M. Brusamento, G. Donetti

Google Cluster-Usage Traces: Jobs and Tasks Duration Analysis

Dip. Elettronica, Informazione e Bioingegneria
Politecnico di Milano, Italy

{mattia.brusamento, gianmarco.donetti}@mail.polimi.it

Summary

I. INTRODUCTION 3
II. DURATION OF JOBS 4
A. May 2nd, 2011 4
B. May 2nd, 2011 with Granularity of 1-hour 6
C. May 2nd, 2011 with different Granularities 8
D. May 3nd, 2011 9
E. From May 2nd to May 3rd, 2011 11
III. LONG JOBS ANALYSIS 13
IV. DURATION OF TASKS 15
A. May 2nd, 2011 15
B. From 19:00 EDT - May 1st to May 2nd, 2011 16
V. FITTING A PARETO DISTRIBUTION 17
VI. CONCLUSIONS 19

Abstract **– Nowadays the quantity of data generated from most various sources is exponentially growing. Recently, the colossus of the Silicon Valley, Google, decided to make publicly available information concerning the usage of a Google compute cell. A cell is a set of machines, typically all in a single cluster that share a common cluster-management system that allocates work to machines. The goal of Google was to make visible many of the scheduling complexity that affect Google's workload and to give a starting point to new researches. In this work, we analyse a little part of the available data just to have a first insight of jobs and tasks duration.**

I. INTRODUCTION

The analysis carried out by us is just a little piece of a bigger puzzle, which should be considered in order to have a more complete view and more reliable information. In any case, we have already succeeded in extracting some interesting results. Let's summarize what does the data concern. The whole set of traces, provided in CSV files, gives a wide range of information, but we focused just on two traces: the one with jobs' events and the other with tasks' events. The traces span 29 days, from 19:00 EDT - May 1st, 2011, but our analysis is reduced to the very first days, until May 3rd, 2011. In particular, our attention has been on the duration of jobs and tasks.

II. DURATION OF JOBS

A. May 2nd, 2011

The first analysis we carried out regards May 2nd, 2011. In the day we analyse, 61172 different events involving jobs happens. These events regard 20790 different unique jobs. Analysing the whole day as a unique window, we have that:

- 12192 jobs start and finish here ("good jobs")
- 58 jobs have only the finish events here (they should have started before the analysed window)
- 8040 jobs end their execution in a different way (evict, fail, kill, lost)
- 501 jobs begin their execution here but no end-event is present in the analysed window

Now, we can study the response time of those jobs that has both the start and the finish event in our window of analysis. Below, a normalized histogram referring to the distribution of the jobs duration.

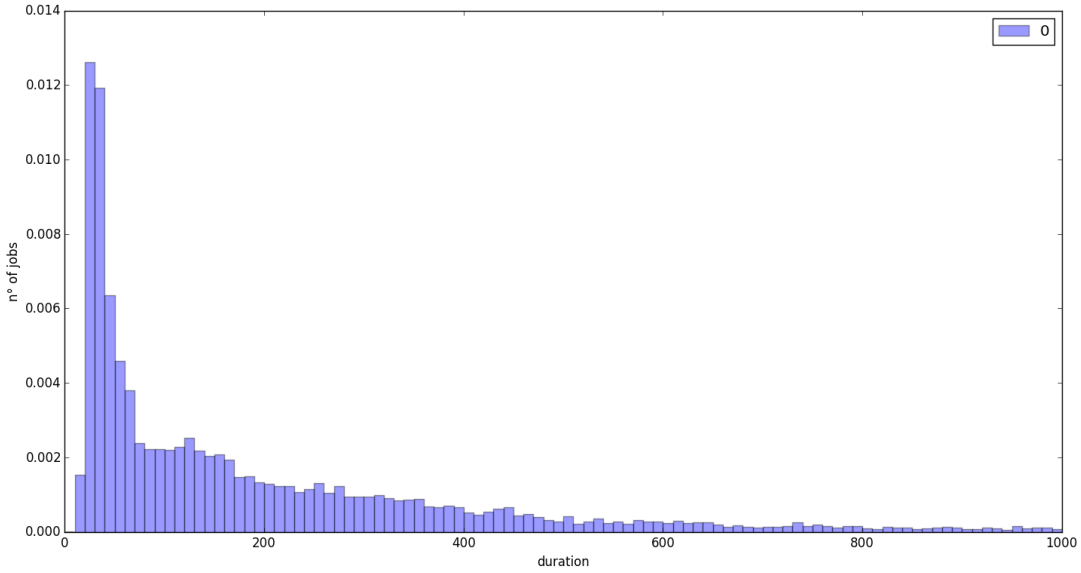


Figure 1. Histogram of jobs durations in May 2nd

The histogram give us a first idea of the distribution of the duration of jobs. It is possible to see that there is a peak around 20~40 seconds of duration and there is a very fat tail, typical of distributions with a high variance, such as Pareto. It is important to say that the x-axis is cut at 1000 sec, but the fat tail is longer.

Some statistical indicators for these response times:

- max: the largest response time is about 57344.96 seconds (very close to 16 hours)
- min: the "luckiest" job has a response time of 14.17 seconds
- **mean**: the sample mean is something around 574 seconds (near to 10 minutes)
- **median**: the durations in the middle of the distribution are of 133.46 seconds
- variance: the distribution variance is 4854075
- **standard deviation**: the standard deviation from the mean value is 2203.2 seconds

Another important thing to say is that the maximum duration is limited by the one-day window, indeed the theoretical maximum possible duration is 24 hours.

In the next page, the box plot can help us visualizing the information. The unmodified box plot shows us the real nature of the distribution, with an enormous quantity of points very far from the Interquartile Range (IQR), while in the zoomed one we are able to notice mean, median and the IQR.

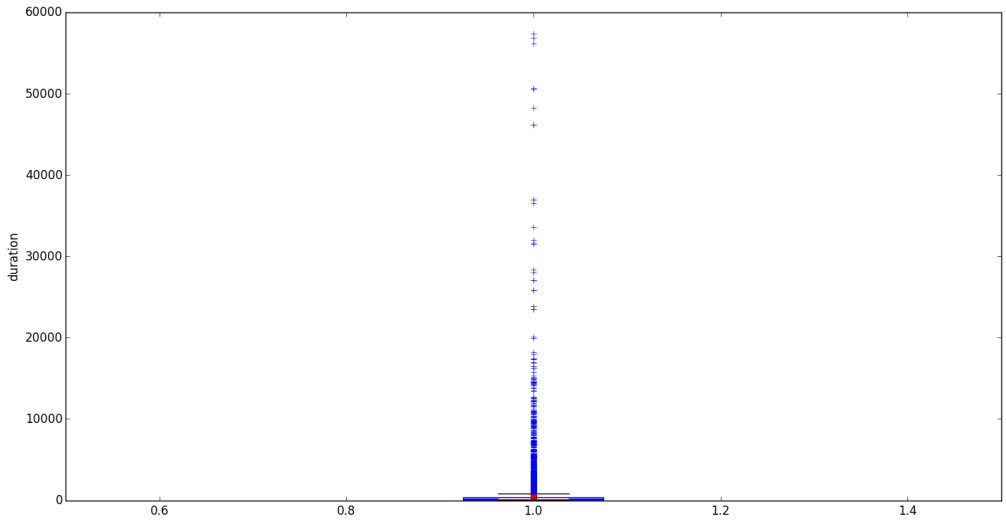


Figure 2. Original box plot of jobs durations (May 2nd)

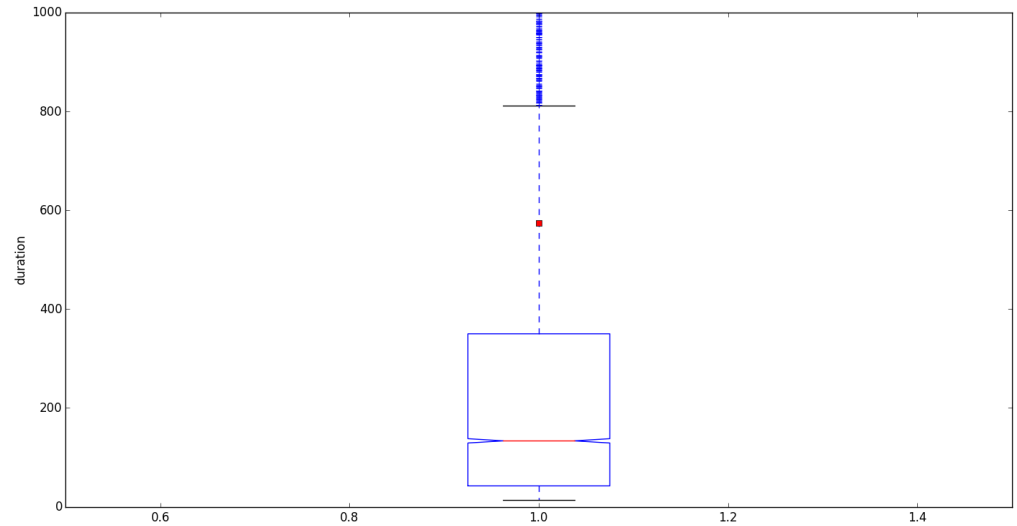


Figure 3. Zoomed box plot of jobs duration (May 2nd)

B. May 2nd, 2011 with Granularity of 1-hour

With a first idea of the distribution, we now move to an analysis that considers different window sizes, in order to see if there are interesting behaviours. Here two approaches have been adopted:

- The first one is to divide the data frame in window and then analyse each window as an independent one (in this case, a job that starts at 2:15 and ends at 3:30 is considered not ended in the window)
- The second one is to compute the duration even if the job ends in another window.

With the first approach, the maximum duration of a job is the window size. Here we can see the mean oscillating around 4 minute, with a much lower median (around 2 minute). Even though we limit the duration to the window size, the high amount of variance pushes the mean almost out of the IQR.

It is also important to notice the proportion of jobs that finish in the window and the ones that does not finish in the window. This to be aware that many jobs will last longer.

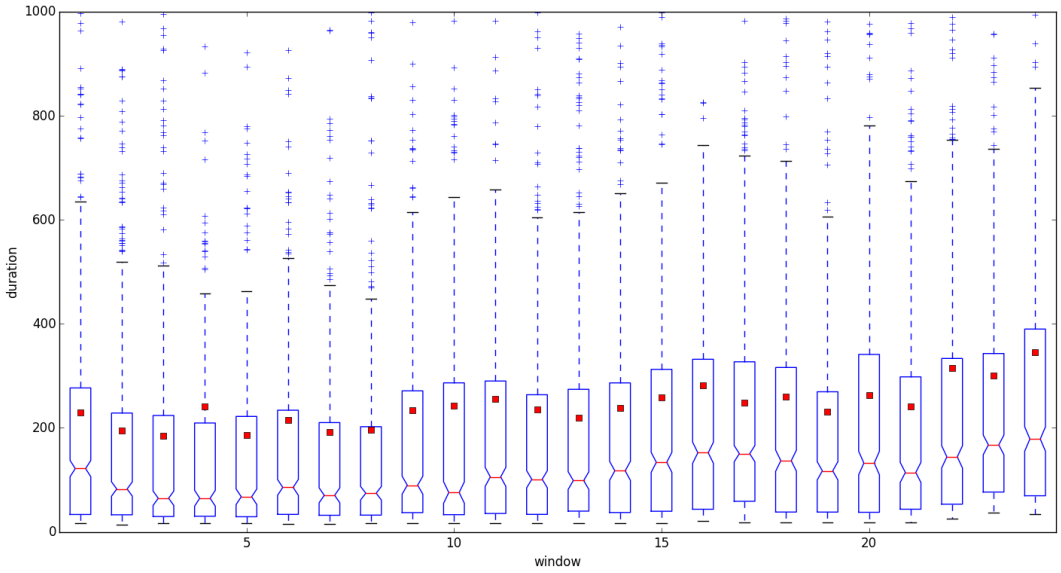
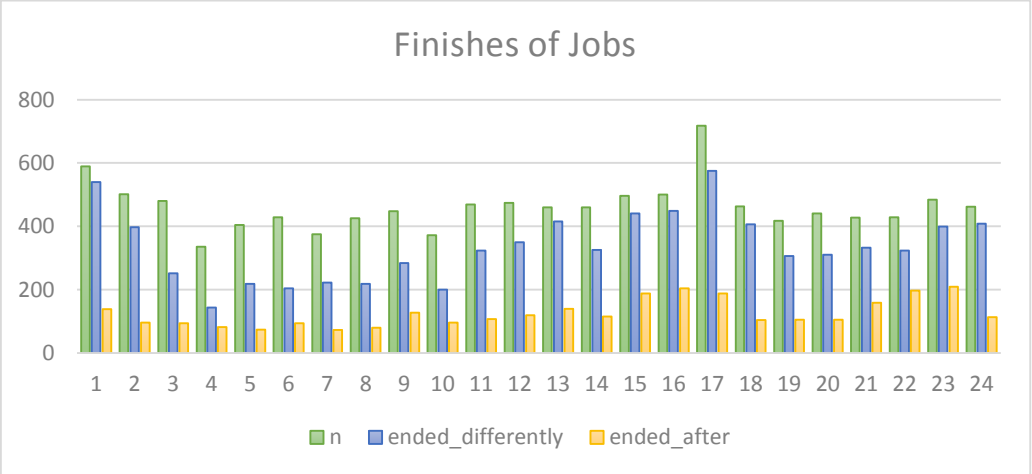


Figure 4. Jobs duration of May 2nd with the first approach and 1-hour window



It is possible to see the difference between the two approaches: here the effects of the fat tail of the distribution are more evident, even if the most of the jobs have again a short duration and the median does not change too much.

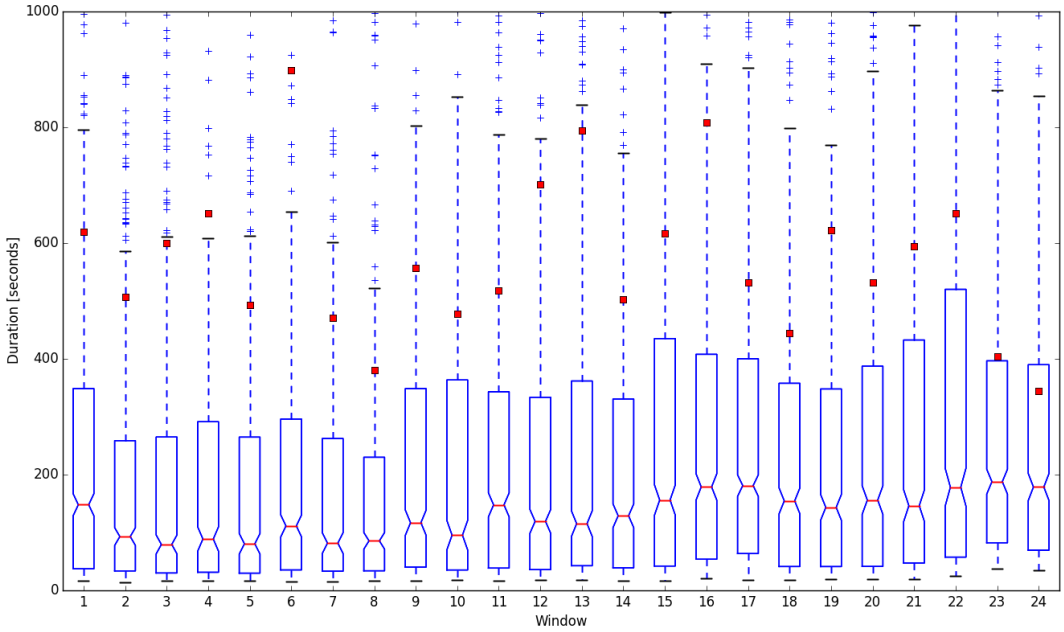
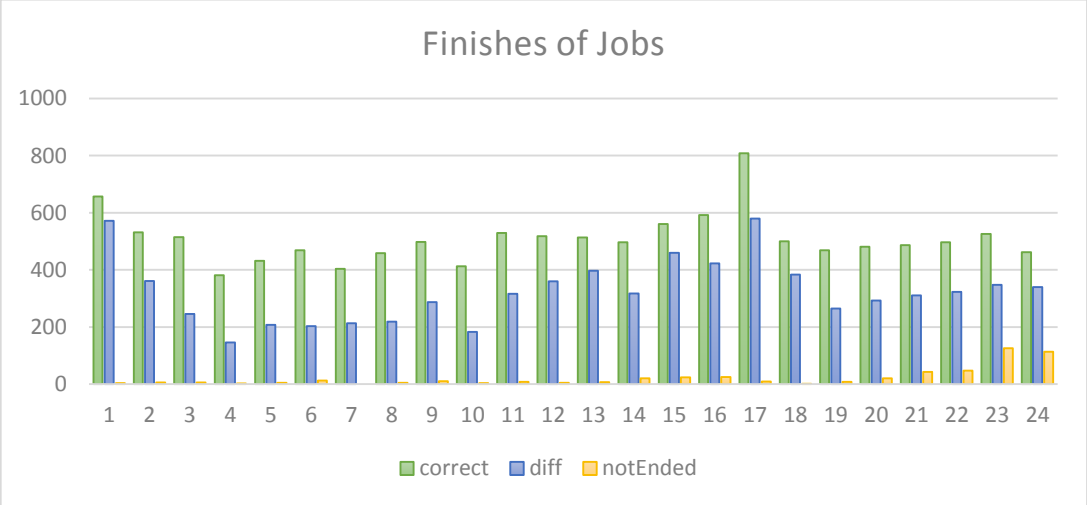


Figure 5. Jobs duration of May 2nd with the second approach and 1-hour window



In the above graph, it is possible to notice that with the second approach, the number of jobs that finish well are slightly higher and the ones that finish after the day are lower. So we can say that the second approach give us a more complete view in terms of summary statistics.

In both the cases, there is a very high percentage of jobs, which ends in unexpected way.

C. May 2nd, 2011 with different Granularities

In the previous case we looked at a window size of 1 hour, now we will have a look to a couple of smaller window sizes. Since the size is smaller, we report the results obtained with the second approach, since with the first many jobs will not end in the window and the data are going to be quite flat.

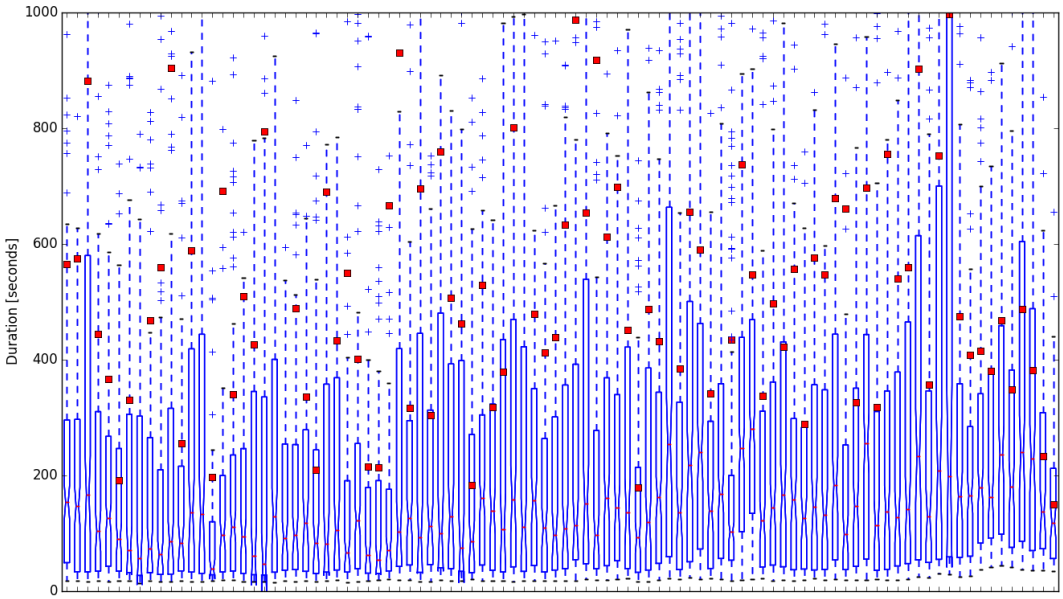


Figure 6. Jobs duration with 15-minutes window

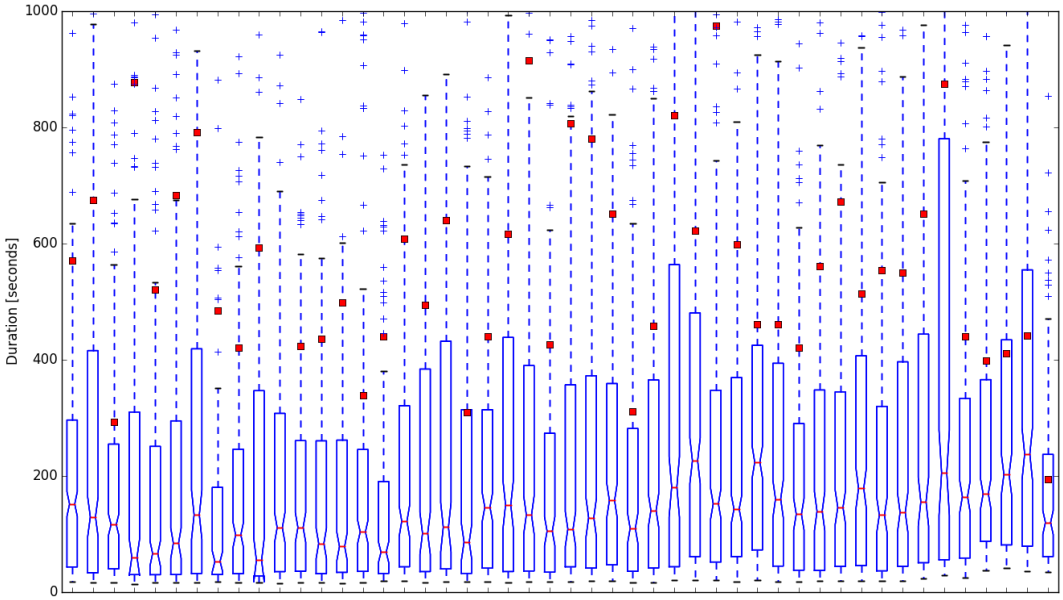
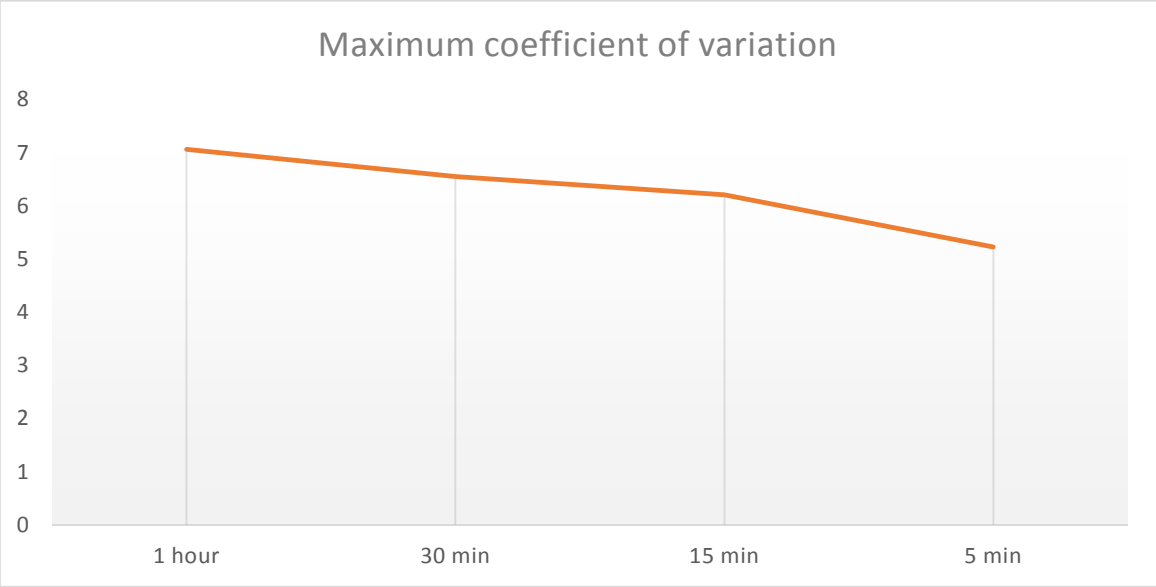


Figure 7. Jobs duration with 30-minutes window

It is possible to see that in any case the median stays low, while the mean oscillates a lot.

We have analysed also other granularities, such as 5 minutes, but with such a small granularity, aggregated measures like mean and median starts to loos significance.

We have also done a little analysis of the evolution of the coefficient of variation with different granularities.



It is quite evident that the maximum coefficient of variation, goes down quite linearly, so there is not an ideal window size, considering also that very small window sizes does not have any sense.

D. May 3nd, 2011

Now, we analyse the day after, May 3rd, in order to do a fast analysis with reference to the previous day of study. Here 70346 different events involving jobs happens. These events regard 24188 different jobs. Analysing the whole day we have that:

- 13487 jobs start and finish here ("good jobs")
- 125 jobs have only the finish events here (they should have started before the analysed window)
- 9964 jobs end their execution in a different way (evict, fail, kill, lost)
- 612 jobs begin their execution here but no end-event is present in the analysed window

Now, we can study the response time of those jobs that has both the start and the finish event in our window of analysis.

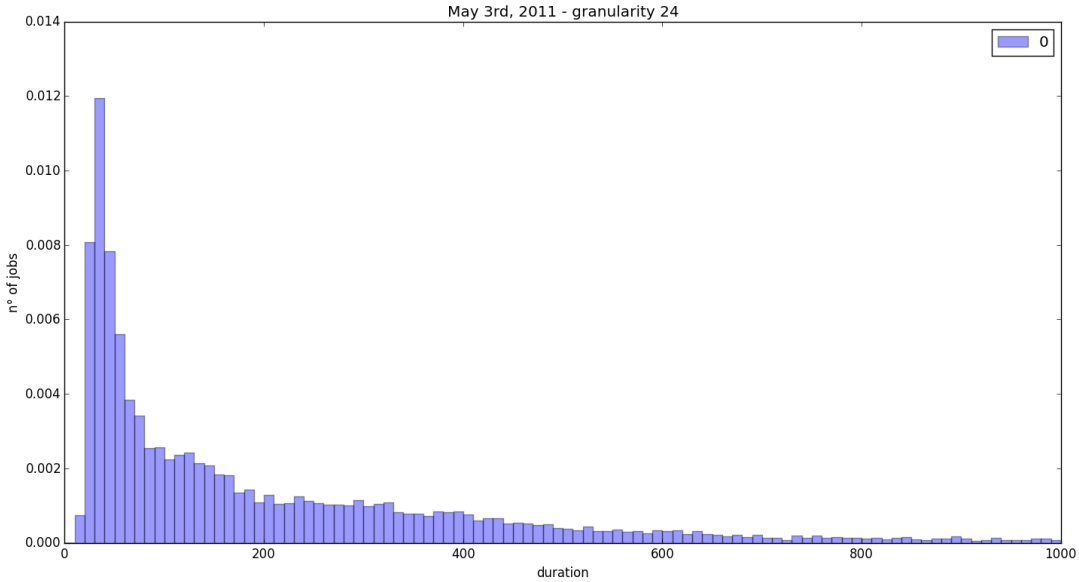


Figure 8.Histogram job duration May 3rd

It is possible to see that there is a peak around 20~40 seconds of duration and there is a very fat tail, very similar to the previous one.

Some statistical indicators for these response times:

- max: the largest response time is about 53085.94 seconds (very close to 16 hours)
- min: the "luckiest" job has a response time of 15.93 seconds
- **mean**: the sample mean is something around 574 seconds (near to 10 minutes)
- **median**: the durations in the middle of the distribution are of 141.57 seconds
- variance: the distribution variance is 3808391
- **standard deviation**: the standard deviation from the mean value is 1951.51 seconds

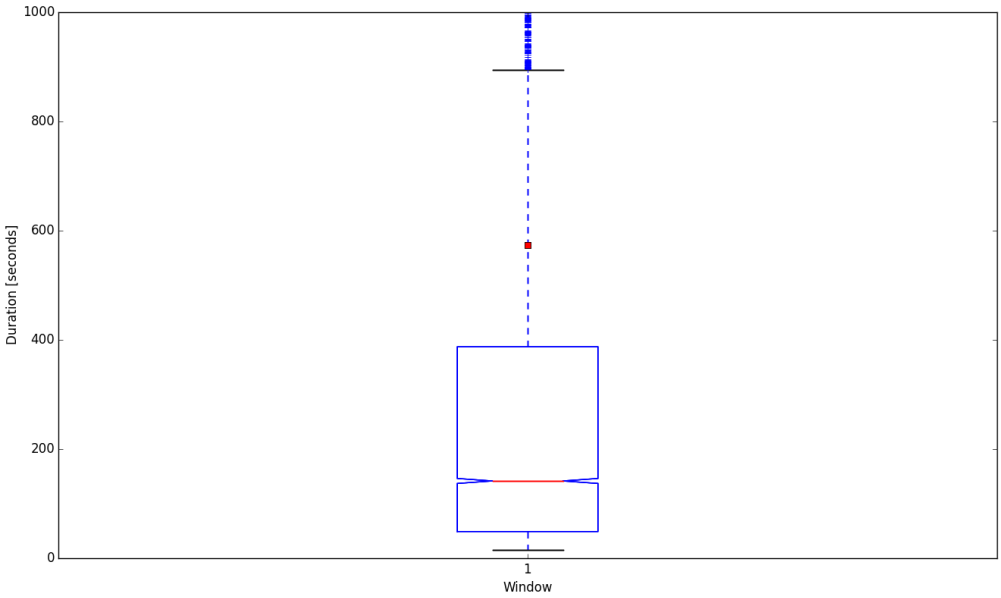


Figure 9. Zoomed box plot May 3rd

The next table summarizes the comparison of the two days of analysis:

window	n	Started_before	ended_differently	ended_after	mean	median	stdDev	var	minimum	maximum
May 2nd	12192	58	8040	501	574,68	133,46	2203,20	4854074,86	14,17	57344,96
May 3rd	13487	125	9964	612	573,59	141,57	1951,51	3808391,44	15,93	53085,94

It is quite evident that there are no big differences between the two days.

E. From May 2nd to May 3rd, 2011

Now, we try to analyse the two days studied before as a single window, with a granularity of 48 hours. Of course, 131518 different events involving jobs happens. These events regard 44600 different jobs.
We have that:

- 25803 jobs start and finish here ("good jobs")
- 59 jobs have only the finish events here (they should have started before the analysed window)
- 18004 jobs end their execution in a different way (evict, fail, kill, lost)
- 735 jobs begin their execution here but no end-event is present in the analysed window

Let's now have a look to the box plot and to some statistics:

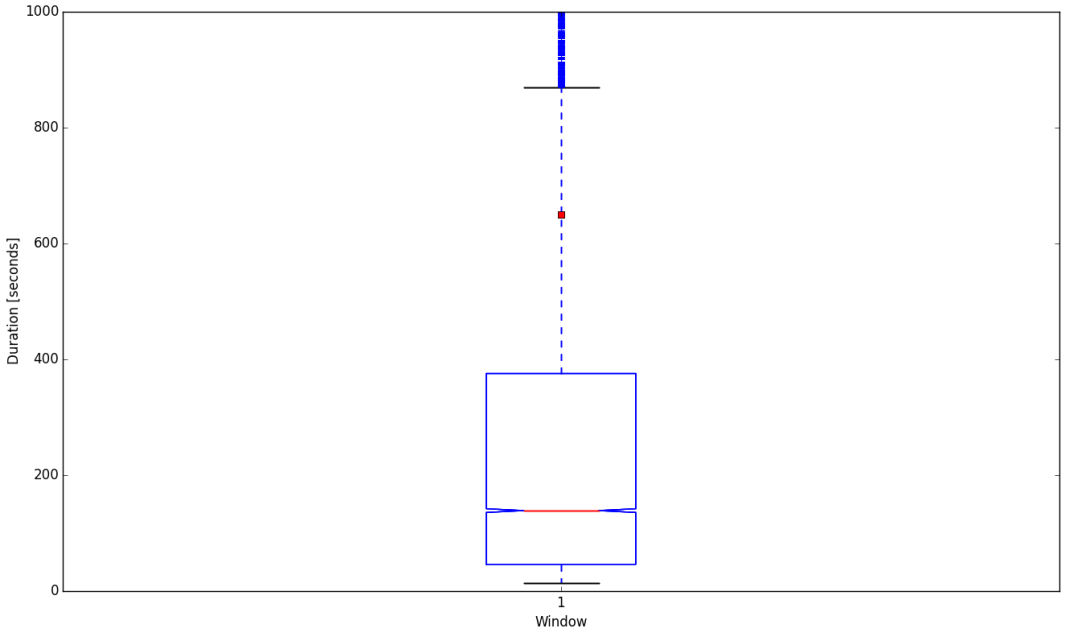


Figure 10. Boxplot from May 2nd to May 3rd

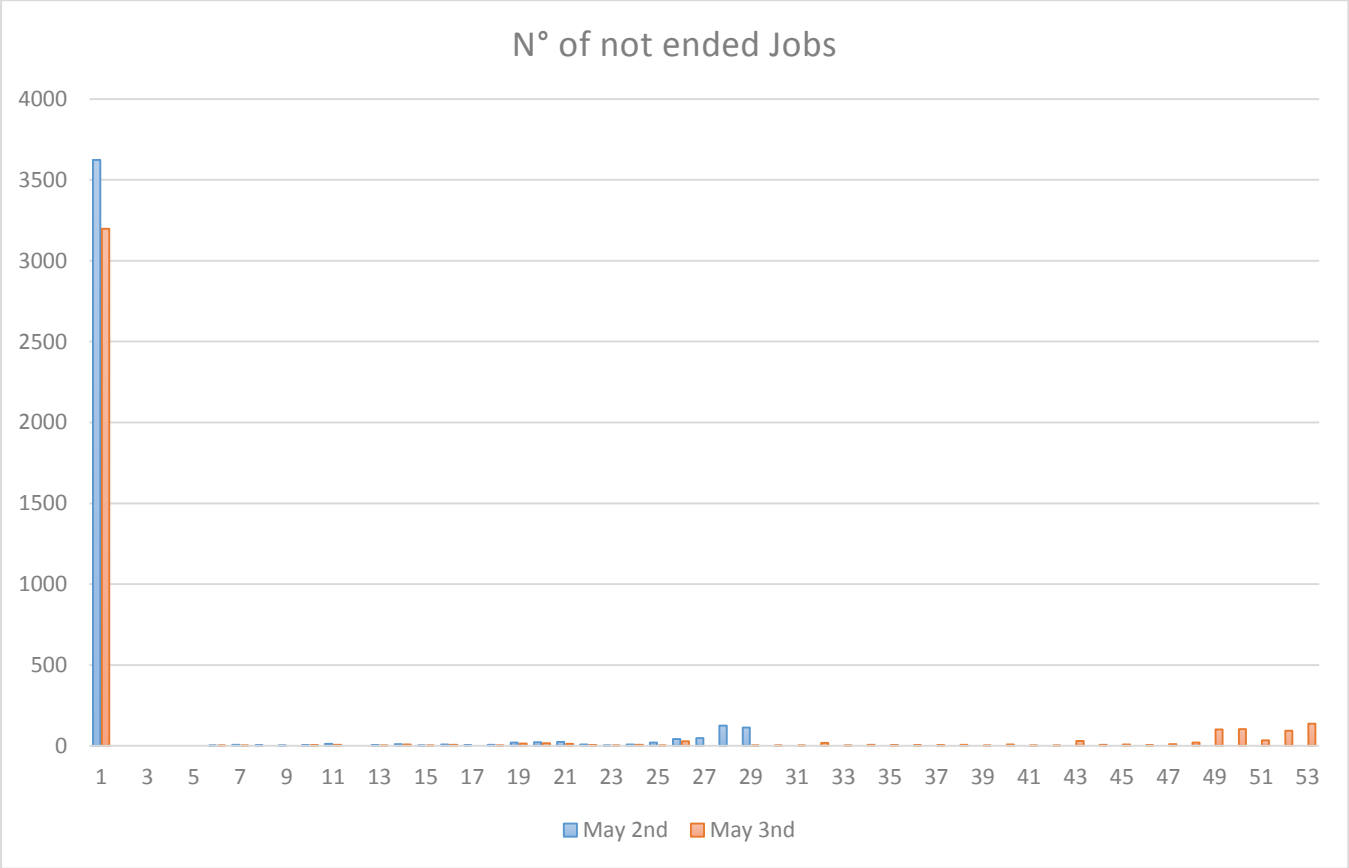
window	n	started_before	ended_differently	ended_after	mean	median	stdDev	var	minimum	maximum
May 2nd-3rd	25803	59	18004	735	650,19	138,93	2745,01	7535068,89	14,17	105128,11

It is interesting to see that there is a visible increase in the mean, the standard deviation and of course in the maximum. This is because there are few very long jobs of May the 2nd that ends in May 3rd and give their contribution. Probably the more we enlarge the window, the larger the mean, the std. deviation and the maximum become. While looking at the median and at the IQR, it is possible to see that they are not pushed so high. This is because the majority of jobs still have a low duration, what changes is the presence of jobs with very large duration.

III. LONG JOBS ANALYSIS

Since an important role is played by very long jobs, it could be interesting to have look at them. It has been noticed that many jobs start in the considered window but end far in the future. Moreover they have been analysed the days of May 2nd and 3rd, but at the beginning of the Google traces, there are a lot of jobs that were already in execution before 19:00 EDT - May 1st, 2011 and it is not known when they started.

Now we will look at the quantities of jobs that does not end within one or two days. The histogram below shows the number of jobs that does not end within the considered window. In both the cases, the starting point of the window is 19:00 EDT - May 1st, 2011, but in one case, it ends with May 2nd and in the other with May 3rd.



The graph above shows a lot of interesting information:

- There is a big number of jobs that started before the trace that does not end within two days;
- Most of the jobs that does not end within May 2nd, does not end neither within May 3rd.

This means that most of the workload is generated by very big jobs that have an enormous duration and that accumulates keeping the resources busy.

As always, a box plot can show us some additional information.

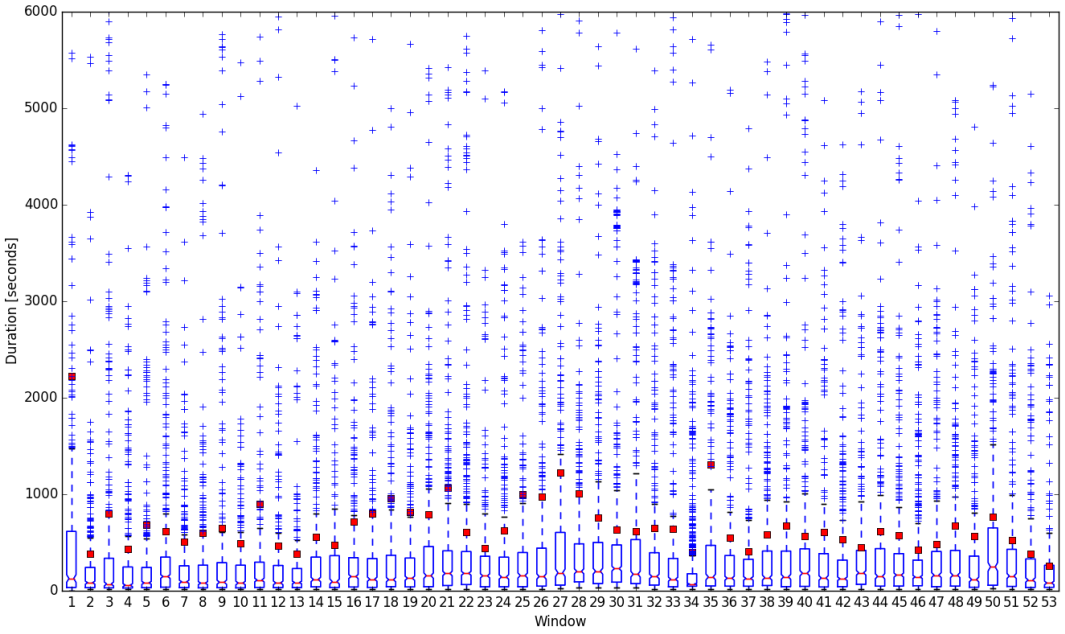


Figure 11. Box plot of jobs duration from the start of the trace to May 3rd. Granularity: 1 hour.

As it is possible to see, the first window, which contains the jobs already started in the past, as a considerable higher mean (around 2228 seconds, so 37 minutes). This means that those jobs, started before and ended within 2 days, give a high contribution. Not only this, indeed we know that those jobs started before, but we do not know how far in the past, this lead us again to the conclusion that there is a big quantity of jobs with a very high duration.

IV. DURATION OF TASKS

Let's move now to the finer grain analysis of the duration of tasks. In this case, the amount of data is much bigger, since there are much more tasks. We will concentrate to a narrower analysis, without enlarging the window to May 3rd.

A. May 2nd, 2011

The first analysis we carried out regards May 2nd, 2011. Here some quantitative information:

- 496208: tasks that finished well within the day
- 93657: tasks that did not finish within the day
- 2244363: tasks that ended in a different way

The first thing we notice is that there is a huge number of events such as kill, lost, evict or fail with respect to the number of correct ends.

A deeper analysis on the distribution shows that tasks have a very different behaviour with respect to jobs.

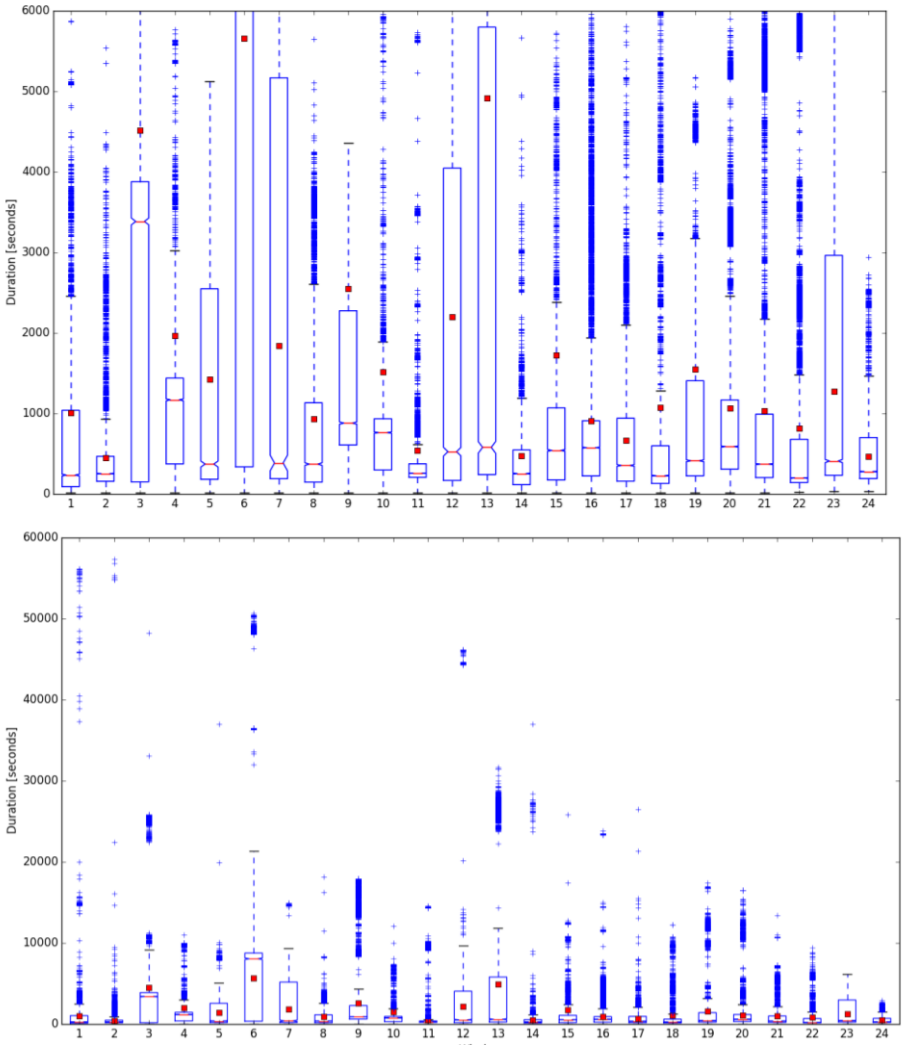


Figure 12. Tasks duration May 2nd with granularity of 1 hour and different zooms.

In this case, we must zoom out a little bit more because there are many variations from an hour to another, moreover the mean is pushed higher and some IQR are stretched towards higher values. In particular, the overall mean is around 30 minutes (against the 10 minutes of jobs).

This behaviour could be explained in the following way: the jobs that last the most creates a higher number of longer tasks.

B. From 19:00 EDT - May 1st to May 2nd, 2011

If we extend the analysis to the starting point of the traces, we can see that there is no more the same behaviour of the jobs. There is no such an anomaly in the first time window.

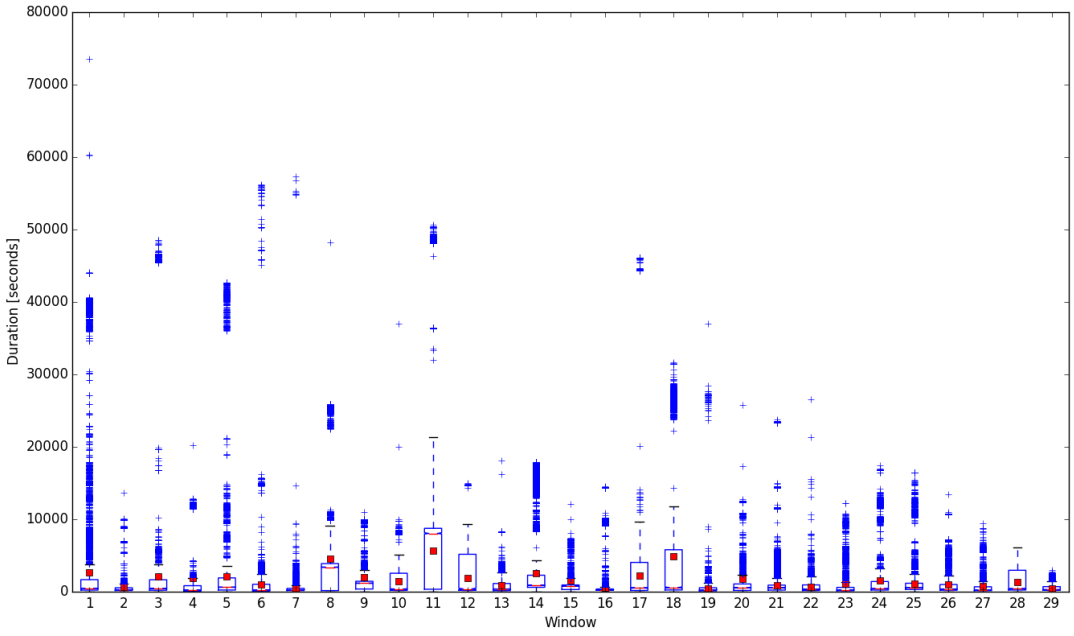


Figure 13. Tasks duration from start to May 2nd with granularity of 1 hour

Looking at the box plot, we can conclude that, differently from the jobs, it is possible to take a better picture of the duration of tasks. This because the presence of some very long tasks is smoothed by the fact that there are already many tasks with quite a long duration. As we can see, the first window is not pushed higher than the others from as before.

V. FITTING A PARETO DISTRIBUTION

The conducted analysis lead us to the conclusion that the duration of jobs is similar to a Pareto distribution.

The probability density for the Pareto distribution is:

$$p(x) = \frac{am^a}{x^{a+1}}$$

where a is the shape and m the scale. Remember that for a less than or equal to 1, the distribution has infinite mean and variance. In our case, we used a equal to 0.3 and m equal to 29. The following histograms show the comparison between the real values (normalized) and values of a Pareto distribution for the day of May 2nd and for May 2nd and 3rd together.

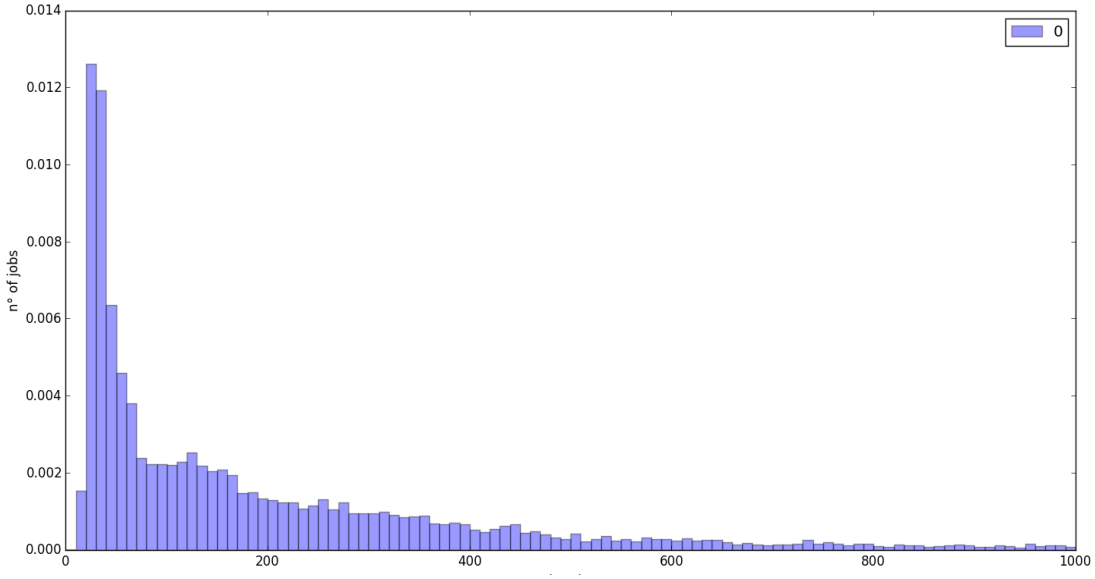


Figure 14. Histogram of job duration May 2nd

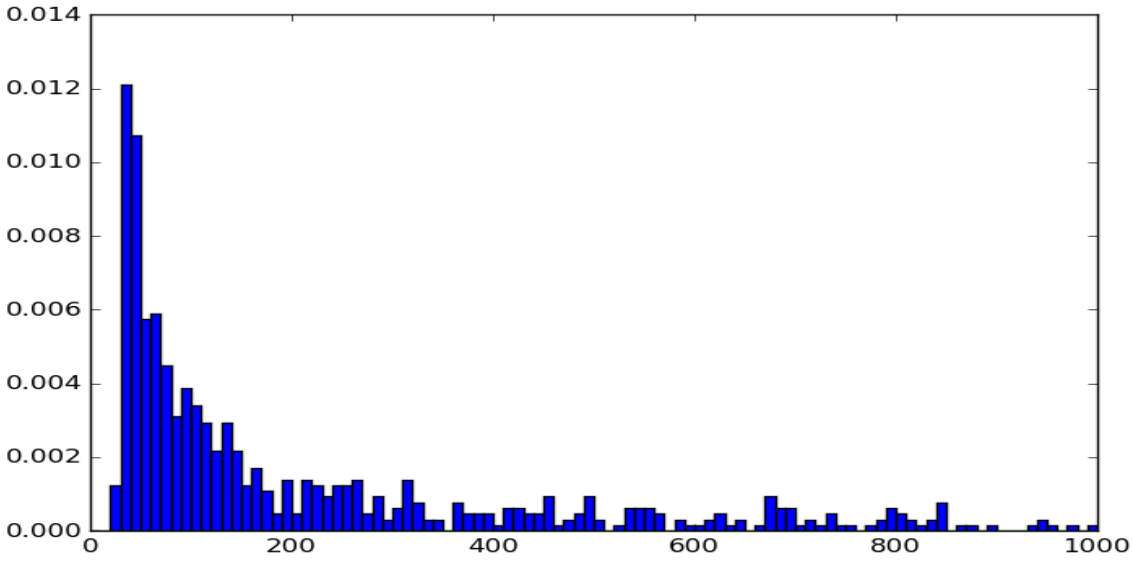


Figure 15. Pareto distribution

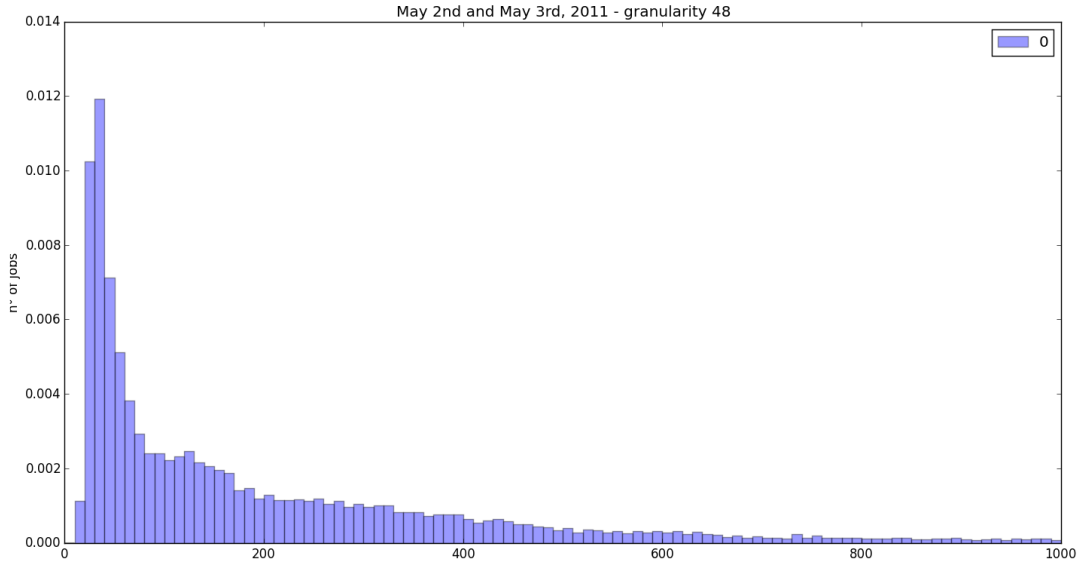


Figure 16. Histogram of job duration May 2nd and 3rd

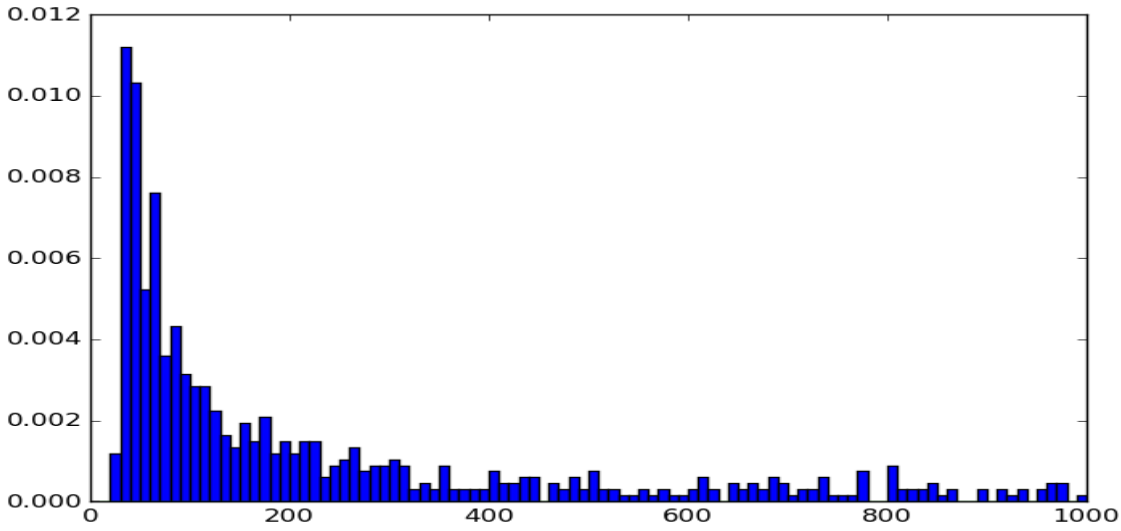


Figure 17. Pareto distribution

As it is possible to see, the behaviour of our data is quite similar to the one of the Pareto. The tail of the distribution of our data decreases more than the one of the Pareto, but it is crucial to remember that we analysed just a window and that many jobs did not ended there. In other words, we do not see such a fat tail because many long jobs were cut.

VI. CONCLUSIONS

The analysis of the duration of jobs and tasks in the Google cluster-usage traces allow us to obtain some interesting result. Even if the focus was just on a little time window of one or two days, it has been possible to notice some nice behaviour. It will follow a brief summary of the main result obtained and previously discussed.

It is interesting to notice that there are a big number of jobs with a short duration, but these jobs cover a little portion of the workload of the cluster. There is also an important number of jobs with a huge duration. Many of them started even before 19:00 EDT - May 1st, 2011 and do not end within 2 days. Probably the biggest part of the workload is generated by this kind of jobs and by the big number of jobs that do not finish as expected.

The analysis moved then to the duration of tasks. This showed that the number of tasks with high duration is much higher, so long jobs produces many long tasks. Another important observation is that also here there is an enormous quantity of tasks that ends in unexpected ways (killed, evicted, lost, failed), so a big part of workload is caused by those.

Finally, the comparison between our distributions and a Pareto showed some similarities and we can conclude that the duration of jobs follow a distribution with a high coefficient of variation.