WIP LIMITS

Outcome

Work in progress (WIP) limits are a technique mainly used in Lean methodologies such as Kanban. A WIP limit is a constraint put on the number of work items to be worked on by a team at any one time, usually within a work process – such as testing or development.

Function

Queues within a workflow often have a high cost, are difficult to reduce and increase overall cycle time. An important goal for the team is to control the number of queues and their size, and ensure a smooth flow of work. The use of a pull system and the reduction of batch sizes aim to help control the flow, but they are not sufficient. Constraints or limits on the work in progress should also be defined. In fact, they are probably the most effective way to smooth flow and reduce cycle times. It's a satisfying paradox: we do less in order to do more.

A side effect is to increase quality. With fewer work items to handle, the team usually make fewer mistakes and overheads of reporting and tracking are reduced. Because WIP limits improve cycle time, there is a shorter feedback loop, which should also improve quality.

Benefit

Implementation

Prerequisite

You need to have a clear view of your current process. The use of a Kanban board is highly recommended.

Putting in place WIP limits

1. Choose one step of the process where there is a big queue (inventory of work items).
2. Define a limit and display it on your Kanban board. To compute your first limit, you can use a basic rule of thumb 'number of people involved in the step times 1.8'. If you have the data, use the historical average WIP observed for the step. Since the limit will be reviewed periodically, you could also simply begin with an arbitrary limit the team are comfortable with (see next steps).

Who

3. When the limit is reached:
   - Do not start any new work in this step.
   - Analyse why the limit was reached.
   - Try to solve the issue, for example by the team swarming to offer help to get the work done and reduce the queue, or purge or put on hold low-value work from the queue if higher value work items are blocked.

Scaling Factors

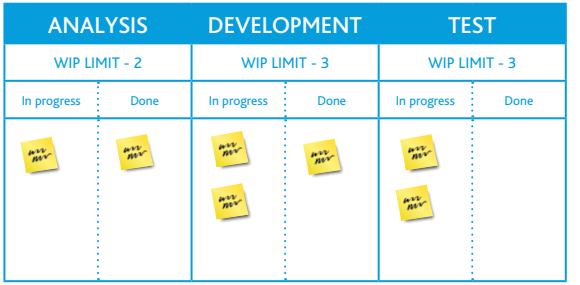
4. If the limit is reached often then:
   - Analyse why the limit is reached so often. Techniques like the 5 Whys may help you find the root causes of the problem.
   - As a team, devise potential solution(s) and take actions.
   - Depending on the type of solution, you may choose to work with a queue in the short-term by raising the WIP limit – for example, the solution might take time to implement. But, do not increase the limit until you have exhausted all other solutions!

Difficulty

5. As shown by Queuing theory, there are links between the cost of capacity, the cost of delay, cycle time and queues. The actions taken in steps 3 or 4 may change these parameters and the team should therefore review the limit periodically.
6. If you never reach your WIP limits, and you have a smooth and sustainable flow, try to reduce your WIP limits slowly to improve your cycle times. But, keep monitoring your costs of capacity and cost of delay to ensure this is still a reasonable action.

Applying WIP limits to all the process steps

Every step in the process can have its own WIP limit. If one of the steps reaches its limit, because the team then stops pulling work, the upstream steps normally also reach their limits quickly. This makes it very clear that an issue needs to be solved, and forces the team to find a solution quickly.



Class of services and WIP limits

If the team has defined class of services, the WIP limit can be different for each class of service in order to ensure that the right mix of work items is preserved even through a limited queue.

Potential pitfalls

- **WIP limits are never reviewed –** WIP limits should be periodically reviewed, in particular whenever there is a change of the capacity of the team or a step is altered.
- **WIP limits per person –** Some teams define WIP limits per team member. This uncommon practice occurs usually in teams where members have a strong 'T Shape' profile (i.e. team members with multiple skills who can work on several if not all steps of the process). However, WIP limits are more effective and their impact easier to monitor when applied to the steps of the process. If you have defined WIP limits for your process, you shouldn't need to define WIP limits for your team members.

If you want to learn more, consider reading:
The Principles of Product Development Flow by Donald Reinertsen
Kanban Successful Evolutionary Change for Your Technology Business by David Anderson