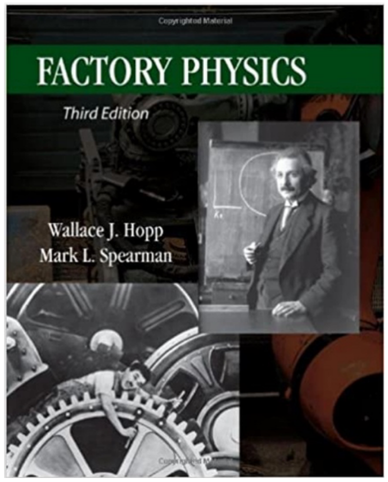
Bottlenecks, part 1

Almost 15 years ago I had the opportunity to attend a 4 day seminar with the authors of the well-known book "Factory Physics" LINK



In the opening session we talked about what is limiting factory performance and sure enough bottlenecks came up. The question was asked, what can be done to improve a bottleneck. After a lively discussion between all attendees about what they have done or what they think should be done, Dr. Mark Spearman stated:

"… I propose you walk on the factory floor and look at the tool or tool group and see if it is indeed running (at full speed and efficiency) …"

I had a pretty big "aha!" moment and I remember this, like it was yesterday. But this proposal comes with another interesting challenge:

How do we know what is the factory bottleneck???

I think to answer this question correctly is the foundation for a lot of things. In its simplest form, the correct answer would lead the folks who want to see the bottleneck on the floor to walk to the right tool/tool group. Obviously, there is much more connected to that, for example:

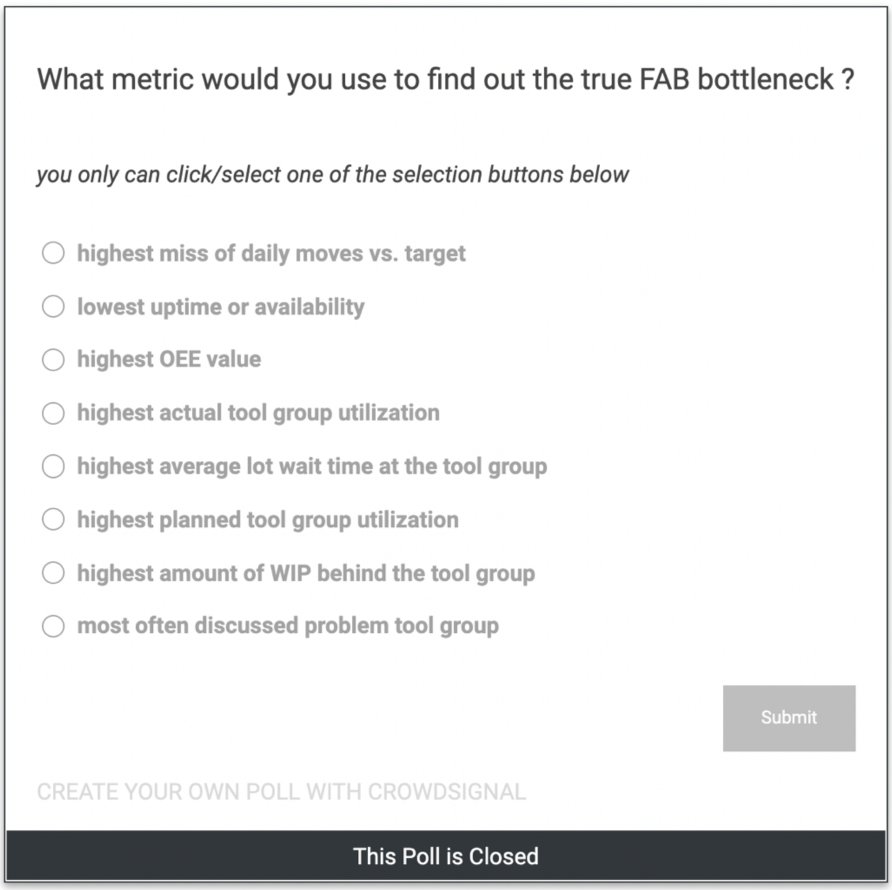
- where to spend resources for improvement activities
- if the bottleneck capacity is used to define the overall FAB capacity, it would be great, if the correct tool/tool group was identified
- where to spend capital to buy another tool

How do we find out, what is the factory bottleneck tool group? One obvious answer is let's look into data – what data – and how do we know it is indeed the bottleneck. The answer becomes quickly" … it depends …"

It depends on what is the definition metric and I have seen a few of them so far:

- highest tool utilization as per capacity planning numbers
- highest tool utilization as per actual numbers (daily, last week, 4 weeks ?)
- highest amount of WIP behind tool group
- highest average lot wait time at the tool group
- highest miss of daily moves vs. target
- frequency / intensity a tool group is discussed in morning meeting as a "problem kid"
- lowest tool group uptime (or availability)
- highest OEE value

I'm pretty sure all these metrics have some value, if used in the right context. I do have my own opinion, what I would select as the key metric, to declare the FAB bottleneck, but I really like to get some discussion going here, therefore I like to run a little poll, to see what the majority would select as the key metric:



What metric would you use to find out the true FAB bottleneck ?

you only can click/select one of the selection buttons below

- highest miss of daily moves vs. target
- lowest uptime or availability
- highest OEE value
- highest actual tool group utilization
- highest average lot wait time at the tool group
- highest planned tool group utilization
- highest amount of WIP behind the tool group
- most often discussed problem tool group

Submit

CREATE YOUR OWN POLL WITH CROWDSIGNAL

This Poll is Closed

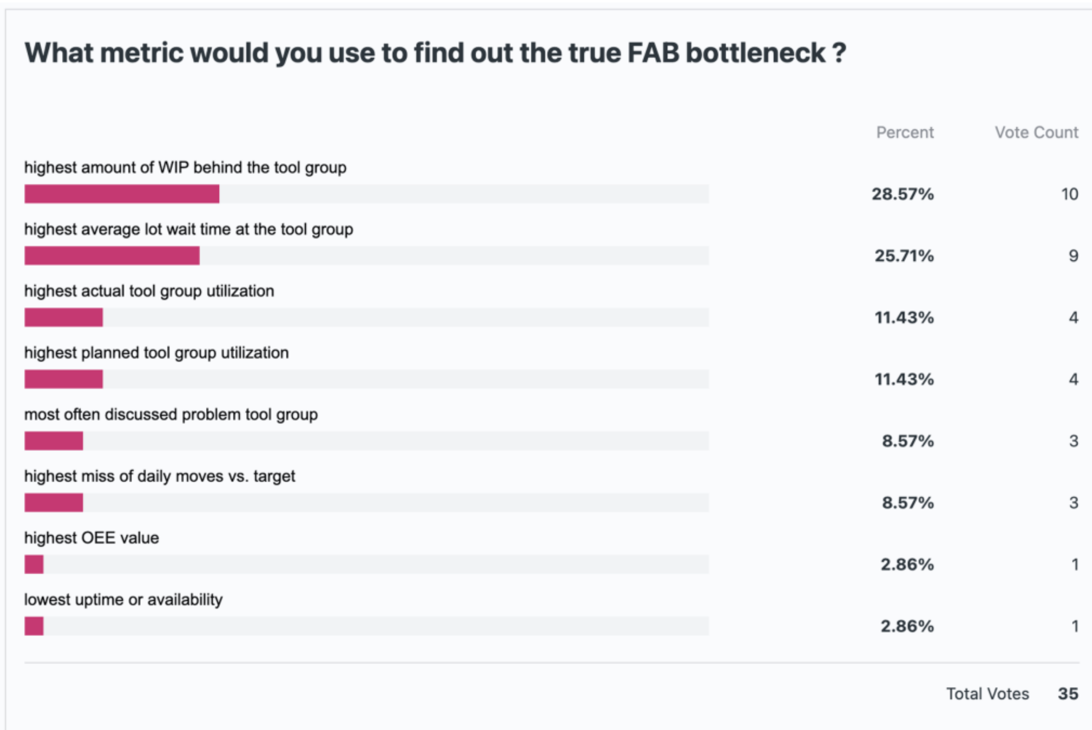
I can't wait to see the results. I'm fully aware that the answer selection is not that straightforward without more content – so if you like to provide thoughts, please use the comment functionally at the bottom.

I will share and discuss the results in my next post, sometime before the holidays

Thomas

Bottlenecks, part 2

A big thank you to everyone who voted in my little poll, here are the results:



I kind of expected a picture like this – but what does this mean? Here is my interpretation:

Bottlenecks are widely known as the one thing one should work on 1st to improve the overall FAB performance. But it seems we have different opinions how to measure and therefore to define what is the bottleneck.

For a real existing FAB, that would mean if different people or groups use a different definition, they would very likely identify different tool groups as the bottleneck – for the very same factory!

Of course, we did not yet discuss what type of bottleneck we are talking about: a short-term current one, a long term planned bottleneck or any other definition. Nevertheless, people would identify very likely different tool groups as the key FAB problem …

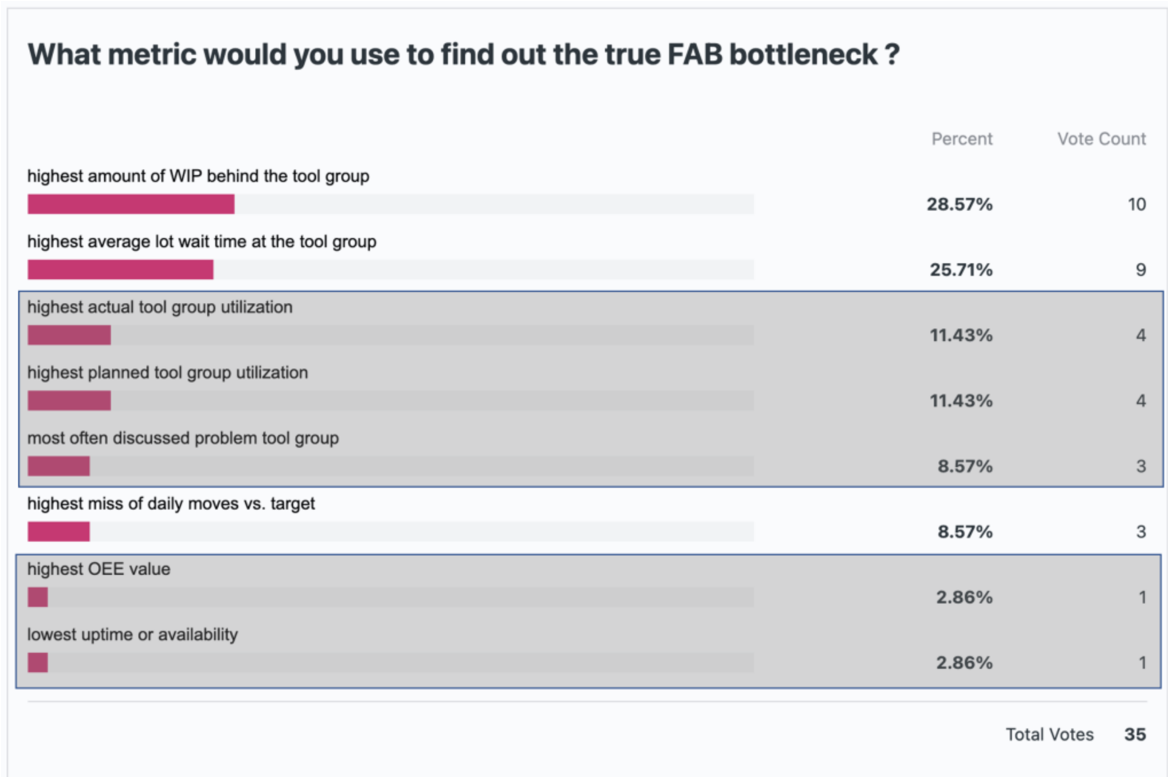
Before we discuss this a bit more, I think we need to clarify what is the meaning of “bottleneck for the FAB”. In my opinion the purpose of a FAB is to make money and in order to do this, wafers need to be delivered to customers in a way that the overall cost is lower than the selling price. Selling price also means one needs to have someone to sell them to – the CUSTOMER. For the purpose of this bottleneck discussion, I exclude topics like yield and quality, assuming these are “o.k. and in control”. I will just focus on the 2 other key metrics for “FAB performance”:

- deliver enough wafers in time –> customer point of view –> cycle time of the FAB
- manufacture enough wafers –> total cost / manufactured wafers –> cost per wafer –> FAB output

So in my opinion, a bottleneck is a tool or tool group which negatively impacts the cycle time of the FAB and therefore the FAB output in general, but more specific the output of the right wafers (products) for the right customers at the right time (aka on-time delivery)

With that in mind, I think we need to define the metric in a way that it measures the impact to these 2 parameters. In a semiconductor FAB the typical unit to track wafer progress through the line is a “lot”. Hence, in order to measure how good or bad a tool group impacts the flow of lots through the line, we need to look a lot related indicator. This disqualifies grey marked ones in the picture below and leaves us with 3 potential candidates

---



Let’s have a look at the greyed-out metrics.

highest planned tool group utilization

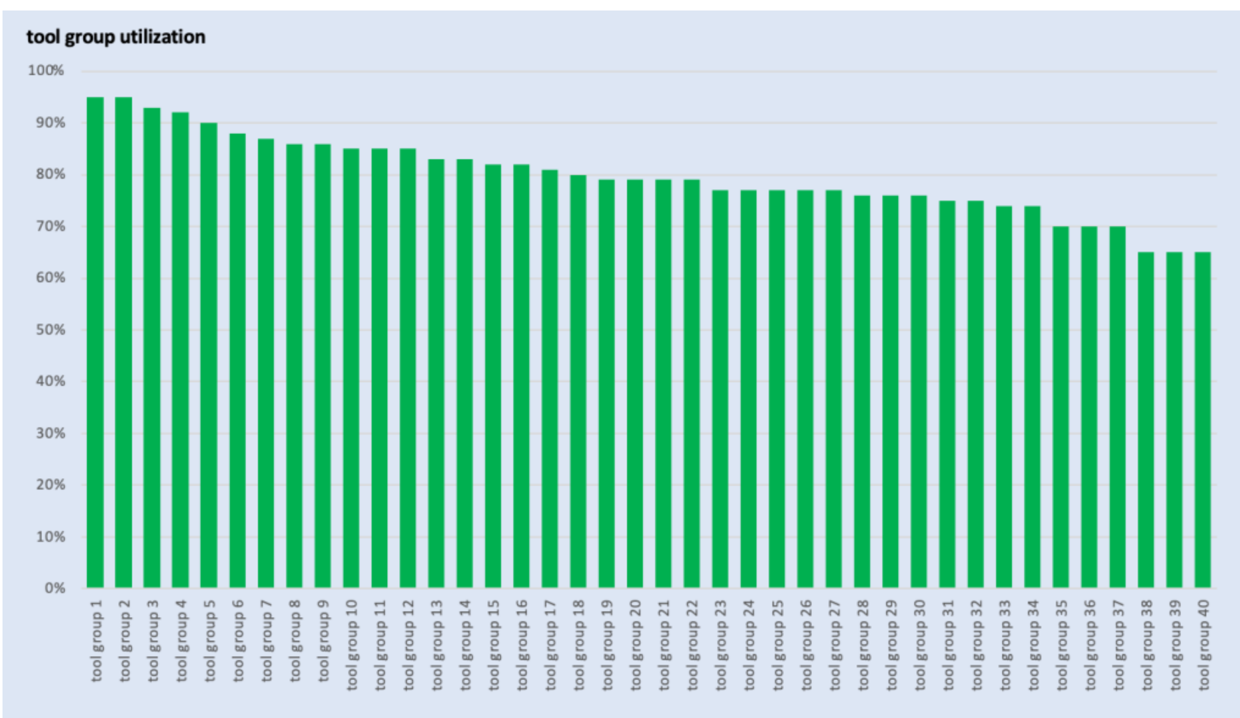
It is very tempting to pick this metric since very high tool utilization signals to some extent, we might reach **capacity** limits soon. Also it is widely known, that tool groups with **high utilization** tend to also generate **high cycle** times. So there is a good chance, that the true FAB bottleneck has a high or the highest utilization – but there is not guarantee – that this is the case. This very much depends also on the overall **utilization profile of the factory**.
Another interesting topic to discuss in a future post is: What means “high” utilization and “high” cycle time? Similar, how to define “FAB capacity”, which I will discuss also in a later post.

Factory Physics And Automation

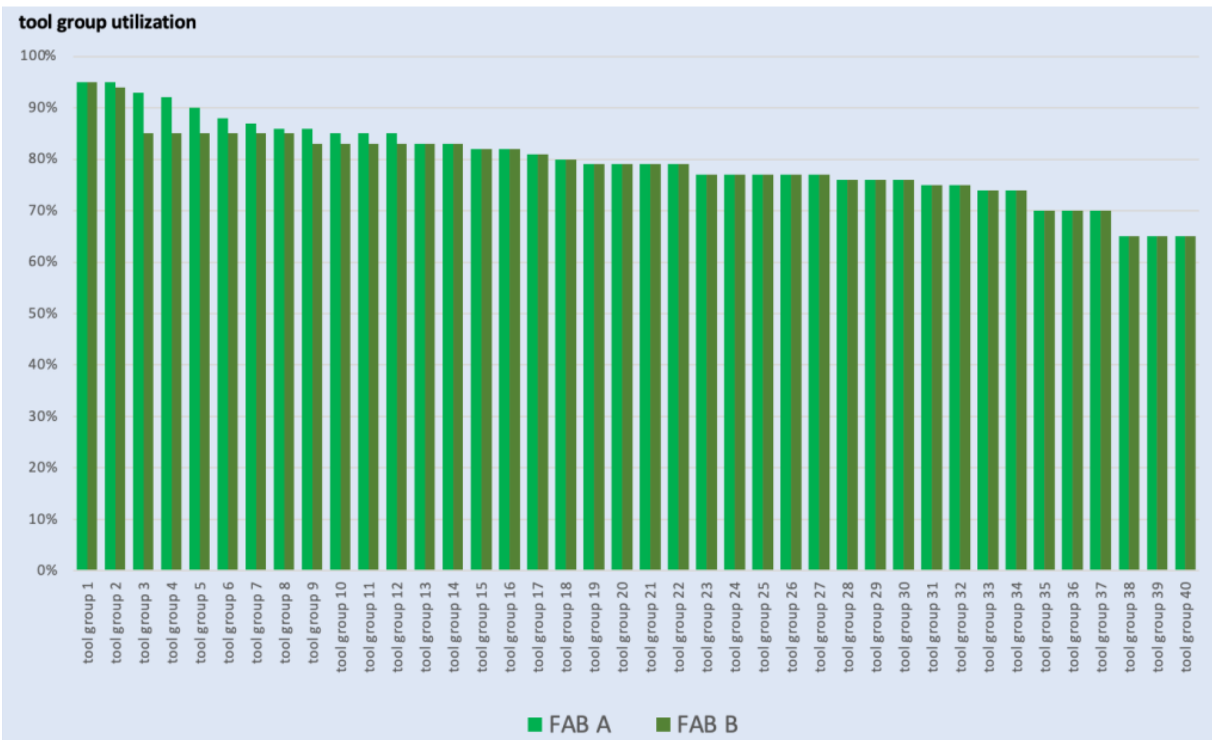
highest actual tool group utilization

Everything I wrote above for the planned high utilization is valid for the actual utilization as well. I just like to add at this point, comparing actual tool group utilization and planned tool group utilization should be a frequent routine, to understand how close or distant the capacity model is able to follow the actual FAB performance – or should I say the actual FAB is able to follow the capacity model ? You guessed it, an interesting topic for another post …

Before we move on into the next metrics, I like to spend a few thoughts on the topic factory utilization profile. The factory utilization profile is a chart of all tool groups, showing their average utilization (planned or actual, for selected time frame, like last 4 weeks or last 8 weeks) and the tool groups are sorted in a way, that the tool group with the highest utilization is on the left and the one with the lowest utilization is on the right. A theoretical example is shown below:



Different factories will have different utilization profiles. Even the very same factory will have different utilization profiles over time if things like wafer starts, product mix, uptime or cycle time change. So, I always thought it is a very good idea, to keep an eye on that and also compare the profile planned data vs. actual data. An example of comparison (with dummy data) is below.



For example: Look at tool group number 3! How likely will become #3 a problem in FAB A vs. in FAB B?

I think you get the general idea, but there is much more interesting stuff to read out of FAB utilization profiles. Before we go there – **have you lately checked / seen your FABs utilization profile?**

most often discussed tool group

This metric has some advantage, since it is not focusing on one specific indicator and if a tool group is very often in focus, it has for sure some problematic impact on the overall line performance. I rather would choose real data-based metric, but for FABs with less developed automatic data generation and data analytics capabilities it is a usable starting point. I also like about this approach – once used for some time – it will inherently drive the demand for a more data-based approach – to find out, why is a tool group discussed so often and where to start with improvement activities – which in today's manufacturing world is an absolute must in my opinion.

highest OEE value

OEE, it feels had its peek time when a lot of people talked about it, but it seems lately the topic became a bit quieter. The OEE method itself has its value, if used on the right tool groups with the right intentions. If applied solely to increase the name plate OEE value of every tool group in the FAB, it can become quickly counterproductive and hinder the overall FAB performance (at least if FAB performance is defined and measured as proposed in this post) In my active days as an FAB Industrial Engineer I often used the slogan:

"… if the OEE method is used the right way, its target should be not to increase the OEE value of the tool group, but increase the tool groups idle time …"

If OEE projects are aiming in that direction, they will for sure help to improve the overall FAB performance, but as the key metric to identify the biggest bottleneck I would not recommend using OEE.

lowest uptime or availability

As mentioned above, uptime is a tool or tool group focused metric and for sure a very important one in every FAB. While low uptime is not desirable, it is not a good indicator if the tool group is indeed a factory bottleneck, since it will not tell us anything about the actual impact on the FAB without other information.

At this point I will stop for today. In my next post I will spend a bit more time on the 3 remaining – lot related – indicators and will also share, which one I think will be the most useful one to use. As always, I would love to hear feedback from you via a comment. One last thing: I will eventually stop announcing every new post via LinkedIn, so if you want to get notified when there is new content here, please use the email subscription form below

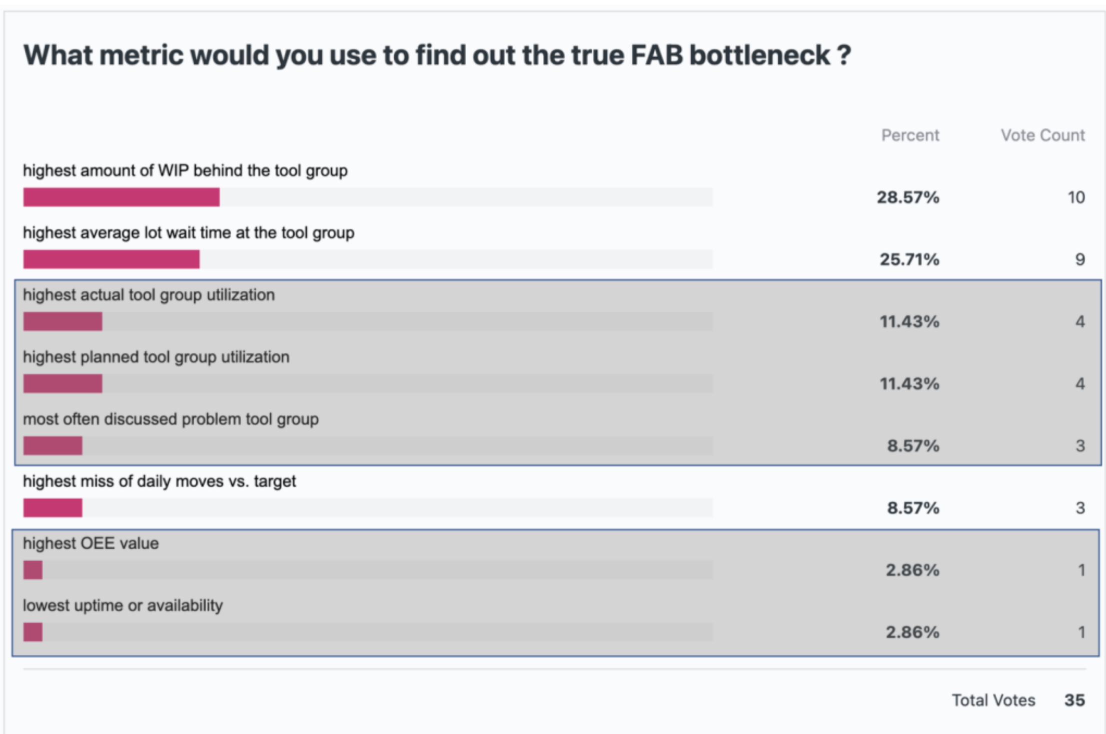
Happy Holidays!

Thomas

Bottlenecks, part 3

Merry Christmas and Happy Holidays!

I hope everybody is having a good time with friends and family and after a lot of good food is ready sit down and discuss more details about factory bottlenecks. In today's post I will start zooming in on the 3 not grayed out metrics from the poll results picture below:



What metric would you use to find out the true FAB bottleneck ?

	Percent	Vote Count
highest amount of WIP behind the tool group	28.57%	10
highest average lot wait time at the tool group	25.71%	9
highest actual tool group utilization	11.43%	4
highest planned tool group utilization	11.43%	4
most often discussed problem tool group	8.57%	3
highest miss of daily moves vs. target	8.57%	3
highest OEE value	2.86%	1
lowest uptime or availability	2.86%	1
Total Votes		35

To disclose my opinion upfront: I think that "highest average lot wait time" (or metrics that are derived from this) is the most objective way to measure and define what is the true factory bottleneck. But let's discuss all 3 of the metrics a bit.

highest miss of daily moves vs. target

I think every factory in the world is measuring and reporting in some way the number of "Moves" – the number of wafers which were processed/completed on a step in a day, a shift, an hour, for the whole FAB or departments and down to individual process flow steps and grouped by equipment or equipment groups.

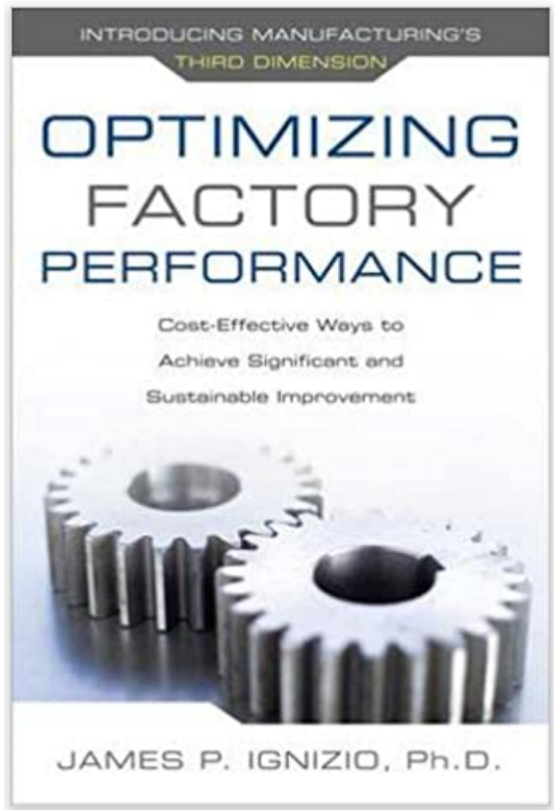
"Moves" is a very attractive and popular metric for a lot of reasons:

- Moves can be easily measured and aggregated in all kind of reporting dimensions
- based on the numbers of steps in a process flow (route) it is clear, how many Moves a wafer needs to complete, to be ready to be shipped
- Moves is a somewhat intuitive metric – humans like to count
- target setting seems to be straight forward – "more is better"

I personally think, measuring a FAB via "Moves" as the universal speedometer can be very mis-leading and might drive behaviors – which are counterproductive – for the overall FAB performance. At the very least a well thought through and dynamic target setting is needed to steer a factory which is mainly measured by the number of Moves. The danger of Moves as the key metric might be less in fully automated factories, since the actual decision making is done by algorithms which usually incorporate a lot of other metrics and targets and therefore Moves are more an outcome of the applied logic, less an overarching input and driver.

In manually operated factories, where operators and technicians make the decisions, which lot to run next and on what equipment, a purely Moves driven mindset can do more harm than good – to the overall FAB performance.

I think a lot has been written and published on this topic and there are strong and different schools of thought out there, but I'm fully on board with **James P. Ignizio's** view in his book



In chapter 8 of his book – titled

"Factory Performance Metrics: The Good, The Bad, and The Ugly"
"Moves" get a nice talk – in the "Bad and Ugly" department – for the very reason, that Moves can drive counterproductive behavior. If you are interested in this topic – I strongly recommend reading the book.

Before I jump to the next metric – I just wanted to say – that I think that Moves are important to understand and is a useful indicator if used within the right context, but not "blindly" as the most important indicator, which drives all decision making.

highest amount of WIP behind a tool group

Almost one third of the voters picked this metric. Similar to Moves there are a lot of advantages to measure WIP:

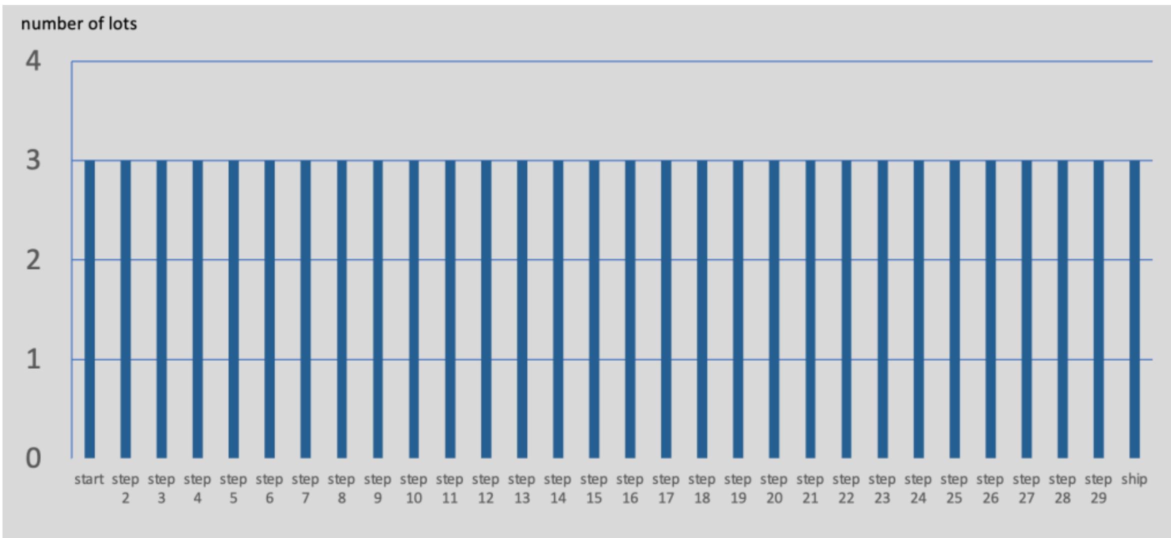
- WIP can be easily measured and aggregated in all kind of reporting dimensions
- using "Little's law" it is easy to define WIP targets
- WIP is a very intuitive metric, especially in manual factories – is my WIP shelf full or empty ?

In general – for daily operations – having a lot of WIP is seen as problematic, since it might lead to lots not moving, starvation of downstream steps and tools, long lot wait times before they can be processed. So high WIP is not a desirable status and very high WIP must be for sure a problem. I think here as well – it depends. For example, it depends on what is the target WIP for the given context (like a tool group) to just try to lower the WIP as much as possible ("at all cost") might lead to generating WIP waves in the factory and to underutilization and lost capacity.

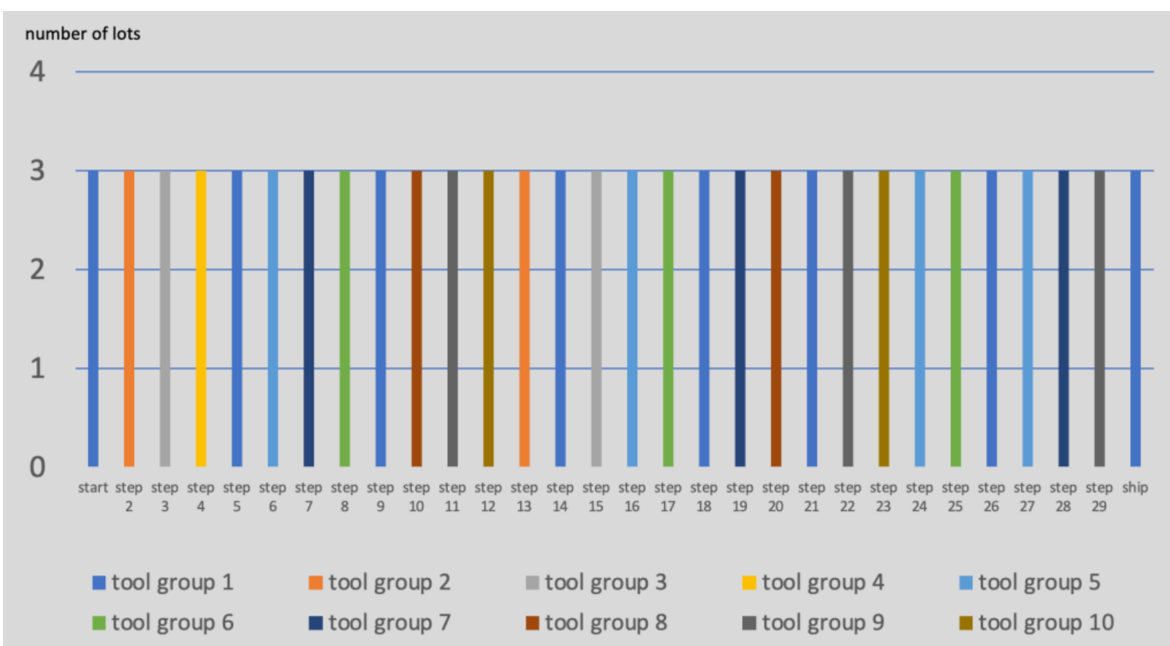
Why do I not 100% subscribe to the highest WIP = the bottleneck? It is simply, that the tool group with the highest WIP not necessarily has the worst impact on the FAB performance. Here are some data points for this:

Let's assume we have a very small factory running a very short route – with only 30 steps. If we plot a chart showing the WIP (in lots) per step for each step and sort the steps in the order of the process flow – meaning lot start on the very left and lot ship on the very right – we get what is typically called a line profile chart.

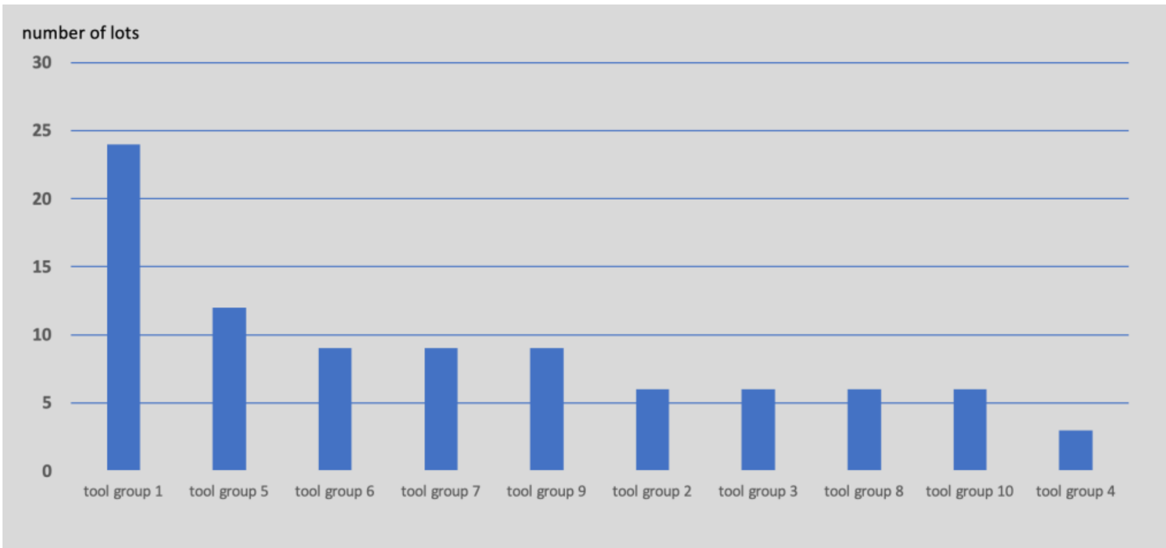
In the picture below our factory is perfectly balanced (if we define balanced as lots per step – another great topic to talk about) because on each step there are currently 3 lots waiting – or processing.



If we look a bit closer, different steps are of course processed on different tool groups, if we add this detail, the same factory profile looks like this:



For example, tool group 2 has 2 steps in the flow and tool group 9 has 3 steps. Our bottleneck metric is the aggregation of the WIP by tool group (“highest WIP behind a tool group”). To find out, which tool group this is, we simply aggregate the same data from the line profile by tool group instead per step:



Tool group number 1 has the highest WIP of all tool groups in this FAB – it clearly must be the number 1 bottleneck – I do not think so. As discussed earlier, there is more content needed. For example, if tool group 1 is a scrubber process, which is typically in the flow a lot of times and it is an uncomplicated very fast process, having the overall highest number of lots there is not necessarily the biggest problem of the factory. Yes, one can argue, still it would be nice to have less WIP sitting at a scrubber tool set, but this is already part of the missing context, I mentioned earlier.

Measuring and reporting WIP is an absolute must in a semiconductor factory but interpreting WIP levels and assigning them attributes like “high”, “normal” or “low” needs a very good reference or target value. Setting WIP targets should be done via math and science, to reflect what is the overall factory desired WIP distribution – in order to achieve the best possible FAB performance.

Before I close this topic for today – let me say: my simple “perfect balanced” line from the pictures above might not be balanced at all, if we incorporate things:

- different steps / different tool groups have very likely different capacities
- different raw processing times
- might be batch or single wafer tools
- might sit inside a nested time link (queue time) chain
- …

At this point I will pause and hope that I could stimulate some thinking and of course would love to hear feedback from the readers out there. The next post will be fully dedicated to the last open metric …

Thomas

Bottlenecks, final part

Happy New Year !

This will be the last part of the Bottleneck discussion. As mentioned in part 3 – I think the most objective and telling indicator to see what the true factory bottleneck is:

highest average lot wait time at a tool group

Wait time or cycle time in general is one of the very few indicators which cannot be easily manipulated or "adjusted" by using different methods of calculation or aggregation. Time never stops and measuring the time between a lot arrives logically at a step and it starts processing at the step (on an equipment) are 2 simple time stamps which are typically recorded in the MES of the factory. For example:

> lot arrived at the step: 01/02/21 4am
> lot started processing: 01/02/21 10am
> The wait time of the lot is super simple –> 6 hours.

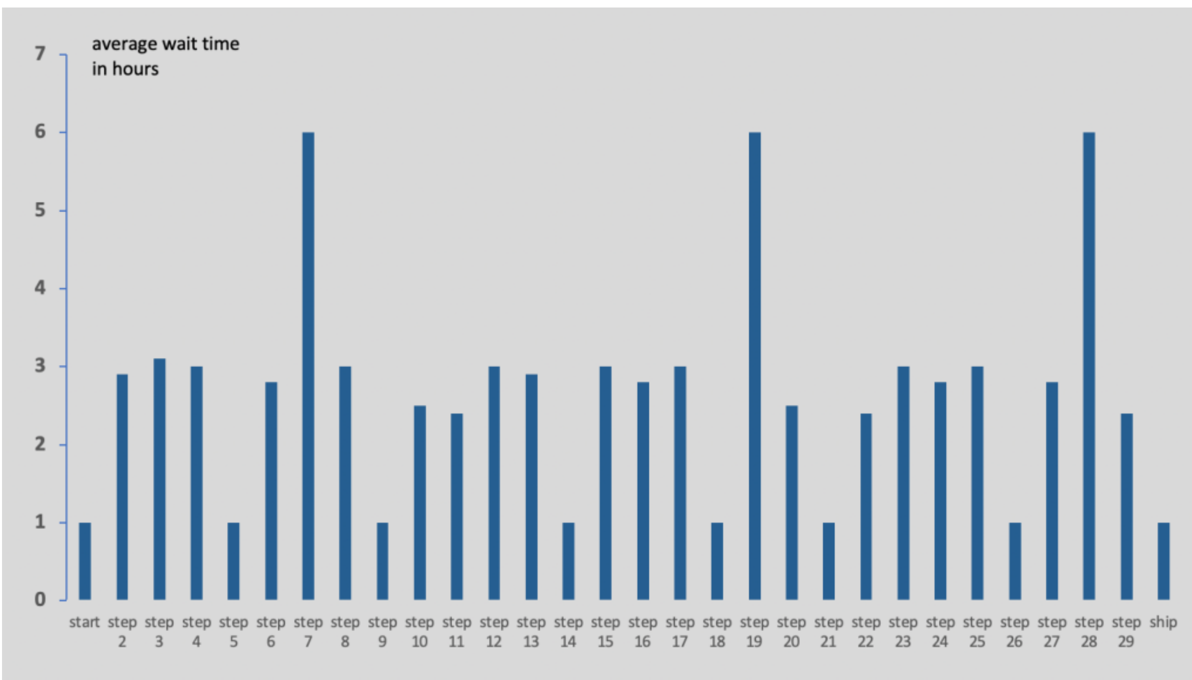
The beauty of this metric is that no other information is needed – just these 2 time stamps. It will cover any possible reason why the lot waited 6hours – no matter what:

- equipment was not available due to down time
- equipment was not available since it was busy running another lot
- lot was not started due to missing recipe
- lot was not started due to no operator available
- lot was not started since operator chose to run another lot
- lot was not started due to too much WIP in time link zone
- lot was not started due to schedule had it planned starting at 10am
- lot was not started due to … "name your reason here"

One key part of the FAB Performance metrics – as discussed in part 2 – is:

- deliver enough wafers in time –> customer point of view –> cycle time of the FAB

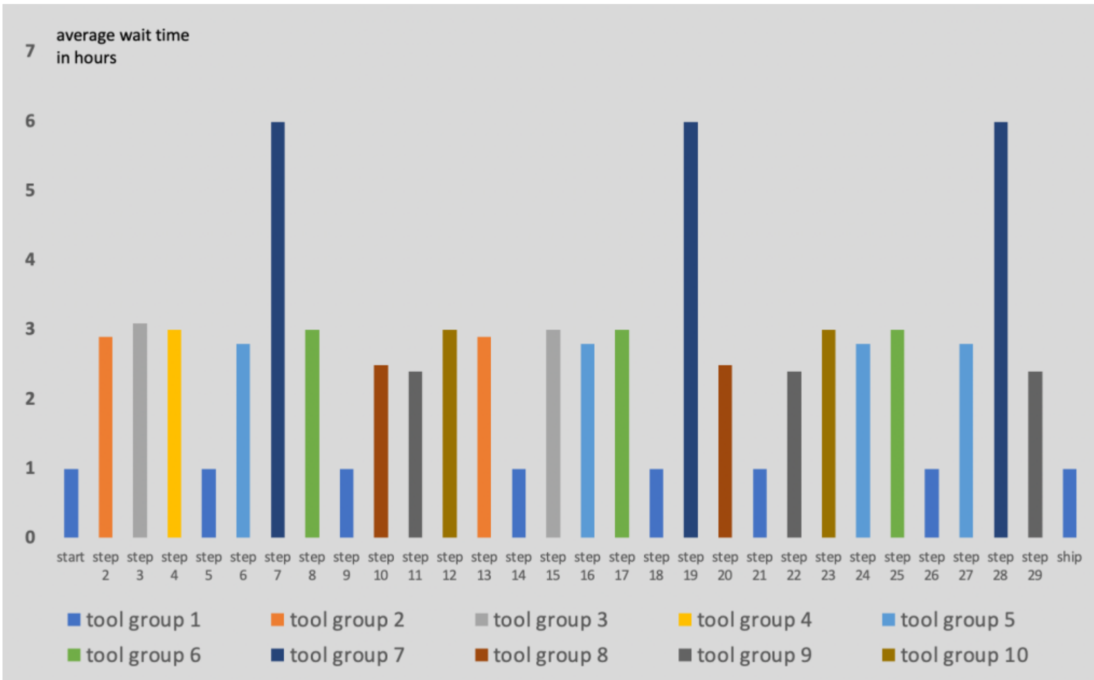
In other words, once the decision was made to start a lot into the factory it has some kind of target date/time by when this lot needs to be finished or shipped. Any wait time is by nature now a "not desired state" especially if the wait time is **"very long"**. That means tool groups which generate the highest average lot wait time will be very likely the biggest problem or bottleneck.
Let's have a look at some example data to illustrate that:



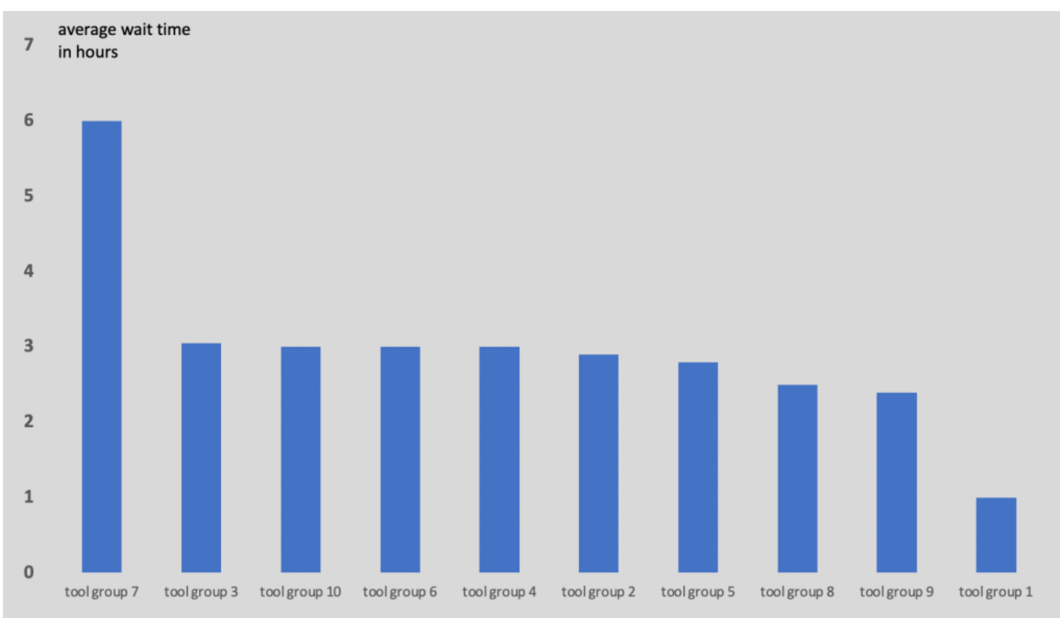
The chart above shows the average lot wait times per step of our complete factory. Some steps have 1h wait time others have up to 6h.

Since this chart shows the data by step in the order of the route or flow it does not tell immediately which tool groups the various steps are running on.

The same data – including the tool group context – will tell this better:

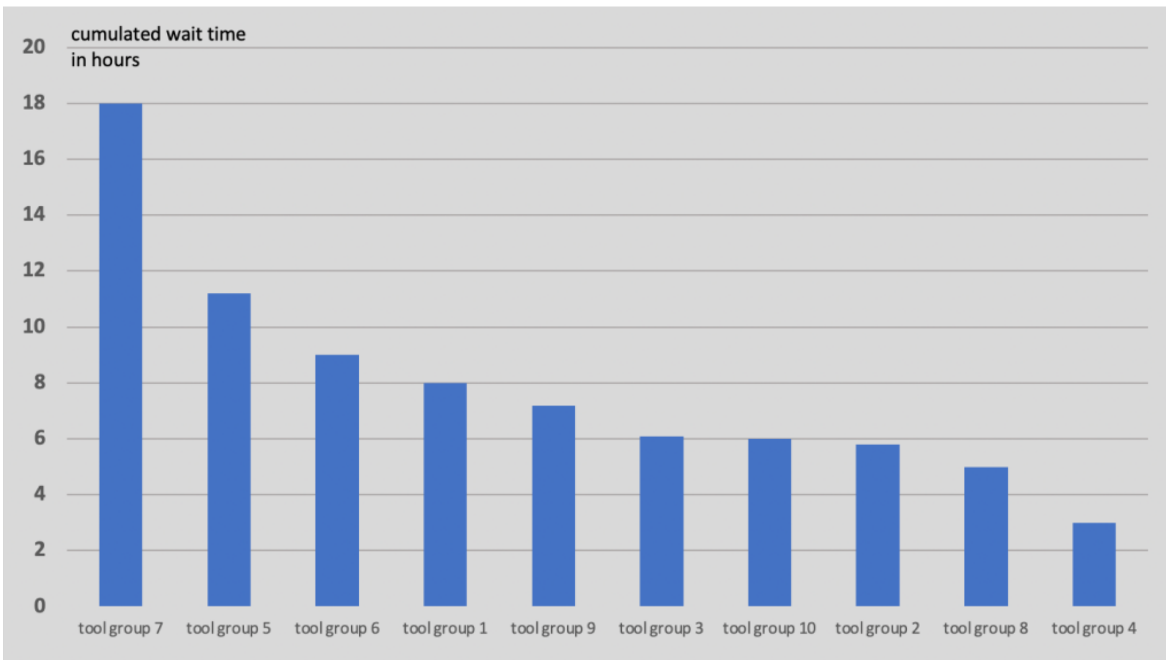


If we now aggregate and sort this by tool group instead of step, we have our bottleneck chart:



From this chart tool group 7 clearly has the greatest average lot wait time of all tool groups. An interesting version of this chart is the “total wait time contribution” chart which shows the sum of the individual step wait times.

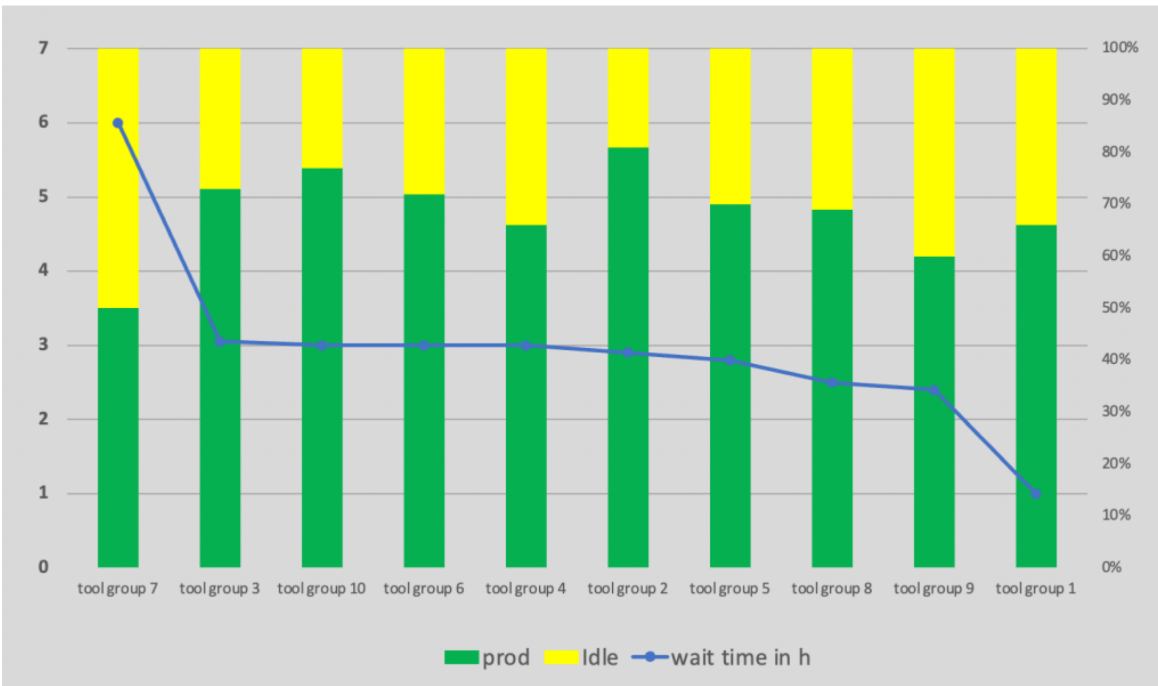
For example, tool group 7 has 3 steps in the route and on average a lot waits on each step 6h. If we plot the same data as “total wait time contribution chart” we will not average the wait time of the individual step but add them: Tool group 7 will show 6h + 6h + 6h = 18h of total wait time for each lot.



Note that the sort order of the tool groups is now different. For example, tool group 1 which on average has the lowest wait time (1h) is now ranked as number 4. From an overall “is this tool group a problem for the factory?” point of view I say no – since lots barely waiting there – it just happens that tool group 1 has a lot of steps in the flow. I strongly lean to the average chart for the overall definition of the FAB bottleneck but recommend always to have a look on the cumulative chart as well.

In part 2 of the Bottleneck blog series, I discussed the "Factory Utilization Profile chart". I think this chart enhanced with the wait time data from above will give the "complete view" what is going on in the factory and will spark enough questions to dig in deeper at the right tool groups.

The chart below shows the data sorted by the highest average cycle time:



Obvious question is: Why is there so much wait time on tool group 7 at such low utilization or asked differently: half of the time the tool group is idle – why do lots wait on average for 6 hours?

Or another one: How is tool group 1 able to achieve such low wait time?

At this point I like to stop for a second and point you to an excellent source of additional discussion on the topic of bottlenecks and cycle time:

FabTime

If you subscribe to the newsletter, you will have access to past editions as well!

Let me get back to the statement: **Any wait time is by nature now a "not desired state" especially if the wait time is "very long"**
Given the nature of the wafer FAB the ideal case of zero wait time at all steps is not very realistic since there are too many sources of variability in a factory. Therefore, experienced capacity planners and production control engineers typically set an expected wait time target per step (and therefore by tool group). Using these expected wait times, the definition of "very long" becomes easier.

For example, if

tool group A has an expected wait time of 2hours
tool group B has an expected wait time of 5hours

An actual achieved wait time of 6 hours would be kind of tolerable on tool group B but clearly seen as very high on tool group A.

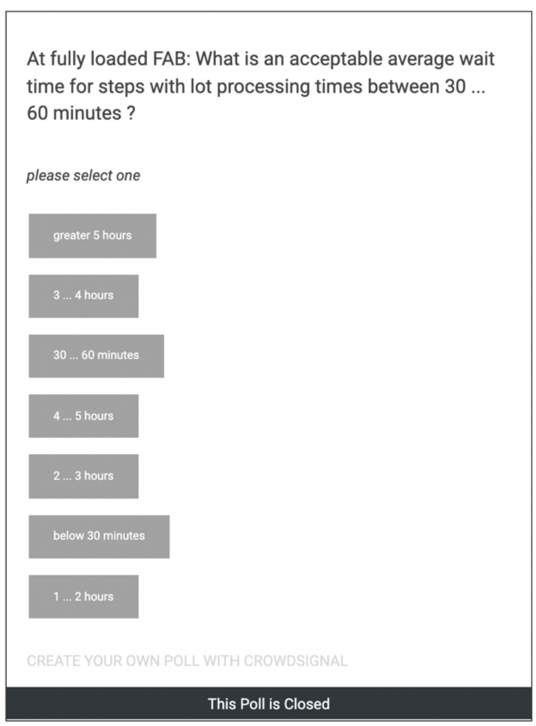
Setting expected wait times per step and/or tool group depends on a lot of parameters, like:

- planned tool group utilization
- number of tools in the tool group
- duration of process time
- batch tool / batching time
- lot arrival time variability
- many others

I'm curious what the readers of this block think would be an acceptable average wait time for non-bottleneck steps in a fully loaded factory.

Let's assume that most steps in the factory have processing times of 30 – 60 minutes, running on non-batch tools, and the factory is fully loaded = the capacity planners tell you, you cannot start more wafers. What would be an acceptable average lot wait time for these steps in your opinion?

Please vote below, what you would see as good / o.k. / acceptable



I will share and review the results in my next post.