POLITECNICO
MILANO 1863

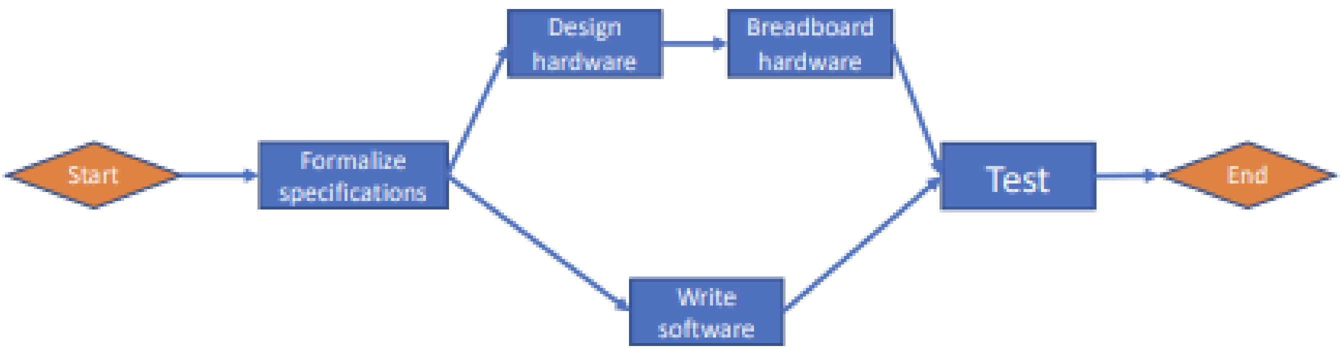
Performance Evaluation and Applications Projects

Project Type C

Giuseppe Calcagno - 10659505

Specification

Given a PERT diagram of a company divided in five different and independent departments and the correspondent service time, we would like to know the best number of project N that they should work to have the best tradeoff between throughput and project completion time



Part 1 – Distribution Fitting

For each department, a file with 5000 samples of the durations of the corresponding phase is given.

I decided to find the distribution of that service times to generalize the problem to be less dependent from the given data.

Thanks to the large number of sample, both methods for fitting (Maximum likelihood and Moments), should give a good result.

I choose to go for method of moments to simplify the calculus for some distribution (i.e. Erlang Distribution)

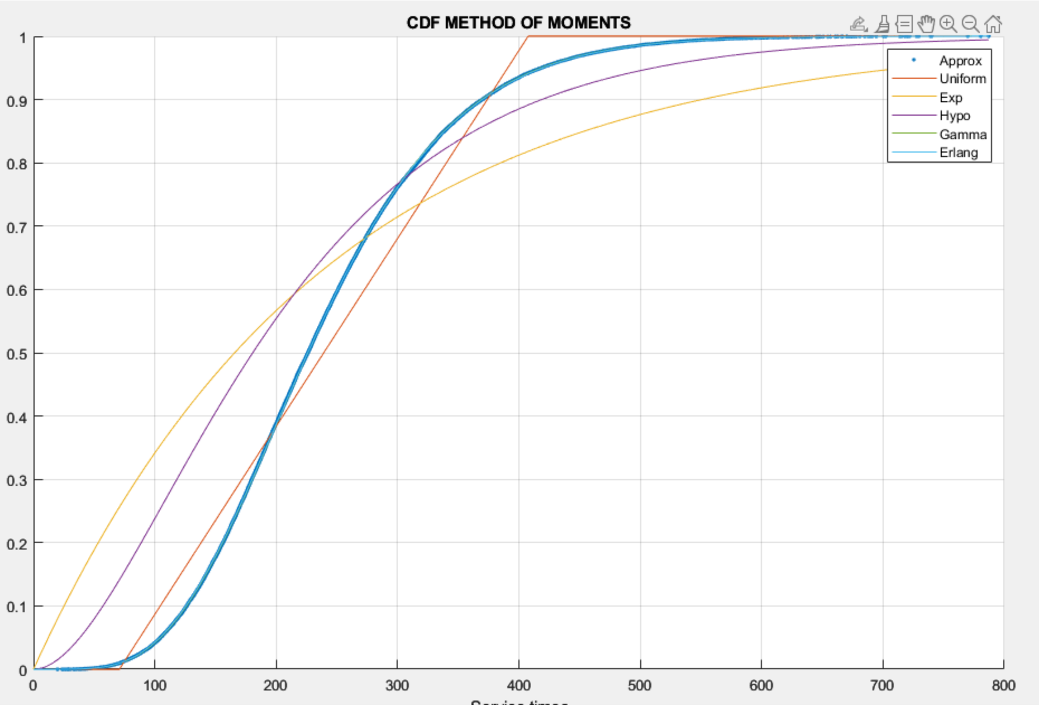
Note that all the computations are done in hour

Part 1 – Distribution Fitting
1.1 Formalize Specification

The coefficient of variation (cv = 0.406520) of this data distribution is less than 1, so a hypo exponential distribution is needed.

Hypo exponential with n=2 is not enough to fit the data .
Analyzing the gamma and Erlang distribution I decide to go for the Erlang because the k parameter of the gamma distribution was quite near an integer (K=6.051)

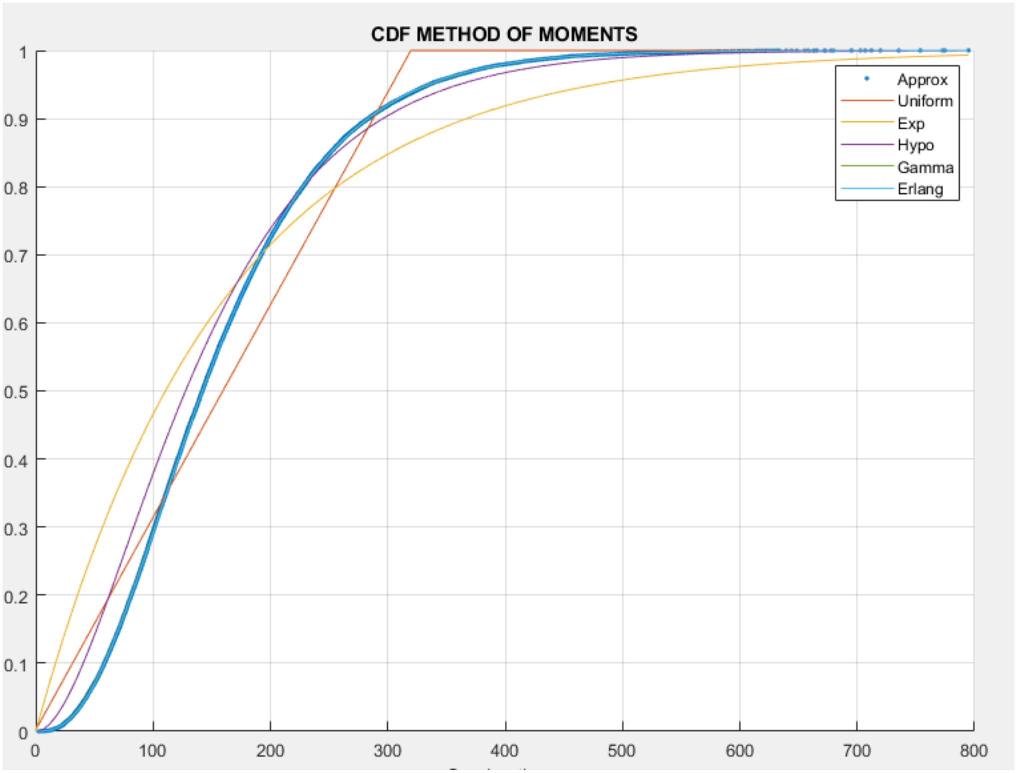
Chosen Distribution:
Erlang Distribution
- K = 6
- λ = 0.0250869



Part 1 – Distribution Fitting
1.2 Design Hardware

The coefficient of variation (cv = 0.579903) of this data distribution is less than 1, so a hypo exponential distribution is needed.

For the same reasoning of the specification part, I decide to go for an Erlang Distribution.



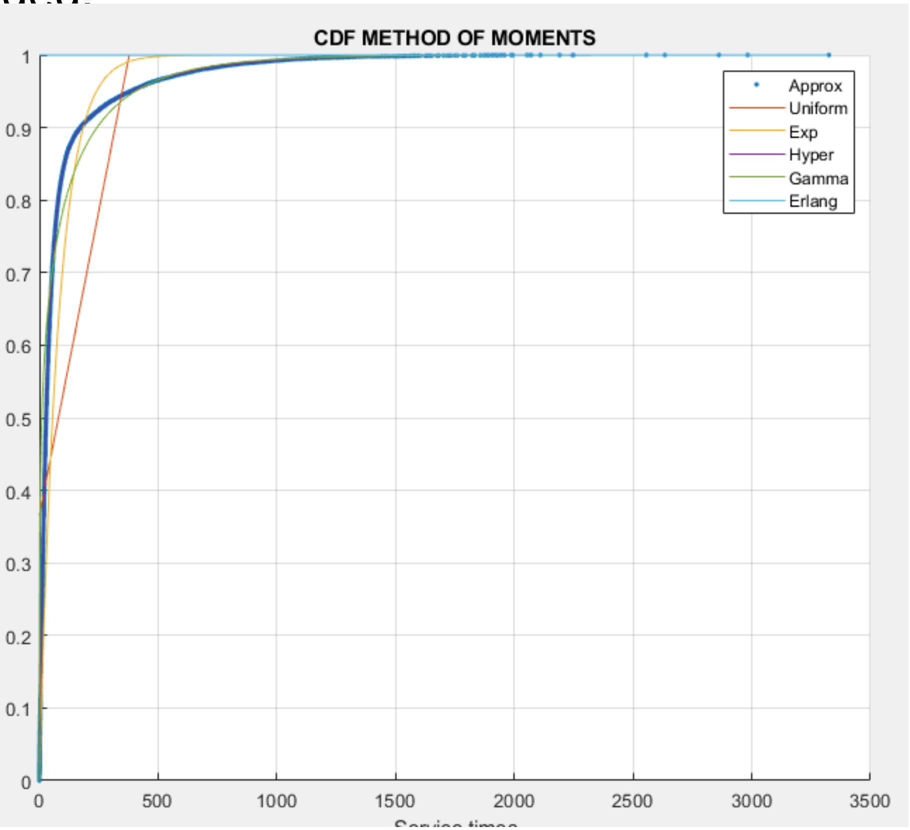
Chosen Distribution:
Erlang Distribution

- K = 3
- λ = 0.0188149

Part 1 – Distribution Fitting
1.3 Breadboard Hardware

The coefficient of variation (cv = 2.146308) of this data distribution is more than 1, so a hypo exponential distribution is needed.

Hyper exponential with n=2 is enough to fit the data .

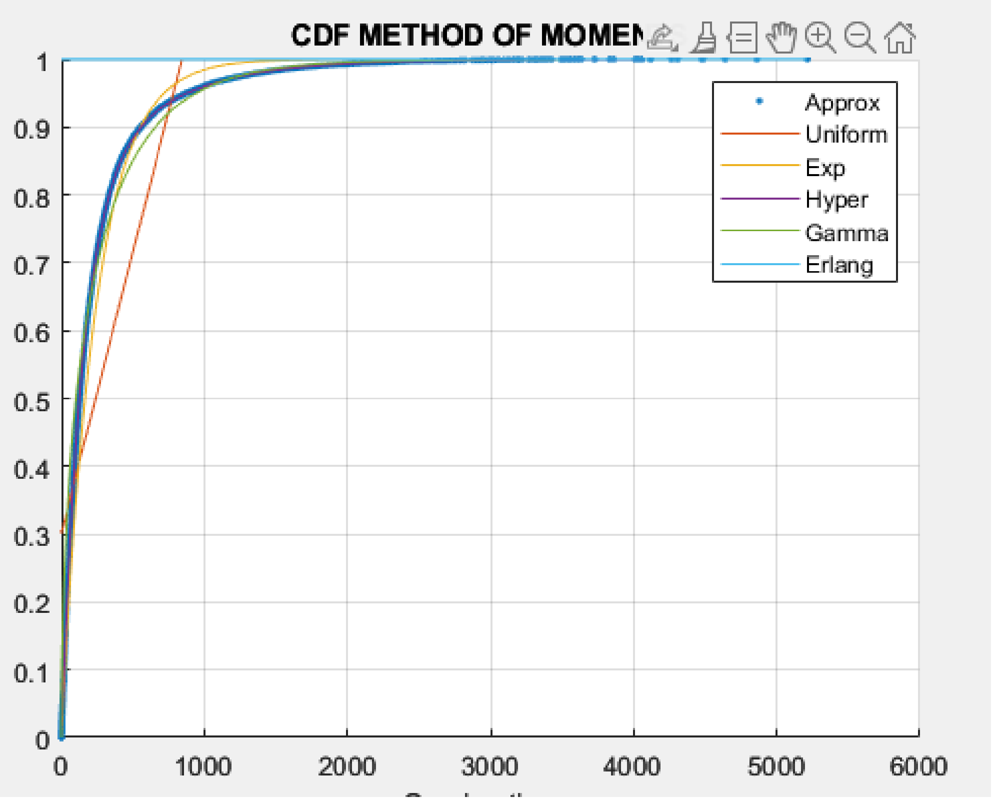


Chosen Distribution:
Hyper exponential Distribution
- p1 = 0.163853
- λ1 = 0.00308955
- λ2 = 0.0307289

Part 1 – Distribution Fitting
1.4 Write Software

The coefficient of variation (cv=1.454127) of this data distribution is more than 1, so a hypo exponential distribution is needed.

Hyper exponential with n=2 is enough to fit the data .



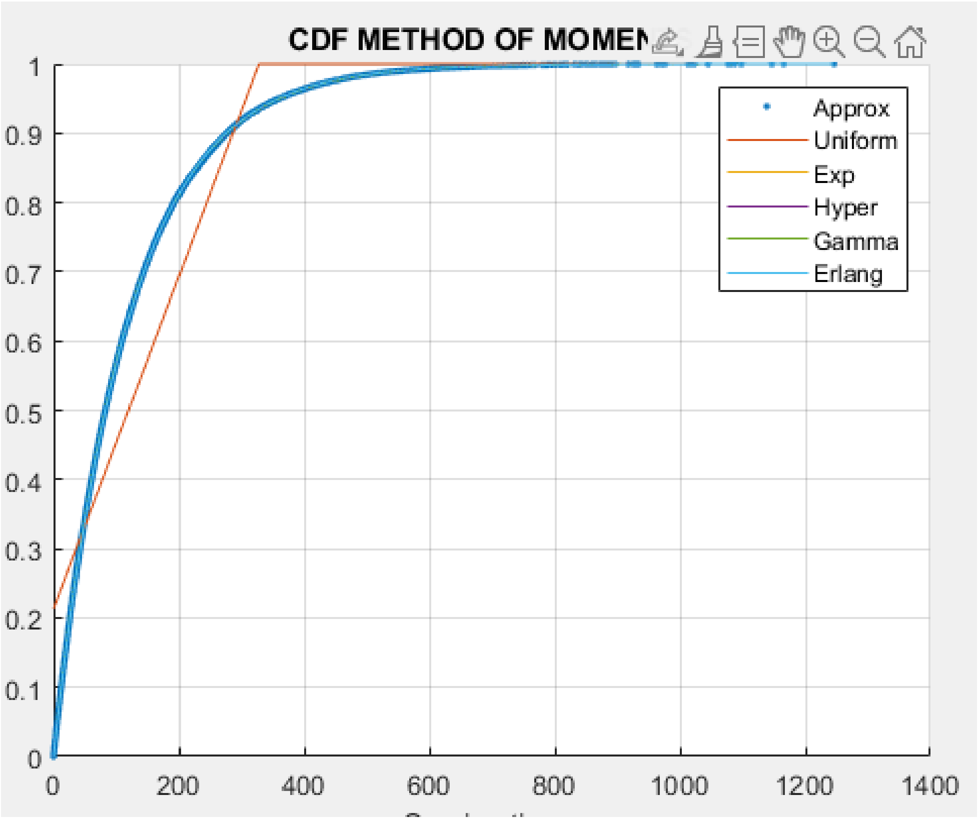
Chosen Distribution:
Hyper exponential Distribution

- $p1 = 0.218902$
- $\lambda 1 = 0.00173056$
- $\lambda 2 = 0.00689595$

Part 1 – Distribution Fitting
1.5 Test

The coefficient of variation (cv = 1.004673) of this data distribution is approximating to 1, so a exponential distribution is needed.

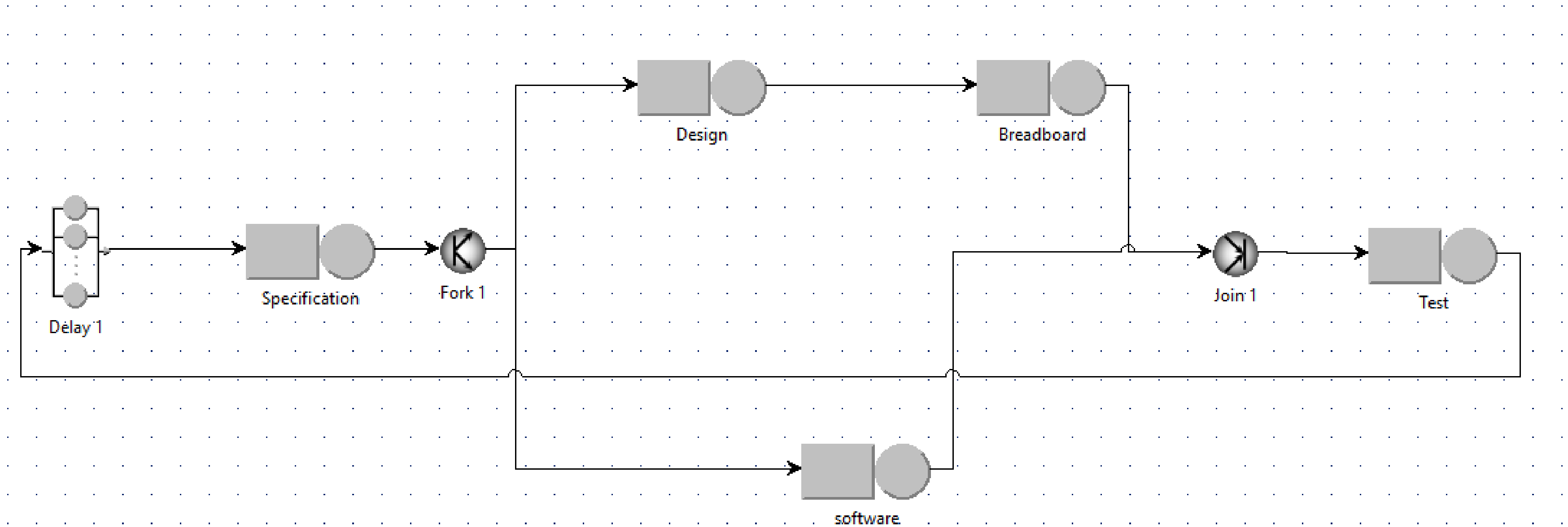
Chosen Distribution:
Exponential Distribution
- λ = 0.00835552



Part 2 – Modeling

JML is the chosen tool to model the system. This tool allows to model with simplicity all characteristic of the model and compute all the relevant performance index with the preferred confidence interval

Model:



Part 2 – Modeling
Design Choice

- The model was designed as a closed model, this choice is done for the nature of the problem and to allow analysis on the number of project N present in the system
- A delay station with 0 delay is used as access point of the system and is used as reference station for the input job
- To model the parallel works on hardware and software of the company, a fork and join node are used. The former spit a single work in 2 parallel work, the latter is used to wait until the 2 sub-job are finished and bring them back to a single work before the test phase
- Each department is modeled as a queue station with infinite queue and a FCFS policy, that seems the most appropriate for the nature of the system. Distribution analysis is exploited to set the service strategy according to the data provided
- No particular routing policy is needed because of the topology of the system
- A single job class (task) is used

Part 3 – Analysis

A “what if” analysis is performed in JMT to analyze the following performance metrics in relation to the number of job in the system N:

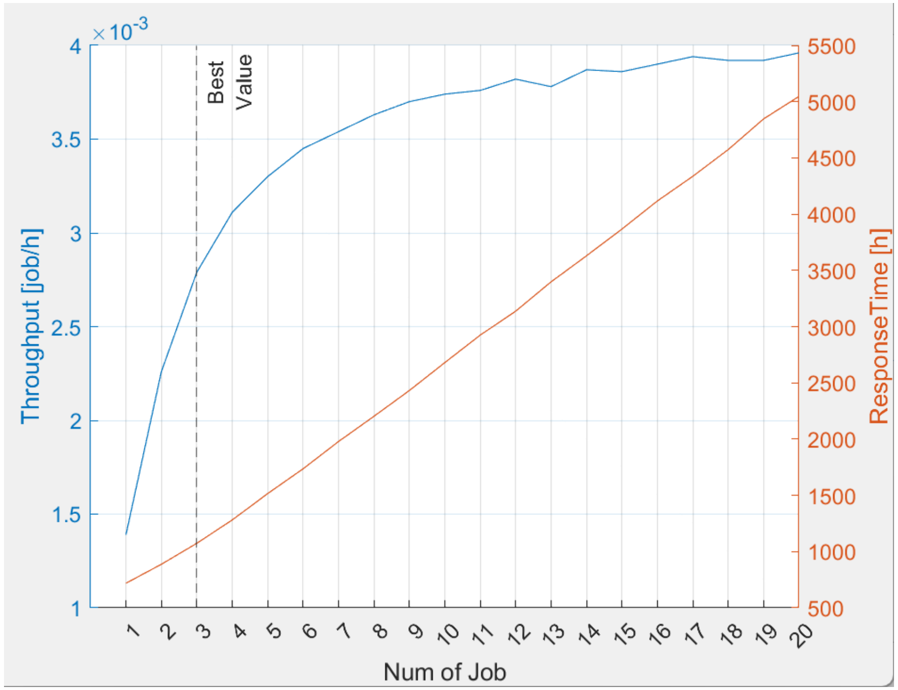
- System throughput
- System Response Time

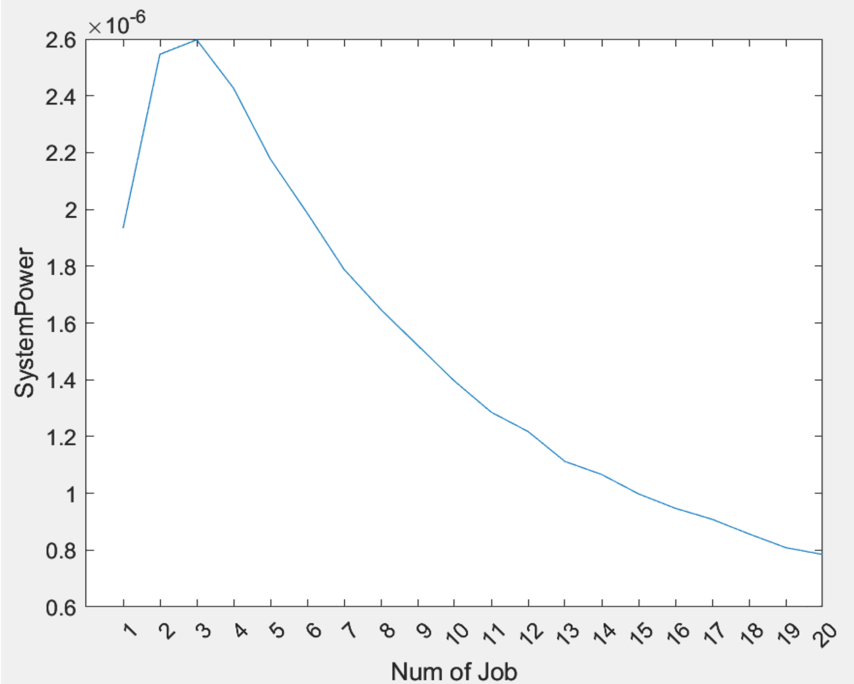
A possible significant range the number of job in the system is choice and correspond to the [1,20] range. All metrics are computed with a confidence interval of 0.99 and a Max relative error of 00.3

We want to find the best tradeoff between this 2 metrics. The System Power has been introduced to better identify this tradeoff. It is defined as the ratio between the 2 quantities.

Part 3 – Analysis

The output of the what if analysis is moved on Matlab, where an analysis of the measurement trend of the system power is done.





The results indicate that the best ratio between System Throughput and Response time is obtained for a number of jobs in the system equal to 3 *

Part 3 – Analysis

Some Additional Observation

- The choice of Matlab for the final analysis was made to have clearer data visualization, the same analysis is possible using JMT and leads to the same results.
- * The result of the project depends on the performance of the stations, indicated with probability distributions. Although very high confidence intervals (0.99) have been chosen, due to the random component, repeating the simulation could lead to different results, arriving in some (rare) cases to have an optimal number of jobs N=2 (with performance very close to N =3). To avoid this phenomenon, a seed for random events was chosen in order to make the simulation reproducible.
- A basic analysis to find the bottleneck of the system has been done. Analyzing the utilization of the stations for n=3 it is noted that software and specification stations have significantly greater utilization than the others. Increasing the number of jobs in the system, this difference increases arriving at a utilization tending to 1 for the 2 stations above. This data is a good starting point to understand how to improve the system, indicating what could be its bottleneck