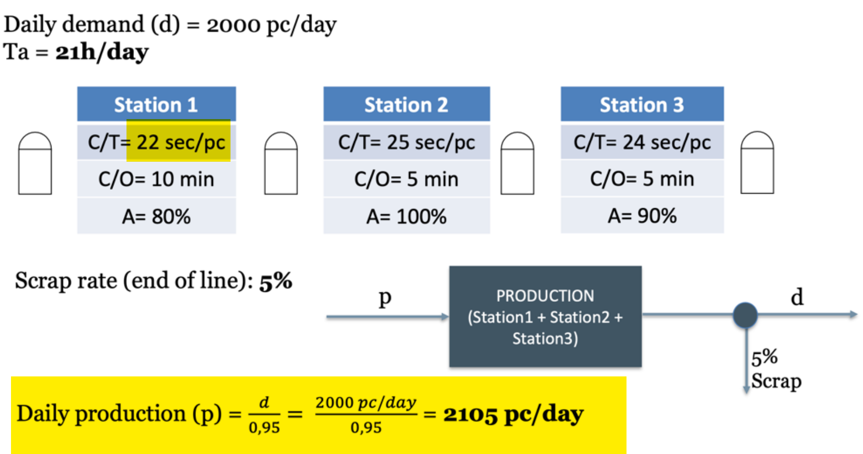
Consider a single-product situation (Washing Spa is producing only 1 product).

1. What is the daily production capacity of Washing Spa?
2. The company is evaluating the possibility to organize its production system according to a coupled serial configuration. What would be the daily production capacity of Washing Spa in this case?

Now, consider a multi-product situation (Washing Spa is producing more than 1 product).

1. Assume that batch size is 60 products and that a setup is required in each station every time the production of a new batch starts. Is Washing Spa able to meet customer demand according to its production system decoupled by stocks?
2. Assume that Washing Spa has a product range of 15 finished products and that it is willing to produce the whole range every day. What would be the maximum setup time if the production system is organized as a coupled one?

Solution:



Q1 - Daily production capacity of a decoupled system

Ta= 21 h/day

Expected output = Daily production capacity = $\frac{Ta * A}{C/T}$

Daily C1: (21*3600 sec) * 0,8/22 = **2749 pc/day**

Daily C2: (21*3600 sec) * 1/25 = **3024 pc/day**

Daily C3: (21*3600 sec) * 0,9/24 = **2835 pc/day**

Daily production capacity of Washing Spa: **2749 pc/day**

Q2 - Daily production capacity of a coupled system

Station 1	Station 2	Station 3
C/T= 22 sec/pc	C/T= 25 sec/pc	C/T= 24 sec/pc
C/O= 10 min	C/O= 5 min	C/O= 5 min
A= 80%	A= 100%	A= 90%

C/T= 25 sec/pc
C/O= 10min
A= 72%

Daily CSystem: (21*3600 sec) * 0,72/25 = 2177 pc/day
As the decoupled system, the coupled system is able to produce what is required daily by the customer.
Δproduction: **-20,8%** compared to the decoupled system

Q3 - Batching in a decoupled system

Each station has its own stocks and therefore they work rather independently one from the other.
Batch size = **60 pc/batch**
#Setups per day = 2105 pc/day/60 pc/batch = 35 setups/day

We must verify for each station whether it is possible to satisfy customer requirements within available time.

$$Tp + Tsu \leq Ta$$

Station 1: 22/0,8*2105 + 35*10*60 ≤ 21*3600
78887,5 sec/day > 75600 sec/day X
Station 2: 25/1*2105 + 35*5*60 ≤ 21*3600
63125 sec/day < 75600 sec/day V
Station 3: 24/0,9*2105 + 35*5*60 ≤ 21*3600
66633,33 sec/day< 75600 sec/day V

The system can work in batches of 60 pieces at phases 2 and 3. It cannot work instead in batches of 60 pieces at phase 1.
What activities would you propose to be able to work with batches of 60 products?
What advantages do you expect from this activity?

Q4 - Maximum setup time in a coupled system

Product range = 15 pc
Whole product range required daily
#Setups per day = 15
NB: In a coupled system, for each change (setup) in one station, the whole line is stopped (all the other stations are stopped)

We can verify if we can perform 15 setups per day according to actual conditions

$$Tp + Tsu \leq Ta$$

System: 25/0,72 * 2105+ 15*10*60 ≤ 21 h/day *3600
82090 sec/day > 75600 sec/day X
In 1 day, according to actual conditions, it is not possible to produce the entire range in the quantities required by the customer during the available production time.

We must find the maximum setup time that enable Washing Spa to satisfy customers requests in terms of volume and variety.

$$Tp + Tsu \leq Ta$$

25/0,72 * 2105+ 15*x*60 ≤ 21 h/day *3600 x= C/O = Ta -T$\frac{Ta-Tp}{\#Setups\ per\ day}$p = $\mathbf{2{,}8\ min/setup}$

The time required to produce the entire range of product in respect of the daily demand is an index of the **system flexibility**.

It consists in the calculation of the time required to produce the entire range of varieties in the quantities required by customers.

16.2 EXERCISE 2 – MTCF SPA

MTCF Spa is a leader in the production of wrist watches.

The company has a daily demand of 3,000 watches. 100% quality control at the end of the line intercepts the 20% of the watches that are not compliant. In each stage, there is one operator that is specialized and full-time dedicated.

The production process consists of three production stages, decoupled by a fair amount of stocks. The company works on two shifts per day, for 200 days a year. Each shift is 8 hours with two 15 min breaks. The daily demand of the customer is extremely constant in volume, and, if not satisfied, overtime production is required, because the company does not want to lose demand.

At the paint stage, the watch hands are painted. Each watch has two hands those are assembled on the dial, in the next production stage. The case assembly stage puts together the painted hands and ensures sealing of the watchcase. At the band mounting stage, the operator puts together the watchcase and the band. The table below shows the quantitative data of the production process.

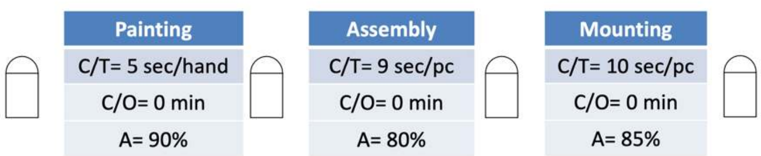
Process	Cycle time (sec)	Changeover (min)	Uptime (%)
Painting of 1 hand	5	0	90%
Case assembly	9	0	80%
Band mounting	10	0	85%

1. With the production parameters in the table, is the company able to avoid the use of overtime if a single product type is produced?
2. Assume now that the setup times at painting and case assembly stages are not null. In particular, at the painting stage it is possible to paint the hands with 4 different colours. In each watch, the hands are same coloured. The colour change at this stage takes 15 minutes. At the case assembly, each pair of watch-hands is mounted whit each of the 3 different cases. The set-up time at the case assembly and closing stage when switching from one case type to another one is 5 minutes. What is the minimum batch size to meet customer demand without using overtime?
3. In the case MTCF changes to a coupled serial production system, how would be the answer to question 1?
4. Assume now that MTCF works according to a coupled serial system configuration with availability equal to 90% for each stage. Same setup times and number of variants explained in question 2 are here valid. Moreover, the whole range is requested every day. With this configuration is the overtime necessary to meet daily customer demand? What is the minimum batch size not to use overtime?

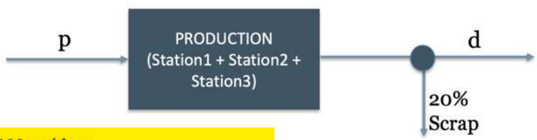
Solution:

Daily demand (d) = **3000 pc/day**

Ta = 7,5h/shifts*2shifts = **900 min/day**



Scrap rate (end of line): **20%**



Daily production (p) = $\frac{d}{0{,}8} = \frac{3000\ pc/day}{0{,}8}$ = **3750 pc/day**

Q1 - Seek overtime in a decoupled system?

We must verify for each stage whether they are able to fulfil customer demand within the available time.

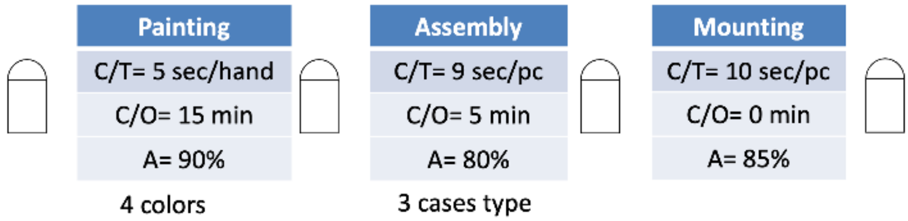
$$Tp + Tsu \leq Ta$$

Painting: (2*5)/0,9*3750 ≤ 900 min/day *60
41667 ≤ 54000 V
Assembly: 9/0,8*3750 ≤ 900 min/day*60
42187,5 ≤ 54000 V
Mounting: 10/0,85*3750 ≤ 900 min/day*60
44118 ≤ 54000 V
No Overtime needed

Q2 - Batching in a decoupled system

Total range = 4 colours * 3 cases type = 12 products

Total range = 4 colors * 3 cases type = **12 products**



The system is decoupled:
- Each phase is then "free" and is unaffected by other phases.
- Each phase has its own batching only depending on the setup of the phase itself.
- Phase 3, having a setup time equal to 0, can always produce a different watch without doing any batching.

We need to compute for each stage the minimum batch size that enable MTCF Spa to avoid overtime.

X = #Setups = $\frac{Ta-Tp}{C/O}$

$$Tp + Tsu \leq Ta$$

Minimum Batch size (MBS) = daily prod./#Setups

Painting: (2*5)/0,9*3750 + x*15*60 ≤ 900 min/day *60
X= 13,7 setups/day
MBS= 3750/13,7 = **274 pc/batch → MBS of MTCF**
Assembly: 9/0,8*3750 + x*5*60 ≤ 900 min/day*60
X= 39,38 setups/day
MBS = 3750/39,38 = **95 pc/ batch**

NB: If the variants requested everyday by the customer are 12, MTCF will use an average batch size of 312,5 pc → 3750/12.
If the variants requested per day increase the batch size is constrained by the **MBS**

Q3 - Seek overtime in a coupled system?

Painting	Assembly	Mounting
C/T= 5sec/hand	C/T= 9sec/pc	C/T= 10sec/pc
A= 90%	A= 80%	A= 85%

C/T= 10 sec/pc
A= 61,2%

$$Tp + Tsu \leq Ta$$

System: 10/0,612*3750 ≤ 900 min/day*60
61274 ≤ 54000 X

Overtime is needed to satisfy average demand volume

Overtime needed = around 2 hours

Q4 - Batching in a coupled system

Painting	Assembly	Mounting
C/T= 5sec/hand	C/T= 9sec/pc	C/T= 10sec/pc
C/O= 15 min	C/O= 5 min	C/O= 0
A= 90%	A= 90%	A= 90%

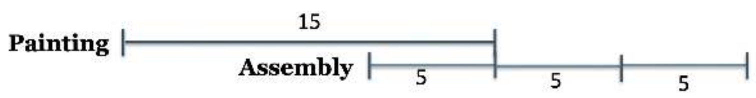
C/T= 10 sec/pc
C/O= 8,3 min
A= 72,9%

For understanding whether overtime is needed, we must compute the maximum production capacity of our system and to compare it with customer requirements and setup time.

System:
900*(0,729/10*60) = 3936 pc/day ≥ daily required capacity (3750 pc/day) V

Since the previous condition was verified, we need to understand whether MTCF Spa is able to also satisfy the range requirements of the clients (12 range).
At system level:
- 4 setups at stage 1
- 3 setups at phase 2
- Each time the system changes the colour of the hands, hands of the same colour must be assembled on 3 kinds of different cases.
- How much is the stop time for the line because of sets up?

Colour	size	Total
Blue 15 min	1 5 min	15
	2 5 min	5
	3 5 min	5
yellow	1	
	2	
	3	
Etc...		



Total time of stop for the line to make setups per each day= 4* (15 + 5 + 5) = **100 min**
#Setups per day = 12
Average C/O= 100 min/12 setups = **8,33 min/setup**

$$Tp+Tsu \leq Ta$$

System:
3750*10/(0,729) + 100*60 ≤ 900 min/day *60
57440 ≤ 54000 X
Overtime = 57440 – 54000 = 3440 sec/day = **57,33min/day**

Batching
Average C/O = Tsu /range = 100/12 = 8,33 min/setup 10/0,729*3750 + x*8,33*60 = 900min/day * 60
x= #Setups = $\frac{Ta-Tp}{C/O}$ = 5,12 setup/day
Minimum Batch size (MBS) = p/#Setups = **732 pc/batch**

17. LEAN MANAGEMENT

One of the highest values is that there are a consistent set of principles and tools that reinforce one another and go in the same direction.
Generally, in order to increase profitability (principal goal), companies can increase sales or reduce costs. The former, it is by now a marginal lever, due to the fact that markets have become mature; as regard the latter, we need to act in order to:

- **Remove functionality**: the number of specifications and the characteristics that allow to sustain lower manufacturing costs;
- **Reduce material used**;
- **Reduce processing time**: it means متحمل شدن incurring in lower labour costs.

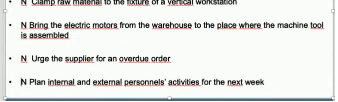
In particular, we have to pay attention to product costs, that are related to what the product does or has and which usually are easily measurable costs. These fields are extensively explored and adopted by enterprises, so they do not lead to a consistent competitive advantage.

Lean management follows a different approach: it is necessary to focus the attention on what is actually realized and to eliminate what the customers do not require and pay for. This is the key points introduced for the first time by Taiichi Ono, a Japanese engineer, who is considered the father of the Toyota production system. Actually, it was not invented in Japan, but in USA: Toyota was the first company to introduce it with success. In fact, after the 2WW Japan was destroyed and had to rebuild the country and its industries. USA taught the Japanese the APDC (a scientific methodology to solve issues) and they were good in applying. They were able to introduce this new approach and had a big growth, while in the US the performances were flat. At the beginning, Toyota produced in an entire year the same volumes produced by USA in a day. After 20 years, Toyota was able to overcome USA.
Only recently western companies have realized the powerfulness of this approach. In the 90s an American researcher made an analysis on Japanese production and noticed that their way of doing things was totally different → this was the result of lean management.
However, the truth is that nonJvalue activities are more than 90% of the overall activities performed inside an organization. So if we can reduce this amount, we will spare a relevant portion of costs.
As a result, the starting point is the classification of activities in:

- **Value adding activities**: those the customer demands, due to the fact that are functional to the satisfaction of his/her need;
- **Not value adding activities but unavoidable in the short term:** customers do not demand them, but we do not find a way to get rid of them yet. They are essential, and we cannot tackle them in the short term.
- **Not value adding activities and avoidable in the short term:** it is waste, they are easy to be eliminated.

Usually, immediate interventions are required for the third class of activities, while the actions of the second class are more complex: in fact, although they are NVA activities, they are essential for the process as it is structured. Their elimination will imply the redesign of the entire system. It is important to underline that, the fact that these activities result in wastes does not mean that they are useless. Only after these improvements, it could be possible to remove them; in fact, their inclusion in a non-optimal configuration represents in effect a waste.
Only after the identification of nonJvalue added activities, it is possible to exploit all the opportunities that may emerge inside a firm: one example could be the establishment of a strong relationship with a supplier, which as a result could guarantees with a relevant certainty, the constant alimentation of the process, minimizing the probability of lack of materials.
Look at the following activities; which are value adding and which are wastes?

- Insert a ball bearing on its shaft
- Test a product before sending it to the customer: if we are sure that all our activities are under control and the quality of the output respects ~~always~~ all requirements, there is no need to test products at the



end; the process is so good and in control that the natural deviation is not enough to compromise the standard.

- Fix a rough material to the fixture of a vertical workstation: from a customer perspective, taking the material out is a value adding activity; however, we can find another way to do that.
- Bring the electric motors from the warehouse to the place where the machine tool is assembled: every transportation of material, unless it is required by the customers, is a waste.
- Urge the supplier for an overdue order: it may be necessary, but it is not value adding, because no customers will be willing to pay people that make phone calls.
- Plan internal and external personnel's activities for the next week: planning activities are waste.

Among all these activities, the only one that is providing value is the first one, that is an assembling activity, because it is changing the value of the product. Planning the activities does not change the value because the customer does not recognize whether the activities are performed in a certain way or in another one. With this perspective, who creates the value? The production worker, people on the field, operators, people that usually are seen as the lowest value of the company. Managers usually do not add value to the customers, because they are typically involved in planning activities, so why the managers are necessary? They are not necessary, because they do not add value. So, an improvement can be to reduce the managers, resulting in less manager paid more. Only the best managers will go on.
We have to keep in mind that the wastes are caused by our activities, and not by people performing them. On the contrary, people are fundamental to understand how to eliminate them. But how can we find non-value-adding activities? Inside the company there are typically seven kinds of wastes: TIM WOOD

1. **Transportation**: as a customer, we do not care whether the product is manufactured in our town or 1000km far away.
2. **Inventory**: the only exception is the case of some foods that increase their quality with aging.
3. **Movement**
4. **Waiting**: the products waiting to be processed are not functional for the creation of value.
5. **Overproduction**: the customers will not pay for the product they have not requested.
6. **Overprocessing**: we are performing some additional activities that provide a value not recognized by the customers, it is useless.
7. **Defects**: not only because we have to eliminate them, but also because they consume activities to do on the defect themselves

Obviously, this classification was elaborated when other themes were not so important as they are today: safety and security, energy consumption, pollution, compliance to laws. However, we can observe how each of the more recent issue can be included in one or more of the previous clusters.

The philosophy that lies behind this approach can be summarized by the following scheme. The overhead architrave represents those objectives that lead to the continuous improvement, with the aim to always guarantee products with a greater quality, with lower costs and that allow shorter lead time. All of this can be reached addressing a unique macro-objective: the reduction of the production flow, which coincides with the interval between the moment in which raw materials are received and the moment in which the finished products are ready to be delivered. In order to achieve this result, it is necessary to eliminate wastes and not to go faster: only in this way, we will have the possibility to build a system that require less time to complete nonvalue added activities or inefficiencies (if this is short, the company has less time to dedicate to waste). As a consequence, the principle underlying the Toyota philosophy identifies a physical (and not monetary) indicator, the so-called **throughput time**, as the driver that describes properly the healthiness of the system.

As we can observe, the architrave lies on two columns: one that relates to the single operator (Jidoka), and one connected to the whole system (Just in time):

- The first one, in the matter of the single operator level, we talk about **Jidoka**, which implies to analyses the single resource and how it works. The key point is that it is not possible to transfer downward an output that is not qualitatively perfect, we need to improve the system in order to prevent those mistakes turn into defeats. A system that is able to intercept مانع شدن errors, gives also to operators the instruments to understand the cause, avoiding that the product proceeds towards the downstream stage and allowing that the event will be repeated in the future. It is fundamental to underline the difference between errors and defeats: the **error** is something that is not working as it was supposed and designed (related to uncertainty), it is the variability in the process, and it is structural in every system. The **defect** is the negative impact of an error on the system, it creates nonvalue added activities, since the customer recognizes it.
  From this consideration, the concepts of **"build in quality"** and **self-control** emerge. Traditionally, it was necessary to divide those people who are responsible for the decisions about what to do, from those who have to perform tasks. In addition, it was required to separate the person responsible for the execution of that task is from the person in charge of the control (check and certify). We can understand immediately how under this logic the self-control is seen in a negative way, because it may influence the judgement. However, the Toyota production system has upset and revolutionized this perspective, proposing that, who performs an activity, should also be able to control it and understand, in case of errors, where the problem lies. Only understanding causes, it is possible to improve the performances of the system and of the operator itself.

- The second pillar is related to how the different resources interact with each other, consisting in the implementation of the **Just in time,** which is fundamental to guarantee the integration between the various phases of the process. It means to have the right component at the right time in the right phase. So, I need a flow that is synchronized and make the system working with only the things that are necessary in that specific moment and place. Under this logic, the system should be able to make the product immediately available (infinite speed), which represents an ideal situation. A good compromise consists in organizing the activities in order to have a certain quantity of components in process, which therefore are immediately available when required. We have to pay attention because stocks are wastes, but in some case essential; however, analysing the specific level of materials to be maintained in process for each typology, there are the possibility to minimize the inventory level. It is evident that the main cause is represented by the reorder lead time. For this reason, we must be as fast as possible, work in strict contact with suppliers in order to have short lead time to receive material and consequently reduce the lead time to respond to the market. From system pushed by production, we move to system pulled by the customer's requirements. Rather than a better planning, we have to rely on a faster and better system in order to react. The goal of the lean production is not to reduce the inventory, but improve the system, the inventory is an indicator of the fact that something is going wrong.

By now, systems have become more and more complex that it is appropriate to create automatisms and delegate and decentralize a portion of decisions. Management has to intervene only centrally for exceptions: in normal condition, each operator should decide what to do and how to intervene without the need of the production manager. Wrong operations lie on the need of a supervisor, without which activities cannot be performed. In the ideal system the production manager is a waste, and it is not needed. This is what makes operations successful and well-functioning: the organization works only if management is not constantly required for intervention, but it is dedicated in helping the system improving and the resources growing continuously. We need to make the system as simple as possible.
It is necessary to note how these two pillars are built on two important prerequisites:

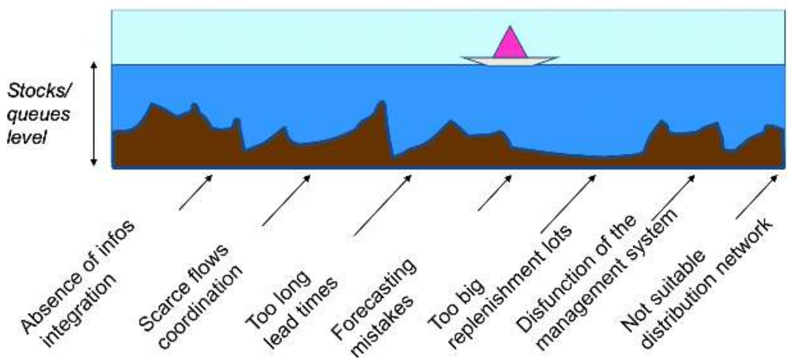
- **Standardization of procedures**: everyone operates in the same way, the large part of the managers spends 98% of their time in running daily activities, while their goal should be to spend the time in improving the system and not in running it, focusing on strategic activities.

- **Continuous improvement** (Kaizen): it means changing the system, so taking the problem with a scientific method, looking at it not just to solve it, but to analyse it, and change the system in order to prevent that it will happen again in the future.

The lean company must work as a human body. What is the main activity to survive? Breath and heart beating. What is the most relevant resource that we have in the company? Creativity and innovation and brain of the people. The basic processes should work without efforts, we should focus on those that create value. We need to eliminate the problems (reducing the weaknesses) and make the production linear, synchronized, without defect. This way the production flows well and you can dedicate your brain and activities on innovating the processes and products and so creating competitive advantage. This is the philosophy.

To summarize:
- Reduce all activities that do not add value
- Eliminate wastes → to have better processes.
- Bring out **weakness**/defects/inefficiencies, in order to solve them: *Do not hide them, realize and solve in the long term all these causes giving you troubles*. Eliminate the cause of the problem, rather than avoiding that it gives troubles. If you cannot see the problem, you cannot solve it.

The small boat and the cliffs: The small boat and the cliffs example is showing us that in order to be faster and reduce cost we need to lower at the minimum possible in order to “let the boat go”. Lowering stocks level is possible to point out causes of inefficiencies, address them, and eliminate them; we can consider this process as an endless process: *The focus is the process, not the result.* It is a metaphor to say that lean manufacturing has a long-term perspective and not a short-term perspective.



Traditional approach: Production is the boat that is moving in the sea. The level of the water is the level of the stocks, the problems are the rocks. We have the production, the boat that is moving, but sometimes you could find the rock. The approach of a traditional manager is that if you see the rock, let us increase the level of stocks. In this way our boat is not anymore hitting the rocks. But in this way, they are continuously raising the level of the water, and this led to problems. The problem is that it comes difficult to navigate a boat in such a high level of water. They brought huge costs in the medium-long term because it comes more difficult to coordinate the stages and to be reactive to market requests. Processes are weak and you need brilliant people to run the production.

Lean approach: We continuously decrease the level of water until we find the highest rock and then we walk together to destroy and remove that rock with a continuously improvement project. Lean management arises the problem and continuously decreases the level of the project and became easy to manage the production. With this approach we have a total change of mind by the management. In this context, when you see a problem, you consider that an opportunity because you define a rock you can work on and make the system better than before. Everyone can run the production because the process is robust, well designed and coordinated.

Iterating this approach, it will change the company and also the people approach.
Another aspect of the lean philosophy is related to the human resources: As the CEO of Toyota said:

"We get brilliant results from average people managing brilliant processes – while our competitors get average or worse results from brilliant people managing broken processes" Toyota is the company where lean was invented and it was exported all over the world, their competitive advantage was not definitely on design (it was according to the Japanese taste and not the one for Europe or USA).

Because of this, they decided to improve the efficiency in the cost and the quality; regarding the cost they reduced the production cost (because the logistic cost was already huge) and regarding the quality they increased the reliability in the manufacturing process. By doing so, their advantage became the capability to transform the production the easiest possible that could have used average (and not highly skilled people). Instead of increasing their employee's capacity, they made the process so simple, that anybody could do it. The sentence is switching the focus compared to the European way, because in the case of Toyota the people used for the lean manufacturing are average/normal.
If you put people in the condition not to have a mistake because the process is simple, then you reduce the risk to have some mistakes.

Lean principles: by Liker (only mentioned because it exists but we do not focus on it)

1. Flow: Any interruption/break that we have in the process that is preventing a smooth flow, should be removed because it is slowing the production.
2. Pace: Keep the pace constant of the system; each station should work at the same frequency of the other.
3. Pull: The lean manufacturing is using the pull system, that is, the activation of the process only if there is a real request from the market. The product/system should be pulled by the market, if this is happening you activate the system only when there is a real market request. If you push the product on the market, you maybe produce more than what is needed from the market and in this way, you create inventories that will constantly increase creating waste.
4. Zero defects: The idea is that the product should have no defects and that the company is producing the right thing the first time, there are no re-working.

1. Identify **Value**
2. Map the Value Stream
3. Create **Flow**
4. Establish **PULL**
5. Seek **Perfection**

Step of the lean principles: by Womack and Jones

1. **Identify Value:** Which are the activities that create value for the customer and classify them. The others are no value added activities. How to identify it?
   The value is where the customer is willing to pay, which means that has a value for you. **Value is defined by the customer: «All that the customer is willing to pay».** We need to shift from the internal perspective to customers' value in order to redesign the manufacturing system according to what is perceived as a value. We need to move from an internal approach to an external one.

How do the companies define the price?
Traditional companies - **From: Cost + mark-up → price**
Lean principles - **To: Price - mark-up → cost** (How much the customer is going to pay? It is a customer-centric approach. The price is the input because is the value for the customer. The design of the processing is done by a target cost and then define a mark-up).

2. **Map the Value Stream**: The Value Stream is the set of activities that are made on the inputs to give to the customer a product/service. In this case we need to see which are the different steps and where the activities are created. It is a tool used to map the current and the future situation. We need to map the value, visualize it and have a clear idea where the value is. Map the flow and the steps in order to generate the value for the customers.
   There are three main flows:
   a) Development and introduction of new products/services: concept of the product, how do you design the product?
   b) Managing information from customer's order to delivery: related with all information flowing from the customer to the company, they have to be considered and to be highlighted
   c) Manage product/service realization from the inputs elements to the final output: making of the product, from input to output, it is after the engineering part. How much effort we will put in this process?

3. **Create Flow:** We should create a smooth flow, so we should identify any cause stopping the flow → Anything that stops the flow is waste, so, identify what stops the flow and remove it. Make the flow more linear.

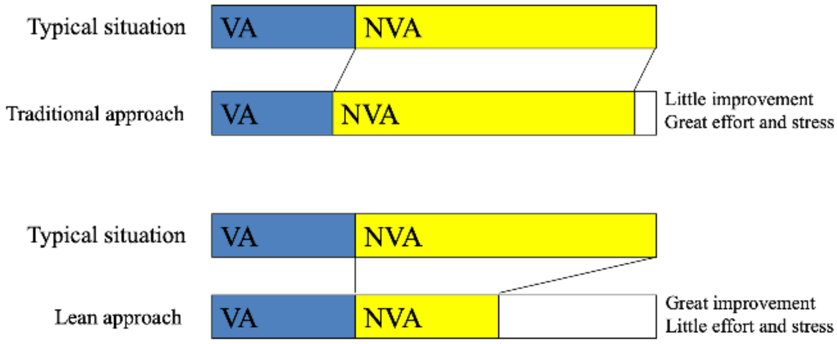
4. **Establish PULL:** Push vs. pull system: it is better the pull system. From the bottom you pull the production, the idea behind the pull system is that we are not able to understand completely the demand, so link your decision with customer behaviour. Usually when we forecast we make mistakes (we forecasted an over-demand or an under-demand. Forecasting is based on push method).

5. **Seek Perfection:** It is not just we want a great quality or efficiency. It is much more than that. When is quality sufficient? When is cost low enough? When are wastes low enough? There is not a threshold for quality and you want the perfection. The target is 100%. But you want also to reduce your costs and you can reduce them until you want.
   The meaning is you must have a journey that will never stop. If perfection is the aim, we should focus on the journey and not on the destination. You must do anything every day in order to improve and to innovate and to reduce costs in the direction of the perfection. Do not focus on the specific result of today. You must be never satisfied.
   This principles along the time will create competitive advantage. The result is once you started lean after some year you see huge changes in productivity, a reduction of lead time, of respost time, of the stops. They are result achieve with a very low level of investment.

If we implement all these concepts/ steps, the results are very important, we will have:
- 50% improvement of work productivity in the whole system
- 90% cut of production lead time
- 90% reduction of stocks
- Halved process defects and wastes
- Halved accidents
- Halved Time to Market.

The tricky point is that these improvements are usually done with a limited capital investment. It means that it is easier to see those improvements in a company that does not have a huge variety of products or a high volume.

The right approach of lean production:



In the first situation we see a company that is not using the lean production, it has an amount of VA (value added activities) and NVA (not value added activities); in this case the company is trying to use the traditional approach by improving the VA instead of diminishing the NVA.
The second situation shows a company that is using the lean management, as we see there are in this case an amount of VA and NVA, what the company is trying to do is diminishing the NVA in the process by eliminating the waste, in order to improve the total efficiency of all the activities.
The best idea is to eliminate the NVA, not improving the VA → Do not work more, work better.
If the lean approach is the best one in order to improve all the activities, then, why this was not done before? Causes:

- Mainly for a cultural aspect. It is used in east-world countries
- Because 50 years ago it was not perceived so urgent
- Because it is necessary to completely change organization and human resources managerial vision: Because lean is working on human capital and resources, in the lean case people should tend to have a high inertia to change, you need to change a bit the HR in order to change the concept of people
- Because the heart is the process and not the final product/service: Lean is putting a focus on the process, not on the product or the service.

There was a credence that lean system could work only on Japanese companies. Because of this credence many companies rejected this approach. But there is an example that affirm that lean principles work also in the western companies.

NUMMI case: An example of the success of the lean approach could be given by the Freemont factory (Freemont was considered as the worst factory). From 1963 to 1982 Freemont (Oakland, CA) was operated by General Motors. The GM management was based on the Taylor approach and they were completely misbehaving with workers, many unions were provided, there were many problems of misunderstanding and threats by the managers. The problems were so many that GM decided to shut the company down.
After 1982, Toyota hesitantly agreed to open a joint venture with GM in the same plant with the name of NUMMI; they decided to re-hire the same workers (but not the managers).
From that moment on the situation completely changed, 60 workers were sent to Japan for trainings, there was a strong employee involvement realized in all activities and NUMMI soon achieved the Japanese quality levels, productivity was 40% better than all GM, workers satisfaction was 90% in 1991. The Freemont became the best factory and achieved the Toyota results. It destroyed the credence of the east countries.
NUMMI was the proof that the guilty party was the traditional management and not the workers of the unions. These improvements were done thanks to some implementation: TPS (lean) work organization, respect and partnership, constant mixing of managers and workers, job security, work decisions done at the lowest possible level and the absence of authoritarian management.
Messages:

1. Stops hidden the problems. Highlight them.
2. In order to make lean system we follow 5 principles. Lean system is a journey not a decision.
3. It is complex to create because it takes time, effort but it is something that is possible to do even in west culture.

Question: How is it going to create a competitive advantage?
Competitive advantage is something that make you different from you competitor and this difference is difficult to replicated and it is align with the customer request.
Time and number of choices create a competitive advantage (much more than just buying a new technology). Lean management is a continuous improvement approach and you continuously innovate and improve you processes. It creates competitive advantage because in order to arrive to performance is a result of a continuously improvement process, it is the final point of a journeys that take to you this time and it is composed by so many choices, improvements, projects. This is important because if another company wants to achieve your same performances that company has to face with: your time.
The other point of doing lean approach has a competitive advantage is the results that come from many project and it comes very difficult for companies to have so many decisions. You change many things and many briefs, you do not have the same people, environment and resources.

18. VALUE STREAM MAPPING

All we are really trying to do in lean manufacturing is to get one process to make only what the next process needs when it needs it. We are trying to link all processes -- from the final consumer back to raw material -- in a smooth flow without detours that generates the shortest lead time, highest quality, and lowest cost.

Value stream mapping is a pencil and paper tool that helps you to see and understand the flow of material and information as a product makes its way through the value stream. What we mean by value stream mapping is simple: follow a product's production path from customer to supplier, and carefully draw a visual representation of every process in the material and information flow. Then ask a set of key questions and draw a "future state" map of how value should flow.
Doing this over and over is the simplest way - and the best way we know - to see value and, especially, the sources of waste.

The first step is drawing the **current state**, which is done by gathering information on the shop floor. This provides the information you need to develop a **future state**. Notice that the arrows between current and future state go both ways, indicating that development of the current and future states are overlapping efforts. Future-state ideas will come up as you are mapping the current state. Likewise, drawing your future state will often point out important current status information you have overlooked.

There are different level of analysis regarding the value stream map, from the most operative to the managerial ones:

- **Process level**: directly in the machine, manufacturing process for instance, and we could improve the process and the performances (shorter set up time, faster machines). We need technicians and mechanical engineers who are directly connected to this knowledge
- **Single plant**: this is the most common level of analysis, since a single plant is composed of different machines. Originally, the value is created on the *door to door*, so the value stream usually begins at plant level
- **Multiple plants**: supply chain of the company
- **Across companies**

Two goals:

- Clarify and identify the real as is condition of the system, where are the opportunities of the system to improve
- Set the new configuration for our supply chain to setup and innovate plants to make them better than before

CASE: **ACME**
They produce many different products. We want to analyse two families (left and right). We have to understand what is really happening in the as is system. Value stream mapping means walking and drawing the processing steps (material and information) for one product family from door to door in your plant. Your customers care about their specific products, not all your products so you will not be mapping everything that goes through the shop floor but focus on the specific family of products.

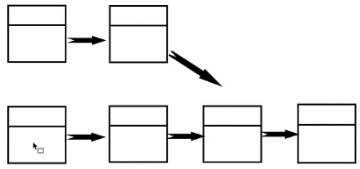
18.1. IDENTIFY A PRODUCT FAMILY

How do we define it? Which are the features that defines a specific product family? It is not about size, colour, shape, but according to the same pattern, which is referred to the resources are used by that product. To do that we can leverage the following matrix, a visual tool that matches two dimensions: the different products (from a to g) and the resources used. We can notice that some products follow the same pattern (A,B,C), so they can be grouped in the same family.

		ASSEMBLY STEPS & EQUIPMENT							
		1	2	3	4	5	6	7	8
PRODUCTS	A	X	X	X		X	X		
	B	X	X	X	X	X	X		
	C	X	X	X		X	X	X	
	D		X	X	X			X	X
	E		X	X	X			X	X
	F	X		X		X	X	X	
	G	X		X		X	X	X	

18.2. MAPPING

We can distinguish 3 parts of the value stream map:



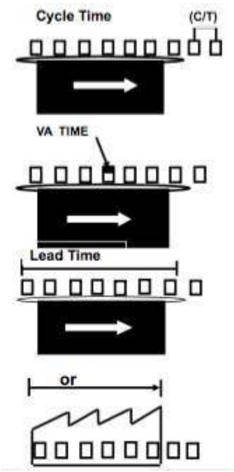
1. **Material flow**
2. **Information flow**: mapping of all info necessary to manage the material flow
3. **Timeline**: it is on the bottom side of the map, it is the measurement of time performances of our system

18.2.1 Material mapping

Then we start to map the sequence of steps followed by products. It is not requested to map all the possible processes, but to focus only on the mainstream of the specific components. The map should in fact be as simple as possible.

We start from the beginning (i.e. raw materials). We would like to have a single flow, but sometimes there can be 2 or 3 flows that converge. More than 3 would be useless, and we must divide them.

The **flow** is the main element of the map, the second one is the **time**. There can be different measures:

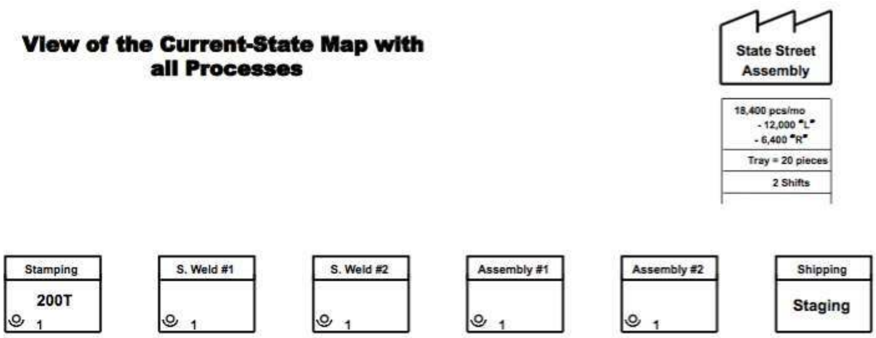


- **Cycle time**: it is the interval time between one exit from a line and the following one. The slowest station (the bottom-neck) gives the cycle time to the whole system. We can improve the cycle time removing the bottom-neck or reducing the time of working of it (balancing the working time between different stations), in order to have a higher production rate.
- **Value added time**: time really related to increasing of the value of the product we are processing. In the value stream mapping usually we do not make difference among value added time and cycle time, they are same.
- **Lead time** (or throughput time): it is the time spent by the product from the beginning of the process to the end of the process.

When we have a buffer upstream and downstream, the movement between different stations takes time (maybe 1 seconds) and the time to take the movement is not dealing with the activity of the workers in the stations (9s + 1s = 10s). We need to see if there is buffer, and we add the movement time! If the product stays in the buffer for 2 hours, the waiting time should be included in the lead time!!

In any organization the value added time is a small part of the cycle time over resources. In general it is true that: VA < C/T < LT

We start from the customer, who makes the order and influences the production. It is represented by two icons. The first one stay for the shape of the factory, used as external entity (state street assembly). The second is a data box, where we put some info: volume of demand, the mix of the demand. Then other info such as: "the customer is receiving the product in trays of 20pc and we can have 2 shifts".



from end

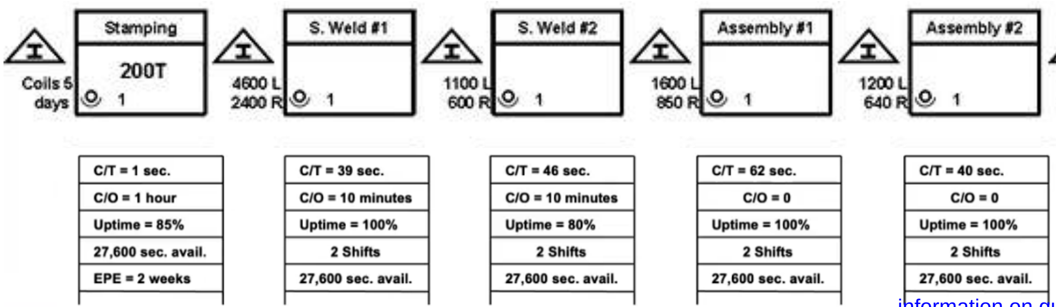
Then we go backward. Let us map the value stream inside the organization, distinguishing 6 different stations. The icon with the semicircle represents the operator working at that stage seen from above. In this case the box is a square, since we are analysing an internal stage, which must be different than an external one. The databox contains additional information about the system. In particular:

- **Cycle time**
- **Change-over time** (c/o): it is the setup time needed by the machine to change the product
- Availability (**uptime**): measure of reliability of resources

If necessary, we can add more info:

- Number of shifts
- EPE (only for stamping): since the press is not dedicated only to a single product, it is the response time that it gives us when we put an order, because there is a queue.

The triangle with "I" is related to the inventory: we count the number of inventory we have that decouple the different steps of the flow → it is important to be highlighted because stocks are wastes. If we have a good availability of historical data, we can put the average number, but sometimes we do not have it.



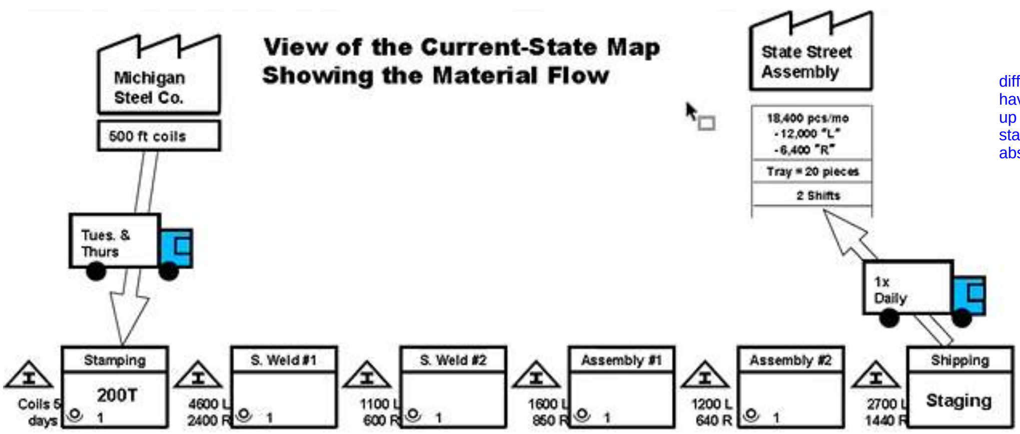
left to right 0

information on quality

every part every:
after finish they product after 2 weeks

How do catch those data? GEMBA is a Japanese word explaining this, which means "when things happen", so the real cycle time is the one directly measured in the box. You see and measure it. In this way we avoid misunderstanding, avoid conflicts.

After Acme's last assembly workstation, products are staged in the shipping area according to the daily shipping schedule and delivered daily by truck to the customer's assembly plant. A truck icon and a broad arrow indicate movement of finished goods to the customer.
At the other end of the map, we will represent the steel supplier with another factory icon. We use the same truck icon and broad arrow to show movement of material from the supplier to Acme.
The steel supplier receives a weekly order from Acme and ships twice a week.



different in cycle time is no good reason to have inventory.
up time of system! no stopping the system.
stamping machine
absorb customers variability

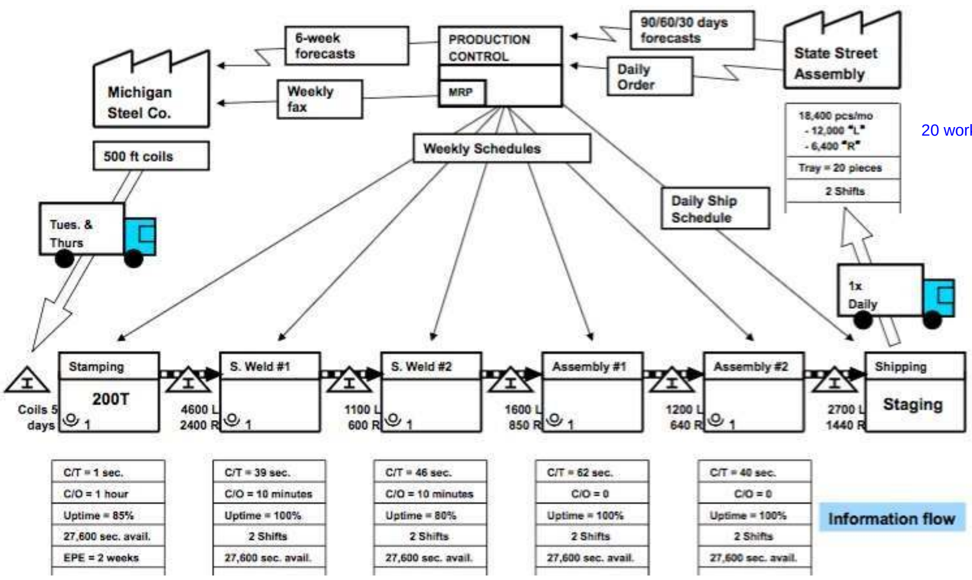
18.2.2 Information flow

We describe which info are necessary and used to manage the process of the order. To do this we will need a few more icons and arrows, in particular a narrow line to show information flows. The mapping icon for push movement of material is a striped arrow. At Acme Stamping, only the shipping department is connected in any way to a "customer." Each of the other processes is producing according to a schedule, so the transfer of material from one process to the next is occurring via push. A push arrow is drawn between each process.

Information flow is drawn from right to left in the top half of the map space. In our Acme Stamping example, we draw the flow of information back from the State Street Assembly Plant to Acme's Production Control department and from there to Acme's steel-coil supplier. Notice there are separate lines for the forecasts and daily orders, as these are different information flow.

- We have a weekly schedule, every single station says what it has to produce, giving them information which are independent from the other station (the MRP runs every week, once in a week, and gives the information to the stations)
- The MRP (Material Requirement Planning) is also calculating the raw materials
- The market issues every day an order, which is used for managing the quantity to deliver every day with the shipment
- We need forecast because otherwise we are not able to understand the raw materials that we will use and need

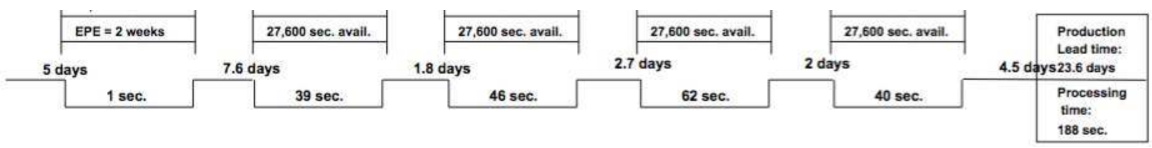
starts from customer
just a forecast



20 working days => 920 daily

18.2.3 Timeline

We measure the time, in particular highlight the difference between how much time we work the item and how much time the product passes inside our system.



How much time we process our product? Sum of all the cycle times

How much time the product passes inside the organization? Throughput time = cycle time + waiting time. In the medium-long term (1 week, 1 moment) what is the rate of the consume? On a stable system, what we produce is given by the rhythm of consumer demand. The consume of the resources in the medium time is led by the consume of the customers. the length of the queue is defined by LT and demand pace.

Value Added Time/Total Lead Time

Approximated by

Processing time /Total Lead Time

As an alternative indicator

Waiting time = Total lead time – Processing time

Little's Law:

LT*lambda = Q

LT = inventory/demand

unit to time!

In the timeline we see two level: throughput (waiting time of each buffer) and the transformation time (value added time) and we sum it for each step. In the end we look at the discrepancy, which is the signal about how much improvements can we do.
How do we define when we have an inventory If there is the warehouse it is simple? If there is not a warehouse, we put inventories every time we have item waiting on the floor.
What happens if we are working on two different items on parallel? Cycle time is half of the processing time. The value added time is the same but he reduces the cycle time.

FUTURE STATE: An innovation of the operation system for the future. The future state map brings the idea of creating the plan of development and the final set up (necessary changes) in order to achieve the future state design. There is a framework for building up the future state composed of 8 questions.

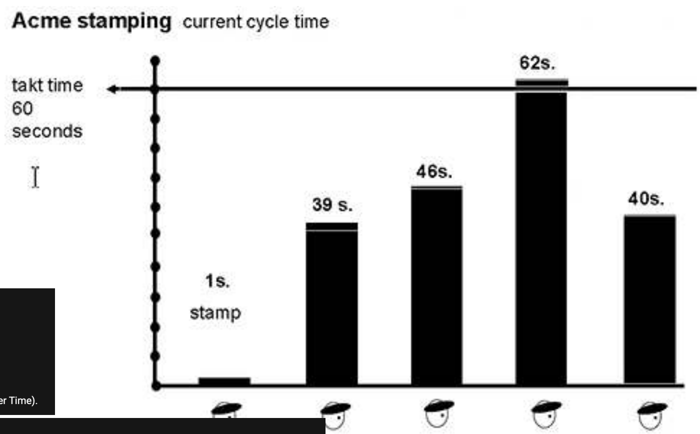
1. What is the demand volume we want to satisfy in this flow?
2. Are we manufacturing a product to be directly shipped to the customer or to supply a finished goods warehouse?
3. Where can we increase the flowing of the system without stops?
4. How do we connect stages that we are not able to put in the continuous flow?
5. How do we plan, schedule the production system?
6. How do we manage the mix of the production?
7. What is the increment that we release in the system, the minimum unit we can manage and move inside the system? Single items or trays of items?
8. With the first 7 we are changing the configuration of the system (the goal), but in order to make it feasible, what are the necessary improvements, operative changes?

1. What is the demand volume we want to satisfy in this flow?

Takt time: rhythm, pace that the company must respect to satisfy the demand volume of the customers. In lean companies it is really an indicators, it guides the operators as a reference. We compute it, taking the time available and distribute this time in the number of item that we have to produce in a day. In the example: the takt time for the company is 60s, then we have the different stages and we can notice that something is not working properly. We have differences in the cycle times, not balanced. Then the fourth stage does not respect the threshold (62>60). Then the time spent for stamping is too little compared to the others.

KEY QUESTIONS FOR FUTURE-STATE DESIGN

1. **What is the takt time?**
2. Will you build to a finished goods supermarket, or directly to shipping?
3. Where can you use continuous flow processing?
4. Where will you need to use supermarket pull systems to control production of upstream processes?
5. At what single point in the production chain (the "pacemaker process") will you schedule production?
6. How will you level the production mix at the pacemaker process?
7. What increment of work will you consistently release and take away at the pacemaker process?
8. What process improvements will be necessary for the value stream to flow as your future-state design specifies?



Available Working Time: 28,800 – 1200 = 27,600 seconds per shift

Available Working Time / Customer Demand = 27,600 sec. / 460 units per shift

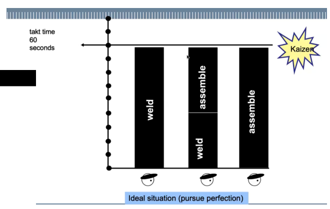
Acme Steering Bracket Assembly Takt Time = 60 seconds

Minimum number of people

Work content / Takt Time

problem: salary, physical situation of doing each and

- What is the Takt Time of the Value Stream?
- Is it possible to build a supermarket where the customer pulls?
- Where in the process can we use flow?
- Where the processes are too far and we need supermarkets?
- Where will we set the PACEMAKER of the flow as a single point of scheduling?
- How will we level the production mix?
- What increment of work will we consistently release?
- What improvements will be necessary? (i.e. Improve Uptime, Improve Quality, Improve Changeover Time).



Principle 1 - Do not overproduce.
Produce at the Takt Time speed to avoid overproduction.
Overproduction is the king of the 7 wastes.

Principle 2 - Create the flow, where is possible.
When the process workstations are far from each other, the material stagnation is natural.
With closer workstations will be easier to have one-piece-flow.

Principle 3 - Use the supermarket to control production.
When is not possible to move workstations closer, use a supermarket and pull system to control the inventory.

Principle 4 - Schedule only at one point.
If scheduling occurs in multiple points the information flow will be messy and the material flow accordingly.
This scheduling point is named the PACEMAKER.

Principle 5 - Level the production mix
At the PACEMAKER try to schedule the production mix at the shortest time interval you can. Distribute the production in the sequence ABCABC rather than AABBCCAABBCC.

Principle 6 - Level the production volume
Release into production small and consistent increments of work.

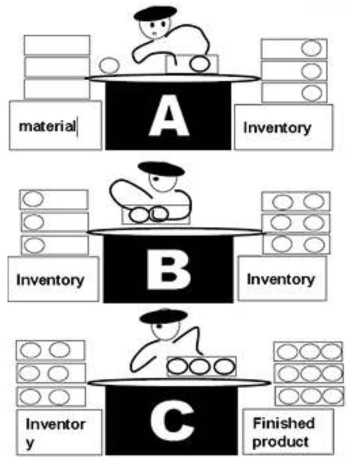
Principle 7 - Develop the organization to avoid large batches of production. The best way to do that is by SMED events.

#1 - TAKT TIME ?
#2 - Customer Supermarket ?
#3 - Flow ?
#4 - Processes Supermarkets ?
#5 - PACEMAKER ?
#6 - Levelling ?
#7 - Work Consistent Release ?
#8 - Necessary Improvements ?

Uptime? - Quality? - Changeover?

2. Are we manufacturing a product to be directly shipped to the customer or to supply a finished goods warehouse?

If we satisfy customers with make to stock logic, we are supplying a warehouse. If we produce with a logic of make to order, this means we want to ship directly to the customer. What is the implication with the satisfaction of the demand mix?



- Make to stock: response is in 0s.
- Make to order: we need more time to produce.

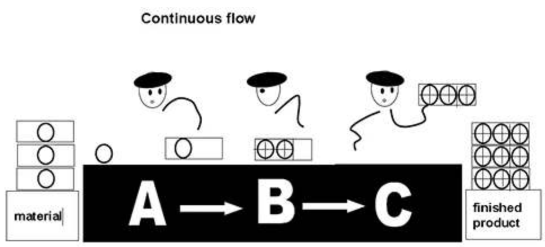
Companies use stocks because they are not able to avoid them, meaning that they cannot produce in a short time. But if they are able to do it, they will avoid them, since they are wastes. Make to order is the preferrable configuration, while make to stock is a backup solution. It depends on the response time requested by the customers.

3. Where can we increase the flowing of the system without stops?

There are often spots in the value stream where continuous flow is not possible and batching is necessary. There can be several reasons for this including:

- Some processes are designed to operate at very fast or slow cycle times and need to change over to service multiple product families (e.g. stamping or injection moulding)
- Some processes, such as those at suppliers, are far away and shipping one piece at a time is not realistic.
- Some processes have too much lead time or are too unreliable to couple directly to other processes in a continuous flow.

Continuous flow refers to producing one piece at a time, with each item passed immediately from one process step to the next without stagnation (and many other wastes) in between. Continuous flow is the most efficient way to produce, and you should use a lot of creativity in trying to achieve it.



If we want to reduce the stops, instead of working with the perspective of the single step, where operator continues to process and puts the items in the downstream buffer, we move to the continuous flow. The stages a coupled, connected, so the output of the stage A is directly the input of stage B and so on. The sage A has no way to produce more because the stage B will not be able to receive it. Removing the opportunity to create stocks, this means that if one of the stages stop the other stages stop.
What are the **benefits** of this configuration?

- Facilitation about communication.
  It is much easier to communicate with the system that are close to each other and where there are no barriers, both physical barriers and mental barriers. For example, if one of the three people has a problem or if he needs a break, it has to communicate to the others because if one stops the other has to stop doing. Communication is an initial improving for collaboration that is the intermediate steps for integration.

- Reducing the stops.
  You reduce the space occupied and the inventory carrying costs because you have less money mobilized on stocks. So, you reduce costs in general.

- Reduction of throughput time
  Strictly linked to the reducing the stops point. We create a continuous flow and we will use the throughput time of the system.

- Defects
  Imagine we have machines that are working, each operator is working on some machines or some tools. The machine is working but after a certain time, the machine might have a fault, but not the fault that makes the machine stop working, but a fault in the sense that it starts to make a difference.
  What usually happens in a company? You make defective items, you produce them, then you start producing another batch, you produce product A which is defective, then you produce product B, then you produce product C. What happens is that after two weeks, you have the defective item that gets to the point where it does not work or cannot be assembled. What happens is that after two weeks, you have the batch of defective items that get to the point where they do not work or cannot be assembled because they are defective. What has happened? If you ask this question after two weeks, you are not able to figure out what happened because maybe too much time has passed. So if you cannot answer the question "what happened", you cannot establish the main cause of the defect. This means that you will never be able to eliminate the possibility of defects recurring.

  What happens in a continuous flow configuration? That the moment when the defective product is put into the system and the moment when it is detected is much shorter. The impact is lower because we do not have a pipeline full of defective items because the pipeline is shorter. Early detection can greatly improve understanding of defects. In a continuous flow, if you have a defect, you immediately identify the defect. So, is it easy to turn your head or to communicate that there is a defect.

LINKING PROCESSES
What are the criteria that guide the choice to one configuration rather than another?
Looking our example, we can recognize some reference criteria, such as **set-up time**, **probability of breakdown**, dedicated resources or not (for example we can adopt it for assembly and welding). However, the shift from one logic to another is not “on-off”, but there is a continuum of configurations.

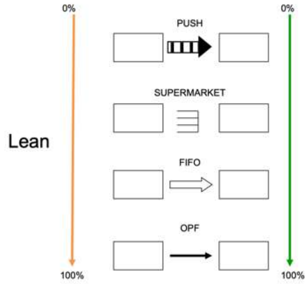
- **Push**: All stages are independent one from another and each of them follows the established schedule, being valued on the basis of the adherence to that plan.
  Push the product: it means that I am in the upstream station and I produce pushing the product downstream and as soon as the other station receives the product, I continue producing. The speed of the system depends on the speed of the pushing: if the second station is slower, the stock will arrive later. The first station keeps producing without caring.

- **Pull** (supermarket): Pull implies to see the production process as operations that goes from downstream to upstream and works pieces only when they are required (so when to replenishment order is received and there was a consume in the downstream stage). Inevitable in this configuration, phases are linked to each other: if the downstream stage stops, also the upstream one will be not able to proceed, after having worked off the queue. In a push system, we cannot put a limit to inventories; instead in supermarket we can control it, although the sequence in production is not constrained, so the stage upstream can replenish the shelf in a different way respect what was done by the upstream one. In the supermarket we can change the sequence, because what is produced depends on what is requested/consumed.

- **FIFO**: we have balanced the line and synchronized the stations, where each station is working with the same sequence. An assembly line is FIFO by definition. It is characterised by:
    - An information flow that maintains the right sequence between the stages;
    - The limitation of the length of the queue: after having reached the maximum value, it is no more possible to introduce further material in the system;
    - The coupling of phases, which is even stronger: only when the downstream phase picks a piece/product/component, the upstream one is activated. As a result, the more the coupling is relevant, the shorter the queue. One peculiarity of this solution respect the previous one is the capability to maintain and respect the right sequence in each phase.

- **OPF**: it is the extreme of FIFO, one-piece flow, in which the length of the queue of the FIFO approach tends to zero: the logic behind is "make one, move one", where stages are exactly in sequence and completely coupled (they are coupled in term of overall amount of inventory, sequence and production rate).

As highlighted by the image, we can observe how, moving from the first to the last configuration, we are reach the ideal system, to which the lean approach tends.



DECAF Condition

Considering **ACMA case**, there is a simple rule to use that has the acronym of **DECAF condition**. DECAF, which is a continuous flow that works if the system that you put in the continuous flow configuration is DEdicated, Capable, Available and Flexible.

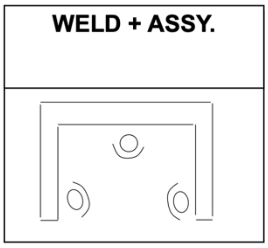
- **Dedicated**: All the stages included in the continuous flow must be the Dedicated to the flow. It is not that they are resources shared with other flows, they must be dedicated to this one, otherwise it does not work. Why if does not work? Because if it is not dedicated, it means that part of the time it works for another flow. So, it means that the entire system will stop.

- **Capable**: The system, which is a complex system as you have seen system physics, must be capable. So, it means that the cycle time must be faster than the tack-time. If the system is not faster, we cannot satisfy the demand.

- **Available**: The cycle time is the ideal capacity. But in the real world resources are affected by failures, breakdowns and quality issues. You must have a system which is available enough, if you adjust your cycle time according to the availability of your resources, you still must be faster than the tack-time.

  reduce set up time.
- **Flexible**: The continuous flow system that you are going to create must be flexible enough in order to satisfy the demand mix that is required by the customer or basically by a part of the system.

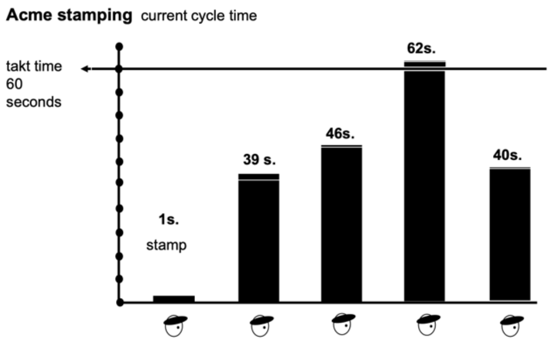
If the system composed by the stages that we put in the continuous flow system respect the DECAF condition we can create a continuous flow. It is important to underline how these conditions must not be intended as an obstacle to the implementation of the lean approach. The right way of thinking is: *"what should we change in order to have a lean system?"*. It is a different perspective respect "can we implement lean?", basing our consideration only on DeCAF criteria. Only in this way, we are able to achieve the real advantages of an interconnected system: it is not necessary to respect the limit of all the DeCAF criteria, but to improve them along time; only in this sense, we have the opportunity to make problems visible and we are obligated to resolve them. The main issue is to define the minimum target of each dimensions, that we want to address; in particular, these minimum target act as intermediate goals that, once they are achieved, must be improved, with the aim to increase the overall performance of the system.



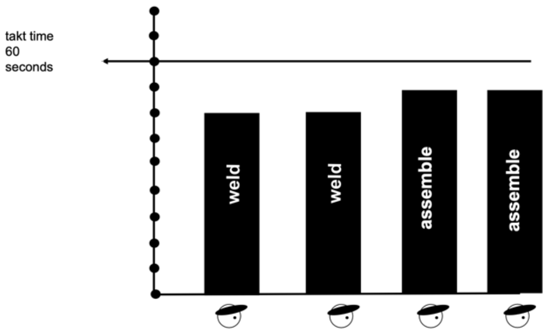
Thinking at the ACMA Case:
Where it is possible to use a continuous flow? Not in the stamping station, because it requires a specific machine, but maybe in **welding** and **assembly**, redesigning the system and training the workers. With set up time this does not work and the ability of the workers should be more or less the same, because we need to balance the line.

The main constraint is to **remove stocks**, we need to remove all the problems (we have seen how in system physics). What is the minimum change needed to achieve a goal?
We need to opt for supermarket when we cannot achieve 100% DeCAF: for example, in stamping, where we have to put shelves in order to allow people to pick product from the shelf.

Therefore, it is evident how, in our case, there are some stages (pre-welding, final welding, final assembly) characterized by a processing time lower than the takt time, while another (pre-assembly) that has a processing time higher than the ideal value. The principle of redesigning the flows derives exactly by the necessity to uniformity the working time at each phase, in order to be aligned with the market's requirements. The answer to this issue does not lie on the creation of stocks; this represents a short term solution, but in the long term, it will represent a waste and even a useless approach, since it will force to recur to overtime.



A first alternative could be the **line balancing**, which implies to redistribute workloads in a uniform way, in order to respect the takt time. However, looking at the lean principles, this is not the right approach: balancing and so trying to be as close as possible to the market is not the correct solution to the problem, because we will progressively saturate all the phases.



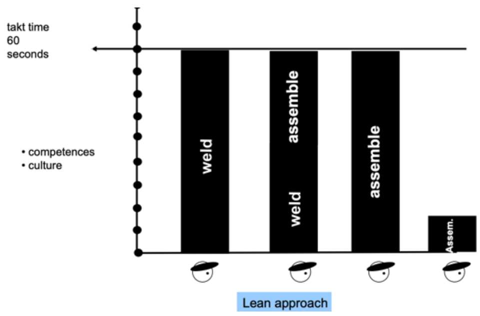
idle is better solution. and it's impossible to utilize everything.

In addition, it takes the perspective of the resource, and not of the product (which represents the customer), like in lean management. In this way, the remaining part would be assigned to the last resource. We can say that this approach is focused on the **product** and on the **client**, rather than on the **resource**.

The **right procedure** is therefore the following:

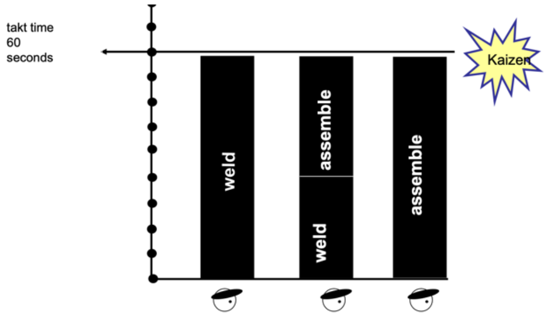
1. We need to define all the activities required, in order to realize the product in the best way possible;
2. We have to compute the takt time;
3. We assign a number of activities equals to the takt time to the first resource, and so on for the other ones, until we reach the required workload. Obviously, as in the reference case, it is easy to assign to the last resource as residual workload, obtaining in this way a non-saturated resource.

The main advantage of this procedure is the **ability to put in evidence all the wastes**: in fact, we are able to separate resources and activities, dividing the value added from the non-value added one. In particular, **it emerges how the first three resources are fully saturated while the last one does not** (we need to note that, in the previous configurations, each phase presents inefficiencies, while in this case only one – the spare time is still the same).



In addition, it is easier to resolve a problem that is circumscribed to only one stage, rather than extended to all of them; but also if only one phase is characterized by an idle time, it is possible to plan other activities (for example the assembly of another product, maintenance and cleaning).

The goal is the **ideal approach of Kaizen**, trying to find a way to do the same product by **using less workload**. In fact, if we look at the first alternative and if we are able to **increase the efficiency** of the first three departments, we can think to remove the unsaturated resource, moving it under the responsibility of the other ones.



Ideal situation (pursue perfection)

Obviously, this transformation is not lacking **difficulty**; in particular, the **main constraint** is represented by the redefinition of the resource's skills, which will imply the inclusion of additional activities (for example welding is a very complex process and requires lots of training). It is necessary to understand that barriers are not linked to the inability to execute the operations, but to the cultural, image and heritage limits of the organization and employees.
Lean approach revolutionizes the traditional wisdom, especially if there is the possibility to modify the relationship with resources, passing from a hostile to a collaborative one.

If those barriers are overcome, we have the opportunity to obtain the wanted improvements: in particular, the ideal situation consists in working with the **minimum number of people**, indispensable to follow the pace of the market:

total processing time

Minimum number of people = work content/takt time

In this exercise 7 seconds is the waste. This indicator has a double meaning, because it measures the degree of waste in term of human resources, compared to the actual state. It is important to note that if we buy a more sophisticated tool, another fixture, we can save one operator. With line balancing we cannot think about this possibility. However, it is mandatory to protect the resource: if one operator is spared, it will not be fired. We can use it to increase volume because we are able to reduce costs and prices. We can insource something that before was outsourced, saving additional money because resource is a fixed cost. It is more convenient to dedicate time to other activities rather than waiting.

The key point of these consideration is the following: when we design a **lean system** (therefore compatible with the takt time required by the customer), it is necessary to be able to structure it in a flexible way in order to adapt to unexpected changes.

4. How do we connect stages that we are not able to put in the continuous flow?

In a **simplified supermarket configuration**, there is customer and a supplier, where the customer has built inventories of boxes, each of them containing ten pieces. When the customers open a box, he/she detached from it a ticket, called **Kanban**, with all the description about the products; for instance, the number of units, the supplier from which it comes from and the customer.

This ticket is then sent to the supplier. And so on for the other boxes, in this way the supplier has an idea of the inventory of the customer and follows the Kanban to start production. This is called **Kanban loop**, because the tickets are sent again to the customers, once the order is fulfilled. In this case we are not constrained by the sequence, but we need only to avoid that customers stop us.

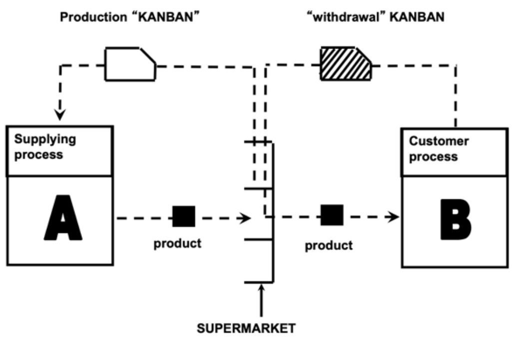
We can classify order on the basis of their urgency.
It is important to keep in mind that this process can be replicated to all the other stages of the production system, and that the supermarket system should be placed in proximity of the supplier, due to the fact that stocks represent the inability of the supplier in following exactly the demand.

The situation is more complex if we consider a **two - Kanban system** or supermarket pull system, because here there are two Kanban processes:

- CUSTOMER PROCESS goes to supermarket and withdraws what is needs when it needs it. ("withdrawal" Kanban)
- SUPPLYING PROCESS produces to replenish what was withdrawn ("production" Kanban)
- PURPOSE: Controls production at supplying process without trying to schedule. Controls production between flows.

Customer, when opens a box, detaches a ticket called withdrawal Kanban that triggers the transport and not the production. From the storehouse, the transport Kanban is substituted by the production Kanban that then is sent to the supplier in order to activate production. As a consequence, we can recognize two different loops: one for the withdrawal and one for the production.

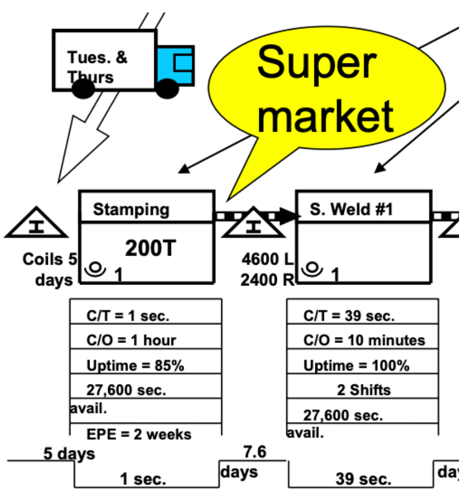
A **"production" Kanban** triggers production of parts, while a **"withdrawal" Kanban** is a shopping list that instructs the material handler to get and transfer parts.



The purpose of placing a **pull system between two processes** is to have a means of giving accurate production instruction to the upstream process, without trying to predict downstream demand and scheduling the upstream process.

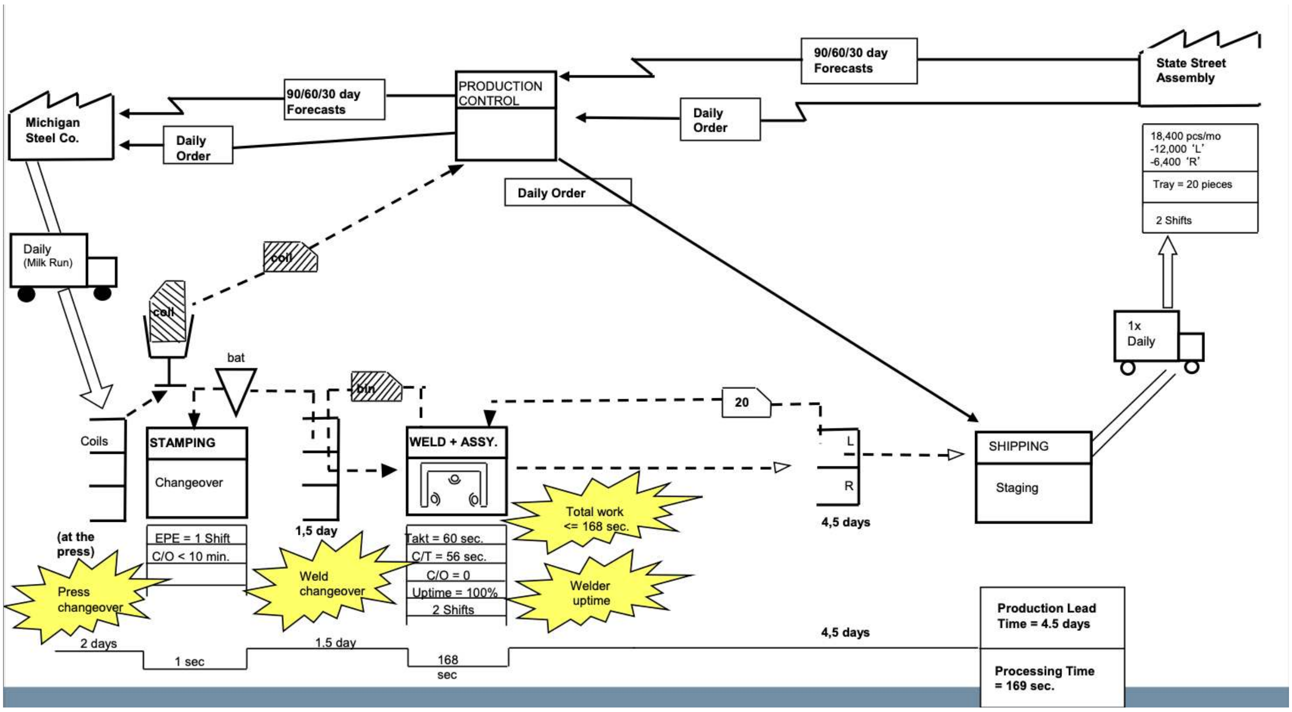
Pull is a method for controlling production between flows. Get rid of those elements of your MRP that try to schedule the different areas of your plant. Let the downstream process' withdrawals out of a supermarket determine what the upstream process produces when and in which quantity.

Where is the best position for the supermarket in the Acme VSM?
The only way to decouple the system in two parts is to use a supermarket, which needs to be placed downstream of the stamping machine, where the set up time is very high due to high changes (the high set up is not a good thing). The lower the set up time, the lower the number of Kanban, so the machine can go from one product to another in an easier and quicker way.



Moreover, DECAF is not possible between stamping and welding due to the fact that stamping is not a dedicated resource. As a consequence, we have to opt for a supermarket configuration, which allows to use upwards productive sequences and management logics that are partially independent respect what it is required by the downstream stage. It is evident that, if the cell starts producing with the same sequence of the stamping phase, it should inevitably manufacture left products for two weeks, and only right products for the other following weeks.

The future state after, having placed the supermarket, will look like: with two shelves for the 2 products and with one shipping per day, without changing the value of the stock. Takt time is now 60 seconds, and C/T is different. C/O is zero now. The total work has changed.



5. How do we plan, schedule the production system? PACEMAKER Process

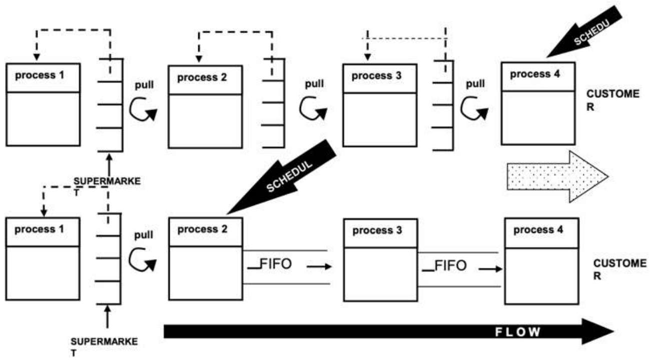
The greater portion of wastes and inefficiencies in the current state derives from the fact that the production planning assigned to each stage a specific schedule, strictly followed without a systemic vision. The result was the creation of stocks and the misalignment of the schedules themselves: in fact, each plan is designed in order to optimize the workload of each stage, protecting it from any negative event. However, these schedule derive from forecasts, which are highly influenced by variability.

All of these have to disappear in the future state: How future state works? We receive the customer order, we sent it to the shipping department from which the order is delivered. Then shipping communicates to weld+ass that it is necessary to replenish the shelf. So they start producing the amount of product that were delivered and they put them in the finished product warehouse of the supermarket. There should be one single **scheduling point**, on the basis of which the other stages have to adapt and react. Only in this way, we have the chance to minimize the probability of misalignment between the phases, building a system that is highly flexible and easily controllable: in the moment in which the central stage is obligated to review its plans, also the other ones have to consequently react, realigning to this change.

The more we reduce the lead time, the more changes we are able to achieve, the more the EPE and the batch size will be reduced.
The pacemaker is where we schedule (scheduling point) and it gives the rhythm to the system.
By using supermarket pull systems, you will typically need to schedule only one point in your door-to-door value stream. This point is called the **pacemaker process**, because how you control production at this process sets the pace for all the upstream processes. For example, fluctuations in production volume at the pacemaker process affect capacity requirements in upstream processes. Your selection of this scheduling point also determines what elements of your value stream become part of the lead time from customer order to finished goods.

Note that material transfers from the pacemaker process downstream to finished goods need to occur as a flow (no supermarkets or pulls downstream of the pacemaker process). Because of this, the pacemaker process is frequently the most downstream continuous-flow process in the door-to-door value stream. On the future-state map the pacemaker is the production process that is controlled by the outside customer's orders.

CASE 1) The first one is directed by the finished products warehouse: the entire process produces for the final buffer, which means that programming is done on the basis of the requirements that come from the customer. As a consequence, it is necessary to put the scheduling point in correspondence to the last stage: every pick to satisfy demand activates the previous stages in order to obtain the replenishment of this pick◊in the case of MTS or supermarket.
EX) If number 1 is controlled by number 2 and number 2 is controlled by number 3 and number 3 is controlled by number 4, the scheduling point is at station number 4.



Schedule production at a single point in the production chain (the "pacemaker process")

FIFO line downstream of scheduling point

CASE 2) The second option consists in fixing the scheduling point in the phase that directly receive the customers' orders. In this case, it is evident that the downstream stages will react to and follow the sequence established by the scheduling point, but also the upstream supplier will act in turn in reaction (in the moment in which the production will start, it will have to restore consumptions). In this way, there will be not the possibility of misalignment: in fact, there will be the exact maximum number of stocks, permitted for a FIFO logic → when we differentiate our products

The second station pushes the other stations (controlling them) so the scheduling point is at station 2. Process 3 is activated by process 2 and process 4 is controlled by process 3, so both 3 and 4 have no freedom. The schedule point is placed where there is the highest degree of freedom.
Anyway, we need to keep in mind that the adoption of one configuration or of the other requires that the stages that have to react should be characterized by a high level of flexibility, due to the fact that they have to react immediately to changes. In addition, the higher the flexibility, the less the necessity to schedule again.

Acme VSM: the schedule is downstream the supermarket. We have obtained a better and more efficient system respect the original one; a further proof is given by the fact that with this configuration it is possible to operate with a lower level of stocks.

6. How do we manage the mix of the production?

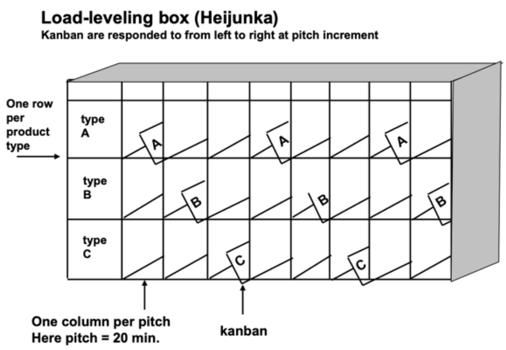
Levelling the product mix means distributing the production of different products evenly over a time period. For example, instead of assembling all the "Type A" products in the morning and all the "Type B" in the afternoon, levelling means alternating repeatedly between smaller batches of "A" and "B". **The more you level** the product mix at the pacemaker process, **the more able you will be to respond to different customer requirements** with a short lead time while holding little finished goods inventory. This also allows your upstream supermarkets to be smaller.

But be warned that levelling the mix requires taking some pains in assembly, such as more changeovers and trying to always keep all component variations at the line (to eliminate changeover time). Your reward is the elimination of large amounts of waste in the value stream. The icon for levelling is the following symbol, which is inserted into an information flow arrow.

How will we level the production mix at the pacemaker process? How do we level production?
We can level in term of **volume** (same amount every day) or level in term of **mix** (it is more complex and it means to change the perspective). The **traditional approach** implies to change as less as possible, from large to average to small. The **lean approach**, instead, says to produce the same every day: we reduce the impact from changing from large to average to small, but we have the advantage to produce the same sequence every day. In other words, it means to product in a short interval the average mix/proportion over a month. We replicate the same proportion on a shorter frame. In order to do that, we can rely on the **load-levelling box**.

LOAD-LEVELING BOX

A good place to start is to regularly release only a small, consistent amount of production instruction (usually between 5-60 minutes worth) at the pacemaker process, and simultaneously take away an equal amount of finished goods. We call this practice a "*paced withdrawal*".



We call the consistent increment of work the **pitch** (pacemaker interval), and often calculate the pitch increment based on packout container quantity (the number of parts one finished-goods container holds), or a multiple or fraction of that quantity.
For example: If your takt time = 30 seconds, and your pack size = 20 pieces, then your pitch = 10 minutes (30 sec x 20 pcs = 10 minutes).
In other words, every 10 minutes:

- give the pacemaker process instruction to produce one pack quantity;
- take away one finished pack quantity.

So in this case pitch means multiplying your takt time upward to a finished-goods transfer quantity at the pacemaker process. This then becomes the basic unit of your production schedule for a product family.

A **load-levelling box** has a column of Kanban slots for each pitch interval, and a row of Kanban slots for each product type.
→ Pitch, which is the interval of control and the lot of information that are sent by the pacemaker (the scheduling point).

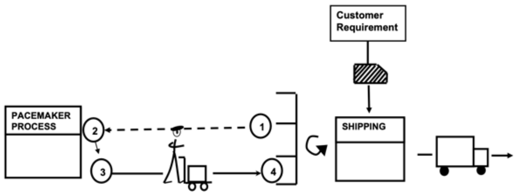
In this system, the Kanban indicates not only the quantity to be produced, but also how long it takes to produce that quantity based on takt time. In fact, at this point it is not only required to define the right production sequence, but also its speed: provide orders at smaller intervals means communicating to proceed in a faster way, while diminish the quantity means proceeding in a slower way.
Kanban are placed (loaded) into the levelling box in the desired mix sequence by product type (see levelling box illustration). The material handler then withdraws those Kanban and brings them to the pacemaker process, one at a time, at the pitch increment.

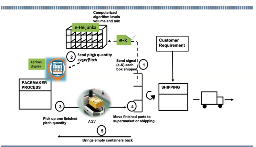
If there are 3 types of products (A, B, C) and three rows and one Kanban means one tray, you take A and you produce A, letting empty the column, so you go to the second column. The process is in sequence from left to right. In this case pitch is equal to 20 min: this guy, every 20 min, goes to the levelling box and takes all the Kanban which he sees in the first column (in this case A, while in a more generic situation we could have A and B and so on). If there is more than 1 Kanban, the peacemaker will take more than 1 Kanban per pitch. One column one pitch, every 20 min. This is the way to synchronize the system. The pacemaker is a point and a physical person.

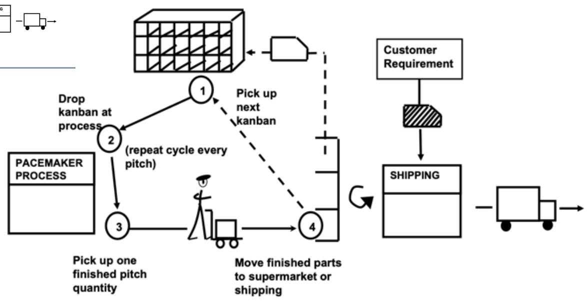
EX) Suppose that a Kanban corresponds to a tray of 20 pieces and that the takt time (final demand) corresponds to 1 piece every minute; this means that one box will be required every 20 minutes. If we decide that this is the control slot, each column will have at maximum one Kanban, due to the fact that one ticket is equal to 20 minutes. Each phase has to produce one Kanban every 20 minutes and not more or less; otherwise, we will incur in overproduction or we will not be able to meet demand.
This is the levelling in term of volume, because every 20 minutes one tray has to be produced.
In this ambit the resource that is responsible for the respect of the takt time is called runner; the operator moves inside the factory, paying attention to the flow of Kanban, as shown by the following figure:

- Shipping picks a tray and detaches a transport Kanban;
- The transport Kanban is then given to the runner, which brings it to the cell, allowing the star of the production;
- If the pace is respected, the cell should provide the exact amount of products to replace the previous consume, which then is transferred to the supermarket;
- Subsequently, the runner will take a new transportation Kanban and will start another tour.

He/she goes through the tour in a defined interval of time: if the frequency is equal to one Kanban at a time, the tour will be equal to 20 minutes. It represents a sort of feedback and a check of the production pace of the stages, acting as a real metronome (in fact we can understand the degree of delay).





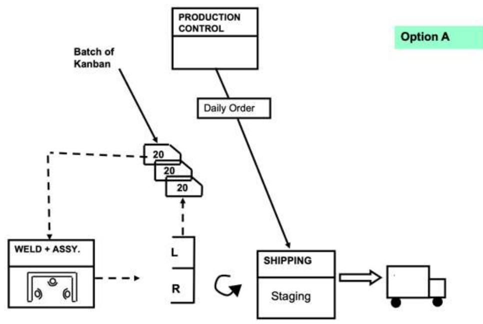


Obviously, we can deal with more sophisticated situation, as shown. In the case we plan the levelling, the transportation Kanban is not directly transferred to the cell, but to the resource that manages the load levelling box, which will insert it in the correct column.

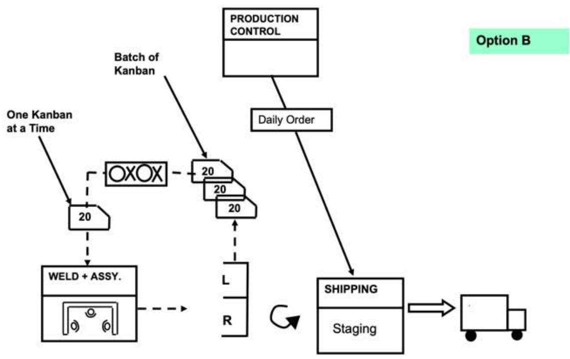
Finally, the runner will pick the Kanban located in the first column on the left side, bringing it to the relative stage.

It is fundamental to keep in mind that we need to schedule and level production always on the basis of the target determined by the takt time: if the demand will be higher than this value, the surplus will be allocated to the following day, in addition to the demand that it will occur in that day. In other words, we produce on the average value, leveraging the inventories in case of variability.
We need to pay attention because these considerations are not valid in the case of a lower value of the demand respect the average. In fact, without Kanban, we are not allowed to produce.
Since the redesign of the process looks at a system that produces an entire family of products, it is appropriate to comprehend which are the logics that lie behind the levelling and sequencing of the mix. In an ideal system, we should produce the exact mix that the customer requires.
However, it is possible to identify many different alternatives.

The **first one** coincides with the no-levelling: the customer sends the order to the warehouse, which picks the required boxes and prepares them for the shipping. Picking a tray, one or more transport Kanban are detached and sent to the cell (as an input for the replacement) without levelling (the cell sees the arrival of the Kanban on the basis of the same logic of picking in the warehouse).

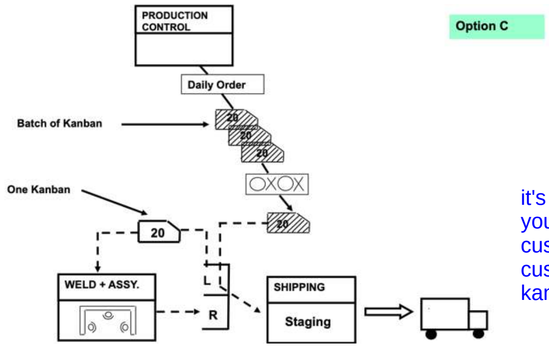


P25 is reviwing



The **second option** requires to level the Kanban before they are transferred to the cell: the system does not receive the entire amount of batches, but one at a time, on the basis of the levelling logics.
This guarantees that on single Kanban for each column of the load levelling box is produced, regardless the type of picking in the warehouse. As a result, the cell is not able to decide to produce a batch that is different from the one defined by the Kanban (increasing in this way the batch size).

The final alternative coincides with possibility to level upward the finished products warehouse, so before the scheduling point:
20 → number of units for each Kanban
OX → icon for levelling, mix and volume
B → we level demand and mix
C → we level the withdrawal Kanban, we do not produce



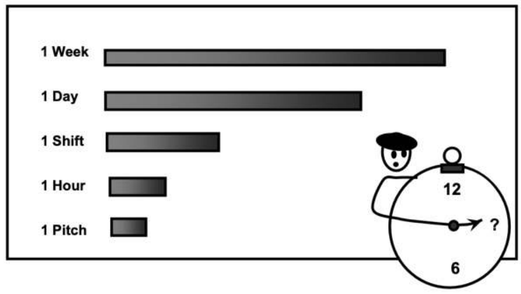
it's not working.
you have to send to customer!
customer send kanban works

In other words, we do not sell to the warehouse the order but a production Kanban: in this way, we are able to level production even before the execution of the shipping, which receives as an input a levelled scheduling.
The main benefit of this solution is to minimize the inventories in the finished products warehouse.

7. What is the increment that we release in the system, the minimum unit we can manage and move inside the system? Single items or trays of items?

This is referred to the management interval, so the pitch. It is normally smaller than 1 hour, because it gives the rhythm of the system and we should decide where to put it.
We have seen that the "*runner*", also called "*water spider*", is the figure that has the task to collect Kanban with a certain pace. It is evident that this pace depends on how the control interval and the transfer of the Kanban have been defined. Consequently, it is possible to schedule in order to transfer a Kanban every 20 minutes, so 3 Kanban per hour (the amount of information and workload that we bring to the pacemaker, and we bring back, is one hour worth) and so on.
However, change the interval means also organize in a different way the columns of the load levelling box, due to the fact that a column could correspond to 20 minutes of the customer demand (extremely punctual), or to 3 hours of demand (with a higher aggregation). It is obvious that a more punctual solution allows to monitor strictly the pace of the production: on the one hand, we are able to check what we have produced every 20 minutes; on the other hand every 3 hours. There is a very relevant difference.
With a control that is exactly equal to the Kanban, the check is more stringent and efficient; however, we must keep in mind that this advantage implies a higher waste in term of runner's movements. But the more the management interval is punctual, the more the production works without problems, respecting the pace. In fact, working in a so small interval, if there would be problems, we will be unable to resolve them. As a consequence, it is appropriate to proceed gradually and having a stricter control only if the system is able to improve.

The concept of **pitch** represents this dimension; it is the managerial interval of control (how often the runner starts his tour and check production). We do not have to forget that the pitch is correlated to the production volumes and workload: the more the product is complex and volumes high, the more the control should be stringent.



Pitch: lot of information

In some case one day is more than enough, in other they work by the week (the due date is set in weeks). This depends on the business. The larger this amount the more difficult is to level demand. It is much easier to produce 24 trays in one day than 1 batch every 20 minutes.

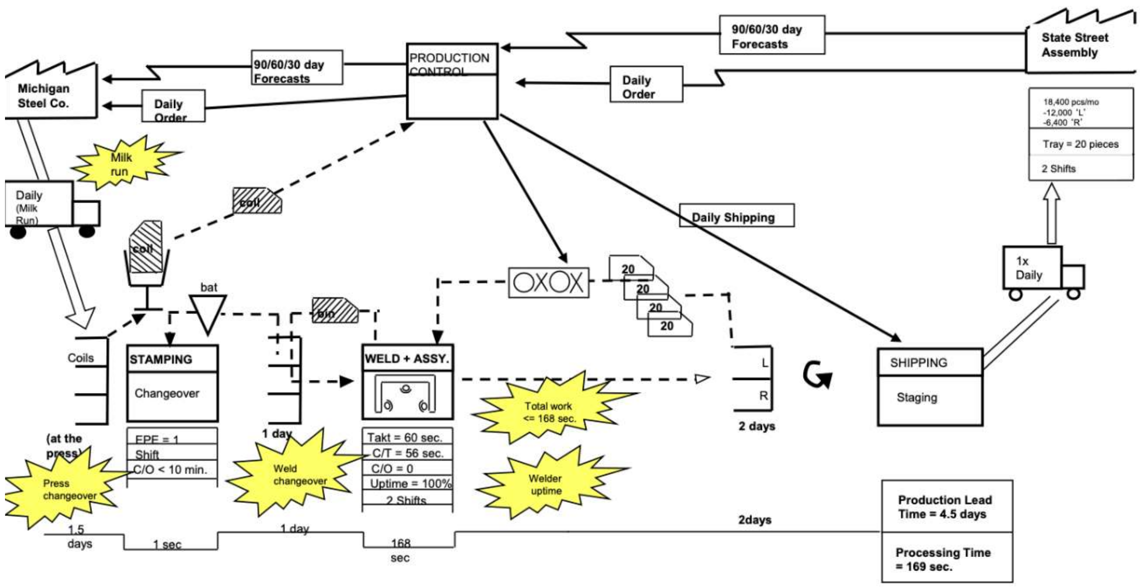
8. With the first 7 we are changing the configuration of the system (the goal), but in order to make it feasible, what are the necessary improvements, operative changes?

Also considering the levelling of the demand we can discover that the lead time decreases to 4.5 days.

- From push to pull → first improvement
- Level demand → second improvement

Acme Stamping Lead-Time Improvement

	Coils	Stamped Parts	Weld/Assy WIP	Finished Goods	Production Lead Time	Total Inventory Turns
Before	5 Days	7.6 Days	6.5 Days	4.5 Days	23.6 Days	10
Continuous Flow & Pull	2 Days	1.5 Days	0	4.5 Days	8 Days	30
With Leveling	1.5 Days	1 Day	0	2 Days	4.5 Days	53



Push or Pull?
pull >> if the downstream stops, upstream stop. but the assembly is push! because won't stop

19. VALUE STREAM MAPPING - EXERCISES

19.1. EXERCISE 1

Demand

- The company works on two shifts, each of eight hours, with a 30 minute break for each shift.
- The production family A is made by a range of 68 different products, from A1 to A68 (products with low volume/high variety).
- The total demand of the production family A is on average 450 pieces per day.

Processes

- There are three main production stages (machining, heat treatment, assembly, in this sequence). In each of these stages, there are one machine and one operator fully dedicated to this product family. Downstream of the assembly, the piece goes to a warehouse, ready for shipment. Before each stage there is a stock where pieces wait to be processed.

Stage	Cycle Time (sec/pc)	Set-up (min/setup)	Availability (%)	Stock (n° of pc. downstream)
Machining	95	10	95	1500
Heat treatment	80	0	95	3200
Assembly	105	0	100	10430

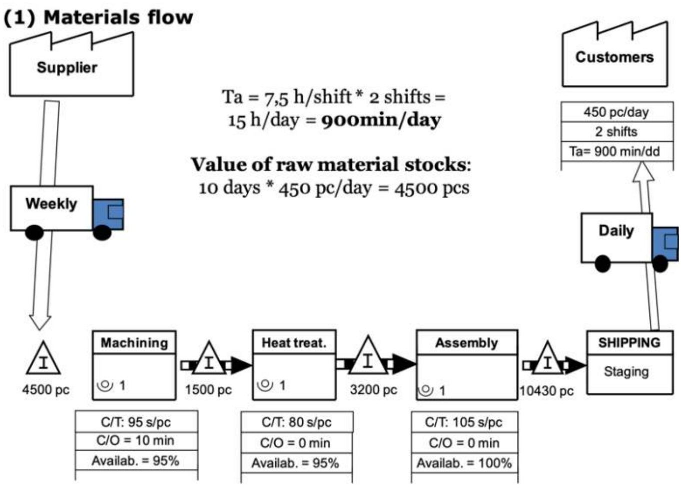
- Cycle times and setups are very similar for each product.
- Upstream of the machining stage there are stocks of raw material for a requirement of 10 days.

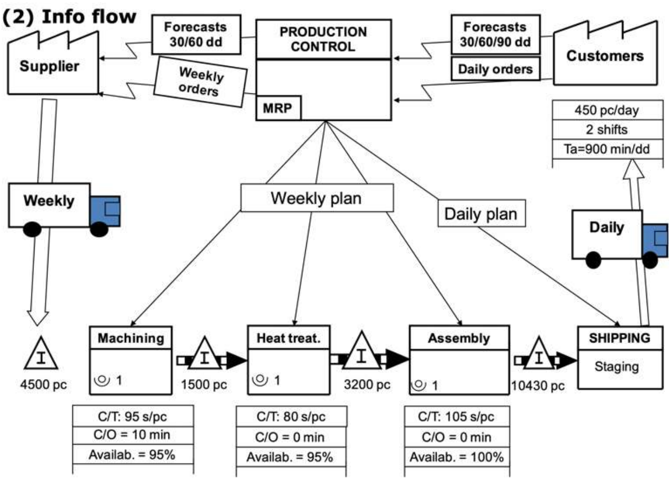
Information flow

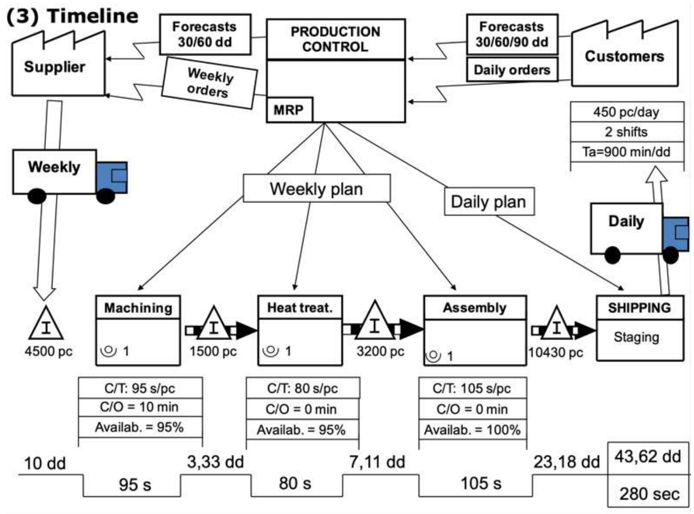
- Customers make a forecast (to 30/60/90 days) and send it monthly to the Production Planning department (PP).
- Customers send orders to Sales daily, with agreed shipping dates, generally equal to 15 days. The demand of the various products of the family is quite similar.
- PP tracks in the system the orders received from Sales and develops a weekly production plan for each production process and a daily scheduling plan for deliveries to Shipment. Company is indeed working with daily shipments.
- PP sends a forecast (30/60 days) to the supplier and confirms orders weekly. The supplier delivers weekly.

Q1: Using icons and criteria seen in class, draw the Current State Map of the company based on the data provided.
Q2: The general manager defines as internal target that each stage of the company must be flexible enough to produce all the variants each day. Which improvement should the company implement?

Q1: 3 steps:
1. Material flow
2. Information flow
3. Timeline







Q2: The general manager defines as internal target that each stage of the company must be flexible enough to produce all the variants each day. Which improvement should the company implement?
EPE (every part ever) = time required to sort the whole product range

EPE*Tp + Ts ≤ EPE*Ta

Being able to daily supply customers with the whole range -> EPE = 1
The stage that creates rigidity to the system is the **machining.**
All other stages have null setup times.

Ta= 900 min/day
Product range= 68 products
Demand= 450 pc/day

EPE*Tp + Ts ≤ EPE*Ta 1*Tp + Ts ≤ 1*Ta

We need to verify for Machining stage if the equation is valid
i.e., whether Machining stage is able to fully satisfy customers requests in terms of volume and mix.

$$1*\frac{95\ sec/pc}{60*95\%} * 450\ \frac{pc}{day} + \overbrace{10\ \frac{min}{setup} * 68 \text{ setups}}^{Ts} \leq 900\ \frac{min}{day}$$

$$1430\ \frac{min}{day} \leq 900\ \frac{min}{day}$$ X

Hence, according to the current situation the company is not able to fully satisfy customers demand daily.
We need therefore to improve the system i.e., redefining setup time for Machining so then it is possible to achieve EPE = 1 day.
1*Tp + Ts ≤ 1*Ta

$$1*\frac{95\ sec/pc}{60*95\%} * 450\ \frac{pc}{day} + 10\ \frac{min}{setup} * \text{x } 68 \text{ setups} \leq 900\ \frac{min}{day}$$

Where x is the setup time.
$\boldsymbol{x} \leq 9000/ 4080 = 2{,}2$ minutes
Reducing the setup time by SMED.

19.2. EXERCISE 2

Demand

1. The company works on two 8-hour shifts. Each shift has a 30 minutes break.
2. The production family B is made by a range of 6 different products, from B1 to B6 (all high-volume products, sold every day).
3. The total demand for the production family B is on average of 1800 units per day.
4. Twice per day products are delivered to customers.

Process

- The company has 5 production stages + shipping department. Press and cutter are parallel stages.
- Afterwards component coming from press and cutter are welded together (welding stage). Then there is drilling and finishing stages.
- The press and cutter are shared resources with other production families. The press is dedicated to this production family for 60% of its time. The cutter for 70% of its time. All other resources are instead fully dedicated to this family.
- The press and cutter are technologically advanced machines, others are quite simple manual machines.
- The press works 6 different types of products for the concerned production family.
- The cutter produces 6 different types of components, one for each of the six finished products offered by the company. On each piece (metal foil) of the press must be welded two identical cut components (pipes).
- One operator conducts each production stage.
- The raw material supplier delivery weekly metal sheets that feed the press and pipes that the cutting needs. There are 13 days of supply for raw material: respectively 13 days of metal foil in front of the press and 13 days of pipes in front of the cutter.

Stage	Cycle Time (sec/pc)	Set-up (min/setup)	Availability (%)	Stock (n° of pc. downstream)
Press	4	120	80	20000
Cutting	10 (sec/component)	15	100	21600
Welding	10	0	100	4000
Drilling	12	100	95	5000
Finishing	23	0	100	30000

- Cycle times and changeover times are very similar for all products.

Information flow

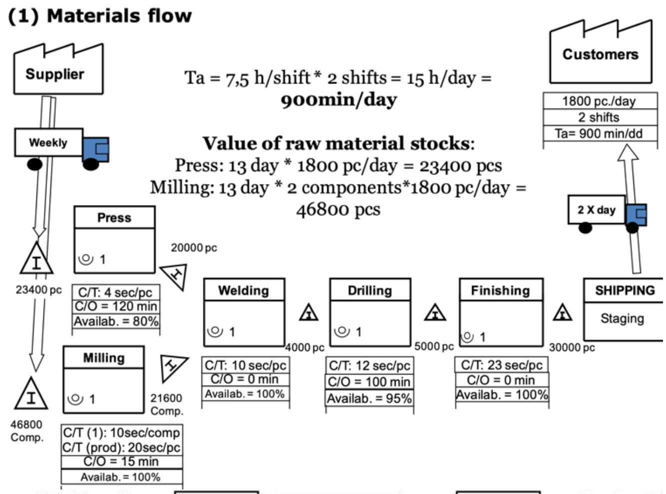
- Customers make a forecast (to 30/60/90 days) they send to the Production Planning (PP).
- Customers issues orders daily to Sales, with expected delivery date the following day.
- PP tracks orders received from Sales in the system and generates a weekly production plan for each process and a daily dispatch plan for Shipment.
- PP sends a forecast (30/60 days) to the supplier and confirms orders weekly.

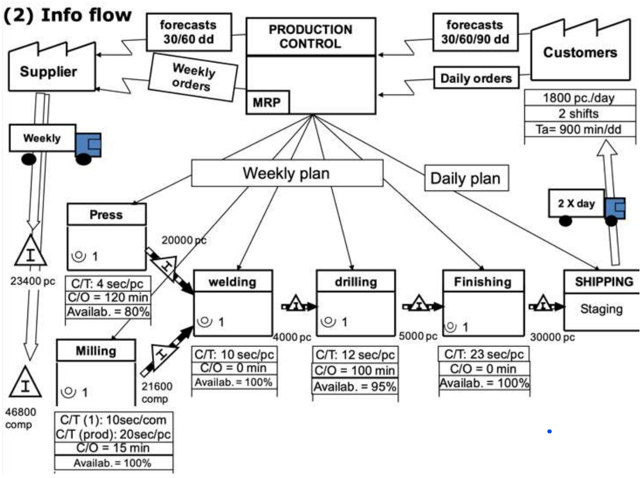
Q1: Using icons and criteria seen in class, draw the Current State Map of the Company based on the data provided.
Q2: In the current situation, which is the EPE of the company to fulfill the requirements of demand (volume and mix)?
Q3: The new plant manager decided to decrease warehouse costs. Therefore, she decided that it is possible to stock at maximum the WIP required for the production of 3600 final products. Draw the Future State of the company, highlighting all suggested improvements in order to ensure the target achievement.

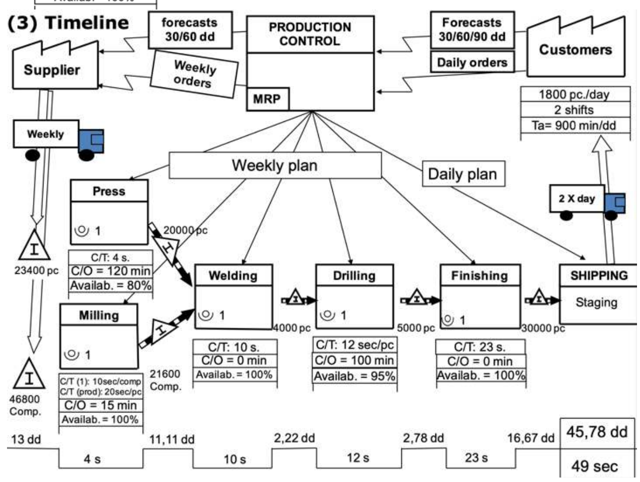
Q1: 3 steps:
1. Material flow
2. Information flow
3. Timeline

2nd inventory: 21600 two unit so u use 3600/d
so 20000 over 1800
so the number is for press

devide by daily demand







Q2: In the current situation, which is the EPE of the company to fulfill the requirements of demand (volume and mix)?

EPE (every part ever) = time required to sort the whole product range

$$EPE*Tp + Ts \leq EPE*Ta$$

$$EPE \geq \frac{Ts}{Ta-Tp}$$

What is the EPE for the company?

It's necessary to calculate the time, in days, necessary to restore the product range, under the constraint of available times and the demand required by the customer.

Press

$4/0{,}8*1800*X + 6*120*60 \leq 900*0{,}6*60*X$

$X \geq 43200/23400 = 1{,}85$ dd

Milling

$(10)*2*1800*X + 6*15*60 \leq 900*0{,}7*60*X$

$X \geq 3$ days ←

Drilling

$12/0.95*1800*X + 6*100*60 \leq 900*60*X$

$X \geq 1.15$ days

Welding and Finishing have C/O null.

COMPANY'S EPE = 3 DAYS

Q3: The new plant manager decided to decrease warehouse costs. Therefore, the maximum size allowed for any stocks or buffer is 2700pc. Draw the Future State of the company, highlighting all suggested improvements in order to ensure the target achievement.

Buffer size = 1,5 * EPE^{target} * D = 2700pc
EPE^{target} = 2700pc / (1,5*D) = 1 day

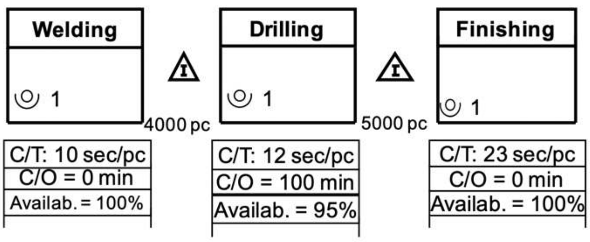
The Future State Map is built by answering 8 key questions:

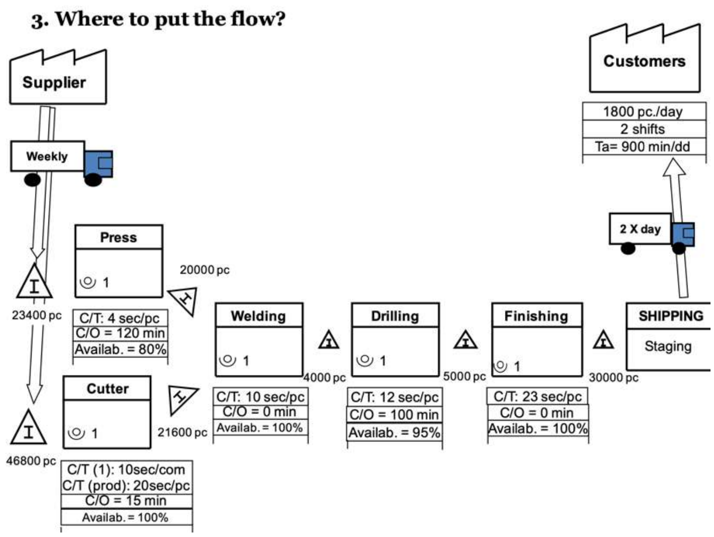
1. **What is the takt time of the production family?**
   Time Available for production (Ta)= Time plant opening– Scheduled stops Customer request = customer demand (d)
   Takt time (TT) = $\frac{\text{Ta}}{\text{d}}$
   Ta = 900min/day
   D = 1800 pc/day
   TT= 900 min/day/1800 pc/day = 30 sec/pc

2. **The company needs to build the product for a finished products supermarket or directly for shipment to the customer?**
   How to organize the production?
   Some drivers:
   - Delivery time to the customer
   - Product features (good value, obsolescence level, standardization level of the product)
   - Demand predictability
   - Demand stability
   - ....

3. **Where to put the flow?**
   <u>General methodology</u>
   - Start from the final stage and go upstream thinking stage by stage where to put CONTINUOUS FLOW and where to decouple (with SUPERMARKET or FIFO).
   - Verify DECAF Conditions.
   - Fix intermediate targets (not necessarily all at once in a continuous flow, but also FIFO and supermarket).

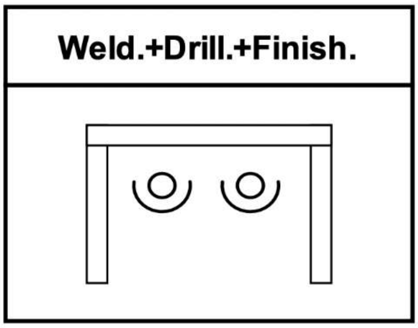
DeCAF condition
Dedicated
Capable
Available
Flexible





You can create a cell for welding, drilling, assembly.
What actions are needed?
Create a cell means to bring machines closer and remove the decoupling stocks.

- How many operators are needed at the cell?
- What is the cell cycle time?
- What is the cell Availability?
- What is the cell setup time?



C/T= 23 sec/pc
C/O= 100 min
A = 95%
2 Operators

$$WCT=(10+12+23)sec = 45\ sec$$

$$\#operators=\left\lceil\frac{WCT}{TT1}\right\rceil=\left\lceil\frac{45}{30}\right\rceil=\lceil 1{,}5\rceil= \mathbf{2\ operators}$$

We can save one operator
Rebalance of work: first operator works on 2 machines for welding and drilling, the second on the finishing machine
(1) Workload of first operator: 22s
(2) Workload of second operator: 23 s

- Dedicated: yes
- Capable: CT<TT

23 sec < 30 sec
yes

- Available: CT/A < TT

23 sec/0,95 < 30 sec
yes

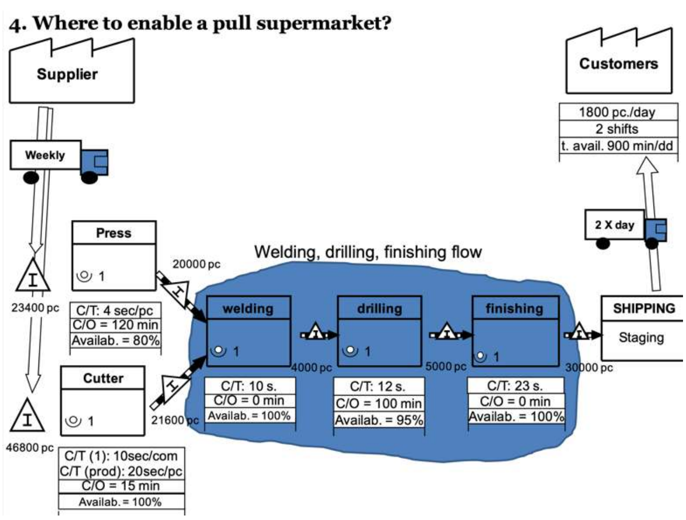
- Flexible

Which is the cell EPE?

$$EPE \geq \frac{Ts}{Ta - Tp} = \frac{6 * 100 * 60}{900min * 60 - \frac{23}{0{,}95} * 1800} = 3{,}45\ \text{days}$$

Improvement:
EPE^{target} = 1 day
EPE*Tp + Ts ≤ EPE*Ta 1*Tp + Ts ≤ 1*Ta
1 ∗ (23 sec/pc)/95% ∗ 1800 pc/day+ $\boldsymbol{x}$ ∗ 6 setups*60 ≤ 900 min/day*60
Where $\boldsymbol{x}$ is the setup time
$\boldsymbol{x}$ ≤ 28,94 min/setups
Reducing the setup time by SMED.

4. **Where to enable a pull supermarket?**
    - High production batches
    - Low machinery reliability
    - Shared resource...



Starting from the initial situation of stage rigidity, identify a first initial goal of improvement (EPE reduction), then proceed to subsequent improvements.

<u>Press stage</u> (initial minimum EPE = 1.85 days)
First improvement goal: EPE^{target} = 1 day
Calculate the first reduction target of the setup time
EPE*Tp + Ts ≤ EPE*Ta 1*Tp + Ts ≤ 1*Ta
1*Tp + Ts ≤ 1*Ta

$$1 * \frac{4sec/pc}{80\%} * 1800\ \frac{\text{pc}}{\text{day}} + \boldsymbol{x} * 6\ \text{setups}\ \leq 900\ \frac{min}{day} * 60 * 60\%$$

Where **x** is the setup time

$\boldsymbol{x} \leq 65$ min/setup

<u>Cutter stage</u> (initial minimum EPE = 3 days)

First improvement goal: EPE^{target} = 1 day Calculate the first reduction target of the setup time

EPE*Tp + Ts ≤ EPE*Ta

1*Tp + Ts ≤ 1*Ta

$$1 * \frac{10sec/pc}{100\%} * 1800\ \frac{\text{pc}}{\text{day}} * 2\ \frac{comp}{pc} + \boldsymbol{x} * 6\ \text{setups}\ \leq 900\ \frac{min}{day} * 60 * 70\%$$

Where **x** is the setup time

$\boldsymbol{x} \leq 5$ min/setup

Size of internal supermarkets downstream Press= 2700 pc

Size of internal supermarket downstream Cutter = 2700pc * 2= 5400 pc

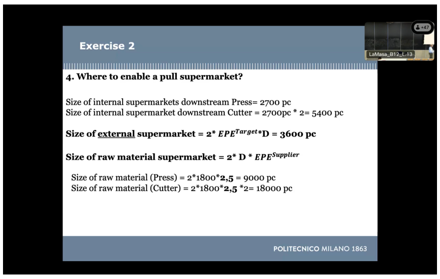
Size of external supermarket = 2* EPE^{target} *D = 3600 pc

Size of raw material supermarket = 2* D * LT supplier

Size of raw material (Press) = 2*1800***3,5** = 12600 pc

Size of raw material (Cutter) = 2*1800***3,5** *2= 25200 pc

incorrect, correct ==>>>>



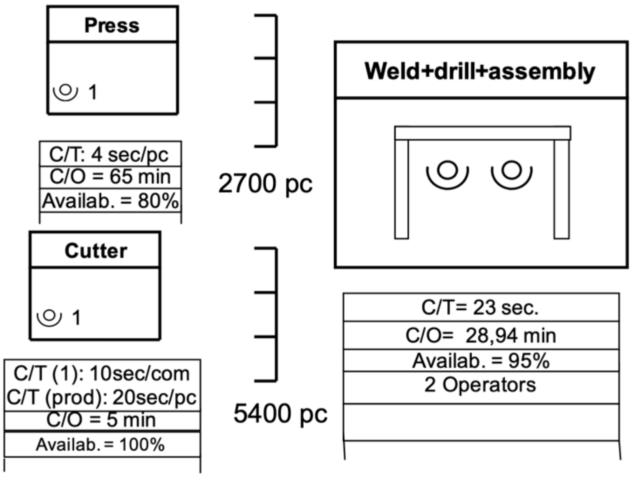
5. What only point in the production chain (the pacemaker process) does the company have to plan?

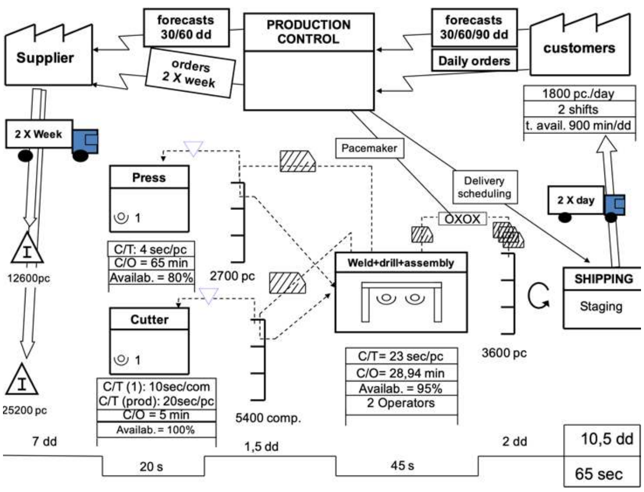
Pacemaker is the cell

6. How should the company level the production mix at pacemaker process?

Kanban System

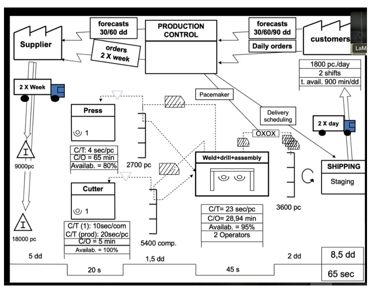
According to shipments that are made, the finished goods warehouse sends upstream the Kanban.



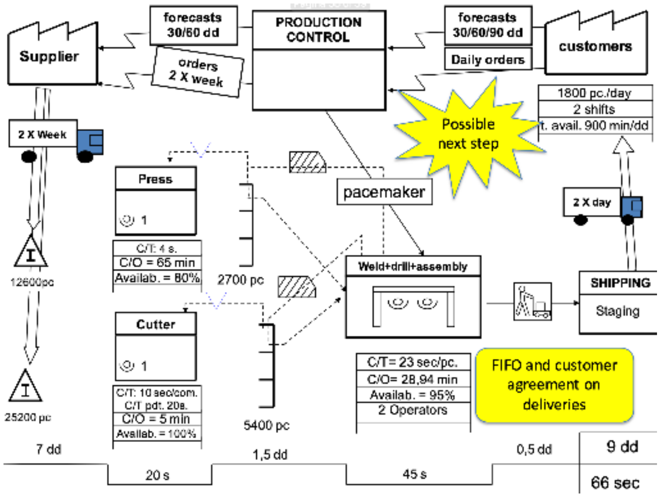


Results comparison

	Current State	Future State	Variation
NVA time	45,78 dd.	10,5 dd. (supermarket)	-77%
#operators	5	4	- 20%



At the end, the company should not stop at this point and continue in improving the operation. The next step of improvement is one piece flow or the FIFO line.



19.3. EXERCISE 3

A company assembles tractors and receives from different suppliers' kits and components needed for assembling finished products. Currently, because of the low production volumes and the enormous range variety (it is assumed that each tractor is different from the other), the production configuration is a fixed position assembly. This strategy takes advantage of the high degree of competence of the operators, who are able to perform all the steps required for assembling a tractor. In the plant, there are 20 operators dedicated to assemble the tractors, two warehousemen in charge of outbound and inbound material handling and one person dedicated to the production planning. The 20 operators are distributed in 10 cells that work in parallel. In each cell two operators work at same time on the same tractor unit. There are two families of tractors in the company, which require different components and tools (mainly due to the technology of the equipment and the size of the components). The total processing time (work content) is 160 minutes for tractors of family 1 and 140 minutes for tractors of family 2.

Each work cell can assembly any version of the product. Every day, basing on the previous day requests, the production planning defines the orders release for the different production cells, ensuring a quite uniform workload for each cell. For this reason, it releases workload to each cell with a mix of 50% family-1 products and 50% of family-2 products. The time available for the daily production is 7.5 hours. The daily demand is constant and equal to 50 units, equally split among the two product families. Moreover, the demand for both families of product is constant in volume. Each cell makes a setup every time it starts new unit production: set-up time is 60 minutes if production moves from family-1 tractor to family-2 tractor and vice-versa; set-up from family-1 tractor to another family-1 tractor is 10 minutes and setup from one family-2 tractor to another family-2 tractor it requires 15 minutes. Both operators in each cell make the setup activities in parallel.

* 20 – 10 cells working in parallel

Overall Demand	50 tractors: - 25 tractors F1 - 25 tractors F2
Work content	- 160min/F1 - 140min/F2
C/O	- F1 to F1 = 10 min - F1 to F2 (or F2 to F1) = 60 min - F2 to F2 = 15 min

Q1. In the current situation, can the company be sure of reaching the requested daily production without overtime?

Q2. During the last production meeting, plant director reported that there is tension and dissatisfaction within operators in the production plant. Indeed, assemblers are complaining that different stations are not well- balanced in their work content. Moreover, logistic operators are having difficulties in handling and in properly feeding (there are issues of mission components) materials to different station. Therefore, the management set the goal to improve efficiency. How will you reconfigure the production?

Q1: In the current situation, can the company be sure of reaching the requested daily production without overtime?

For each cell: **Tp + Ts ≤ Ta**

- Ta = 7,5 h * 60 min/h * 2 operators = **900 min/day**
- Demand for each cell = Overall demand/#cells = 50/10 = **5 tractors per day**
- Tp = worst case = 3F1 + 2F2 = **760 min**
- Ts = 2F1 + F1F2 + 2F2 = 2*10min + 60min+ 2*15min = **110 min**

870 min/day ≤ **900** min/day
Company does not need overtime

An alternative method consists in using the **Work Pace**

For each operator: **Tp + Ts ≤ Ta**

- C/T if workload is perfectly balanced between operators → C/T = WKC / #operators
- C/O if workload is perfectly balanced between operators → C/O = Ts/#operators

380 min/day + 55min/day ≤450 min/day
Company doesn't need overtime

Q2: During the last production meeting, plant director reported that there is tension and dissatisfaction within production plant. Indeed, assemblers are complaining that different stations are not well-balanced in their work content. Moreover, logistic operators are having difficulties in handling and in properly feeding (there are issues of mission components) materials to different station. Therefore, the management set the goal to improve efficiency. How will you reconfigure the production?

We have to improve the performances of each stages → It's as we are focusing on step 3 of future state creation. I can create/dedicate lines for the two product families.

Pros:

- I decrease the C/O because I no longer have to change from one family to the other.
- I can exploit specialization: increased efficiency and reduced costs.

Cons

- The change in mix becomes very critical because it could create unbalance. You must also take into account demand variability.

In this case we can separate the two lines because the demand is constant.

Takt time

- TT (F1) = 450 minutes / 25 units = 18 min/u
- TT (F2) = 450 minutes / 25 units = 18 min/u

Produce for supermarkets or for shipping?

First step: Verify the characteristics of the product and the market

- Cumbersome product
- Product of infinite range
- Product to order
- Constant volume of demand

Here we want to produce for **shipping** (**MTO**)

Therefore: 25 F1 per day + 25 F2 per day → each line has a target of 25 setups per day

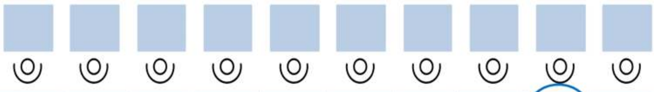
Creation of a continuous flow

- How to create the continuous flow?
- How **many resources** are needed and how to allocate them to the various product families?
- What actions are needed? To what extent will these **interventions** be?

<u>FAMILY 1– FLOW CREATION</u>

- TT (F1)= 18 min/u
- WKC= 160 minutes (to be divided between the different stations)
- CO= 10 minutes workload for each production change (to be divided between the different stations)

Since the setup time is not null, <u>it is not possible to allocate a workload equal to the takt time</u>



Allocated time	**18**	**18**	**18**	**18**	**18**	**18**	**18**	**18**	**16**	
Cumulated	**18**	**36**	**54**	**72**	**90**	**108**	**126**	**144**	**160**	

1 Resource is no longer required → increase in Productivity

The last station results to be less loaded. In case of further improvements, it may be possible to move the resource to other value-added activities. Unfortunately, until then we can only re-balance the load to save more time for setup.

Needed interventions:

- #operators = 9
- WKC: 160min/u
- C/T= 160/9 = 17,78 minutes/u
- C/O=10minutes(setup work content) –to see if compatible with given objectives and constraint
- Daily demand: 25 tractors F1
- 25 setups per day

DeCAF condition

Dedicated
Capable
Available
Flexible

- **Dedicated**: yes
- **Capable**: CT<TT
  17,78 min/u < 18 min/u
  yes
- **Available**: CT/A < TT
  17,78 min/u / 100% < 18 min/u
  yes
- **Flexible**

Target is to carry out **25 setups per day**.

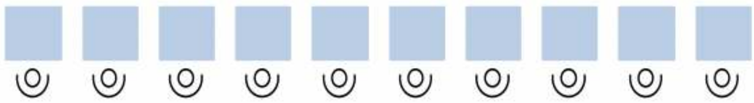
(17,78min/u)∗25 u/day+$\boldsymbol{x}$∗25 setups ≤ 450 min/day
Where $\boldsymbol{x}$ is the setup time
$\boldsymbol{x}$ ≤ 0,22 min/setups
Reducing the setup time by SMED

It means that 0.22 min is the maximum time of C/O that a station may have. If we bring the setup time to less than 0.22 minutes we can work in a flow with 9 resources on family 1 line. It means that if the setup load is perfectly balanced on operators (10min/9 op. = 1.1 min/op) from 1,1 min should reach 0.22 min.

FAMILY 2– FLOW CREATION

- TT (F2)= 18 min/u
- WKC= 140 minutes (to be divided between the different stations)
- CO= 15 minutes workload for each production change (to be divided between the different stations)

Since the setup time is not null, it is not possible to allocate a workload equal to the takt time



Allocate time	18	18	18	18	18	18	18	14		
Cumulated	18	36	54	72	90	108	126	140		

2 Resources are no longer required → Increase in Productivity

Needed interventions:

- #operators=8
- WKC:140min/u
- C/T= 140/8 = 17,5 minutes/u
- C/O=15minutes(setup work content) – to see if compatible with given objectives and constraints
- Daily demand: 25 tractors F2
- 25 setups per day

DeCAF condition

- **Dedicated**: yes
- **Capable**: CT<TT
  17,5 min/u < 18 min/u
  yes
- **Available**: CT/A < TT
  17,5 min/u / 100% < 18 min/u
  yes
- **Flexible**
  Target is to carry out 25 setups per day.
  $(17{,}5 min/u) * 25\ u/day + \boldsymbol{x} * 25\ setups \leq 450\ min/day$
  Where $\boldsymbol{x}$ is the setup time
  $\boldsymbol{x} \leq 0{,}5$ min/setups
  Reducing the setup time by SMED.

  It means that 0.5 min is the maximum time of C/O that a station may have
  If we take the setup time to less than 0.5 min, we can work in a flow with 8 resources in family 2 assembly line.

In case we were able to bring the change-over times of the two lines below the defined maximum time:

- Material handling simplifies (shift to one point for each line) ...
  Maybe one handler can also be enough.
- The space required in the factory would be reduced (less space for buffers).
- The planning and control of production and coordination activities will drastically simplify, as well as the work of operators (free time to resources, increase labor productivity).
- Productivity improves.

Where is the company single scheduling point?
The company schedules only one station per line (first stage of transformation).

19.4. EXERCISE 4:

A company manufactures mechanical products. The company works two different product types: standard product, managed with make-to-stock policy, and special product, managed with make-to-order policy. The company produces standard units based on demand forecasts. Differently, since special units are highly customized products and every unit is different from another one, the production is based on customers' orders. In the actual situation, stages are decoupled by inventory stocks. The company works two shifts every day, except in the fourth department where it works three shifts per day. Each shift is 8 hours. The average demand for standard products is equal to 85 units per day. There are four types of standard products. The average demand of special products is equal to 15 units per day and each special unit is different from the others. The range of special products comes from a base of 5 raw material components, which are managed based on forecast in stage 1. The full customization of a special product is therefore made in downstream stages (stages 2, 3 and 4). Stage 1 works also other 4 raw material components (different than those for special products): one component for each type of standard product. In production stages 2, 3 and 4 it is necessary to make 1 setup every time it starts a new special unit production. Every time company changes product type production a setup is necessary. The production flow passes through four stages, each led by one operator. In each stage there is only one machine. The data for these stages are given below.

	Stage 1	Stage 2	Stage 3	Stage 4
Standard CT (min/unit)	5	7	7	12
Special CT (min/unit)	7	8	11	10
Setup time (min/setup)	25	5	10	5
Availability	90%	85%	95%	90%
Dedicated	75%	100%	100%	100%
Number of shifts	2	2	2	3

Q1. In the initial situation described in the text, calculate the EPE for standard parts for each stage and the minimum batching for standard units so that the company produces every day 15 special units in unitary batches on stages 2, 3 and 4.

Q2. The company is willing to implement Lean techniques in order to improve the production process, keeping the same approach of make to stock for standard products and make to order to special ones. Support the company in drawing its future state, highlighting information and material flows. In sizing the improvements, you must ensure the company the ability to both deliver 15 special parts every day in unitary batch and have an EPE for standard products at maximum to 4 days in each stage.

Q3. How long does special order take to be delivered to the customer?

Initial configuration

	Stage 1	Stage 2	Stage 3	Stage 4
Standard parts cycle time (min)	5	7	7	12
Special parts cycle time (min)	7	8	11	10
Setup time (min)	25	5	10	5
Availability	90%	85%	95%	90%
Dedicated	75%	100%	100%	100%
Number of shifts	2	2	2	3

Ta= 8 hours net per shift

Standard products	Special products
D= 85u/day	D= 15 u/day
Variants= 4	Variants=each one is different

Production stages decoupled by stocks

Q1: In the initial situation described in the text, calculate the EPE for standard parts for each stage and the minimum batching for standard units so that the company can produce every day 15 special units in unitary batches on stages 2, 3 and 4.

EPE (every part ever) = time required to sort the whole product range

$$EPE*Tp + Ts \leq EPE*Ta$$

$$EPE \geq \frac{Ts}{Ta - Tp}$$

Stage 1

On the first stage also the 5 standard components, necessary for the production of special products, can be produced in batches (the 5 variants are produced based on forecast).

Stage 1 EPE needs to take into account:
- Time required to produce **85 standard units** per day in **4 variants**
- Time required to produce **15 special products** units per day in **5 variants**
- **Setup times** for standard and special products

Ta= 8 h * 2 shifts * 60 min/shift * 75% = **720 minutes**

$$EPE \geq \frac{Ts}{Ta - Tp} = \frac{25\frac{min}{setup}*(4+5)}{720\ min - \frac{(7*15+5*85)}{90\%}} = \mathbf{1{,}716}$$

Stage 2, 3, 4

- Whenever you change a special product version, a setup is needed.
- Every day for each department must be produced on average 15 special products.
- Special products are produced in unitary batches.
- The time available for the production of standard units is equal to the available time for the stage decreased by the time necessary for the production of special units
  Ta (STD) = Ta – Tp (SPE) - Tsu (SPE)

$$EPE_{STD} \geq \frac{Tsu_{STD}}{(T_a - Tp_{SPE} - Tsu_{SPE}) - Tp_{STD}}$$

	Stage 2	Stage 3	Stage 4
Available time (shift time*n of shift*%dedicated)	960	960	1440
Special processing time [(C/T spec / A)*D. spec]	141.18	173.68	166.67
Setup time (15 special)	75	150	75
Available time standard products	744	636	1198
EPE (STD) - days	0.45	4	0.31

Minimum batch size for standard products

$$MBS = \frac{D * EPE}{\#variants}$$

	Stage 1	Stage 2	Stage 3	Stage 4
EPE	1.72	0.45	4	0.31
Average batch size of a standard products	**36.55**	**9.6**	**85**	**6.6**
Average batch size of a special products	5.16	1	1	1

Q2: The company is willing to implement Lean techniques in order to improve the production process, keeping the same approach of make to stock for standard products and make to order to special ones. Support the company in drawing its future state, highlighting information and material flows. In sizing the improvements, you must ensure the company the ability to both deliver 15 special parts every day in unitary batch and have an EPE for standard products at maximum to 4 days in each stage.

FROM PRESENT STATE TO FUTURE STATEV→ 8 Questions

1. What is the takt time of the production family?
- TT (Stage 2-3) = 960 min/ 100 units = 9,60 min/u
- TT (Stage 4) = 1440 minutes / 100 units = 14,40 min/u
- TT (Stage 1) = 960 min *Dedication/100 = 7,2 min/u

2. Produce for supermarkets or for shipping?

First step: Verify the characteristics of the product and the market
- **Standard units**: few variants (4), very short delivery times, size is not critical, costs, perishability, etc.. They are produced for supermarket.
- **Special units**: many variations, greater time allowed. The aim is to produce for shipping (in the text it says they are already produced for shipping)

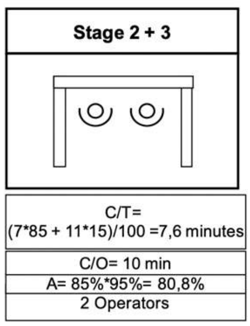
3. Where to put the flow?

General methodology
- Start from the **final stage and go upstream** thinking stage by stage where to put CONTINUOUS FLOW and where to decouple (with SUPERMARKET or FIFO).
- Verify **DECAF Conditions**.
- Fix intermediate targets (not necessarily all at once in a continuous flow, but also FIFO and supermarket).

Start from department 4 then move upstream

The department 4 works on 3 shifts, while dep. 3 works on 2 shifts, so it's not possible to introduce the continuous flow. It should be decided between FIFO and SUPERMARKET.

Stage 2 + 3



DeCAF condition
- Dedicated
- Capable
- Available
- Flexible

- **Dedicated:** yes
- **Capable:** CT<TT
  7,6 min/u < 9,6 min/u
  yes
- **Available**: CT/A < TT
  7,6 min/0,808 < 9,6 min/u
  yes

- **Flexible**

It's requested that the company can deliver 15 special pieces every day in a batch of one and have an EPE for standard products of maximum 4 days

Which is the cell EPE?

$$EPE\ STD \geq \frac{Ts}{Ta\ STD - Tp\ STD} = \frac{10\ min/setup * 4}{\left(960min - 10*15 - \frac{11*15}{0,808}\right) - \frac{7*85}{0,808}}$$

This way you can not even meet customer demand.

→
- **Flexible**

EPE^{Target} = **4 days**

EPE*Tp + Ts ≤ EPE*Ta

4* (7/80,8%)*85 +*x**4 ≤(960 -*x**15-(11/80,8%*15))*4

Where *x* is the average setup time for standard and special

***x* ≤ 1,18 min/setups**

If you can bring the setup time to 1.18 min. then you can couple the two stages (2 + 3) in a continuous flow.

Stage 1 is not dedicated to the product family, so it can not be put into a one piece flow.

4. Where to put the pull-supermarket?
STD products
For STD products there's a finished good supermarket. Moreover, between dep. 4 and 3, being a supermarket downstream of 4, it is necessary put a supermarket. Finally, there is a supermarket upstream of department 2. The pull supermarket will be inserted upstream of Department 2.

?

Supermarket sizing = 1,5 *EPE* Ddd
<u>Main rule: EPE to consider when sizing the supermarket is the one of the stage upstream of the same supermarket</u>

SPE products
- Stage 2 → Stage 3 OPF (continuous flow in the cell)
- Stage 3 → Stage 4 FIFO

5. Where is the company single scheduling point? STD products
The only scheduling point is Stage 4, upstream the finished goods supermarket.
- SPE production
  The only scheduling point is the cell that includes stages 2-3.

6. How should the company level the product mix to pacemaker process?
- STD products
  According to shipments that are made, the finished goods warehouse sends upstream the Kanban, and these are smoothed (with Heijunka) based on volume and mix.
- SPE products
  You take leveled (volume and mix) from the Pre Shop Pool.

Q3: How long does special order take to be delivered to the customer?
Delivery time = PSP Time + Processing time (cell) + FIFO Time + Processing Time (stage 4)

PSP Time → How to size it?
Total Processing Time = 8+ 11 + 10 min = 29 min

How to size FIFO?
- FIFO Size = Δshifts*D*1,5 = 1*15*1,5 = 22,5 pc → more conservative
  You could also consider the demand in one shiftà15/3=5 pc
- FIFO Time = 22,5 pc/15pc/d =1,5 days

How to dimension PSP?
Assuming a significant variability in the volume of SPE demand (+/- 60% d/d)
Variability is absorbed with PSP (pre-shop-pool) → 3σ
- **PSP size** = Average demand + (3*% Variability * Demand) = 15+ 3*60%*15 =42pc
- **PSP Average Time** = 15pc/15pc/d = 1 day
- **PSP Maximum Time** = 42pc/15pc/d = 2,8 days

Considering this average level, it is good to give the customer a time indication that **considers 3 days as throughput time of the PSP** (1 on average, with an oscillation between 0 and 3 days)
[for the exam only → PSP throughput time for the customer= PSP Maximum Time]

- **Average Delivery time** = PSP Average Time + Total Processing time + FIFO Time = = about 1 day + 29 minutes (negligible) + 1,5 days= **c.a. 2,5 days**
- **Maximum Estimated Delivery time** = PSP Maximum Time + Total Processing time + FIFO Time = = about 2,8 day + 29 minutes (negligible) + 1,5 days = **c.a. 4,3 days**

20. LEAN TOOLS

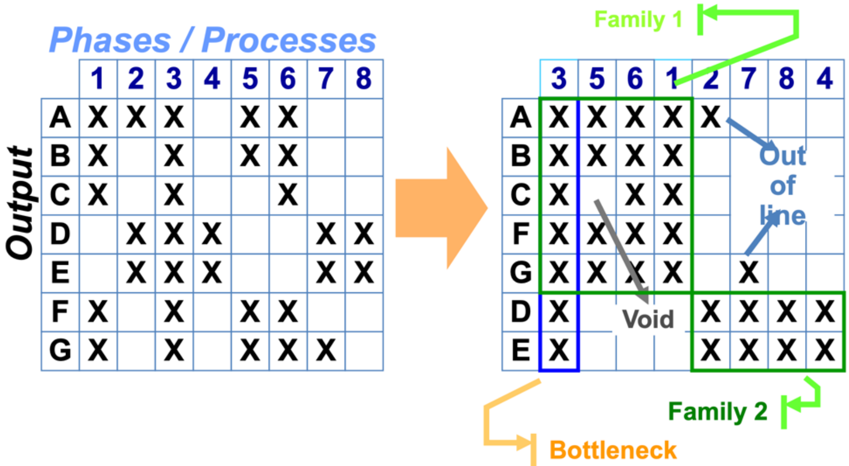
Value stream map is the mapping of the flow of a family of product, the definition of the family is according to the characteristic of the product. In an operation system if we use the lean approach we take the perspective of the flow characteristic, the product family must be defined by the flow characteristic. It is the pattern inside the operation system. the same pattern means that you have your item that passes on the same resources of another product. The value stream map is not enough to implement all the improvements, for this reason we can use other tools to increase our performance.

List of different tools:

Production flow analysis

Before doing the product flow analysis we really need to understand this consideration:

1. We are going to do the single product analysis: analysis of a single product. If the volume is high enough to dedicate resources to it.
2. Production flow analysis: Analysis done by identifying the product/service family. If variety is wide and single product volume is low, it is needed to group different products together to reach enough volume to dedicate resources.



In order to do the product flow analysis, we need to create the matrix of the product and processes; by doing this, we can identify the families of the product.

How to identify a family?

We need to understand which are the products that are sharing the same machines and cluster them together.

In doing so, we have a better mapping of all the families of the products that are sharing the same machines, for example all the products are sharing the 3rd process, product A, B F and G are sharing the 5th process and so on.

By doing this we can decouple the system and simplify the problem.

After clustering the products in the same families, we then need to do the rank order clustering.

20.1. ROC

ROC: The rank order clustering allows us to rank the different rows and the different columns according to their family; the procedure is done by using specific criteria.

In the table we need to fill in the boxes by putting the numbers 1 or 0.

1: This product has to go on this machine

0: This product is not going on this machine

We need to consider the 0 and 1 as a binary digit number and we transform it into a decimal number, for example:

Product B: 2^0+2^4= 16+1=17

OUTPUT	PROCESSES								value
	1	2	3	4	5	6	7	8	
A	1	1	0	0	1	0	0	0	200
B	0	0	0	1	0	0	0	1	17
C	0	1	1	0	0	1	1	0	102
D	0	0	0	1	0	0	0	1	17
E	0	0	1	1	0	1	1	0	54
F	1	1	0	0	1	0	0	0	200
value									

Once we transformed them we need to rank the rows from the highest one to the lowest:

OUTPUT	PROCESSES								value
	1	2	3	4	5	6	7	8	
A	1	1	0	0	1	0	0	0	200
F	1	1	0	0	1	0	0	0	200
C	0	1	1	0	0	1	1	0	102
E	0	0	1	1	0	1	1	0	54
B	0	0	0	1	0	0	0	1	17
D	0	0	0	1	0	0	0	1	17
value	48	56	12	7	48	12	12	3	

We need to reshuffle the lines according to the value that we have here. Once we have done this we do the same approach but now we go on the columns, so we need to consider the binary number of the columns and we need to calculate them. Also, in this case we need to shuffle the columns from the highest one to the lowest one, the final solution will be:

OUTPUT	PROCESSES								value
	2	1	5	3	6	7	4	8	
A	1	1	1	0	0	0	0	0	224
F	1	1	1	0	0	0	0	0	224
C	1	0	0	1	1	1	0	0	156
E	0	0	0	1	1	1	1	0	30
B	0	0	0	0	0	0	1	1	3
D	0	0	0	0	0	0	1	1	3
value	56	48	48	12	12	12	7	3	

We need to do the same from the rows again, by doing it we discover that they are already ranked properly and we can consider the end of the solution.
We have one cluster made by product A and F, the second cluster is C and E and the third cluster is B and D.
Thanks to this methodology we are moving the lines till we get to a clustering: The result is that we have 3 families.

20.2. 5S

The second approach we could use is by removing every cause that is stopping the flow. It is an iterative methodologist that works on the shop floor, we have 5 steps. You adopt 5S in the working station.
It is lead to the concept "quick mean solution: low effort and great result": it is an approach that means that we do not have to plan everything in one shot but just start to do one implementation and some approaches in a way that we are able to deliver what we have to. It is an approach that follows the continuous improvement done step by step.

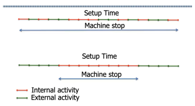
1. Sorting → it means selecting only the necessary tools and materials that are necessary for your working activity. Generally, nobody wants to throw stuff away. Keep in the station only the stuff you are using. In this way you avoid congestions.
2. Set in order → give a place to items that you want to have in this station. Select the item that you want to keep in the station and find a place according to the frequency of use. Golder area of a working station is the area in which you arrive with your hands. Put there only the thing you use every day and so on with items that you don't use so frequently.

3. Shine → keep the things clean. Not only for hygiene, but if you have a clean area it is easy to identify anomalism.
4. Standardize → when you create a certain order, create a standard that you can replicate in every station. It is important because:
   - It is easier to control
   - You have an environment that should be more stable
   - You are creating flexibility for your system because you can move people to one to station to another one so economies of learning are much faster.
5. Sustain → repeat the 5S every year. The first time you will spend a lot of time, then sustain the process.
   - Firstly, do 5S once a year because the company is changing but also the market.
   - Do 5S every week/month

Results: avoid the congestion because the area is not congested and it is easy to move and improve productivity because you spent less time for looking for tools. It is cheap so you can have benefit with no investment.

Quality: scraps are causes of stops and should be eliminated because they are only stopping the flow, since producing scraps means having a missing production and we are stopping the station. In order to eliminate the scraps, we need to change our approach: better quality is not necessary a cost, in the long turn it decreases the overall cost, most of all, if we implement the 5S and we standardized the process. A big part of the quality cost is not that "I'm losing material", but "I'm missing an item because I move all the supply chain", the activity that you did more time because you have a defect. It generates a lot of variability inside of the system and, as we know, it is the nightmare of operation.

SMED (single minute exchange die): another aspect that is causing a problem within the processes is the set up time. The set up time is usually stopping the flows because it includes all those moments in which the machine is not properly working and instead we are repairing it. When we do a setup there are internal and external activities:



- Internal: activity done directly on the machine for which we need to stop it.
- External: activity that can be done while the machine is working.

The first step is try to separate internal with external activities. In this phase, internal and external operations are often mixed. Move all the external activities at the beginning or at the end of the changeover, so you concentrate the internal activities in the middle of the sequence, so the machine can stop only while I'm doing internal activities. While the machine is running you should prepare the equipment before the machine is stopping, you take the equipment close to the machine and then you stop it. This is the main step in terms of saving time.
The second step is try to shorten both of the activities.

TPM: total productive maintenance.
The usable capacity is net of breakdowns, failures, and other stoppages for maintenance. Compared to the potential rate, the system can actually produce much less.
Actual capacity= Theoretical Capacity * Availability

$$A = \frac{MTBF}{MTBF+MTTR}$$

MTBF: Meantime between failure
MTTR: Meantime to repair

Failures occur at unpredictable times and, according to Murphy's Law, in the worst moments.
As it is shown in the formula, availability is the mean time between failure, in order to improve availability, we can have two approaches:

1. *Increase the meantime between failure*: We could increase the meantime between failure by giving higher reliability to the component of our machine.
   - **Meantime between failure**: When one of the system stops working
   - **Reliability:** Probability of every single component to work till it gets broken
   - **Availability:** Reliability of the complex system made by system component; it depends on the reliability of every single component of the system.

   In order to increase the meantime, we need to increase the reliability of our components. To improve the availability we need to improve the mean time between failure by improving the reliability of every single component.
2. *Reduce the mean time to repair*: If the meantime to repair is decreasing then the dominator is decreasing to, which means that the availability is increasing. The meantime to repair measures the time to repair the system. By reducing the meantime to repair we minimize the time in which the machine is not working, by increasing the time in which the machine works.

   In this way we increase the availability by squeezing the time to repair the system.
   Of course, if we increase our availability we increase the total actual capacity.
   During the years the system used in order to decrease the maintenance time has been improved, from a failure base maintenance to a predictive maintenance we have increased our availability; we went from:
   - **Failure based maintenance**: we do the maintenance once we have failure, if there is a failure we intervene by changing the components → reactive maintenance
   - VARIABILITY **Scheduled maintenance:** we estimate the failure time and before having it we prevent it by changing the component → preventive maintenance
   - **Predictive maintenance**: we do the maintenance of the system by a prediction of the failure by sensors used to analyze the vibration; if the vibration is getting to a specific value we need to intervene before we get it.

Invest now a short amount of time in order to avoid bigger downtime in the future. It is much better to have a downtime that is planned that one that is unforeseen.
Of course, the predictive maintenance is more effective because the scheduled maintenance by changing the component even though they are not going to fail. You don't know when a failure will come and according to the Murphy's law it comes in the worst moments. While with the predictive one, we analyze the specific components one by one and we detect the failure.

Levelling: The batch can be processed in two different ways:
Batch processing
AAAAA-setup-BBBBB-setup-CCCCC -setup-DDDDD
Mixed model processing
ABCADCDBACDABACADBACDADCAABDABB

Traditional companies have batches. Lean companies try to make the batch as small as possible and try to mix the mix as much as possible. Why? In this way they level the mix and they support the flow creation.
Market demand is often more leveled than you believe (80% of mix and volume variations are determined by the company itself). Companies, in order to respond to the demand, are usually determining the volume by aggregating the demand and setting a setup time; unfortunately, this approach is not really following the market demand, plus, the aggregation of the demand and the setup time are creating discontinuity.

Product	1	2	3	4	
Weekly demand (units)	75	50	50	25	Average service factor
Part 1- service factor	2	-	-	1	0.875
Part 2- service factor	1	-	3	-	1.125
Part 3- service factor	-	1	-	1	0.375
Part 4- service factor	3	2	2	2	2.373
Part 5- service factor	2	4	1	-	2.0
Part 6- service factor	-	1	-	2	0.5
Part 7- service factor	-	-	1	2	0.5
Part 8- service factor	-	2	-	1	0.625
Part 9- service factor	-	-	3	-	0.750
	8	10	10	9	

In fact, the more the company aggregates the demand, the more leveled it looks till it gets to a point that is non-reversible → the solution would be to eliminate the set up time.
In the picture, each product is made by a number of components, for example product 1 is made of 2 units of part 1, 1 unit of part 2, 0 units of part 3 and so on till 8. The *average service factor* is the weighted average of the number of parts that we need in the weekly total demand of the units.
The average service factor is calculated by the total number of units of part 1 that we need in a week divided by the total amount of unit of the weekly demand:

Total number of units of part 1: 2*75 (weekly demand of the product)+0*50+1*25 = 170

$$\text{Average service factor} = \frac{170}{75+50+50+25} = 0.875$$

We need 170 units of part 1 in order to fulfill the production. If we divide this by the total number of products that we need to produce, we find the 0.875, which is, on average, every unit that we need to produce during the week.
We now need to level the production:

Product	Mon	Tue	Wed	Thu	Fri	W1
1	40	35	0	0	0	75
2	0	5	40	5	0	50
3	0	0	0	35	15	50
4	0	0	0	0	25	25
Part			**Supplier 1**			
1	80	70	0	0	25	
9	0	0	0	105	45	
4	120	115	80	80	80	
5	80	90	160	55	15	
8	0	10	80	10	25	
Total	280	285	320	250	165	
Part			**Supplier 2**			
2	40	35	0	105	45	
3	0	5	40	5	25	
6	0	5	40	5	50	
7	0	0	0	35	65	
Total	40	45	80	150	185	

The picture is showing the total number of units of all the parts that we need to have each day in order to fulfill the weekly demand. As it is shown, the easiest schedule would be to start on Monday morning and produce product 1 →since we have to produce 200 units in a week and we have 5 days, the average capacity would be 40 units in a day.
To start with product one: Capacity on Monday is fulfilled (40), the remaining part is produced the next day (35).
But, by producing 35 of the product 1 the daily capacity is not saturated, so I put 5 units of product 2.
We continue then with the production of product 2, I put 40 units of Wednesday, the other 5 are going on Thursday and we produce 35 of product 3 and the 15 remaining are produced on Friday. Once the production of product 3 is over I produce the product 4. The overall weekly production is fulfilled and the sum of all the capacities gives us the possibility to understand if we saturated completely the demand. This would be finished, but there is a problem, which is the fact that we need the components in order to produce them.
Let's suppose the components are supplied by 2 different suppliers. If we use this matrix, on Monday, what is the requirements of material that we have?
Schedule that the first supplier has to take into consideration for the week: looking at the suppliers' table, we know that in order to make product 1 (for example) we need 2 units of part 1, which means that in order to fulfill the material requirement planning we need 80 units to fulfill the production of product 1. The same on Tuesday because product 1 is produced also on Tuesday. Part 1 is also needed when we produce product 4, this means that we need 25 units on Friday as well in order to support the production of product 4.

Product	Mon	Tue	Wed	Thu	Fri	W2 Morn.	W2 Aftern.	W3 Morn.	W3 Aftern.
1	15	15	15	15	15	15	0	8	7
2	10	10	10	10	10	5	5	5	5
3	10	10	10	10	10	0	10	5	5
4	5	5	5	5	5	0	5	2	3
Part					**Supplier 1**				
1	35	35	35	35	35	30	5	18	17
9	30	30	30	30	30	0	30	15	15
4	95	95	95	95	95	55	40	48	47
5	80	80	80	80	80	50	30	41	39
8	25	25	25	25	25	10	15	12	13
Total	265	265	265	265	265	145	120	134	131
Part					**Supplier 2**				
2	45	45	45	45	45	15	30	23	22
3	15	15	15	15	15	5	10	7	8
6	20	20	20	20	20	5	15	9	11
7	20	20	20	20	20	0	20	9	11
Total	100	100	100	100	100	25	75	48	52

- The system must see the variety that the market demands as it were a single product

high variability => high inventory

Let's make another example:
Part 9 is only produced for product 3. Product 3 is produced on Thursday and on Friday. 35 multiplied by 3 is 105 and 15 *3 is 45 → Monday, Tuesday and Wednesday product 9 is not needed, on Thursday and Friday it is needed with the quantity we just said. Since on Monday we need other components, the schedule we did on the product is amplifying on supplier 1 and 2 the discontinuity, it creates, in

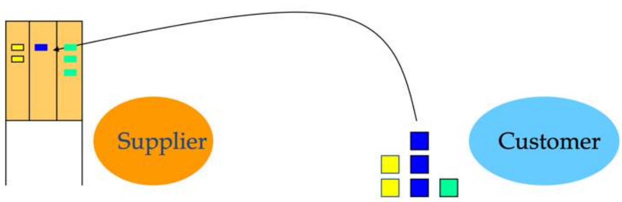
problem of not this: mindset / big : effort is high and advantage is supplier > lower cost and less inventory / over all system improves (reason of doing it) so more competitive advantage

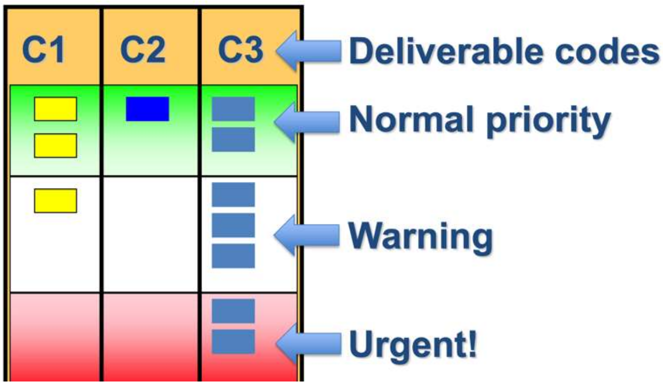
fact, a huge amplification for our suppliers → there is a discontinuity for supplier 1 and most of all for supplier 2, which on Monday and Tuesday it doesn't need to distribute so many products for us but on Thursday and Friday it has to distribute more than the double compared to the day before.
How to overcome the problem? We can use the mixed model instead of the batch model.

Instead of batching the production we have used a mixed model: for 5 days we produce the same sequence of different products. As we can see all the days have the same production plan and the same saturation of the capacity fulfilled by the production of all the products every day.
After doing this, we need to split the days in two parts: we batch the production in the morning and in the afternoon, we have 20 units as capacity in the morning and 20 in the afternoon. The sequence would be 15 and 5 in the morning and the remaining production (5+10+5) is done in the afternoon.
What is, now, the result for our supplier?
In the morning we produce 15 of product 1 which requires part 1, in the afternoon, part 1 will be instead used for product 4 → In this sense, the total production for supplier 1 in the morning would be 145 and 120 and 25 and 75 for supplier 2, as we can see, the difference is still quite high.
In order to solve this problem, we need to level the production between the morning and the afternoon. If 15 is the overall production we can do 8 in the morning and 7 in the afternoon for product 1 and 5 in the morning and 5 in the afternoon for product 2 and 3 and so on for the other products. In doing so, our supplier will have a better balance between the morning and the afternoon, in fact, the units of products will just differ of 3 parts.
In leveling our production process, implementing a mixed model processing and level the demand in the morning and in the afternoon we reached a better stability for our company and for our suppliers.

Levelling means levelling the mix, creating a production that is as much stable as possible. What happens if we distribute our mix instead of in 1 week we distribute it in 1 day? We have a capacity of 40 u/day, we distribute it according to the level mix. Or in this case the mix is: 15, 10, 10, 5. And we produce every day the same mix. And so, we do 4 setups per day. This choice impact us, because the consume of the components comes to be different. The consume of our component is every day the same. Now it is linear and variability is deleted. What is the benefit? The levelling support the creation of the flow upstream the supply chain. It is better also for the internal supply chain and for the partners that belongs to the same companies.

Kanban: in order to allow the company not to over-produce products we must link as much as possible the orders with the customer.
First of all, we activate the production when there is a real need of unit (if we do the make to stock we produce in advance and the risk of overproducing is increasing); so, the only way to avoid overproduction is by the pull approach.
In fact:
- Forecast is always wrong
- To be efficient you have to work on reliable information
- Market is not willing to wait
- If we want to implement ~~pool~~ pull we need to be very reactive and the system must be interlinked; the tool in order to implement the pull approach is the Kanban:





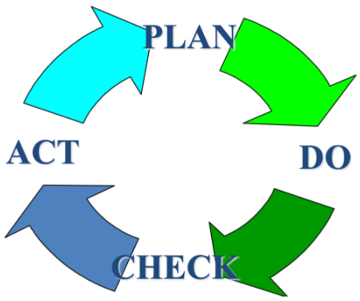
As soon as the customer is picking up a unit of the product from the supermarket/warehouse, the production Kanban has to be sent to the supplier (nowadays, the production Kanban is an electronic one).
The IT system of the customer and of the supplier are connected and the customer (the company) sends a message to the IT system of the supplier that has to produce the unit missing. It can happen that we end up in a situation in which the suppler has many Kanban and doesn't know what to produce first.
Because of this reason, we should implement what is called a Kanban board:

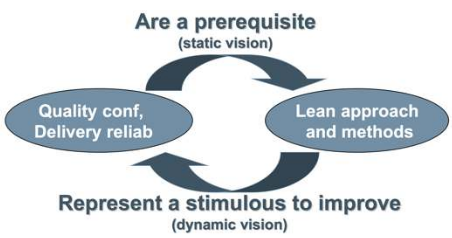
The Kanban board is the tool that the suppliers have to receive from the customer in order to understand the Kanban's priority. The Kanban board is split in 3 lines and must be organized in columns, which are representing a particular product. The green, the white and the red color are representing the priority, which is the urgency that the supplier has to take into consideration in order to fulfill the demand.

What are the benefits?

- You limit the stocks. Suppliers produce only what has been consumed. You avoid overproduction
- You have solid information. You follow market request
- It is a simple system that doesn't need so much effort. It can be managed by everyone. You don't waste resource on planning
- It linked the stages. The buffers create the link and so on the network. If they are not connected it gives problems
- The stages are linked you have that the variability that enters from the customer is exactly the one that is transmitted to the supplier. Because we don't have amplification. Because you must follow what the customer does.

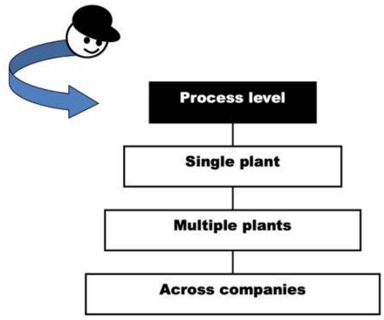
Each step is a circle of PDCA and it means that for each target that you identify necessary for make the improvement step. With <u>plan</u> we map and analyze the system. with <u>do</u> we do activities and experiments and, for example, we do 5S, then there is <u>check</u> in which we ask if we achieved the target of improvement of the previous steps, and then <u>act</u> so consolidation, we monitor the check and consolidate the position we reached.





21. CREATING CONTINUOUS FLOW

After the future state we for sure got improvements on the single plant side. Now, the question is, **are these improvements enough?** Are there any other activities in order to get more level of performance? To answer this question, we have to go to a higher level of detail, working at **process level**, looking at the performances of the machines in order to **try to reduce C/O** (if we have machines with high C/O of course we will have stocks because we are not able to have continuous flow).



Creating a continuous flow means putting in place all the possible actions to make a continuous flow, **to avoid having the need of stocks**, because **LT is the lowest possible** and if it is equal to working time, stocks are equal to zero (*Little's Law*).

21.1. HOW TO DESIGN A CELL?

EXAMPLE

- Apex makes tubular products for different applications: automotive industries, trucks S, L, A, heavy trucks and heavy equipment
- The Light Truck families are the most important ones (those that we consider): S, L and A belong to the same family of products, like heavy types.
- Current state map

Apex's Product Family Matrix

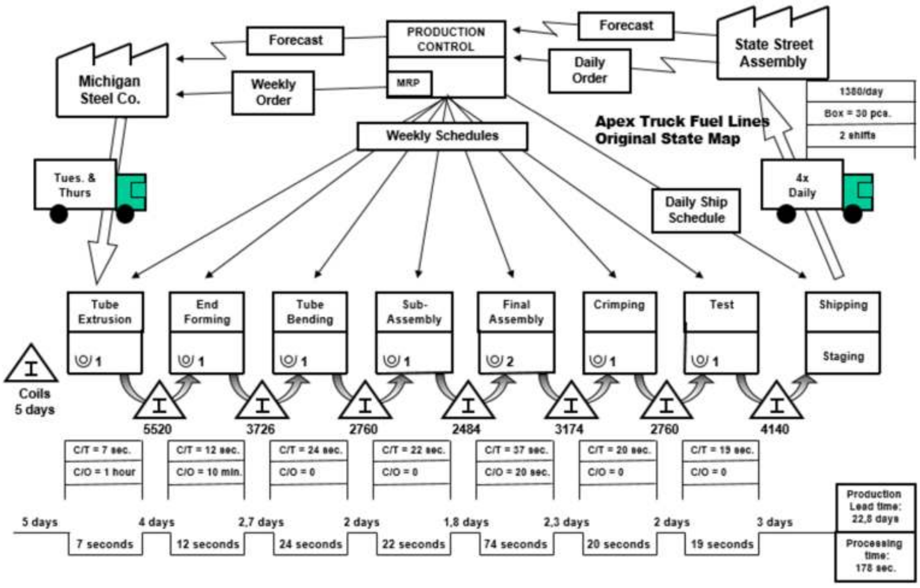
PRODUCTS	Assembly Steps and Machines							
	end form	pierce	braze	bend	sub-assembly	final assembly	crimp	test
automotive	X				X	X	X	X
truck S	X			X	X	X	X	X
truck L	X			X	X	X	X	X
truck A	X			X	X	X	X	X
heavy truck		X	X	X				X
heavy equipment	X	X	X	X		X		X

We ALWAYS start reading it from the right, the demand (what we need to produce comes from the quantity required by the market). If we do not understand the demand, we do not understand what the company is doing (the factory is only a tool used to answer the demand).
Pieces are delivered using boxes with 30 pc each. There are 2 shifts and 4 deliveries for day. The shipping unit does the shipping. Before shipping there are all the phases and processes, which can be read from left to right:

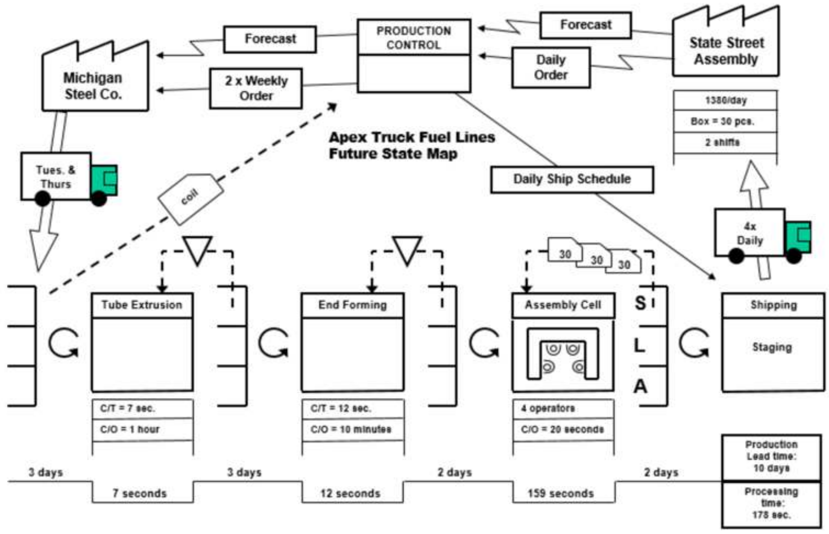
- Tube extrusion: C/O of 1 hour is right because of the different dimensions of the tubes.

Since the schedule goes to each phase according to MRP, stocks are needed.

Original state map:



Current state map:



Where to put supermarket? At the end we have 3 types of products to be shipped in boxes with 30 products each, with the Kanban equal to 30 and there are different Kanban for different products, so before shipping we need to put supermarket for definition.

In this way all the stocks in these phases (from bending, with all assembly phases till the end) are eliminated and LT is shorter. The cell is characterised by:

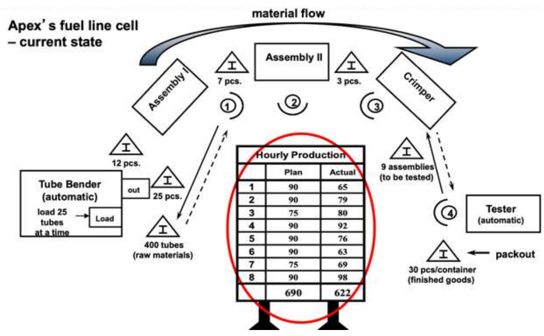
- 4 operators
- C/O 20 sec (the worst)
- LT 159 seconds (sum of values)

NB) Do not pretend to do all the improvements in one shot, because the best way to do it is continuous improvement.

This improvement is not enough

- Final results are not as good as on paper
- Redesign the value stream is good but it is only the starting point
- VSM is no solutions, only problems highlight

Tube bender is a phase with an automatic machine that needs to be manually loaded with 25 tubes at a time. One worker (worker 1) is dedicated to assembly 1 and once in a while he has to go to also load the bending machine: the flow is interrupted, because if he goes to bending machine in that moment no one is working in the assembly 1. To avoid assembly 2 to be interrupted or waiting, 7 pieces of inventories must be present after assembly 1. The same is done by worker number 4, who is dedicated to the tester machine but once in a while he has to go to the crimper to take products to be worked.



Apex's fuel line cell – current state

APEX was working as in the current state map, but they decided to do some improvements.
Firstly, they added cells. This is what an external consultant sees after the last improvement: in some cases, the actual production is lower or higher than the planned production. He has to investigate the reasons. Some of the reasons are due to the fact that the continuous flow is not there. The sources of not continuous flow are:

1. Workers have to move between different phases
2. Workers number 4 has to move to get materials to work. Having people walking is a waste

Cells need to be redesigned in order to put stations closer one to the other (1st improvement), to avoid walking and moving material around. The 2nd improvement consists in changing the way worker does the loading activity 1.

Causes: Why there is so much variation?

- Is the machine incapable (scrap/rework)?
- Is it the machine that won't run?
- Are parts missing?
- Is the supplier shipping defective parts?
- Who reacts when these problems occur?

Apex's Progress with Continuous Flow

	Original State	Current State
Continuous Flow	No	No
Production per Shift (actual/target)	unstable ≈ 622 / 690	unstable ≈ 622 / 690
Space (sq. feet)	1130	580
Assembly Lead Time (WIP x Takt)	11 days	37 min.
Number of Operators	6	4
Productivity (pieces/associate/hr)	13.05	20
Functions Effectively as Pacemaker	No	No

ACTIONS: We should analyze SIGNALS, possible indicators of variation:

- First operator leaves his/her area every 25 pieces (flow stops)
- Inventories between operations
- Production associates are anchored to their machine
- Very wide U

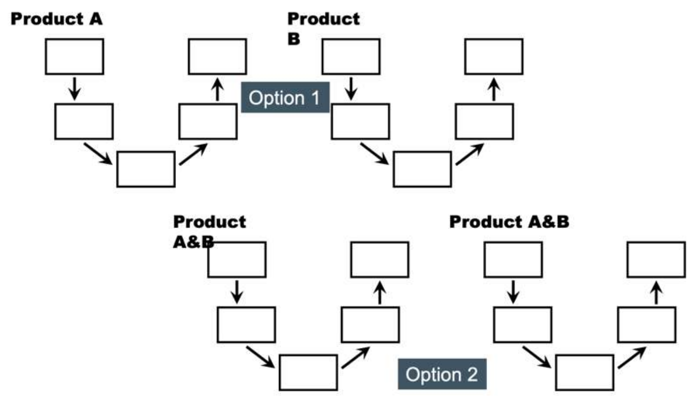
Is there any better way to organize workers' work?
We need to eliminate the differences between actual and plant values. After the first revision, with the introduction of cells, the results are:

TARGETS:

1. To produce exactly what is scheduled to be produced (stability means continuous low)
2. Less workers, less space, less costs
3. We need to identify a pacemaker!

Targets for Apex's Fuel Line Cell

	Original State	Current State	Target
Continuous Flow	No	No	Yes
Production per Shift (actual/target)	unstable ≈ 622 / 690	unstable ≈ 622 / 690	690 / 690
Space (sq. feet)	1130	580	252
Assembly Lead Time (WIP x Takt)	11 days	37 min.	200 sec.
Number of Operators	6	4	2
Productivity (pieces/associate/hr)	13.05	20	40
Functions Effectively as Pacemaker	No	No	Yes



First of all, we need to consider the 3 different products. We have 2 options:

- **Option 1**: creating ad hoc cells for each product
- **Option 2**: creating mixed model cells where the same cell is producing all the products. A and B normally are better balanced, the flow is balanced and it is easier to create continuous flow. For example, 2 products are different and maybe with C/T different but mixing them it is easier to manage them.

How to choose what items in the cell

less variability! shorter queues more saturation

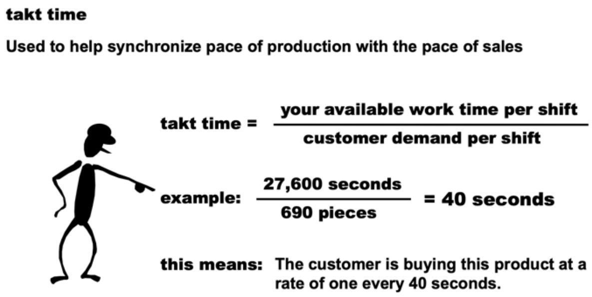
- **Flexibility**: A+B more flexible than A and B
- **Variation in total work content**: Max around 30%: it is just a rule of thumb (if variation is higher than 30% we risk creating work not balanced)
- **Similarity in processing steps and equipment**
- **Takt Time**: Optimal between 10 secs and 120 secs.
  - If it is too low, it means that workers have to repeat the same sequence of operations in a very short time (more than 6 times in a minute).
  - If it is higher than 2 minutes, the drawback is related to a slower learning curve. It could be that workers' tasks have inefficiencies due to the fact that in 2 minutes repetition is not enough and maybe the task is too complex.
- **Customer Location**: Close to the customer

customer demand

TAKT TIME: Expected production rate to deliver the product to the customer. Takt time is a goal, a target.

Available time of workers: we need to take into account breaks.

- Stoppages of machines due to failures should NOT be included in the available time, because we do not want the machine to stop, so we consider it as a non-expected event. It is not a constraint we need to consider!



Se

Because Takt Time represents the demand rate, you do not subtract time for changeover, unplanned downtime, and other internal problems

Takt Time is based on customer demand (which you cannot change) and available production time (which you can change). Your levers are:

- Number and type of items in a cell
- Available production time
- Number of cells making a specific item

Since the Takt time is an average calculation of the demand, another problem we could face is when the demand is changing frequently. This situation could create some problematic since it won't be completely visible from the average takt time. In fact, in order to calculate the Takt time, we take the monthly demand and we divide it by the number of hours, but there could be a variation of the demand within the week; for instance, it is quite normal that the demand is higher at the end of the month.

What if demand rate changes?

Takt Time refers to average demand. Changing Tack Time frequently yields inefficiency.

- In **MTS** use **Buffer Stock** to protect the production system. The buffer stock will decouple the production system, so it is perceiving a smooth uniform demand.
- In **MTO** use **Backlog** to protect the production system, which means collecting more orders and putting them in a batch or inform the customer that we deliver the product in a different time. In this case we need to pay attention, because the backlog could be not good in the long term: If the backlog is too high the customer is not so happy and they could change the supplier.

Cycling much faster than TT

- May require more people
- Increase the chances of over producing
- May conceal production problems
- Eases tension to improve

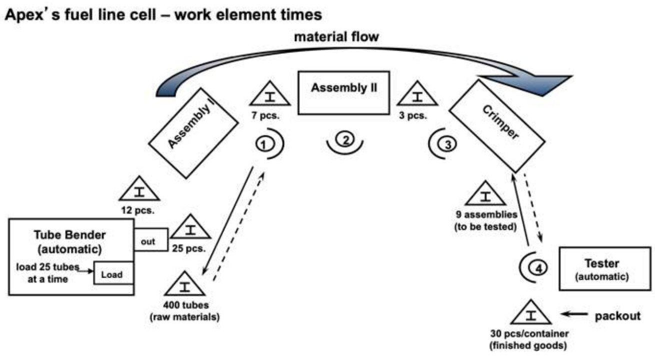
TIMING TIPS

- Collect real time at the processes (**do not rely on standard time**)

- Position yourself where you can **see the operator's hand** motions: e.g., we could use cameras.
- Time **each work element separately** (going is an action, taking is another one, etc.). We need to go in depth and be precise in order to measure them since we need to use the data for our analysis.
- Time several cycles of each work element (we calculate the mean of different shots)
- Observe an operator who is qualified to perform the job
- Always **separate operator time and machine time** (because when the machine is working, normally the operator is not working). We need to separate what is done manually by the worker and what is done by the machine, because the machine can be redesigned in a different way than the workers.
- Select the lowest repeatable time for each element
- <u>Remember shop floor courtesy</u> (be kind)

21.2. PAPER KAIZEN

If we apply this approach in our example: we want to improve the performance of the cell which is done by a bender, we have one assembly station a second one, we have a crimper and a tester.
Furthermore, we have 4 workers, while worker 3 does also internal logistics activity, worker 4 has to take the good from the crimper, load it in the tester and pack the finished good. In order to analysis each work done in a single cell element, we need to use a tool called **APEX PAPER KAIZEN**.



We should use **PAPER KAIZEN**, going inside **looking at process** and activity details (elementary activities performed by workers). We have to record all the work elements and the time needed to do them. DO NOT TRUST PAPER, we need to go and measure them! In the column "work elements" we put all the machines, the workers and the other elements; next to it we put the time, which is the result of our timing (that we measure many) expressing the average of what we measure.
We need to put in the able all the timing and the elements for all the machines and we must stress the separation between the automatic activities and the manual activities.

Bender		Assembly I		Assembly II		Crimper		Tester	
Work Elements	Time	Work Elements	Time	Work Elements	Time	Work Elements	Time	Work Elements	Time
Before Kaizen:		Get bent tube, place	5	Get tube & place	5	Get assembly & place RH side to fixture	5	Get assembly & place to fixture	5
Load auto feeder (batch 25 pcs)	3 min.	Get connector, place & clamp	4	Get teflon hose & right ferrule to hose	4	Place & clamp LH side	4	Remove hose cap & attach hose to fixture	4
Cycle	16 sec.	Get hose & place	4	Place to fixture & clamp	4	Start cycle	1	Start cycle	1
Auto eject		Start cycle	1	Get left ferrule & assemble to hose	4	Cycle (wait)	5	Cycle (wait)	5
		Cycle (wait)	4	Place & clamp	4	Unclamp & remove	3	Remove	4
		Unclamp & remove	2	Get valve & place to fixture	4	Inspect crimps	3	Re-attach cap	1
		Attach convolute	6	Start cycle	1	Pull convolute over crimps	6	Inspect	3
		Aside		Cycle (wait)	7	Aside		Place to shipping container	4
				Unclamp & remove	4				
				Aside					
Machine Cycle = 16 sec.		Machine Cycle = 4 sec.		Machine Cycle = 7 sec.		Machine Cycle = 5 sec.		Machine Cycle = 5 sec.	

For example: In the assembly 1 of the table -> From the machine perspective, the working time is 4 seconds, while from the worker perspective, the waiting time is 4 seconds.
After clustering all the activities, we need to look for **waste**, **inefficiency** and **time**.

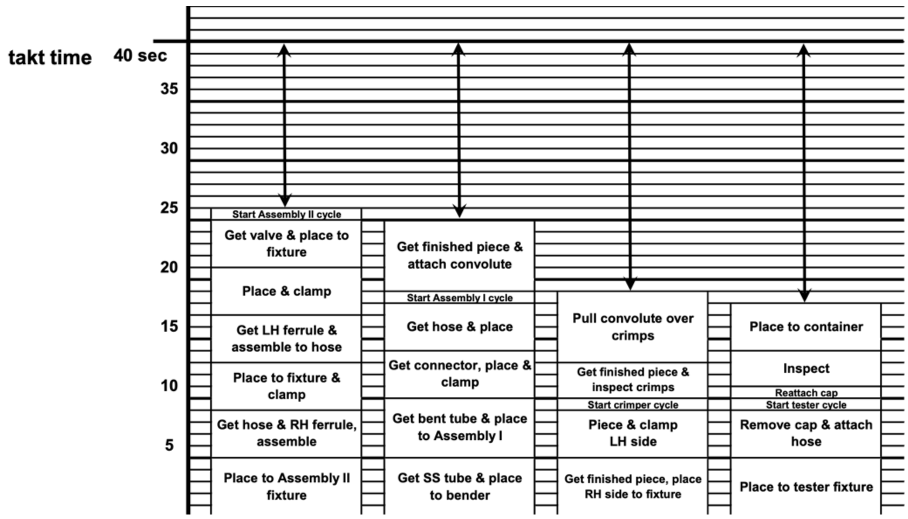
Cycle (wait): cycle of the machine that supports the worker. Highlighting the **machine cycle** is used to understand how much time the worker has available to do something else

We should detect the activities done by workers and also the activities of the machine which is supporting the worker in the process.

- **C/O** is done only when we change the type of product, while the cycle (wait) is the sum of the time needed to do the work of one product.

APEX's operator balance chart (OBC): Current situation

The table is showing the operator balance chart. We need to take each of the 4 columns (stations) and we go one by one and consider only the activities performed at the manual level, the purpose is to **see if the workload of each operator is somehow balanced** and if it is below the takt time.



Example with the tester machine: From the tester the first activity is to get the assembly from the crimper and place it to the fixture which requires 5 seconds, the second operation is to remove cup and attach the holes which takes 4 seconds. We the put 9 seconds since 5+4=9.

The next operation is to start the cycle, which takes 1 second. The cycle then is up to the machine, the operators is doing nothing, for this reason we should not include this cycle in the OBC (we can't put it since we are summarizing only the human work, we are just looking whether workers are balanced or not). We then need to put all the other activities that are performed by the workers in all the stations, except for the bender, which is a fully automatic machine that doesn't require workers.

After clustering all the activities done by each operator in each station, we can now **think of removing the waste**. For example, **is there any activity that we can try to make more automated than it already is?** Can we have some of the activities performed by the workers which can be performed by the machine and thus reducing the number of workers?

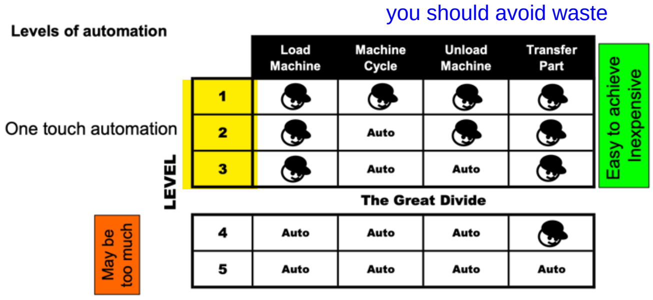
In order to answer to this question, we need to analyze and consider that there is a kind of repetition and a kind of cycle of the activities related between the worker and the machine. The cycle, in fact, is quite standardize: the operator has to set the new work piece aside, remove the finished work piece and put it aside, pick up the new work piece and start the machine, while the machine is working the operator will have to go to the finished workpiece and I bring it to the next machine. This cycle is going to be repeated in an automatic way by each operator. Can we try to automatize some of these activities? **Is there a better level of automation?**

Level of automation:

- **Level 0**: only human work
- **Level 1**: only the cycle of the machine is automated
- **Level 2**: also the unload of the machine is automated (first this one, because the level of precision required by the unload is less than the load, the positioning requires more precision)

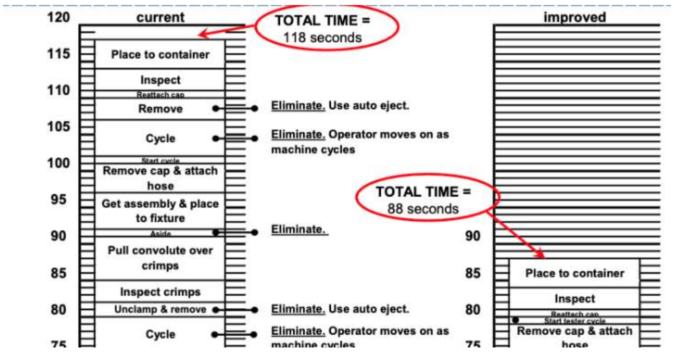
- **Level 3**: also the load of the machine
- **Level 4**: also the transfer of the machine
- **Level 5**: everything is fully automated (cycle, load, unload, transfer)

automate without improvement is risk
you should avoid waste



How we use the analysis done till now? To improve the situation **reducing the time of the workers**, trying to:

- **Eliminate wait while machine cycles** (Operators wait for machine to cycle 21 s)
- **Introduce auto-eject at Ass I, Ass II, Crimper, Tester** (unclamp, remove, set aside): unload the machine in an automatic way.
- **Convert out-of-cycle work** (loading the bender with a batch of 25 tubes) **to in-cycle-work** of loading one tube every cycle

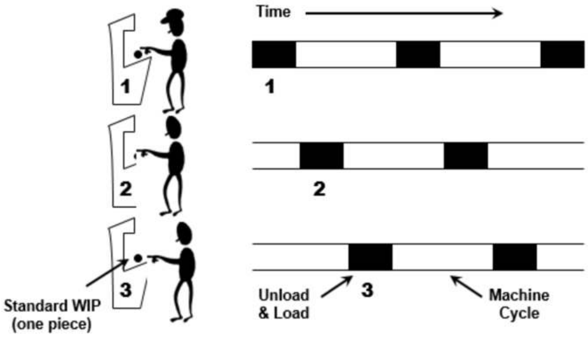


from piece perspective
alocate to different people

Bender		Assembly I		Assembly II		Crimper		Tester	
Work Elements	Time	Work Elements	Time	Work Elements	Time	Work Elements	Time	Work Elements	Time
Before Kaizen:		Get bent tube, place	5	Get tube & place	5	Get assembly & place RH side to fixture	5	Get assembly & place to fixture	5
Load auto feeder ~~(batch 25 pcs)~~	3 min.	Get connector, place & clamp	4	Get teflon hose & right ferrule to hose	4	Place & clamp LH side	4	Remove hose cap & attach hose to fixture	4
Cycle	16 sec.	Get hose & place	4	Place to fixture & clamp	4	Start cycle	1	Start cycle	1
Auto eject		Start cycle	1	Get left ferrule & assemble to hose	4	~~Cycle (wait)~~		~~Cycle (wait)~~	
After Kaizen:		~~Cycle (wait)~~		Place & clamp	4	~~Unclamp & remove~~		~~Remove~~	
Get SS tube & place to bender	5	~~Unclamp & remove~~		Get valve & place to fixture	4	Inspect crimps	3	Re-attach cap	1
Cycle	16	Attach convolute	6	Start cycle	1	Pull convolute over crimps	6	Inspect	3
Auto eject	1	Aside		Cycle (wait)		Aside		Place to shipping container	4
				Unclamp & remove					
				Aside					
Machine Cycle = 17 sec.		Machine Cycle = 4 sec.		Machine Cycle = 7 sec.		Machine Cycle = 5 sec.		Machine Cycle = 5 sec.	

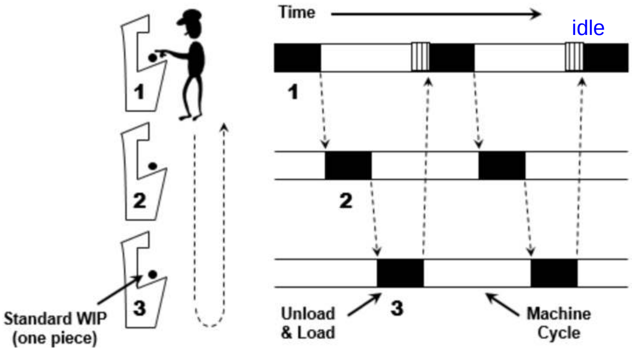
Considering this, how can we apply this to our example? We should:

- **Eliminate wait while machine cycles**: In the moment workers are waiting at the machine, we could try to remove 21 seconds by redesign the system in such a way the worker is not waiting anymore:



In the **traditional approach**: the worker is loading, starting the cycle, waiting and removing the good (the white line is expressing the waiting time); this means that we need three workers, they are shifted and at some point in the production process they are waiting.

We can use a different methodology, which is called the **multi-process**:



we get the perspective of demand meaning market

Instead of waiting, the worker could do his operation, move to the second station while the machine is working, do activities there, start the cycle on the second machine, move to the third station, start the cycle and then go back to the first station. It could happen that the cycle of the machine is shorter than the loading and unloading, this doesn't matter, the machine would be a bit unsaturated because it is waiting for the worker to come back BUT **the worker is fully saturated because he is just one**.

- **Introduce auto-eject at Ass 1 and 2, crimper and tester**: We move from level 2 of automation to level 3, if we introduce auto-ejection system we can improve the unloading part done by the machine. By doing so, some activities in the APEX PAPER KAIZEN could be eliminated since the operator is not doing them anymore.

- **Convert out-of-cycle work to in-cycle work of loading on tube every cycle**: Allow to load the machine while the machine is working, for example we could create an autoloading system like a buffer.

If we do this, we are removing some of the workload done by operators and the activities in the APEX PAPER KAIZEN. The final result would be to eliminate the time of 100 seconds. From 188 second to removed 100 and we now have 88 seconds. Once we have done this, we always need to double check the activities we decided to squeeze, in order to understand if the activities we have done so far are able **to meet the takt time**.
The final result will be the one in the table →

In a world where machine are not completely capable, or completely available, Effective TT <= 0,8*minTT

Machine	Machine Cycle	Load, Start & Unload Time	c/o Time/ Batch Size	Effective Machine Cycle
Bender	16 sec	5 sec	0	21 sec
Assembly I	4 sec	16 sec	0	20 sec
Assembly II	7 sec	28 sec	1*	36 sec
Crimper	5 sec	12 sec	0	17 sec
Tester	5 sec	12 sec	0	17 sec

Sometimes, it can happen that we are not meeting the takt time, if this occurs, we must look for more efficiency and revise some other operations:

- Reduce the setup time.
- Increase availability.
- Kaizen the load, start, unload process: increase the quality and decrease the number of scraps.
- Eliminate waste in the machine cycle itself.
- Split apart some of the tasks a bottleneck machine is performing and use more than one machine.
- Install 2 machines of the same type in parallel: instead of producing 1 piece you produce 20 pieces, so the production rate is still 10 seconds for each machine and overall.
- Create 2 cells instead of one: if there are more than one stations that is over the tt is better if we split the two cells.
- If everything else fails remove the bottleneck equipment from the cell, decouple it and operate in batch.

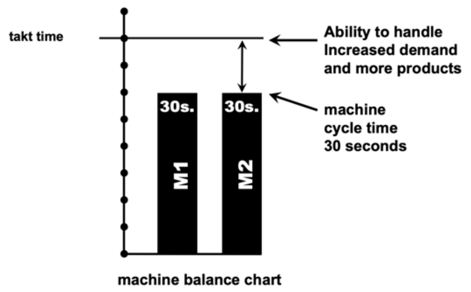
Question: what kind of machine is it better to use for each production process?

We can have two options:

1) **FEW MULTIFUNCTIONAL MACHINES**: In this case, the machines will be few, but they will be able to perform complex processes.

   Pros: Economy of scale.

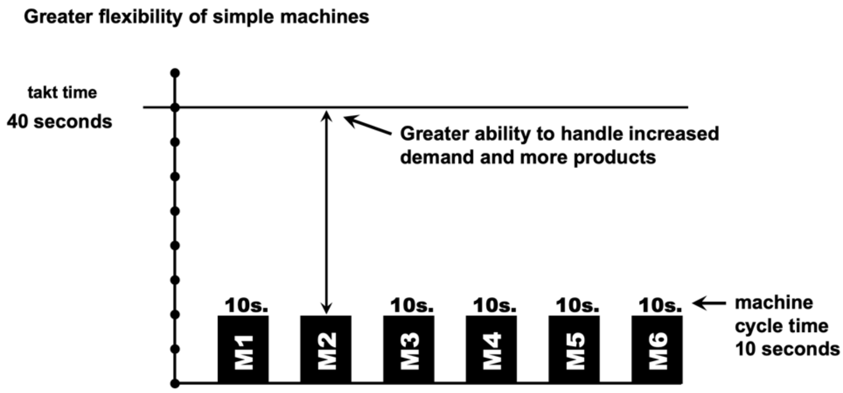
   Risk: If we have an over demand we can be too close from the takt time.



2) **MANY AND SPECIALIZED MACHINES:** In this case, each machine would be specialized in a specific procedure.

   Pros: Higher room to absorb demand variation.

   Cons: Reliability (because if one machine stops working, the entire system stops working), mix flexibility is lower because each machine is specialized.



Guidelines for supporting in choosing between multifunctional or specialized machines:

1) Small equipment for specialized machines or multi-task equipment for multifunctional
2) Introduce auto-eject (Level 3 Automation) whenever operators must use both hands to handle the part.
3) Install one-touch automation where possible.
4) Avoid batching.
5) Incorporate sensors to signal abnormal conditions and even automatically stop machines if necessary, so operators do not need to watch machines during their cycle.
6) Design in maintainability.
7) At the pacemaker process, strive to device machine changeovers between different end items that take less than one takt time cycle.

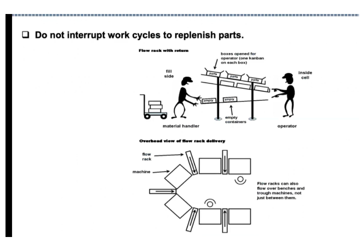
Can we comment on the overall layout of the cell?

1) Place machines and workstations close together to minimize walking distance
2) Remove obstacles from the efficient operator walking path
3) Eliminates spaces and surfaces where work-in process inventory can be accumulated: if you don't give space to people, they will be forced to find solutions
4) Avoid-up and down and front to back transfers

Guidelines for cell layout:

Arrange machines, workstations, and material presentation devices as **if only one operator makes the product from beginning to end.**

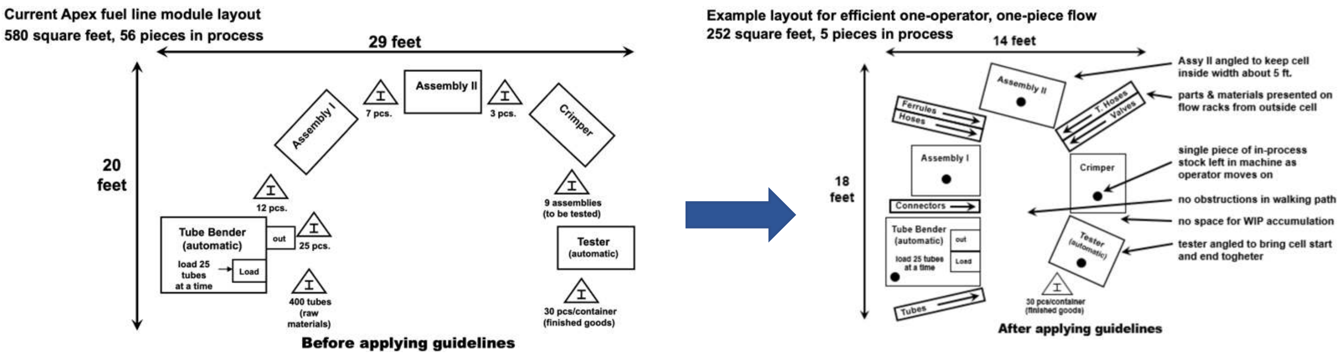
1) Absolutely ensure safety and good ergonomics
2) Segregate level 5 automation and continuous cycle operations from manual operators or operator-based workflow as shown in the diagram below.
3) We prevent incident and increasing the safety of the worker.



Guidelines for material management:

1) Keep no more than two hours of material at the point of use, do not interrupt work cycles to replenish parts.
2) Display parts in order to make operators able to use both hands simultaneously.
3) Try to keep all part variations at the operators' fingertips at all times to eliminate changeover time.
4) Do not have operators get or restock their own parts.

Before implementing all of these suggestions, the layout is as follow



As we can see there is still a problem, the space is too much; in fact, the tube bender needs modification since the operator needs to move really far.

By implementing the suggestions:
We can remove a lot of inventories in a way that machines can get closer and worker from assembly 1 has no obstacles, the beginning is close to the end, so the U shape is closer. Dealing with batch-oriented equipment:

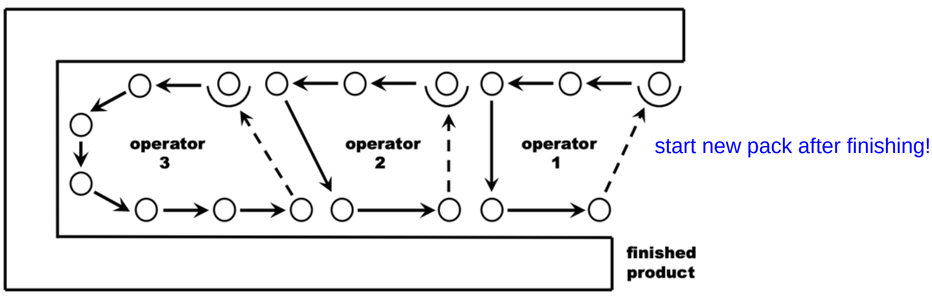
- Separate batch process from continuous flow with a supermarket or a FIFO lane
- Keep in the continuous flow if equipment is moving-conveyor type (and operator can drop a single piece at start and pick up one at end within TT).
- Ignore equipment batch capability and use it single piece.
- Transform equipment from batch to single.
- Replace batch equipment with one or more single piece inexpensive equipment.

How many operators should we have after the improvements?
We got to 88 seconds, so the total work content is 88 secs and takt time is 40. The required operators are: 88/40 = 2.2 operators.

Remainder in # of operators calculation (after paper kaizen)	Guideline / Target
< .3	Do not add an extra operator. Further reduce waste & incidental work.
.3 - .5	Do not add an extra operator yet. After two weeks of cell operation & kaizen, carefully evaluate if enough waste & incidental work can be taken out.
> .5	Add an extra operator if necessary and keep reducing waste & incidental work to eventually eliminate the need for that operator in the cell.

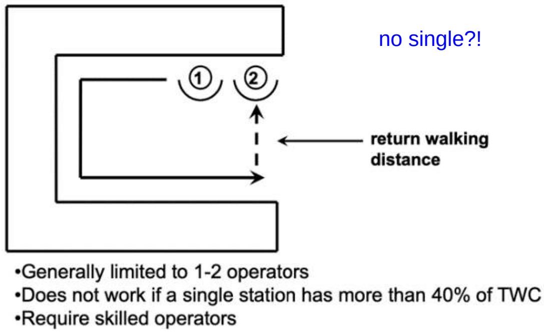
But how can 3 workers work with 4 stations?
We could solve this problems in different ways:

1) SPLIT THE WORK:
Solution: we have a line there are many stations in the line, there are in fact 17 tasks to be performed ONLY with three operators. We distributed the work among them, the first operator does the work in on the first stations then he does one of the last one. We can decouple the number of station by the number of workers. The takt time can be covered by a smaller number of workers.



2) THE CIRCUIT:

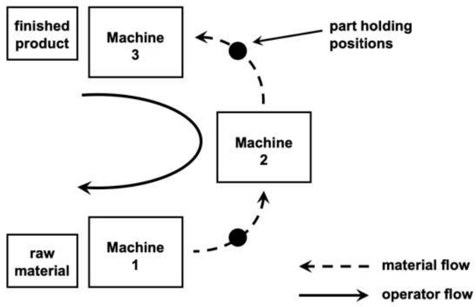
Each operator is performing all the tasks and moving all around; in this case we need multifunctional operators and it could be a bit less efficient. We cannot put too many people because otherwise there would be confusion.



no single?!

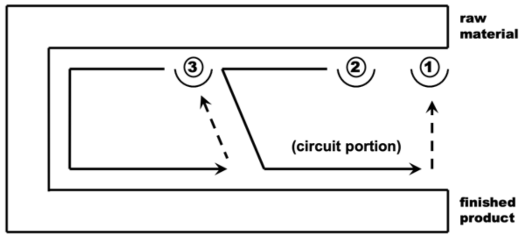
3) REVERSE FLOW:

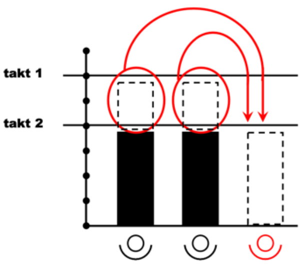
Workers are going the other way round. The operator flow is the straight line and the material flow is the dotted line. Material flow goes to machine one to machine 2 but the worker goes the other way around, we need higher flexibility of the workers and they are working on the same product, so we can increase the quantity.



4) COMBINATION:

Worker 1 and 2 are going in circuit and worker 3 is working by himself.





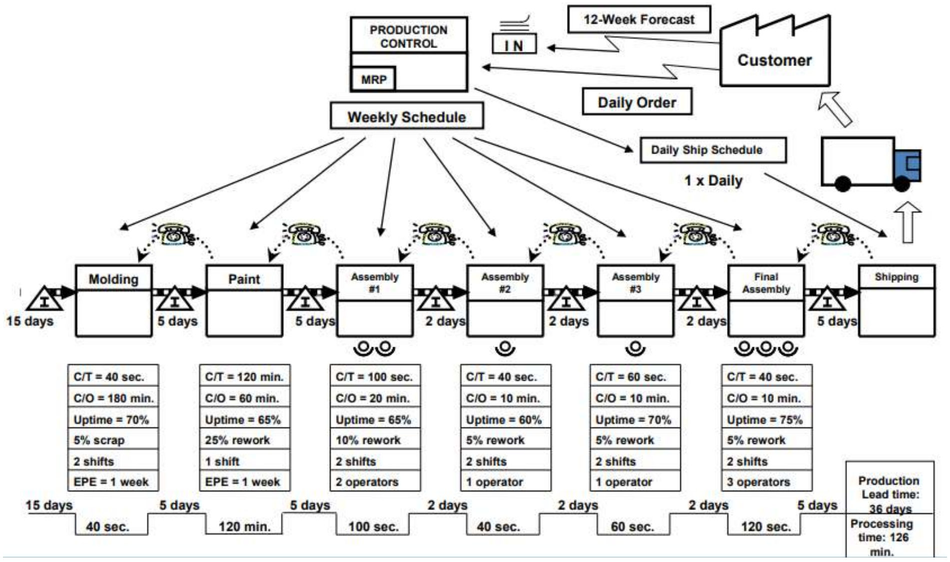
How will the pacemaker react to changes in customer demand?

When designing cells, engineers should prepare one-up and one-down scenarios for responding to volume changes. If the demand increases, the takt time decreases and then you need to move part of the work to another operator.

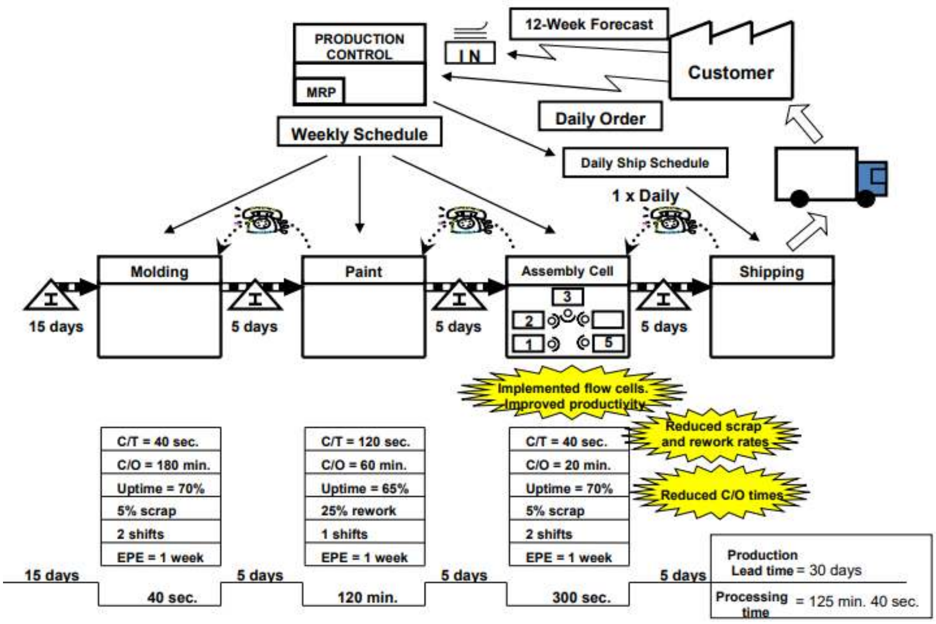
design for 20% increase of demand!

22. LEVEL AND PULL

If we want to increase the performance of a system, one of the main improvements is to have a pacemaker that works in a stable way. What can happen to our system, considering it from a higher level? We are going to discuss how to configure, and plan the pacemaker; which are the strategies used to match our capacity and the demand.



In the current state map above we can see many stages, from the molding to the shipping. Moreover, we can notice that the lead time (36 days) is very high compared to the 2 minutes of processing time. We can try to improve the situation



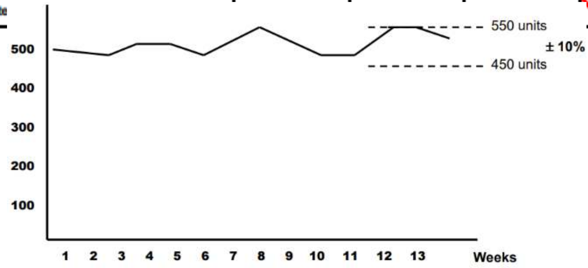
This is the future state, where we find an assembly cell after the molding and paint stages. All the stages/cells are decoupled by buffers. Anyway, we start the system but after a short time the performance worsen: they are better than before, but they are still not enough, not as desired.

We can try to understand what happened inside the pacemaker. The reason for the problems is due to the fact that actually, the company was not able to sustain the logic of the supermarket, resulting in continuous and repetitive actions of the different departments trying to solve the issues, consisting of people soliciting the supplier to provide the materials. And the most used tool nowadays is still the phone. This continuous changing of the schedule from the single department resulted in the misalignment with the requests from the supermarket designed before. This is reflected by the performance table:

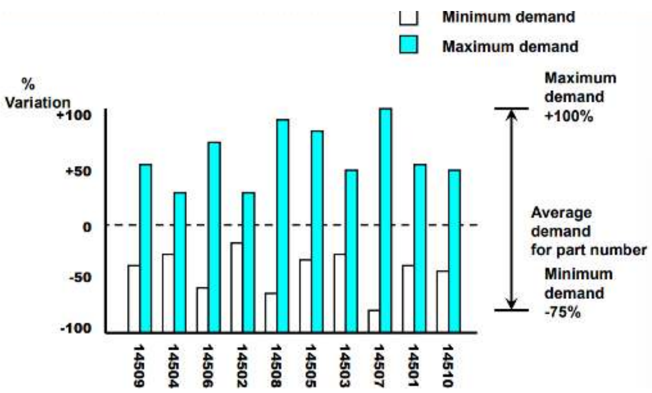
- The number of people necessary for internal logistics increased
- Decrease in productivity in assembling
- Increase in the inventory
- Very bad performance on the on-time delivery, with lots of late deliveries

	Original state	After basic stability	After cell flow kaizen	Current state
Productivity				
Direct labor (pieces per person per hr.)	9.0	10.0	11.5	11.0
Material handlers supporting value stream	3	3	3	4
Quality				
Scrap	5%	4%	3%	2%
Rework**	25%	20%	15%	15%
External (ppm)	500	250	125	105
Downtime***				
Assembly (min. per shift)	40 min.	30 min.	10 min.	20 min.
Paint (min. per shift)	30 min.	20 min.	15 min.	15 min.
Molding (min. per shift)	50 min.	25 min.	25 min.	10 min.
Inventory turns				
Total	8	11	14	12
On-time delivery				
To assembly	65%	68%	80%	75%
To shipping	80%	92%	95%	85%
To customer	100%	100%	100%	100%
Door-to-door lead time				
Processing time (min.)	126.0	126.0	125.7	125.7
Production time (days)	36	34	28	30
Costs				
Overtime costs per week	$6,000	$5,000	$4,000	$5,000
Expedite				$2,000

What can we do?
This situation indicates that the problem is not due to a specific issue in the stages, but how they are integrated. So, we start from the customer demand, and we see that the shape is like this one. Fluctuation of the demand is very slow, while if we go on the floor, it is perceived as very higher. Where is the problem?

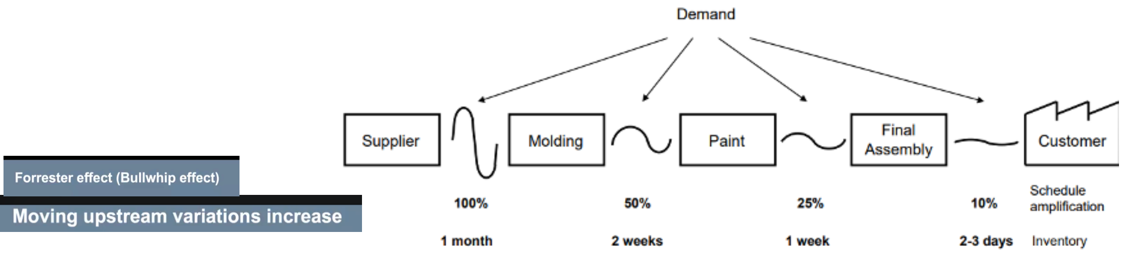


Overall demand variation is low



If we go deeper into the demand for a single item, we notice some alarms: the demand is actually very variable within the same single order. This is critical for the system because if the variability is not properly managed, it could create a lot of issues in the supply chain because variability is the starting point for the **bullwhip effect** in the supply chain.

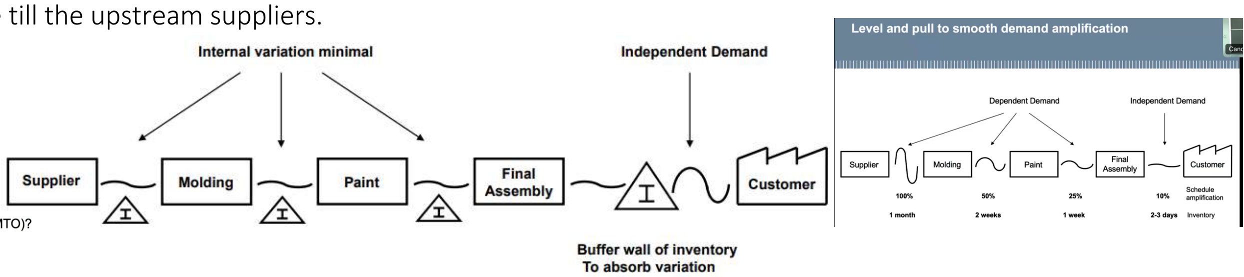
BULLWHIP EFFECT: amplification of the variability when you move upstream in the supply chain. If you go to the most downstream of the supply chain, the variability is very small. Because of how the systems are managed, that variability is amplified every time you move upstream to your supplier. It increases in each step of the supply chain. The most common example is the supply chain of slippers. Children consume them very stably (4-5 times a day) for a lot of months and this is the final consumer. How is the consumption of the parent? Twice a week, for a total of 7*%=35 slippers per week. The problem is that the supermarket doesn't see consumption of 5 per day, but 35 every 7 days, and so the variability starts increasing. At the same time, the supermarket itself purchases them from the central warehouse every 2 weeks, which sees 35000 slippers per 2 weeks, so the variability increases even more and so on.



How many stocks do you need to absorb the variability? It is easier downstream because you only need lower stocks.
What are the sources of this effect?

- **Lead time**: the longer, the bigger the amplification and also how long is the supplier, because it affects the lead time itself.
- **Batch size**: if you increase the batch, you increase the amplification. Why should someone do it? Optimization of the number of setups, higher bargain power
- **Unpredictability**: the higher the probability to have an unforeseen event, the higher is the amplification
- **Wrong expectations, fears**: because you tend to be more cautelative

This is actually what happened in our case. The continuous rescheduling caused a very high bullwhip effect. The strategy is to decouple with a buffer, which consists of the finished good warehouse. This has the role of absorbing the variability from the market and then transmitting stable demand to the supply chain, which has to remain stable till the upstream suppliers.



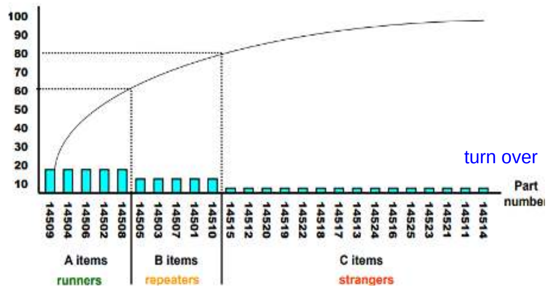
1. What production strategy for each product (MTS,ATO,MTO)?
2. How much inventory for each MTS product?
3. How should you organise and control the finished goods store?

22.1 MATCHING CAPACITY TO DEMAND

1. What production strategy? (MTS, ATO, MTO)

In many cases you cannot use the same logic for all the items of the production range. First of all, you have to analyse the range of product according to the demand. Not all the items are sold in the same volume, some are more relevant, some less.

You can do a Pareto analysis and classify the products in 3 categories: **runners** (higher volume), **repeaters** and **strangers** (smallest part). This means that a small part of items is responsible for the biggest volume ordered, while the rest of the volume is due tou the lare majority of products sold in lower quantity.

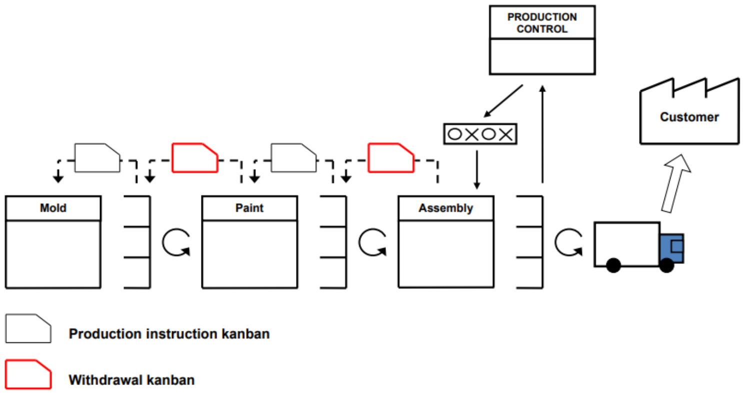


turn over

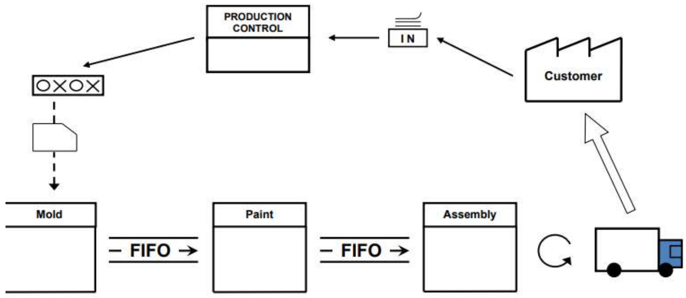
Options	Pros	Cons	Apogee situation	Apogee decision
1. Hold finished-goods inventory of all products (As, Bs, and Cs) and make all to stock–**replenishment pull system.**	Ready to ship all items on short notice	Requires inventory for each part number and much space	Finished-goods stores and shipping unable to hold all items	Not practical due to physical layout constraints and number of end items
2. Hold no finished-goods inventory and make all products to order–**sequential pull system.**	Less inventory and associated waste	Requires high process stability and short lead time to produce	Production lead time too long and paint process too unstable	Not pratical with current lead time and capability
3a. Hold only Cs in inventory and make A and B products to order daily–**mixed pull system.**	Less inventory	Requires mixed production control and daily stability	Daily stability a concern	Possible second step for future
3b. Hold A and B products in finished-goods inventory. Make Cs to order from semifinished components–**mixed pull system.**	Moderate inventory	Requires mixed production control and visibility on C items	Most applicable to current situation	Best fit for today

According to this we can define 4 strategies:

- **Replenish pull system**: I satisfy with a rigid MTS for all the clusters. The issue is that we have huge inventories, also for the low demanded products. Finished good warehouses that absorb the variability, then Kanban that transmit the demand upstream, controlling and limiting the variability.



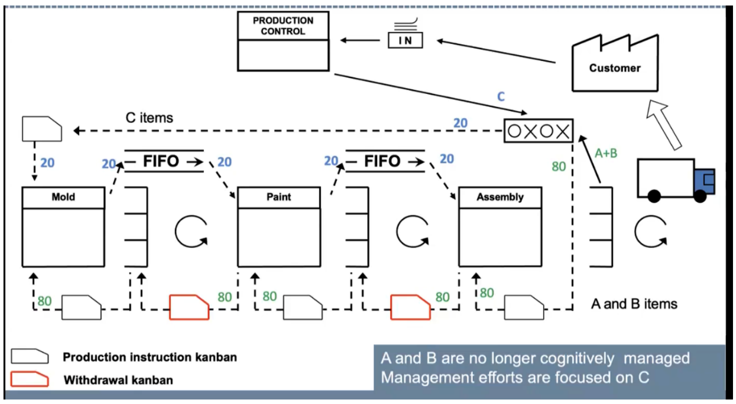
- **Pure sequential pull system**: all the items are satisfied with MTO, the stocks are low, but the response time is high and you need to be very flexible and fast in the response. Not so easy. The order is entering the system from the beginning and moves downstream. Don't confuse it with push system: it is a pull system, because the item advances only if the painting consumes the material, otherwise the moulding stops producing. The advancing of production is linked to the advancing of production of the downstream stages.



c is mto

2nd step for future
not now
- **Hybrid strategy**: mixed pull where we have runners and repeaters (MTS), while cluster c (strangers) is satisfied with MTO if you can do something quickly for a and b, you can do it for c as well.

best
- **Mixed pull strategy**: complementary, cluster A and B in MTO, while the c cluster is satisfied with MTS. It is very rare to find a company with this approach. Finished good ware (A, B) and the order of C that arrives with a sequential pull logic. c: MTO



What are the results?
Pro: level of the stock lower than in the pure replenish system, but the items in stock have a high level of turnover ratio
Cons: you are absorbing the variability (with stocks) of the clusters with lower variability (runners)

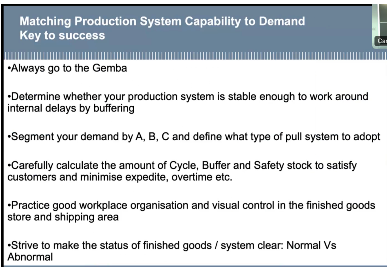
2. How many inventories?

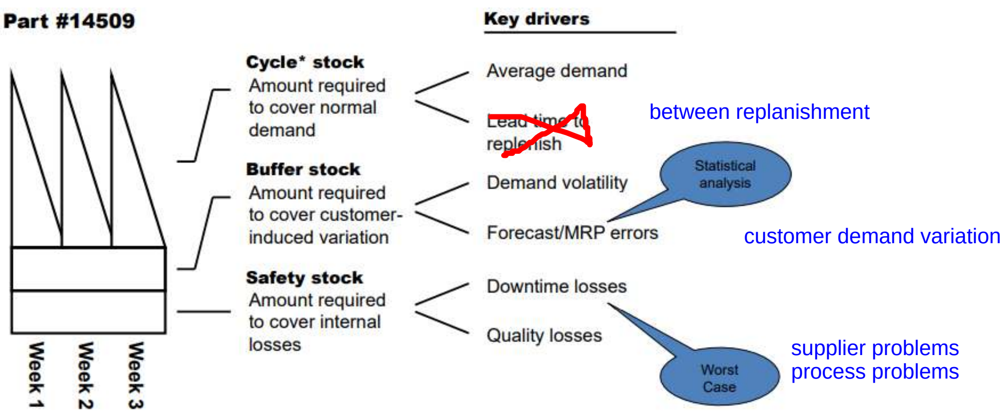
We start from the finished goods inventory. We have a rigid classification between the inventories:

- **Cycle stock** supply and delivery misalignment
- **Buffer stock**: it covers the downstream from the demand variability and from the errors in the forecast of MRP (exogenous causes)
- **Safety stock**: it covers the internal losses, scraps, quality losses, and downtime losses, computed by statistical analysis. You take into consideration the scrap rate of your upstream stage.

Finished-Goods Calculation

	Average daily demand x Lead time to replenish (days)	Cycle stock
+	Demand variation as % of Cycle stock	Buffer stock
+	Safety factor as % of (Cycle stock + Buffer stock)	Safety stock
=		Finished-goods inventory





The simplified formula is EPE*D*1.5 (**coefficient of security** equal to 1.5 or 2 depending on the situation), the increment of buffer stock + safety stock.

3. How will you organize and control the finished goods store?
   - Cycle stocks are free to be used
   - Buffer stocks are locked: if you need to use them you need to ask the permission of the head of the department
   - Safety stocks are locked: the key is in the hands of an even higher-level manager than the previous one. In this way, you are sure that the manager is informed because it means that something strange happened. We need to understand if the cause is structural or contingent, and so if we need to react changing something or not.

4. Where will you schedule the value stream?
5. How will you level production at the pacemaker?
6. How will you convey demand to the pacemaker to create pull?

22.2 22.2 CREATING THE PACEMAKER

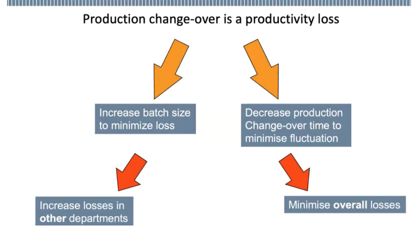
4. Where will you schedule the value stream?

Referring to the case, as we said before, we are going to set a mixed pull strategy, so we need to schedule two points, according to the general rule:

- For the products that go for the **replenish pull**: the pacemaker goes before the finished good warehouse. In the example, it goes in the final assembly.
- Products that go for the **sequential pull**: it is often the first process at the beginning of the value stream

In order to create a levelling of variability, how do we smooth the production?

			Customer requirements	Final-assembly build schedule (left-hand cell)	
Part #	Demand category	Mirror description	Monday orders	Monday 1st shift	Monday 2nd shift
14509	A	Black heated	140	500	0
14504	A	Black unheated	110	0	500
14506	A	White heated	120	0	0
14502	A	White unheated	120	0	0
14508	A	Red heated	110	0	0
14505	B	Silver heated	70	0	0
14503	B	Yellow heated	60	0	0
14507	B	Bronze heated	70	0	0
14511	C	Purple unheated	100	0	0
14512	C	Gold unheated	100	0	0
		Total	1,000	500	500



in upstream size of batches increase

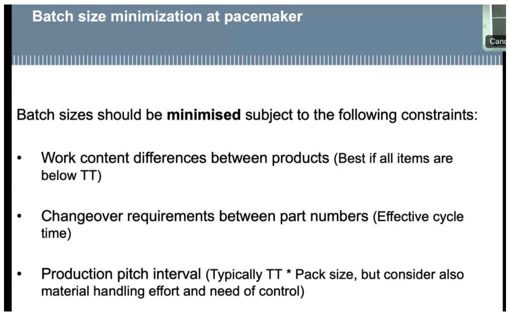
5. How will you level production at the pacemaker?

We change the perspective, instead of increasing the batch size, to minimize losses in the local dep, we work to decrease the C/O by increasing the number of setups, reducing the batch size, and minimizing the overall losses. How do we do it? First, let's analyse our capacity.

What is the takt time? 54 sec * 10 pieces = 9 min

What is the pitch interval? 450 min / 9 min = 50 pitches

Now we distribute the pitches according to the demand we have:



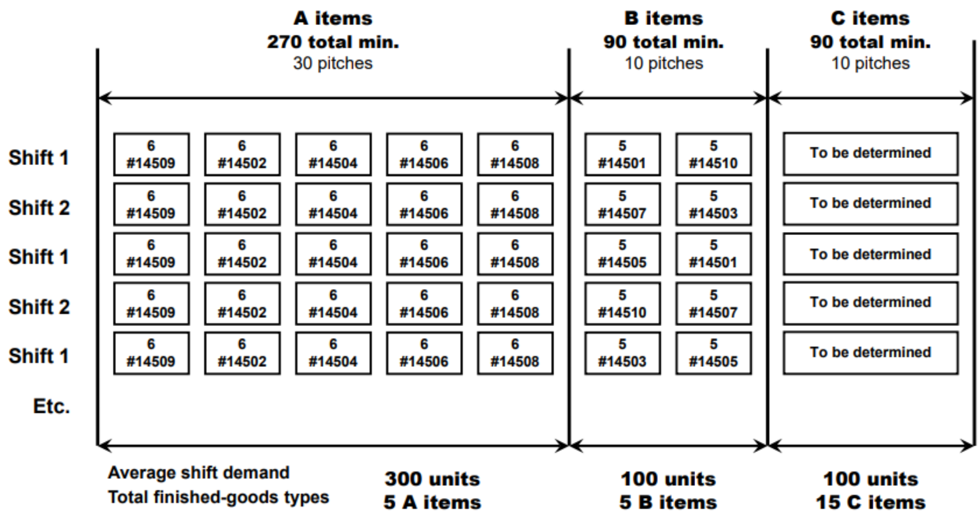
Total interval	x	% of production mix	=	Intervals per item	(equivalent time)
50 intervals	x	60%	=	30 reserved for As	(9 min. x 30 = 270 min.)
50 intervals	x	20%	=	10 reserved for Bs	(9 min. x 10 = 90 min.)
50 intervals	x	20%	=	10 reserved for Cs	(9 min. x 10 = 90 min.)

For product A, since it is demanded every day, we produce and distribute the 30 pitches on the 5 products.

Product A: 30 pitches (300 pieces) / 5 items = 6 pitches (60 pieces) per item

Product B: 10 pitches (100 pieces) / 5 items = 2 pitches (20 pieces) per item

About product C we only know that we dedicate 10 pitches (100 pieces) to produce c items, which are affected by a very variable demand, so you dedicate that time to produce what is actually demanded. In this case, the batch is not fixed, we are flexible on that.



leveling the mix can improve it.
sequential pool ?

3 3 3 5 c 333 5 c

more level because every 4 hours you have it.

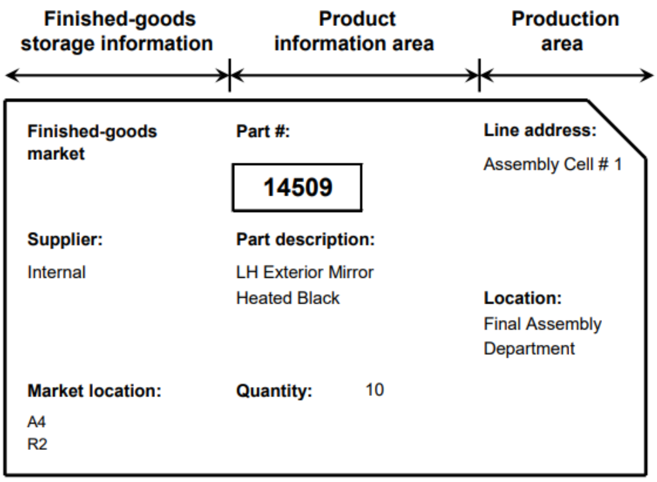
shorter epe! => super market has less inventory

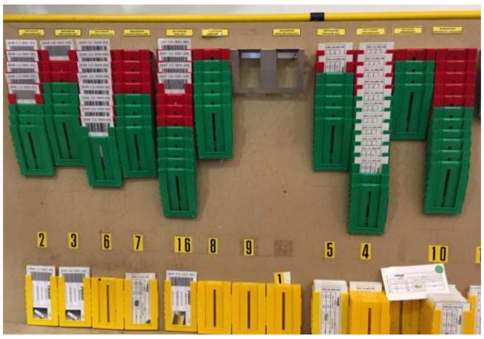
goal is to have reactive system

6. How will you convey demand to the pacemaker to create pull?

We do it through Kanban. We can distinguish two types:

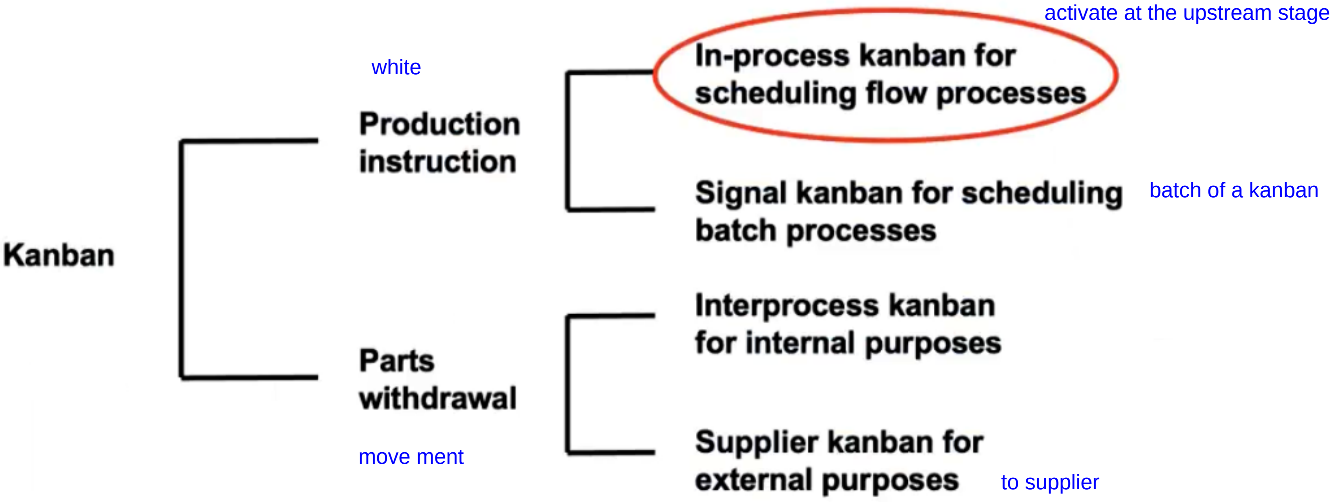
white
- **Production Kanban**: it signals the necessity to produce when the shave is empty. The Kanban is detached, and the runner is in charge to bring it to the central warehouse, it takes the components of this Kanban, then it goes back and replenishes the shelves of components. Then he attaches a new Kanban (the withdrawal Kanban) to the new tray.
- **Withdrawal Kanban**: it triggers the transportation of the stocks from one warehouse to another



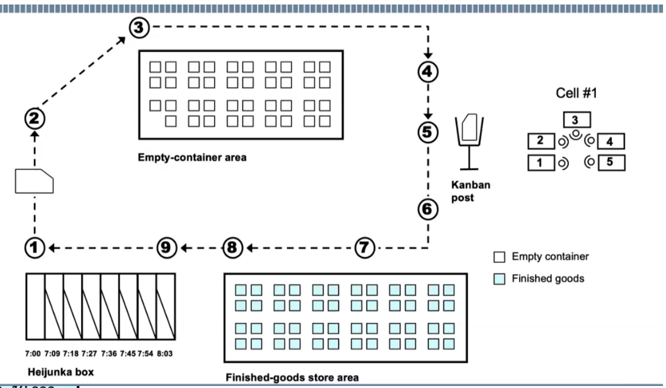


Those people are specifically hired for this purpose: to replenish the component without stopping production. The **waterspider** gives the rhythm to the pacemaker and it is very important because he avoids overproduction upstream.

The **Heijunka Box** has the particular role of not saying only what to replenish, but also to give the rhythm to the pacemaker. In fact, it is present the scheduling time of each Kanban.



Step	Activity	
1.	Pick up instruction kanban from the heijunka box.	
2.	Travel to obtain empty finished-goods container.	
3.	Obtain empty finished-goods container for delivery to the assembly cell.	
4.	Travel time to the assembly cell.	
5.	Drop off an instruction kanban and an empty finished-goods. container at the cell, pick up finished goods from the previous cycle.	
6.	Travel time to the finished-goods market.	
7.	Drop off finished goods to their correct storage location.	
8.	Log product into finished-goods inventory (e.g., scan bar code).	
9.	Return to heijunka box for next instruction.	
	Total time	6 min. 10 sec.



Read file level and pull

23. MANAGE VARIABILITY

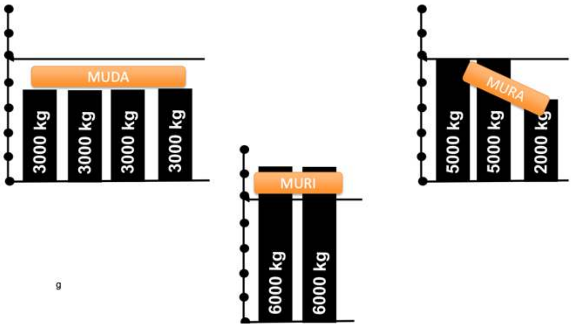
23.1. MUDA MURA MURI

Eliminating wasteful activities is one of the most important prerequisites for building a successful company. This concept is an integral part of **Lean thinking**, and it helps you increase profitability.

Lean is mostly known for **Muda** (Wastes) hunting, but the concepts of **Mura** (Variability) and **Muri** (overload) are equally important

Example: Let's clarify the 3 categories considering the transport of 12000 kg of materials through a pickup truck that can move 5000 Kg: how is it possible?
Which option do you have how?

- 4 trips of 3000 kg → MUDA
- 2 trips of 6000 kg → MURI
- 2 trips of 5000 kg + 1trip of 2000 kg → MURA

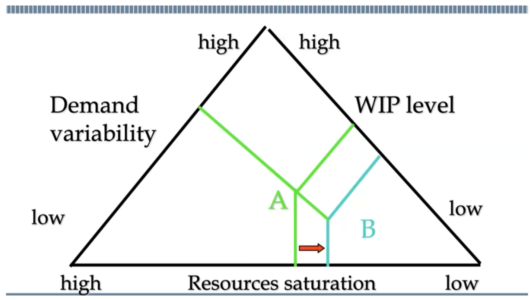


MURI problem is clear
→ Overload of resources, beyond their technical limits

- Safety Issues
- Quality Issues

MURA: what are the problems?
→ Fluctuations, <u>variability</u>, irregularity of the workload

- Inefficiencies
- Queues/stocks



23.2. KINGMAN'S FORMULA

The Kingman formula approximates the relationship between utilisation, fluctuation and waiting time. Now a dice game is explained where we simulate a supermarket checkout so that participants can experience the impact of utilisation, turnover and the (worst) combined impact of both. The waiting time in front of a process is strongly influenced by two factors: the utilisation of a process (how busy is it?) and the fluctuations in arrival and process times (how stable is the system - or not?). The Kingman equation and other equations show this. When the utilisation reaches 100 per cent, the waiting time approaches infinity. When the fluctuations in arrival and processing times increase, the waiting time also increases. This game will demonstrate this effect.

Dice Game Queue AllAboutLean

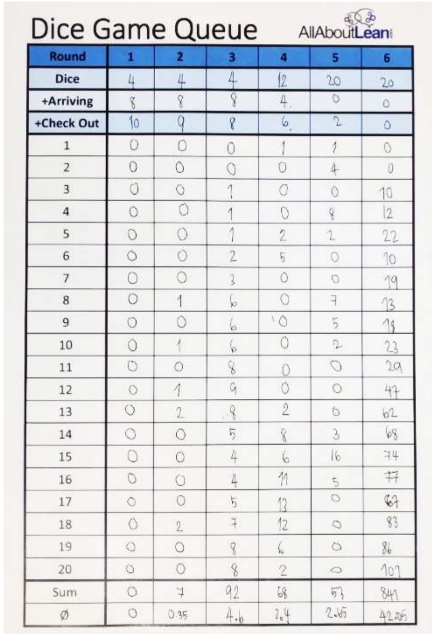
Round	1	2	3	4	5	6
Dice						
+Arriving						
+Check Out						
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
Sum						
Ø						

What you need are dice! In particular, you need three different types of dice - a small one (D4 or D6), a medium-sized one (D10 or D12), and a large one (D20 or D30). Then you would need to print out the data sheets (one for every two players, plus maybe a few for backup). You will create groups of two people, each simulating a supermarket checkout. One will determine the number of people arriving, the other the capacity of the checkout. They get one dice-game queue-length sheet as shown above. Each then gets a die, which will determine the number of customers arriving at the checkout and the number of customers processed at the checkout. They may have to share one die if you are short on dice. There will be six rounds in the game, and each round consists of twenty iterations. In the first three rounds we vary the utilization.

VARYING UTILIZATION

Round 1 – D4 with Offset 2

In the first round, each player gets a four-sided die. For the number of people arriving, we throw the die and add 8, to reduce the 'weight' of the die. For the capacity of the checkout, we also throw a die, but this time add 10. Therefore the checkout has on average two more capacity (Offset 2) than the number of people arriving. This information would also be the first three empty rows on the data sheet, where we write 4 (for the sides of the die), 8 (for the customer), and 10 (for the supplier).



Dice Game Queue

AllAboutLean

Round	1	2	3	4	5	6
Dice	4	4	4	12	20	20
+Arriving	8	8	8	4	0	0
+Check Out	10	9	8	6	2	0
1	0	0	0	1	1	0
2	0	0	0	0	4	0
3	0	0	1	0	0	10
4	0	0	1	0	8	12
5	0	0	1	2	2	22
6	0	0	2	5	0	10
7	0	0	3	0	0	19
8	0	1	6	0	7	13
9	0	0	6	0	5	18
10	0	1	6	0	2	23
11	0	0	8	0	0	29
12	0	1	9	0	0	47
13	0	2	8	2	6	62
14	0	0	5	8	3	68
15	0	0	4	6	16	74
16	0	0	4	11	5	77
17	0	0	5	13	0	67
18	0	2	7	12	0	83
19	0	0	8	6	0	86
20	0	0	8	2	0	107
Sum	0	7	92	68	57	841
Ø	0	0.35	4.6	3.4	2.85	42.05

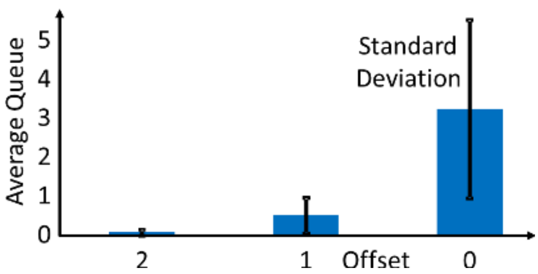
Both throw the die and add to determine the arrivals and capacity. If the capacity matches or exceeds the demand, the two players write a zero in the next cell below. If the arrivals exceed the capacity, the number of people waiting is noted down. These people waiting in the queue have to be processed in the next iteration. They repeat this for twenty iterations, always keeping track of the number of people waiting in line. Afterward they calculate the total and average queue length by summing up the twenty entries and dividing by 20. This is written in the last two rows of the sheet. The results are added to an overview statistic for all teams. In all likelihood there will be no queue, as the expected outcome is 0.07 people waiting in line with a standard deviation of 0.08 people (these averages are obtained if we had 20 tables like the one above).

Round 2 – D4 with Offset 1: The next round we reduce the offset. The arrivals keep on adding 8 to the die, but the checkout now adds 9 (instead of 10) to his throw. Therefore, the first three rows of round 2 are 4, 8, and 9. Again they play for twenty iterations and then determine the average queue length. These results will be added to the overview statistics for all teams. The expected outcome is now an average queue of 0.5 people with a standard deviation of 0.4 people, so there will be more people waiting in line.

Round 3 – D4 with Offset 0: The next round we eliminate the difference between the arrivals and checkouts. Both can add only 8 to their die throw. Now the queue starts to heat up. The expected outcome is an average queue length of 3.1 people with a standard deviation of 2.3 people.

THEORY ON UTILIZATION

By reducing the excess capacity per iteration from 2 to 1 to 0, we have actually increased the utilisation of the cash register. When the checkout had on average two more capacity than the people arriving, he had in effect a utilization of 88.57%. With an offset of 1, the utilization was 93.945%. Without any offset, the utilization was 100%. In the long run, an utilization of 100% means a queue length of infinity, but since we did only twenty iterations, we did not get that far. The graphs shows the expected average outcomes with the standard deviations for a D4 with different offsets.



VARYING FLUCTUATION

In the next two rounds, we vary the fluctuations while keeping the offset constant. This means the utilization will always be at 88.57%.

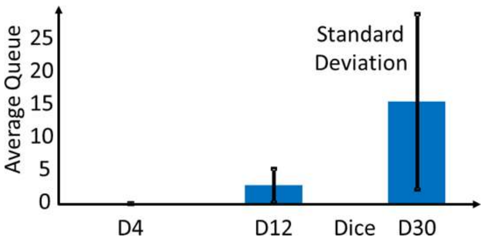
Round 4 – D12 with Offset 2: Now we use a twelve-sided die. The arrivals add 4 to his throw, and the checkout adds 6 to his throw. Again, after twenty iterations, we add the averages to the overview sheet. The expected outcome is an average queue length of 3.4 people.

Round 5 – D20 with Offset 2: Finally, we get to use the twenty-sided die. To maintain the same utilization and offset, the arrivals add nothing to his die throw, while the checkout adds 2 to every die throw. The expected outcome is an average queue length of 2.65

THEORY ON FLUCTUATION

By changing the die from a D4 to a D12 to a D30 while keeping the average arrivals and checkout capacity constant, we in effect increased the fluctuation.

- The D4 +8 arrivals means that the number of arrivals fluctuated between 9 and 12 with an average of 10.5.
- A D12 +4 keeps the same average of 10.5, but now with a fluctuation from 5 to 16.
- A D20 plus nothing again has the same average of 10.5, but this fluctuates now between 1 and 20.



Hence, we increased the utilization. Similar to the utilization, increasing the fluctuation also increased the average backlog as shown in the chart above.

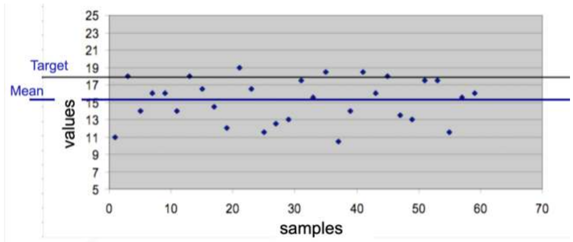
VARYING UTILIZATION AND FLUCTUATION

Round 6 – D30 with Offset 0: For the next round we increase both fluctuation and utilization. Both the arrivals and the checkout get a thirty-sided die, and neither gets to add anything to the die throw.
The expected queue length is 42.05 people. Hence it is much bigger than expected. Emphasize that this is not additive but multiplicative as per the Kingman equation above.

	Mean Queue Length			Standard Deviation of Queue Length		
Dice	Offset 2	Offset 1	Offset 0	Offset 2	Offset 1	Offset 0
2	0,00	0,00	1,21	0,00	0,00	0,88
4	0,07	0,49	3,07	0,08	0,42	2,28
6	0,47	1,35	4,80	0,39	1,11	3,59
8	1,01	2,89	5,80	0,79	2,99	4,03
10	1,88	3,86	7,70	1,50	3,35	5,48
12	2,98	5,39	10,26	2,44	4,45	7,25
14	3,91	6,83	11,71	3,33	5,75	7,90
16	5,49	7,33	13,24	4,88	5,32	8,60
18	6,04	9,56	14,62	4,93	7,42	10,35
20	8,22	11,99	15,67	6,95	9,26	11,83
22	9,85	12,33	19,17	8,23	9,19	13,67
24	10,92	14,19	20,51	9,39	10,56	14,30
26	12,44	16,99	23,02	10,92	11,70	16,17
28	11,89	18,65	22,83	9,13	13,44	15,75
30	14,70	20,36	24,18	11,02	14,49	17,64

23.3. HOW TO SET TARGETS

- **Numerical targets** without specifying the method of achieving them are only exhortations and can be demotivating.
- **Time standards** (or output standards) are negative: those in a favorable situation slow down because they think they have done what is needed; those at a disadvantage, the opposite.



Complex systems have to be analysed in a systemic way

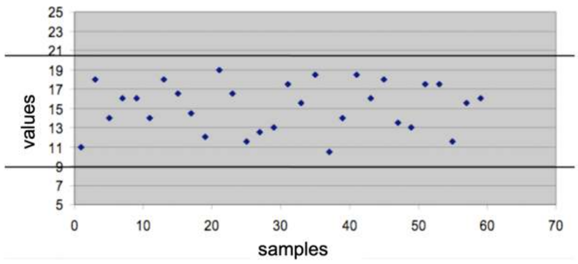
All systems have an output variability (especially in services). Variability is not all the same:

- Special causes: special events that influence the behavior of the system
  - Make the system unstable
  - They are good occasions to identify waste and opportunities for improvement
  - May be the responsibility of the operators

- Common Causes: random events that are part of the normal operation of the system
  - They are the responsibility of management
  - The operating personnel must not set up measures to compensate them (it would only increase the variability)
  - They cover the behaviour of the system (and therefore make it difficult to improve it)

How to recognize one from the other:

1. Statistical control   Normal distribution
2. Limits of natural variability: +/- 3 σ



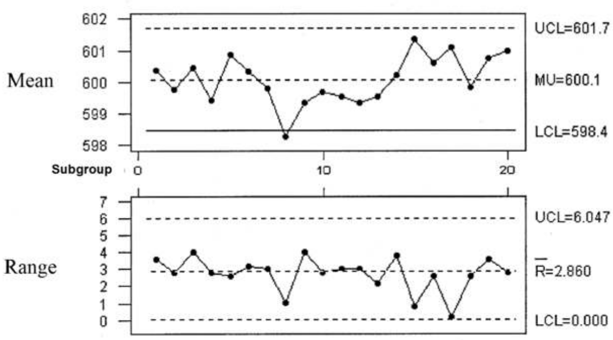
3. Sampling

If you take samples from a population:

- The sample mean is a proper estimator of the population mean
- The distribution of the averages of the samples tends to a normal distribution
- The variance of the averages of the samples is equal to the population variance divided by the sample size

→ The variance of the sample can be used to obtain more information

4. Statistical Process Control charts
   - If there are points outside the limits the system is not in control. It is not stable
   - If an item is outside the limits there are less than 0.15% chance that it is a random effect -> worth investigating
   - Although the system is not in control, you can identify the limits and they show the points out of control.



highest - lowest = R

Control limits - **X bar chart**

X bar chart and R

X = mean of sample

R = range of sample (max – min)

UCLx=X+A2 *R

LCLx=X-A2 *R

UCLr = D4 * R

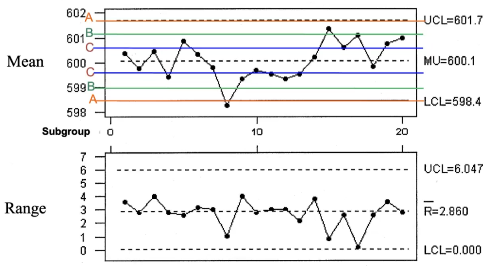
LCLr = D3 * R

Sample size	Mean factor A2	Upper Range D4	Lowe Range D3
2	1.880	3.268	0
3	1.023	2.574	0
4	0.729	2.282	0
5	0.577	2.114	0
6	0.483	2.004	0
7	0.419	1.924	0.076
8	0.373	1.864	0.136
9	0.337	1.816	0.184
10	0.308	1.777	0.223

1. **Special causes: → Zone A -> ± 3σ**
2. **Make the system unstable. → Zone B -> ± 2σ**
3. **Common Causes → Zone C → within ± 1σ**

This type of chart is useful in these conditions:

- The sample size is relatively small (say, n ≤ 10; for larger sample sizes◊x-bar and s charts are typically used)
- The sample size is constant
- Humans performing calculations for the chart



Processes in statistical control

1. 9 point in a row in zone C or beyond (all above or below central line)
2. 6 points in a row steadely increasing or decreasing
3. 14 point in a row alternating up and down
4. 2 out of 3 points in a row in zone A or beyond
5. 4 out of 5 points in a row in zone B or beyond
6. 15 points in a row in zone C (above and below central line)
7. 8 points in a row not folling in zone C (all are in zone B or A on both side of central line)

In general look for patterns

- X outside the limits are evidence of a general change affecting the sistem
- R outside the limits are evidence that the uniformity of the process has changed. Typical causes are change in personnel, increased variability of material, attempts to compensate natural variability

Responsabilty

- If a system is in control, the responsibility for the improvement lies with the management (as well as the responsibility for the quality in output)
- If the system is in control it does not make sense to give improvement targets to the operators, if they cannot act on the method
- If operators have improvement objectives, without being able to act on the method: they either play with the numbers, or they try to intervene and compensate for the variability, causing an increase of it.
- 95% of the ~~errors~~ defects depend on the system, only 5% is caused by the operators
- 95% of managers' attention is towards operators (and their management), and only 5% to the management of the system

Managers usually manage workload, sometime people, but rarely manage the system. Focus on measures rather than methods is a primary source of stress. Make the end-to-end performance visible to the operators and give them the chance to change the system

▪A process with lower variability allows better forecasts on its performance (it is more predictable)

▪A process with lower variability makes easier to understand the problems

▪A process with lower variability is easier to understand and to manage

Focus on measures rather than methods is a primary source of stress

Make the end-to-end performance visible to the operators and give them the chance to change the system

Measurement systems

▪Measures should help in understanding and in improving performances

▪Measures should relate to purpose

▪Measures must be integrated with work

Failure demand vs. Value demand

is in the file 9.0 pdf

In a service systems variations have to be managed at the front office

The higher the variability, the higher the competences and authonomy of the front office

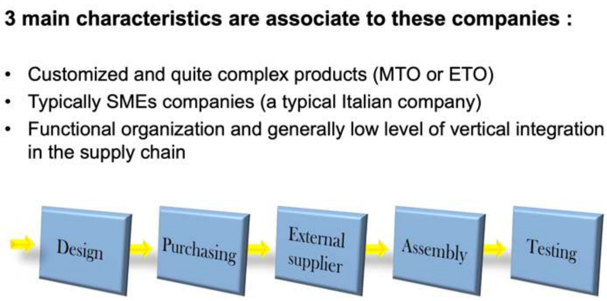
A command and control system is highly inefficient (an often ineffective) in managing service variations

24. LEAN PRODUCTION IN NON-REPETITIVE COMPANIES

Why were those companies neglected at the beginning by the lean approach?
Because speaking about **waste**, we speak about **repeat activities** that you do many times per day and then if you remove that wasting activities you will save, but if we have a low level of repeatability this could be difficult to do. However, we have the same problem that affect the repetitive company.

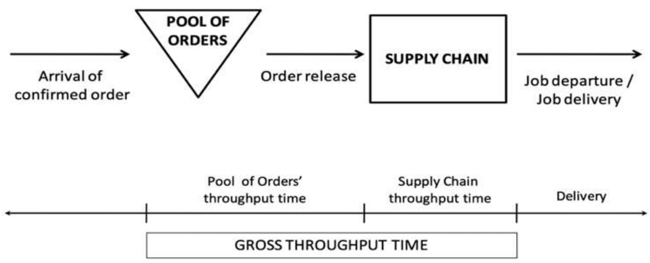
Characteristic of low repetitive companies:

- Low volume of production
- They produce customized and usually very complex products (ETO companies)
- They are SME (small medium enterprise)
- Compared to other companies, they are very low vertically integrated
- In order to deliver the product you have to do both the office activities and shop floor activities.
- The uncertainty and the variability are very high
- Long queues: in the offices, because many projects; in the factory, because you have many machines lying in the shop floor
- Local efficiency logic due to the functional organization
- You need huge efforts to coordinate the entire system, you have people spending a lot of time → in many cases there are things that are not working properly.
- High frequency of rescheduling.
- Huge stocks
- Every order is urgent
- There is no coherence between stages planning → an order is blocked for components missing.



The non-repetitive companies production planning and control framework:

- Pre-shop pool decouples the company from market variability.
- Pre-shop pool contains all the confirmed orders that can be released to the supply chain.
- Usually the release of orders is carried out with a load limiting logic (upper bound) with a very high upper bound.

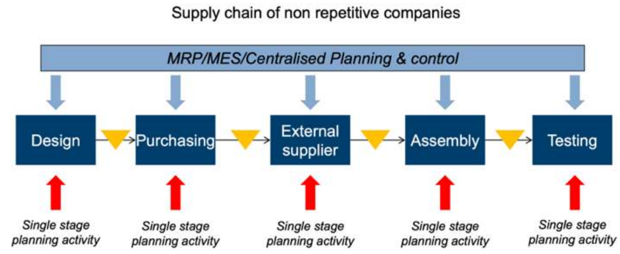


Lean can also help in this situation. How?
We will exploit the **PULL principle** and the **flow principles**.
Pre-shop pull is a buffer. If we are in a **MTO/ETO** we have no stock and we use it to absorb variability. In this buffer we put the customer order. When the order is confirmed, it enters not in the system bu7t in the buffer. Then there will be a person that usually belong to the supply chain management that bring the order from a buffer to the first stage of the supply chain. This person works as valve that is usually closed. It releases the order in a control way. Then the order passes the stages of the supply chain and then it is delivered to the customer.

Why should we have benefits?

- We can use aggregation rules that will lead us to pursue economies of scale.
- We have the pushing of the order by the sales manager with the variability of the market, and now we can eliminate this.

This has a big consequence on the system:
CONS:
Customers need to wait more. The gross throughput time is the sum of the pool of orders' throughput time and the supply chain throughput time. How usually companies operate? All the companies have a pre-shop pull. At least logically speaking.
What change is the logic of opening-closing of the valve → How do you decide to release the orders?
The decision is given not by the sequence, is related on opening or closing it.



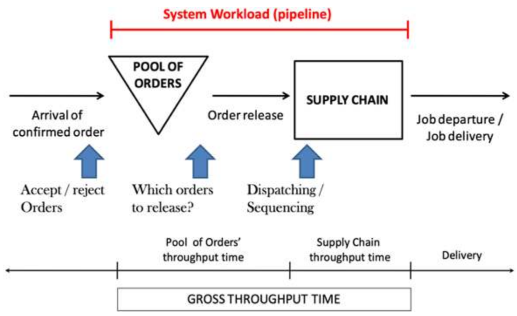
- Long Lead times
- Variable lead times
- High coordination efforts and low integration

Guidelines for achieving Lean in non-repetitive companies:

- Protect the system from external variability with a buffer upstream the first stage (in this case Design).
- Collect orders' processing times (if the precise information is not available, use classes of times (e.g. low/medium-low/medium/medium-high/high).
- Plan the flow in a single point (1 single actor plans the system).
- Level production by releasing constant amounts of workload for the different stages: Balance the Flow.
- Regulate the release of orders depending on the rhythm of the activities in the supply chain in order to maintain the workload level along the supply chain stable (pull production).
- Maintain the flow between the stages of the supply chain. In order to maintain the sequence of orders, work to remove the causes of blocks and the causes of variation in the sequencing of the orders (e.g. set-ups).

The idea is: you have a supply chain, you decide to open or close basing your decision on looking at how much workload is present in the supply chain. If the workload exceed a certain limit you should stop to release more work. The valve is always open because the underline limit of pre-shop pull has the limit that is unlimited.

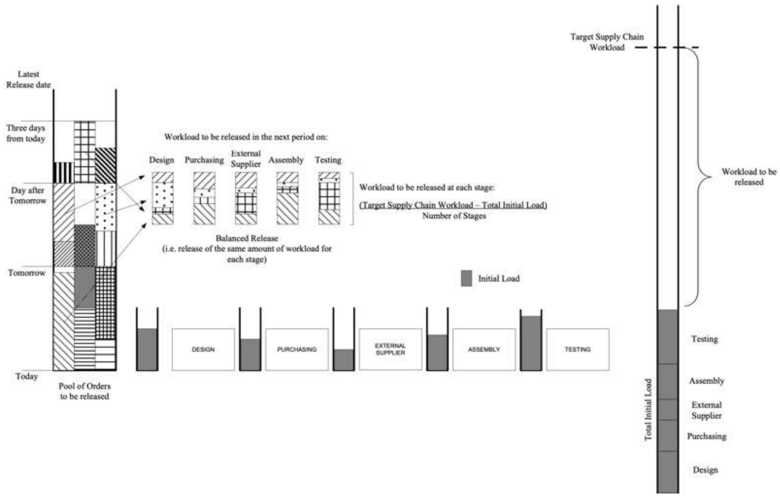
Lean approach can give a huge benefits transforming this current state map in the future state map. Pre-shop pool is the tool that create the level and then all stages can be controlled by the control limited buffers.



We have a system that is more stable because the working time is more under control. You can plan the supply chain in a single point so you decrease a lot the coordination effort. You can make a release which is coherent with the rhythm that you want that the supply chain respect. You can make all the stages working at the same rhythm. You create stability and because of this you have a system with standard levels.
This approach can be applied in linear and more complex systems, as job shops.

What is the impact of introducing the workload control? Creating the queue, that means creating waiting time for the order, that means an increasing of the gross throughput time.
We have benefits because the increase of the waiting time in. the pull of order has a decrease of supply chain thruput time.

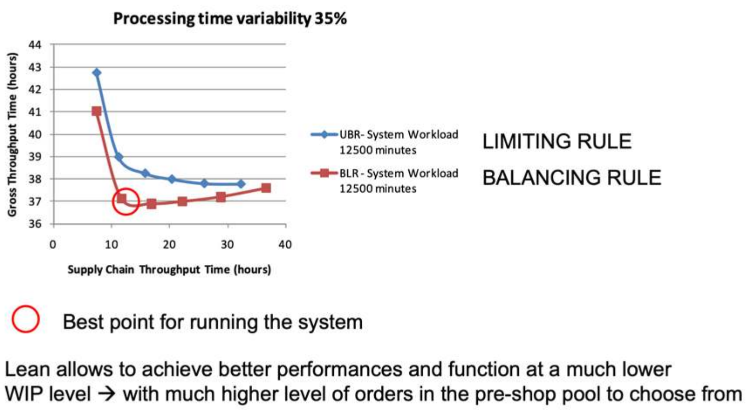
EXAMPLE:
We have the workload of the supply chain. Each stage has a certain queue, then we have orders, this entire pipeline is for us a measure of how much time we are going to take in order to satisfy the customer. Compared to the push system, we have that the supply chain is more stable, so the flow exiting from the supply chain is stable. So if the exit flow is stable, the throughput time is stable. The predictability of the throughput time of the respost time of the customer is stable and under control.



Compared to the traditional companies:
- it is more stable because you don't go over the limits,
- higher predictability, because the flow is constant that means facilitation in the computation of the throughout time of the single order,
- you have under control what you have in the queue.

24.1. IMPACT ON PERFORMANCES

There is a trade off:
There is a waiting time that we add to the order. **The more the waiting time we put here, the lower the time we will have here**.
The division is given by the limit or norm.
The blue line represents the results exploiting workload control. According to the different levels of norm, you can have different values.
How changing the norm at different level, you change the trade off about the gross and supply chain throughput times. If you put the norm very high, you have that he supplies chain TT (throughput time) is very high and the gross TT is at the minimum and vice versa.



Making an increase of 2 days of the entire gross TT can give you a decrease of the supply chain TT from 34 to 18. If we translate it means half of stocks. Worsening 3% the gross TT for having more that 50 % of saving of space, stocks and so on. It Is a lever that can have a very huge impact.

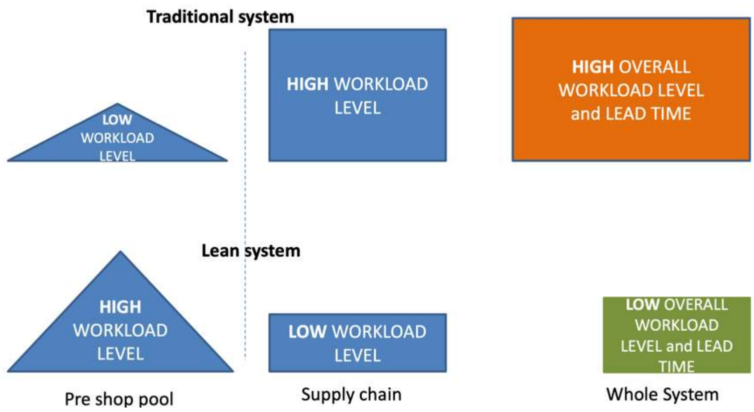
The first point blue is a very strong limit: if you put a limit very low you have a huge waiting of the orders here, and increase of the waiting of the order is not balanced by the reduction of the supply chain TT so you have a small improve of supply chain TT but a huge increase of the entire gross TT. In that case doesn't make sense if you exaggerate outing the limit too low.

The queue is useful to absorb the variability of the single process. If you reduce the queue you reduce the supply chain TT. If you reduce too much the queue, you don't have enough queue to absorb the inside variability of the single stages, so you will be negative impacted by that.

Lean companies are much more on the left than the traditional companies. With traditional we have the unlimited norm or close to it. From the basics of workload control we have the development of many different rules that could be more complex than just a simple limit rule. The simple limit rule is if the sum of the stages exceeds the norm you stop and the dispatching rule, so what do you take from the queue is a simple FIFO line.

Another example:
We have different stages. We have the pull of orders; we release the orders in FIFO logic. Each order has a certain impact, in terms of workload, on the different stations.

Balancing rule: we have the limit on the same norm but the choice of which item we release here is not done by FIFO logic, you pick up the order that optimize the fact that you have a balance on the workload of each station. The workload is more similar on all the stages and this is why balancing rule has chosen it. So it means that you reduce the probability to have a stage that is stacking.
Because of these, you reduce the efficiencies. For the balancing, there is usually a part of the queue where you can find the optimum, there are a lot of rules and some of them are under development.
The FIFO logic has the strength that it is simple to be followed by the people.
This is the sum up.



Lean system is more linear. Fast respost time, higher efficiency, higher stability.
Lean managerial framework is particularly suitable for non-repetitive companies.

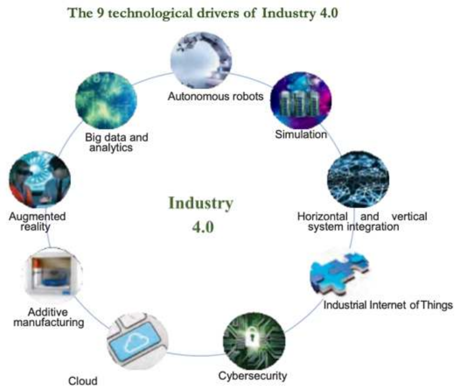
- Focus on the flow and on considering the entire supply chain as a unique system to be optimized at a global level – a flow shop in this case –. Balance the flow!
- Focus on maintaining the flow among the different stages and on avoiding the causes of blocks and the causes of variation.
- Managers in Lean companies can exploit the System Workload (in particular the extra-amount of non-released orders) to choose the best mix of performances for their company.
- System Workload in Lean non-repetitive companies can be seen at some extent as a lever, not only a mere consequence of the variability as it is in the Traditional non-repetitive companies.

25. INDUSTRY 4.0 & LEAN

The **first industrial revolution** used water and steam power to mechanise production. The **second** used electrical power to enable mass production. The **third** used electronics and information technology to automate production. Now a **fourth industrial revolution** is building on the third, the *digital revolution*, which has been taking place since the middle of the last century. It is characterised by a fusion of technologies that blurs the boundaries between the physical, digital and biological spheres.

What does it mean?

Industry 4.0 consists in the industrial application of 9 main technological drivers. None of the 9 drivers is new for manufacturing world. Each of these drivers carry significant benefits both in terms of productivity and of profitability. Greater benefits can be achieved using them in an integrated way.



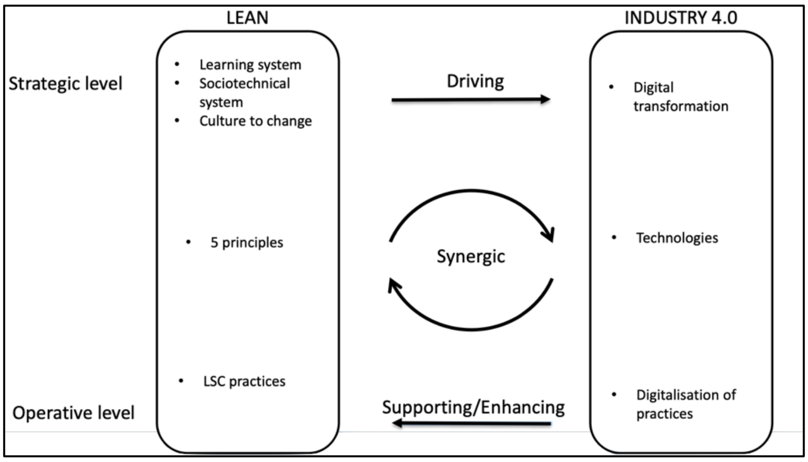
Fil rouge of Industry 4.0 is represented by **digital availability of information**, as the main and fundamental enabling factor of the fourth industrial revolution. The four clusters can be seen as main directions that development should follow

The interrelation between Industry 4.0 and Lean Production

→ Lean acts as enabling factor for successful digital transformation

Lean principles are vital to successfully implement Industry 4.0 related projects.

- Identify Value
- Map the Value Stream Create Flow
- Establish Pull
- Seek Perfection



Industry 4.0: Opportunity and development for LEAN

<u>Technological facilitation of Lean tools</u>

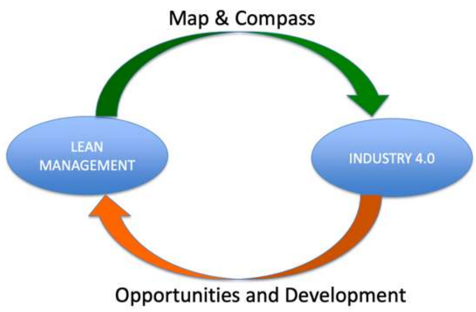
- e-KANBAN
- Digital adjustments of machines: quicker setup
- E-spaghetti chart
- E-Value Stream Mapping
- Human-Machine Interface: easier work on technological different machines
- Sensors, Big Data, Machine learning: Predictive Maitenance

Lean as a guide for Industry 4.0 – Map & Compass: Organization, Processes, People
Rely only on technological change brings limited benefits:
→ Transformation must involve technological aspects, but also Organization, Processes and People (OPP).

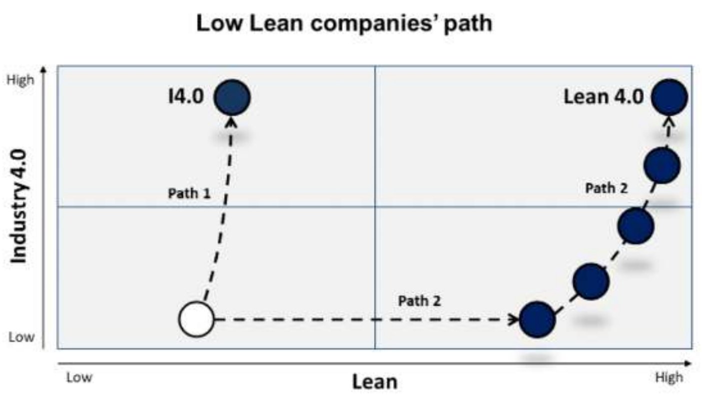
- Industry 4.0 is an **opportunity to change OPP**
- LEAN is the approach to lead the change

Overall view, but gradual changing :

- Fragmented implementations bring limited advantages;
- Learning is the key element: Industry 4.0, as LEAN, involve a huge number of people. The effect is unpredictable and difficulties are many and unkown.
- Pilot project and then a roll-out is the best way to success. Scheduling technolgical implementation, alongside with the development of competences and attitude.



How Lean leads Industry 4.0 implementation



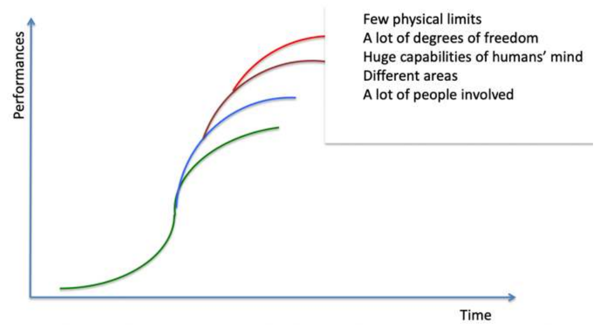
Drivers path-1:

- Productivity
- Production
  Capacity Increase

Drivers path-2:

- Flexibility
- Flow
  Improvement
- Pull Production

The heart of **LEAN** is **Learning** and then improving or innovating
With Industry 4.0 tech....
.... Greater opportunities for diagnosis
.... Quicker identification of problems
.... Quicker identification of root causes
..... Greater interventions' areas



Lean/TPS invests in people → *Thinking People System*
Most companies invest very little in training hours dedicated to their human resources (and the main reasons are not related to financial issues)

Technology **complements people**, **support them**, **help them.** It does not substitute them, nor guide them.
Technology is a tool. **Value is in people**.

"Someday machines will be able to solve all problems, but they will not be able to pose one"
(Alberto Einstein)

26. STARM INDUSTRIES EXERCISE

STARM Industries produces several components for tractors. This case concerns one product family- Steering Arms- which are produced in many configurations. STARM's customers for this product family are both original-equipment tractor builders and the aftermarket repair business. Because of the wide variety of product configurations and the fact that customer configuration requirements vary from order to order, steering arms are a "make-to-order" business. It currently takes a customer order 27 days to get through STARM's production processes. This long lead time and a significant order backlog have prompted STARM to quote a 60-day lead time to customers. However, STARM's customers cannot accurately predict their size requirements more than 2 weeks out, and thus they make adjustments to their orders 2 weeks before shipment. These order adjustments lead to order expediting on the shop floor at STARM. Although STARM Production Control releases customer orders to production roughly in the order that they are received, orders are batched by product configuration on the shop floor to reduce the number of time-consuming changeovers. This also creates a need for order expediting.

<u>The Product</u>

- A steering arm is a metal rod with a forged fitting welded to each end.
- STARM's steering arms are available in 20 different lengths, 2 diameters, and with 3 different types of end fittings. (Each end of the steering arm can have a different fitting.) This means there are 240 different steering arm part numbers that STARM supplies.

<u>Customer Requirements</u>

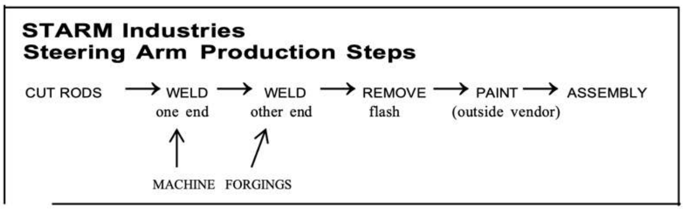
- 24,000 pieces per month.
- A customer order ranges from 25 to 200 pieces, with an average of 50 pieces.
- Customer requires 1 configuration per order.
- Corrugated-box packaging with up to 5 steering arms in a box.
- Several daily shipments per day by truck to various customers.
- Each customer's configuration requirements vary greatly from order to order.
- STARM requires orders to arrive 60 days before shipping date.
- Customers often adjust their size mix 2 weeks before the shipping date.

<u>Production Processes (see diagram)</u>

- STARM's processes for the steering arm product family involve cutting a metal rod followed by welding end fittings to the rod, deflash (machine removal of excess weldment), painting at an outside vendor, and subsequent assembly of the end fittings. The forged end-fitting sockets are also machined at STARM. Finished steering arms are staged and shipped to customers on a daily basis.
- Switching between rod lengths requires a 15 minute changeover at the cutting, welding, and deflash operations.
- Switching between rod diameters takes a 1 hour changeover at the cutting, welding, and deflash operations. The longer change-over for diameters is due mostly to an increased quality-control inspection requirement.
- Switching between the three types of forged end fittings takes a 2 hour changeover at the machining operation.
- Steel rods are supplied by Michigan Steel Co. The lead time for obtaining rods is 16 weeks. There are two shipments per month.
- Raw forgings for the end fittings are supplied by Indiana Castings. The lead time for obtaining forgings is 12 weeks. There are two shipments per month.
- STARM reaches high quality levels and hasn’t any scraps in the production processes.

Work Time

- 20 days in a month.
- Two shift operation in all production departments.
- 8 hours every shift, with overtime if necessary.
- Two 15-minute breaks during each shift.
- Unpaid lunch.



STARM Production Control Department

- Receives customer orders 60 days out and enters them to MRP.
- Generates one "shop order" per customer, which follows the order through the entire production process.
- Releases shop orders to production 6 weeks before shipment to accelerate MRP's procurement of rods and forgings.
- Issues daily "priority" list to production supervisors. Supervisors sequence shop orders through their departments according to this list.
- Receives customer size-changes 2 weeks before shipment and advises supervisors to expedite these orders.
- Issues daily shipping schedule to Shipping Department.

Process Information

1. **Cutting** (The saw cuts rods for many products, it dedicates 50% of time to this product family)

- Manual process with 1 operator.
- Cycle Time: 15 seconds.
- Changeover time: 15 minutes (for length) and 1 hour (for diameter).
- Reliability: 100%.
- Observed Inventory: 24000 uncut rods before the saw, 6000 cut rod.

2. **Machining of Forgings** (dedicated to this product family)

- Automatic machining process with one machine attendant
- Cycle Time: 20 seconds.
- Changeover time: 2 hours.
- Reliability: 100%.
- Observed Inventory: 20 days of raw forgings from the supplier and 4 days of machined forgings

3. **Welding Workstation I** (dedicated to this product family)

- This operation welds the first machined forging to the rod.
- Automatic process, with operator load external to machine cycle, automatic unloading too.
- Cycle Time: Operator = 10 seconds, Machine = 30 seconds.
- Changeover time: 15 minutes (for length) and 1 hour (for diameter).
- Reliability: 90%.
- Observed Inventory: 3600 welded arms.

4. **Welding Workstation II** (dedicated to this product family)

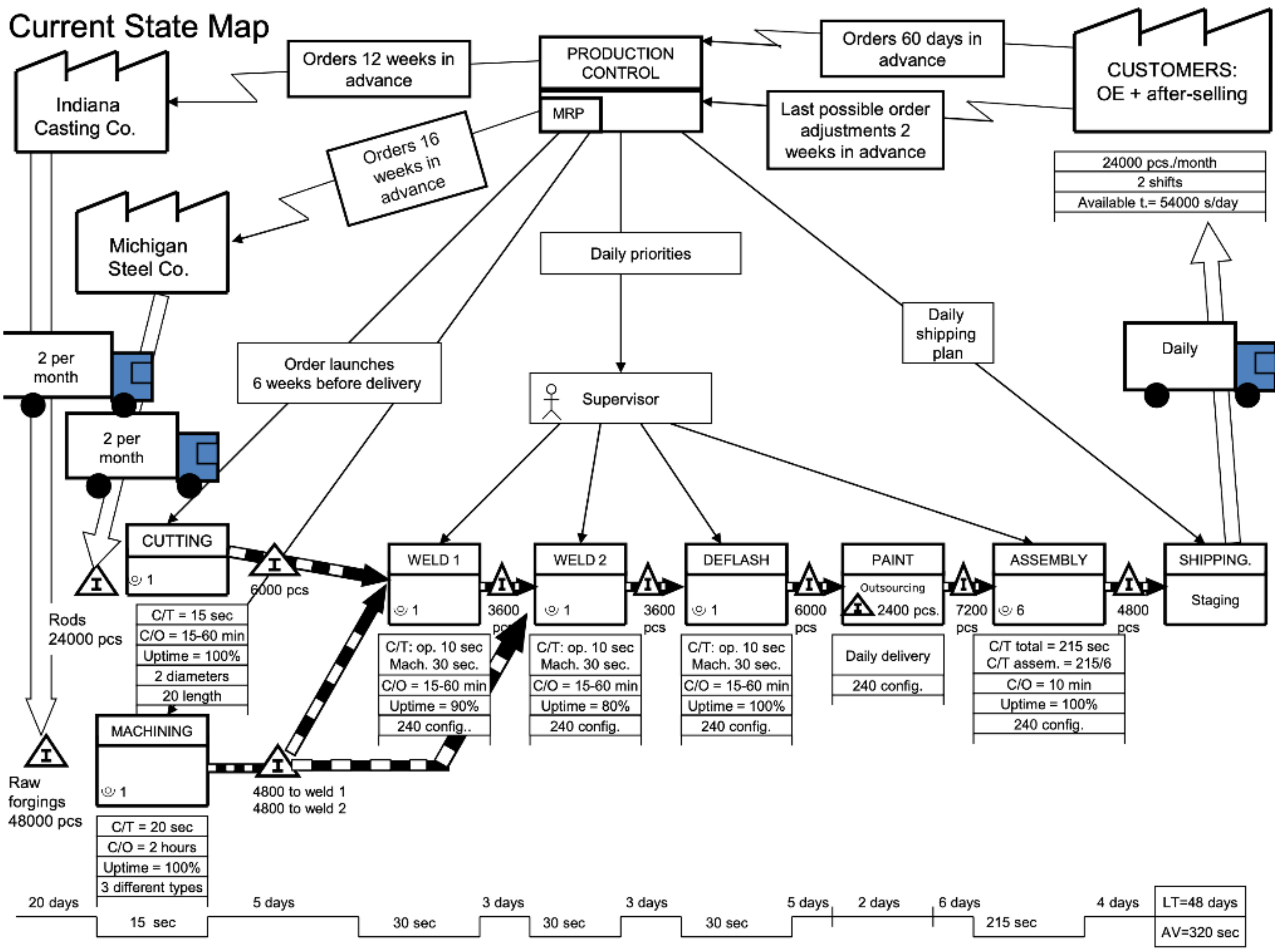
- This operation welds the second machined forging to the rod.
- Automatic process, with operator load external to machine cycle, automatic unloading too.
- Cycle Time: Operator = 10 seconds, Machine = 30 seconds.
- Changeover time: 15 minutes (for length) and 1 hour (for diameter).
- Reliability: 80%.
- Observed Inventory: 3600 welded arms.

5. **Deflash Workstation** (dedicated to this product family)
- Automatic process, with operator load external to machine cycle, automatic unloading too.
- Cycle Time: Operator = 10 seconds, machine = 30 seconds.
- Changeover time: 15 minutes (for length) and 1 hour (for diameter).
- Reliability: 100%.
- Observed Inventory: 6000 deflashed arms.

6. **Painting** (steering arms are shipped to an outside vendor for painting)
- Painting lead time = 2 days.
- One daily truck pickup of unpainted arms and drop-off painted arms.
- Observed Inventory: 2400 arms at the painter 7200 painted arms at STARM.

7. **End-fitting Assembly** (dedicated to this product family)
- Manual process with six operators.
- Total Work Time Per Piece: 215 seconds.
- Changeover time: 10-minute fixture swap.
- Reliability: 100%.
- Observed Finished-Goods Inventory in Warehouse: 4800 finished steering arms.

8. **Shipping Department**
- Removes parts from finished goods warehouse and stages them for truck shipment to customer.



In case of parallel processes: identify critical path and insert it in the timeline

In Starm S.p.A. critical path pass through rod cutting stage and cut rods buffer. Timeline has to consider 15 seconds of added value time and 5 day of not-added value time.
When a product is composed by 2 or more identical elements the added value time is the sum of times to process all the components.
In Starm S.p.A., machining added value time is equal to 40 seconds (2 forgings/product *20 seconds/forging).

To consider the critical paths:
Cutting: 24.000[pcs]/1.200 [pcs/day] + 15[s]/(15[h/day]*0,5*3.600[s/h])+6.000[pcs]/1.200[pcs/day] ≈ 25 days
Mechanical machining: 48.000[pcs]/2.400 [pcs/day] + 40[s]/(15[h/day]*3.600[s/h]) +
9.600[pcs]/2.400[pcs/day] ≈ 24 days
The critical path to consider is passing from the stage of cutting.

<u>Future state map</u>

1. **What is the takt time of the product family?**

Time available: T. opening system – scheduled stops

- scheduled stops:
  - YES scheduled maintenance, lunch (if stipulated in the contract), coffee breaks (stipulated in the contract), ...
  - NO set-up time, downtime, ... Customer request

Takt time = Time available / Customer request

Time available:

- 8 hours per turn
- 2 turns per day
- 2 breaks of 15 minutes for each turn

Time available = (8 – 2 * 15) h/turn * 2 shifts/day = 15 h/day

Customer request:

- 24.000 steering arms per month
- 20 working days per month

Customer request = 24.000 pcs/month / 20 days/month = 1200 pcs/day

Takt Time = (15 h/day *3600 sec/h) / 1200 pcs/day = 45 sec/pcs

2. **Produce for FGW or for shipping?**

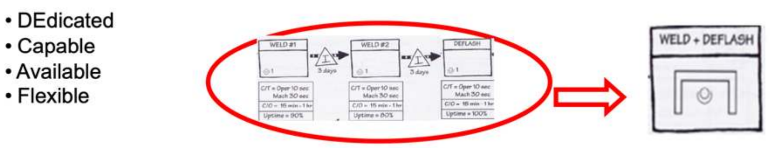
How to organize production?
Some drivers for your choice:

- time of delivery to the customer
- product features (good value, level obsolescence, product standardization level)
- demand predictability
- demand stability
- ....

Make to order - SHIPPING

3. **The continuous flow**

To support the decision of where to implement a continuous flow, think about DECAF:



To put continuous flow the final assembly it is necessary to put in line the operators.

- Benefits: The line configuration has many benefits, including greater efficiency (it is easier to manage) and a lower supply required (not always need material to supply 6 islands).
- Disadvantages: The line configuration, however, requires a larger amount of set-up. In particular, each operator on the line should run 24 production changes every day.

With the current set-up time the final assembly cannot be organize as a line. It is needed to act on the set-up time! At this stage each fitter assembles completely a finished product.
C/T for each operator is 215 sec. Average order: 50 pcs.
With an average order of 50 pcs, they are required 1200/50 = 24 set-up.
6 operators work completely finished products. For this reason, each operator must carry daily 24/6 = 4 set-up/operator.

The total time (set-up + processing) required daily for each operator is then amounted to:
4 set-up/operator* 10 + 1200/6 * 215/60 = 40 + 717= 757 min = 12,62 h
In the stage of final assembly, C/T is 215 seconds for the finished product. The Takt time is 45 seconds/finished product.

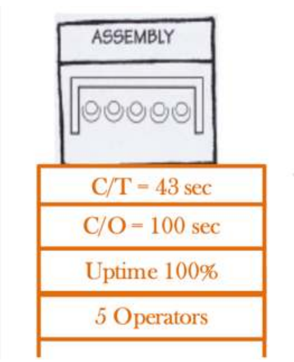
You may calculate the minimum number of operators required.
Content of work / Takt time = minimum number of workers
215 sec/pcs / 45 sec/pcs = 4,77 operators → **5 operators**

In the current situation, with 6 operators working in 6 islands assembly, you cannot work with 5 operators working in 5 islands.
In this situation, each worker should do 24/5 = 4.8 set-up on average. The total time required would be:
4,8 * 10 + 215*(1200/5)/60 = 48 min + 860 = 908 min = 15,13 h > Ta

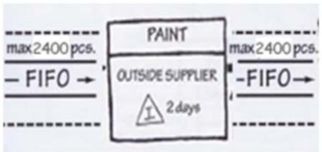
In the current situation, even putting continuous flow final assembly will be able to work with 5 operators.
Put in line operators means allocate to each operator a package of activities with a cycle time of less than but close to takt time.

The possibility of putting continuous flow the final assembly and the possibility to save an operator is in any case subject to the possibility to reduce the time to set- up.
We assume that activities are composed of 'packages' uniforms activities and then load each operator as follows:
215/5 = 43 sec

All five workers will perform 43 seconds of work on an individual piece.
In the situation described, what it is the minimum target of reducing the time to set- up for each station **assembly**?

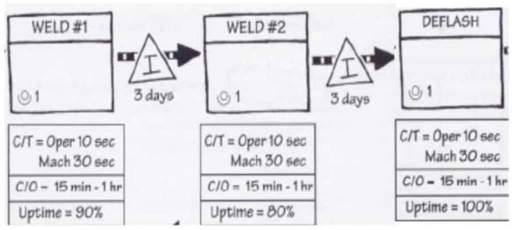


t. production + t. set-up ≤ t. available
43/1 * 1200 + 24 * C/O unitary ≤ 15 h * 3600 sec/h
52800 + 24 * C/O unitary ≤ 54000 sec
C/O unitary ≤ 100 sec → 1,67 min (total time ENTIRE LINE will stop for retooling)
In the case that it is possible to reduce the time to set-up for each station less than 1,67 min, then it will be possible to save one operator (6 to 5), which may go to conduct different tasks.
Supplier takes arms unpainted once a day and return them painted once a day.
In order to absorb the variability in the rates of withdrawal and deposit of the products in the two stages, and prevent blocking events of the stages, Starm must decouple the processes using FIFO queues. FIFO queues allow a constant supply to the stage downstream of the coating and the ability to stage upstream of the painting not hang uselessly.



The use of FIFO queues allows to avoid overproduction in cases of problems in the production process and allows a constant movement of the material!
In order to prevent events of starving and blocking of the stages, Starm must decouple the processes using FIFO queues dimensioned at least on a production day.

Caution when sizing:
1. The supplier may have some variability in time delivery of the goods.
2. Production volumes may fluctuate daily. It may sometimes not be producing 1,200 products a day.
Starm should slightly over-dimensioned FIFO queues, for example of about 2 days of production in order not to lose production and output.

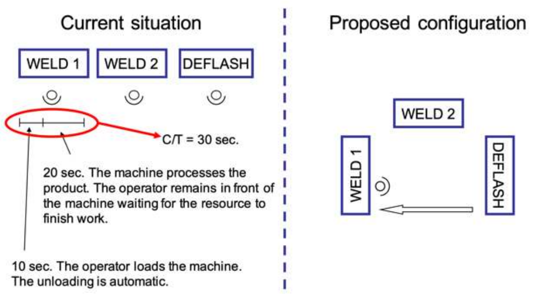


Takt Time: 45 sec/pcs
C/T: 30 sec/pcs

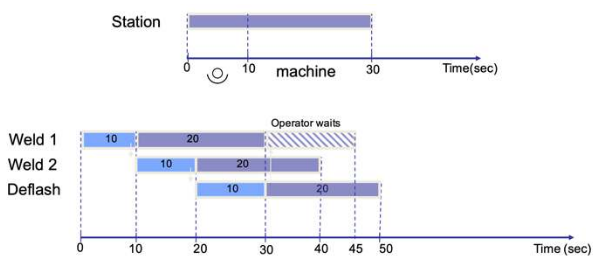
Freeing the operator from the single machine:
You can organize a cell in U and manage it to a single operator that moves between machines.
C/T Oper.: 10 sec/pcs for machine → 30 sec/pcs for 3 machines < TT
C/T machine: 30 sec/pcs < TT
TO MAKE FLOW

1. DEdicate: Yes, all three stages
2. Capable: C/T for all three < TT (30<45)
3. Available?
   - Connect stages also means linking the availability!
   - You need to check if the availability of the stages is a restriction to the possibility of making flow: (C/T) / A < TT?
   - 30 / (0,9*0,8) < 45 à the machine availability is not a problem.
4. Flexible?
   - You can work with an allotment of 50 average products? We work for shipping and must be able to be flexible enough to meet the customer's request. The client requires 24 different versions (1200/50) of product per day on 240 available in the catalog. In this cell it is held set-up only for the length and the diameter then the range which 'sees' the cell is 40 (and not 240), or in the cell there is set-up for welding different types of fittings.
   - So of the 24 variants requested, probably only a part depend on the diameter and length.

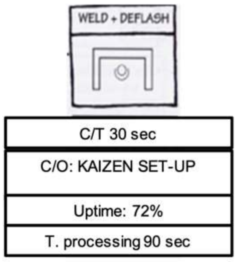
Example:
Product requested by client1: D1 L1 RG RP (diameter 1, 1 length, great fitting, fitting small)
Product requested by Client2: D1 L1 RG RM (diameter 1, length1, great fitting, coupling medium)
They are two different products, but in the cell is not expected to set-up to switch between them. So to make sure that the cell is flexible does not mean that you should take 24 set-ups a day.
For sizing we consider however the WORST CASE, in which 24 set - up a day in the cell are required.



In the current situation, without changing any parameters, it is possible to carry out 24 set-up every day (situation corresponding to work with batches of average size 50 products)?
Available time every day to do set-up:
t. available – t. production = 15 h/day – [30/0.72 * 1200]/3600 h/day = 1.11 h/day
In 1.11 hours it is not possible to carry 24 set-up! → Kaizen set-up
The flow weld 1 → Deflash



- You can save 2 resources (labor), but
- It is necessary to reduce activities not-value added (set-up)



T. C/O + C/T * dailydemand ≤ 15 h
T. C/O ≤ 15 h – 30/0,72 * 1200 / 3600 = 1,11 h/day
With the objective to work with batches of 50 products and considering a daily demand than 1200 products, the number of set-up required daily would be equal to 1200/50 = 24 set-up.
With this objective, the average time for the set-up on each machine should be equal to: 1,11 h * 60 min/h / 24 s-up = 2,78 min.
This means that each machine can be stopped to carry one turn production at most 2,78 minutes.
Having one single operator, this is the maximum time that he has available to retool the whole line. If you can do the set-up for the whole cell in less than 2,78 minutes, then just one operator is needed.

If you cannot reach this level, or you require the use of multiple operators, or you must do a number of set-up lower, or you have to increase the availability of the cell.
In the event that it were possible to retool the cell in less than 2,78 minutes, you can get:

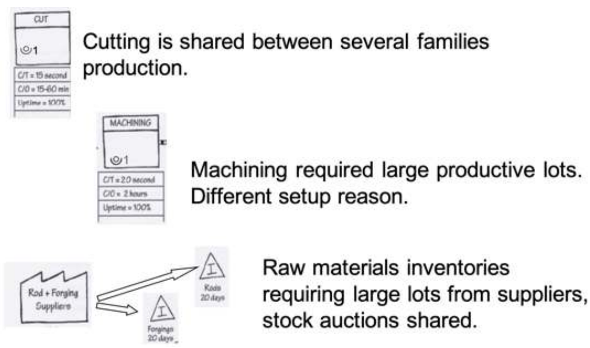
- A saving of 6 days in lead time
- A savings of 2 operators that can be assigned to other value-added activities

Cost of:

- Reduce activities non-value added

4. Where to insert a supermarket pull system?

- Large production lots
- Low machine availability
- Shared resource
- .....



CUTTING: The stage of cutting is shared between different families product.
All production flows starting from the stage of cutting, must be free to take what they want when they want, even when the stage cut is working on another product family.
The stage of cutting cannot therefore be put to flow with the downstream stages.

- At the stage of cutting, for the family of the steering arms, there are 20 different lengths and two different diameters.
- The machine for cutting works on the family of steering arms for 50% of its total available time.
- The total time of set-up required for re-assort the whole production range is equal to:

2 diameters * 60 min + 2 * 20 lengths * 15 min = 12 h
Each of the 20 lengths of the bar can in fact be of two dimensions of different diameters.

In the current situation, the EPE (Every Part Every) on the stage of cutting is:
C/O + (C/T)/A * # pdt required in the time of EPE ≤ time available in the range of EPE
12h + 15/1 s/pcs * 1200 pcs/day / 3600 s/h * X day ≤ 0,5 * 15 h/day + * X day
X ≥ 5 days → EPE = 5 days
The reduction of EPE allows increased flexibility of the machine, allowing a reduction of stock downstream necessary because of the ability to restock the range much faster.
The reduction of EPE is pursued by the goal of reducing the time to set-up.
A gradual improvement in the time to set-up allows a constant increase in flexibility and a reduction in stocks stored downstream.

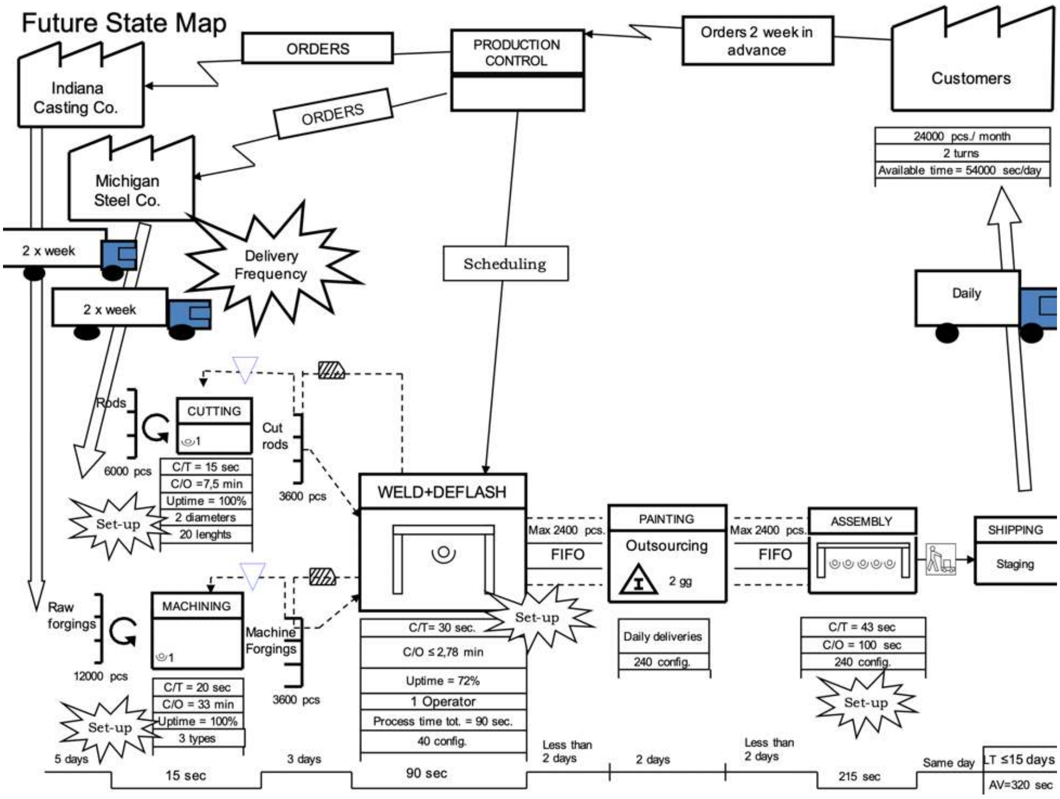
MACHINING: 2 hours spent for each set-up, the stage of machining requires very large productive lots. This production stage works three different forged fittings.
With current technical parameters, the stadium has to do lots.
EPE current:
(C/T/A) * daily demand * EPE + C/O * number variants ≤ T. available * EPE
20/1 * 2 * 1200 * EPE + 120 * 60 * 3 ≤ 15 * 3600 * EPE
48000 * EPE + 21600 ≤ 54000 * EPE
EPE ≥ 3,6 days
In the current situation, Machining requires batches of about 4 days, that is necessary to aggregate average demand of 4 days of production for each product code.
What said for cutting stage, it can be said also for machining stage.
Reducing the time of set-up will increase the flexibility and reduce stock levels.
Stock: 20 days.
Shipping frequency supplier: 2 times a month.
Acting on the supplier can increase the frequency of delivery, coming (for example) to weekly deliveries.
The stock could be organized with supermarket and controlled by Kanban that will be transferred to the control of production. This last will send orders to suppliers at fixed time intervals.

5. **In what single point will be the production schedule?**

The weld cell+ deflash cell needs to work with customer orders as:
- At this point the customer can choose one of 240 different variations of the finished product.
- The mover person picking of the material to supermarket must have an indication of what is necessary to take.

SIZING THE SUPERMARKET:
The stage of cut is shared among the various families productive and therefore must have a supermarket in downstream.
At the stage of cutting, for the family of the steering arms, there are 20 different lengths and two different diameters. The machine works on the family of steering arms for 50% of its total available time.
Now the EPE on the machine is 5 days.
To respect the time available is therefore necessary to parcel out five days of production on the stage of cutting.
To reduce the supply of rods cut and increase the flexibility you need to reduce time to C / O.
A first goal of improvement might be to obtain an EPE of 2 days, given the high impact of the set-up of the current situation.
Daily, the time allowed the set-up is:
C/O + (C/T)/A * daily demand * EPE ≤ (50% * 15 h/day) * EPE
C/O ≤ (50% * 15 h/day) * 2 – 15 * 1200 * 2/ 3600 = 5 h
If the goal is EPE = 2 days, the time of C/O target for product would be:
5 h * 60 min / 40 set-up = 7,5 min
In 7,5 minutes, the operator must be able to prepare the machine for cut any one of the 40 versions of the rods.
To size the supermarket, you need to think about the stocks cycle and safety stock.
In the current situation, Starm stock with EPE = 5 days 6000 products (1200 * 5) for 40 variants of rods.
With EPE = 2 days, you will have two days of stocks (2400 products) as stock cycle.
In addition to this stock of the cycle there will be a stock of safety, for example equal to 1200 products (half EPE).

The stage cutting will be managed by Kanban signal. Withdrawals from the pacemaker will accumulate Kanban in a special table When the accumulation of Kanban will reach a threshold level, the stage of cutting will produce the complete lot of products taken. The threshold is reached in the number of Kanban taken will activate a signal, which requires the cut stage of starting the production.



<u>The objective on the reduction of EPE improves considerably the size of the store of rods cut and activates a process of continuous improvement that aims to raise the regularity of production on the stage of cutting.</u>

1. What said for the cutting stage can be replicated for machining stage.
2. Reducing the time of set-up will increase the flexibility and reduce stock levels downstream.
3. I assume its objective of initial EPE = 1 day for machining.

	Current State	Future State	Variation
Lead time	48 days	<15 days	-69%
# Operators	11	8	- 27%

27. LEAN INNOVATION

According to Forbes, 90% of startups fail → Nowadays, more than ever, we need more innovation and more entrepreneurial approache by managers. However, innovators and entrepreneurs are not common, and we cannot build them, right? Actually, there is a way to foster innovation and, most importantly, become a successful Innovator: **LEAN INNOVATION**
Derived from Lean Management, scientific testing and fast learning

Why innovations fail?.. Look at what goes on in the world
The pace of introducing new technologies has become exponential resulting in an exponential change in customers behavior and expectations. What used to be an order winner, is now an order qualifier. We are entering in the world of uncertainty...

Look at what companies should do... **They should innovate to adapt**
It is not enough to be intelligent and rigorous planner to be the strongest in the market. One should be able to respond to change and adapt to it by stepping into it and not be a third-party observer. True Innovators are innovators who think outside the box, they distance themselves from the old-fashion thinking of "Just-do-it" following what the boss says or the investor asks. And test their hypothesis

Webster dictionary defines **scientific method** as a set of "principles and procedures for the systematic pursuit of knowledge involving the recognition and formulation of a problem, the collection of data through observation and experiment, and the formulation and testing of hypotheses".
A scientific method helps :

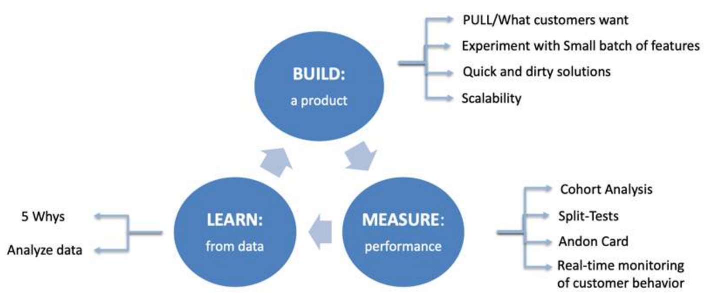
- Tuning down the bias and the prejudgment
- Test the hypothesis in real-time
- Get validated feedback

Can we apply this when we innovate (developing new products, new processes, new ventures)?...
...YES, WE CAN.
Lean Innovation is a scientific based methodology that bursts innovation rate and capabilities of individuals and organisations

Lean Innovation as a **Process**

1. **Set hypothesis**
   Just like any scientific experiment, you need a hypothesis, a vision, upon which you base your predictions about your product value (value hypothesis) and customers (growth hypothesis). In a scientific experimentation, a theory guides us, whereas in a startup, the startup's vision guides the experimentation. An innovator's objective is always to "discover how to build a sustainable business around that vision".

2. **Test hypothesis**
   Start the first BML loop with early adopters and an MVP. This is when you start testing your hypothesis, by identifying your early adopters through a market research, this is your target customers.
   What is an MVP?: It is that version of the product ,with minimum features, that "enables a full turn of the Build-Measure-Learn loop with a minimum amount of effort" ...Think Lean. It is an experiment that requires the joint efforts of all backgrounds. It is not as a result of the usual rigid sequence of product development.



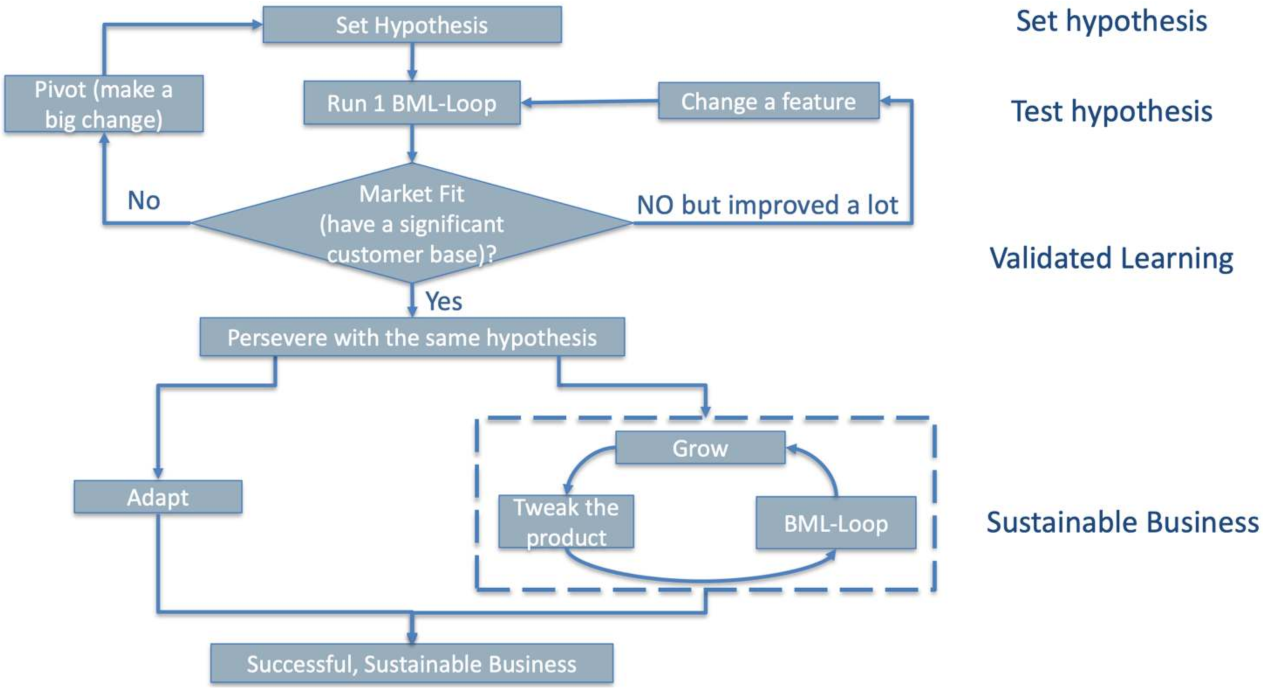
3. **Validated Learning**

- **React rather than forecast**: instead of spending resources (manpower, money and time) on forecasting what the product could be with all its specifications, you release an MVP to the market that allows you to "pull" the actual product requested
- **Get rid of the old-experience-based decision**: entering the market confident of yourself based on old data is far less effective than entering it based on real-time data from your customers
- **From a seat on the bench to a seat in the game**: The world is innovating at a fast pace, to keep up with it you simply cannot wait, not to introduce a new product, not to update your product (imagine facebook updates every year instead of several times a year).
- **Sense and response**: From the spoken need to the observed behavior of the customer. Facebook displays products related posts based on clicks done few days ago or even few minutes before you even access your account.
- **Market fit**: If you acquire a large market, it means you achieved product /market fit. Assessing whether the innovation is getting closer to the product/market fit, is done as it tunes its engine by evaluating each round through the Build-Measure-Learn feedback loop using innovation accounting. It is about the direction and degree of progress rather than raw numbers or vanity metrics.
- **Change a feature**: In this phase, based on the learning and the fact that there shall be a market fit if you improve your MVP, you experiment with a feature. You can add or remove a feature; either way you should start the BML loop again. It is no longer an MVP, it is a product.
  The faster you iterate the faster you learn, the faster you get to your market fit and eventually have a sustainable business.
- **The Pivot** is a structured course correction designed to test a new fundamental hypothesis about the product, strategy, and engine of growth. "A pivot requires that we keep one foot rooted in what we've learned so far, while making a fundamental change in strategy in order to seek even greater validated learning". There are several types of pivots including.
- When you have a market fit, you are confident that your hypothesis is validated, and you can now continue within the same hypothesis by tweaking your product, that is you **Persevere** in the same direction

4. Sustainable Business
   - It is essential **to build an adaptive organization**, one that automatically adjusts its process and performance to current conditions, by training new employees and be attentive to required investments. To build an adaptive organization, Five Whys is to be used with a proportional investment to be made for each of the five levels of the hierarchy, that is proportional to the size of the symptom the company is facing. Large investments can be made in prevention only when dealing with large problems. The Five Whys ties the rate of progress to learning, not just execution. Five Whys should be applied whenever the team faces any kind of failure, including technical faults, failures to achieve business results, or unexpected changes in customer behavior
   - Sustainable growth is characterized by one simple rule: New customers come from the actions of past customers. You are no longer adding major features, you are "tweaking your product" through faster BML loops There exists three types of engine growth:
     - **The sticky engine of growth**: The speed of growth depends on the rate of compounding.
       Rate of compounding = Natural growth rate (rate of new customer acquisition) – Churn rate (the fraction of customers in any period who fail to remain engaged with the company's product)
       Focus on RETENTION rather than marketing to get new customers.

- **The Viral Engine of Growth**: The speed of growth is determined by the Use Viral coefficient, that is how many new customers will use a product as a consequence of each new customer who signs up (person-to-person effect). In the viral engine of growth, monetary exchange does not drive new growth (Facebook). Focus on activities AFFECTING customer behavior

- **Paid engine growth**: its feedback loop is: Each customer pays a certain amount of money for the product over his or her "lifetime" as a customer. Once variable costs are deducted, this usually is called the customer lifetime value (LTV). This revenue can be invested in growth by buying advertising or customers.



28. LEAN AND SUSTAINABILITY

A key lever for improving sustainability is to fight waste:

- 50% of water captured in Italy is wasted along distribution and does not reach the end users
- 30% of food produced in the world is wasted before reaching the end user, or wasted by the end user
- G20-If we reduced energy consumption by 15% we could reduce fossile produced energy by 30%

A framework for modelling energy consumption within manufacturing systems

- **Theoretical Energy (TE):** refers to the minimum energy required to achieve (physically, or chemically) the final result. For example the energy required to melt a specific amount of metal during casting, or removing a specific amount of material during machining operations, or energy to make a chemical reaction happen;

- **Auxiliary Energy (AE):** energy required by the supporting activities and auxiliary equipment of the process, for example generation of vacuum for sand casting, or pumping of coolant for machining. The AE also includes non-productive energy as for example the heat generated by the electric motors of a machine tool and, in general, all energy that the machine is not converting into the result of the process. And the one used for the machine tool start-up, set-up, stand by and cleaning;

- **Indirect Energy (IE):** energy consumed to ensure the correct workplace conditions for the productive processes such as lighting, heating and ventilation;

- **Value Adding Energy (VAE)** = Theoretical Energy Energy used for doing all the activities that create value for customers;

- **Non Value Adding Energy (NVAE)** = Auxiliary Energy + Indirect Energy
  Energy consumption related to the Non Value Adding Activities.

Energy Value Ratio = Value Adding Energy [TE] / Total Energy Used [TE+AE+IE]

The **T.D.U. framework** is based on an important assumption: the utilities used inside a manufacturing plant pass through three phases **Transformation**, **Distribution** and **Utilization**.