A Study of Surfnet Tickets and Hardware Logs

Jorik Oostenbrink Fernando A. Kuipers

15 December, 2017

Contents

Chapter 1

Introduction

SURFnet operates a network connecting Dutch research and education institutions to each other and to other networks. Currently, they are using an issue tracking system called JIRA to keep track of problems and incidents.

At the moment, SURFnet only uses their (problem and incident) tickets to keep track of, and keep their employees in the loop of, current problems and incidents. They are interested in the possibility of analyzing past tickets to gain insight in incident trends and properties.

To this end, we have been given access to their archive of JIRA issues. This report contains our analysis of this issue set. The purpose of this analysis was to (1) discover any relevant insights this data can give us and (2) to research how SURFnet's current system can be improved to become more usable for future more exhaustive analysis.

In the context of our geographical research [1], we were interested in obtaining a set of characteristic geographically correlated failures based on past events. These could then be used to compute the vulnerability of the network, and to predict possible future events. For example, by analyzing past failures during maintenance, it might be possible to predict the impact of other future maintenance operations. Unfortunately, it turned out that the current ticket system does not store enough accurate and precise machine-readable information about past failures to make this possible. At least not without a significant time investment.

After presenting our results on the ticket dataset, we were given access to logs of the state of each enabled port of the SURFnet network.

These logs have the potential to give very precise details about when which devices were malfunctioning. Unfortunately, the logs have their own issues which makes a geographical analysis much more difficult.

Although the data was not perfect for our own purposes, we were still able to make use of them to gain some insights into SURFnet's past network issues.

In chapter 2 we present the results of our analysis on the ticket dataset, as well as some recommendations on how to extend the system to be more suitable for analysis. Next, in chapter 3, we discuss our analysis of and recommendations on the port state logs dataset. We give our conclusions and recommendations on future analysis possibilities in chapter 4.

[1] J. Oostenbrink and F.A. Kuipers, Computing the Impact of Disasters on Networks, ACM SIGMETRICS Performance Evaluation Review (special issue from the 1st ACM SIGMETRICS International Workshop on Critical Infrastructure Network Security), vol. 45, no. 2, pp. 107-110, September 2017.

Chapter 2

Analysis of SURFnet Problem and Incident Tickets

Since 2014 SURFnet has used JIRA for issue tracking.

The main use of the ticket system is to keep track of, and keep employees up to date about, current issues and problems with the network. It is not setup to be a monitoring or archiving tool, or to be machine-readable.

Nevertheless, we were able to gain some rudimentary insights into these tickets, and by extension, in the problems and incidents of the SURFnet network.

Section 2.1 contains the results of our analysis, and in section 2.2 we give our recommendations based on this analysis. Section 2.3 gives the results of our analysis of an updated version of the ticket dataset.

2.1 Analysis

Tickets created in their old system were imported as "ARS-S Tickets", while new tickets are called "SURFnet Trouble Tickets". We are not interested in all issues, only in tickets. As such, we limit our analysis to these "SURFnet Trouble Tickets" and "ARS-S Tickets". [1]

SURFnet has provided us with their complete dataset of issues spanning from January 2010 to mid-March 2017. We further limited our dataset to only include those tickets which were already closed, and were not declined, canceled, or duplicates. The resulting tickets were filtered on ticket type (keeping only incidents and problems) where possible (not all imported tickets have a ticket type).

The final dataset consists of 3725 "SURFnet Trouble Tickets" and 6209 "ARS-S Tickets", for a total of 9934 tickets.

2.1.1 Analysis of Ticket Creation Times

At the suggestion of SURFnet itself, we start our analysis by studying the creation dates and times of tickets. In this section we consider the amount of created tickets over time, number of tickets created in each month, and number of tickets created in each hour of the day.

Tickets over Time

We have grouped all tickets by creation month. For the purpose of this analysis, the first and last month of our dataset were excluded.

In figures 2.1 and 2.2 one can see the amount of tickets created per month. One can clearly see that the amount of tickets created each month was relatively stable until June 2016. From this month onwards there is a sharp increase in the amount of tickets created. This coincides with, and is probably indirectly caused by, the change in operational network management at June 15th 2016 [2].

Unfortunately, it is difficult to evaluate if the increase is caused by an increase of problems in the network, or if the new network management just makes more use of the ticket system. The amount of tickets does not correspond directly with the amount of problems in the network, or even the amount of

[1] Some ARS-S tickets are incorrectly labeled as "SURFnet Trouble Tickets". In our analysis, we treat these as "ARS-S Tickets", not as "SURFnet Trouble Tickets".

[2] `www.surf.nl/nieuws/2016/06/succesvolle-overdracht-operationeel-netwerkbeheer.html`

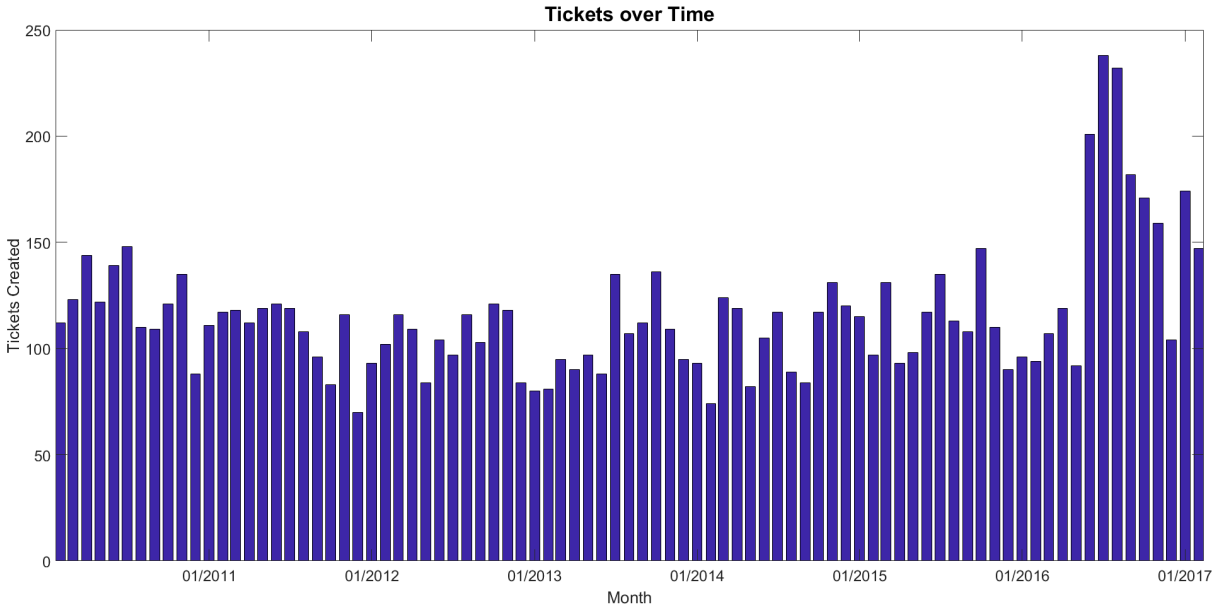

Figure 2.1: Tickets created per month.

issues raised by customers. For example, even after filtering on ticket type, the dataset contains numerous test tickets, tickets about internal issues (e.g. with JIRA itself), and requests (e.g. to test if an e-mail address change has been properly processed)

Splitting tickets by impact should give at least some insight in what kind of tickets have seen an increase since June 2016. In figure 2.3 tickets creation amounts are plotted separately for each impact category [3].

Typically, impact categories range from P3 (small or no impact) to P1 (large impact). However, issues with an international impact and major outages get assigned to "P1 Internationaal" and "Grote Storing" respectively.

While all impact categories saw an increase in tickets around June 2016, the largest increase by far was that of the P3 tickets, the lowest impact a ticket can be assigned. Furthermore, since 2010 P2 tickets have been slowly decreasing in frequency. This could imply SURFnet has been improving robustness somewhat, but it could also mean that tickets get assigned to other categories more often instead.

Interestingly, P3 tickets have seen an increase in frequency since well before June 2016, this increase has been mostly compensated by a decrease in other impact categories.

Monthly Ticket Frequency

It might be interesting to know which months see an increase or decrease in ticket frequency. For example for more efficiently assigning resources.

Table 2.1 shows tickets frequencies for each month. While there clearly are months with more issues (e.g. July) and months with less issues (e.g. December), there does not seem to be a clear pattern.

In table 2.2 we only considered high impact tickets ("P1", "P1 Internationaal", or "Grote Storing"). December has a relatively large amount of high impact tickets compared to the overall amount of tickets. Interestingly, July has both a very high percentage of overall tickets and high impact tickets.

There is not much we can conclude from these tables. Although it is clear that July is a month in which many issues occur. Additionally, in 2016 the summer saw a much larger amount of tickets than one would expect based on historical data.

Hourly Ticket Frequency

Finally, we consider how many tickets get created on average for each hour of the day. As can be seen in figures 2.4 and 2.5 this mostly coincides with working time.

[3]Some ARS-S tickets are not assigned an impact category, thus before 2015 the sum of tickets over all impact categories is not necessarily equal to the total amount of tickets.

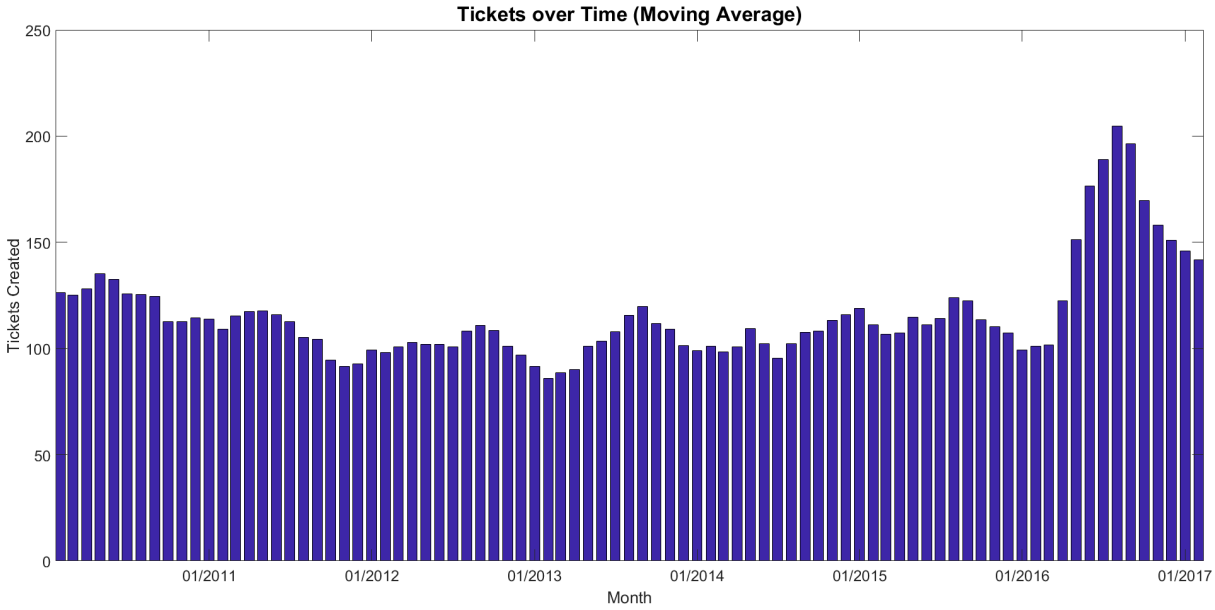

Figure 2.2: Central moving average (with a window size of 5 months) of tickets created per month.

Table 2.1: Percentage of tickets created in each month

Month	**2010-2016**	**2010-2015**	**2016**
January	7.25%	7.68%	5.35%
February	7.06%	7.45%	5.30%
March	8.48%	9.04%	5.98%
April	8.18%	8.53%	6.61%
May	7.26%	7.72%	5.24%
June	9.09%	8.62%	11.16%
July	10.30%	9.59%	13.44%
August	9.11%	8.22%	13.04%
September	8.22%	7.80%	10.08%
October	9.25%	9.25%	9.28%
November	9.08%	9.14%	8.77%
December	6.73%	6.95%	5.75%

2.1.2 Resolution Times and Open Tickets

In this section we consider the time it takes to resolve a ticket after its creation. We measure this by comparing the "resolutiondate" of a ticket with its creation time. We filter out all ARS-S tickets, as this field has only been filled in for ARS-S tickets with a very long resolution time (resolved after the transfer to JIRA). Out of 3725 remaining tickets, 3697 are assigned a resolution date.

Table 2.3 shows the median and maximum resolution times by impact category. A typical resolution time is 16 hours. Median resolution times are lower for non-P3 tickets, with the exception of "P1 Internationaal" and "Grote Storing" tickets, presumably because these are more difficult to resolve. For example, "P1 Internationaal" issues involve more international cooperation and often require SURFnet to rely on another party to resolve an issue (i.e. sub-oceanic fiber cuts).

It is surprising how many tickets have a resolution time of more than 30 days. Most of these tickets are "P3" tickets, and probably had such a low impact other tickets were given a higer priority. However, 3 "P1" tickets took more than 30 days to resolve.

Note that this means there is a distinct possibility that a ticket opened more than a few months ago is still open. We filter out all open tickets. This means that we are underestimating the number of tickets created in some months (especially those in 2017) in section 2.1.1.

In addition, we are also underestimating median (and potentially maximum) resolution times, as recent tickets with a potentially high resolution time have not been closed yet.

Table 2.4 shows the number of open tickets. Most of the 64 open tickets currently have a "Waiting"

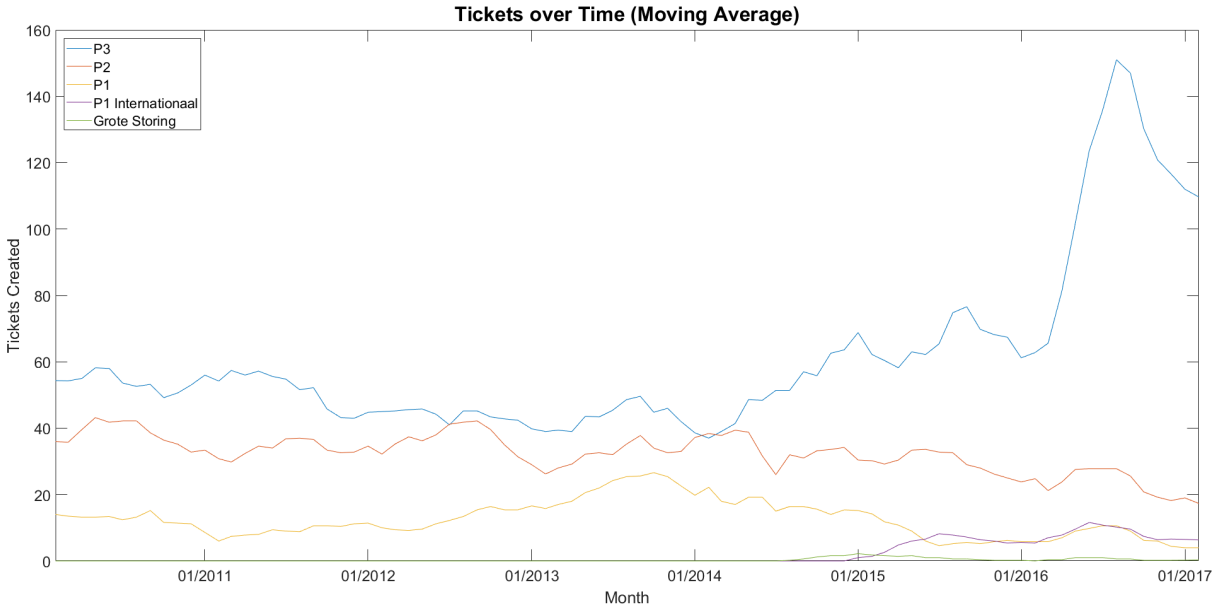

Figure 2.3: Central moving average (with a window size of 5 months) of tickets created per month per impact category.

Table 2.2: Percentage of high impact ("P1", "P1 Internationaal", or "Grote Storing") tickets created in each month

Month	2010-2016	2010-2015	2016
January	6.85%	6.95%	6.28%
February	7.65%	7.89%	6.28%
March	7.41%	7.61%	6.28%
April	8.21%	8.55%	6.28%
May	7.73%	7.42%	9.42%
June	7.81%	7.14%	11.52%
July	11.39%	10.24%	17.80%
August	9.16%	8.36%	13.61%
September	7.33%	7.52%	6.28%
October	9.08%	9.49%	6.81%
November	9.40%	10.06%	5.76%
December	7.97%	8.74%	3.66%

status. No tickets created before 2013 are still open.

2.1.3 Causes

In this section we try to analyze the causes of tickets. This information can be very helpful in reducing the number of problems in the future. As mentioned in the introduction, the ticket tracking system is not setup to be a monitoring or archiving tool, or to be machine-readable. Unfortunately, this makes processing the causes of tickets very difficult, if not impossible.

Tickets in SURFnet's setup can have two fields which could potentially be of help: "Solution" and "Root cause". Unfortunately, both fields are not mandatory, so many tickets do not include them. Out of all 9934 tickets, only 2707 include a solution field, and only 807 a root cause field.

Employees can freely fill in the solution and the root cause of a ticket themselves in these fields, and there is no standard format they have to adhere to. Thus the fields are not easily machine-interpretable.

Finally, the fields are not always used as they are supposed to. Especially the solution field often contains remarks such as "Probleem is opgelost" (meaning problem has been solved).

It is fair to say that automatically (and accurately) assigning causes to tickets in the dataset is very difficult, or even impossible.

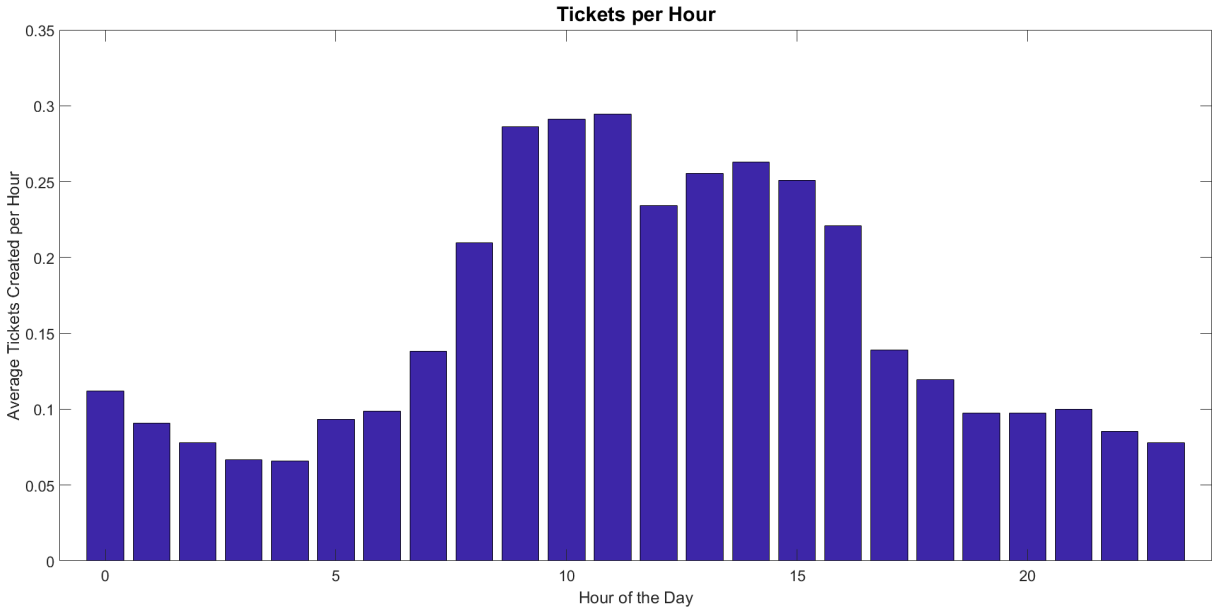

Figure 2.4: Average amount of tickets created per hour per day.

Table 2.3: Resolution Time in Hours

Impact	**Median Resolution Time**	**Maximum Resolution Time**	**# > 30 Days**
P3	18	11496	177
P2	9	5569	19
P1	9	1101	3
P1 Internationaal	20	429	0
Grote Storing	45	455	0
All tickets	16	11496	199

Thus we first consider only tickets assigned the worst possible impact category: "Grote Storing" (Major Failure). This subset only contains 27 tickets, so we can manually assign each of them a cause (category), based on their summary, description, comments, solution, and root cause.

The result can be found in table 2.5. 1 ticket contains no mention of the cause (and is closed as the customer decided to fix the problem himself). 2 tickets are duplicates of other tickets. As mentioned before, tickets which are closed as duplicates or clones are automatically filtered out. Unfortunately, this does not include tickets which were purposefully created to be duplicates (to continue discussing a previously 'resolved' issue).

8 out of 27 major failures have been caused by a single node failure. This seems very high, as one would expect a network to be resilient enough to be able to endure a single failure.

4 of the outages were caused by fiber failures. All fiber failures resulting in "grote storingen" were the result of mistakes by a fiber provider during maintenance or repairs.

Of special interest to us are losses of power, as these kind of failures are geographically correlated. 5 out of 27 major failures were caused by a power cut. Additionally, 3 out of 27 were caused by a loss of power due to a hardware failure at the point of presence.

Almost 30% of "grote storingen" are geographically correlated. This gives an increased incentive to continue our research on this subject.

Loss of power is also perhaps the only category (of those in table 2.5) we can automatically assign tickets to by searching for key words. To get an idea of the total amount of issues caused by a loss of power we counted all tickets containing one of the following keywords: "powerfailure", "powercut", "poweroutage", "powerloss", "blackout", "stroomstoring", "stroomuitval", "power failure", "power-failure", etc.

Table 2.6 shows the amount of 'power loss tickets' per impact category. As these tickets have been labeled automatically the result is probably not completely correct, but it should be fairly accurate. In particular, we can see that 8 out of 27 "Grote Storing" tickets have been labeled as being caused by a loss of power. This matches the results of our manual cause assignments.

With the exception of "P1 Internationaal", higher impact categories contain a larger percentage of

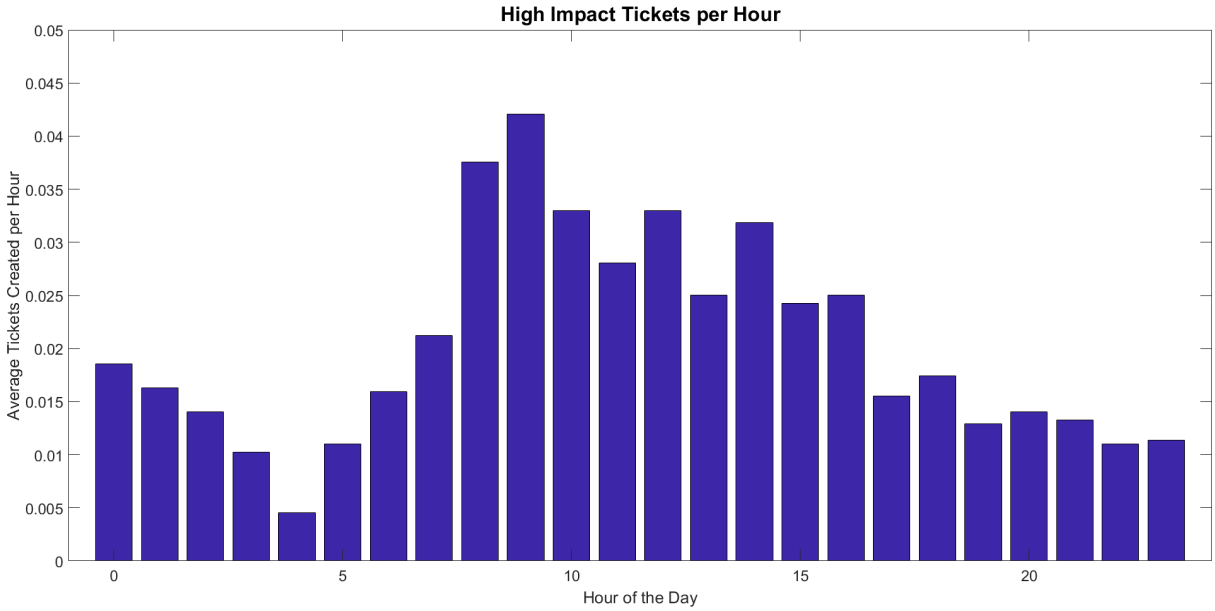

Figure 2.5: Average amount of high impact ("P1", "P1 Internationaal", or "Grote Storing") tickets created per hour per day.

Table 2.4: Open Tickets

Creation Date	Open Tickets
2013-2017	64
2013-2015	5
2016	27
01/2017	9
02/2017	6
03/2017	17

tickets caused by a loss of power. This is to be expected, as a loss of power generally affects a lot of hardware, and thus has a large impact on the network. 45.54% of tickets caused by a loss of power have an impact of "P1" or worse.

We can conclude that power outages have had a significant impact on the SURFnet network. As many as 15.88% of high impact tickets ("P1", "P1 Internationaal", or "Grote Storing") have been caused by a loss of power.

2.1.4 Hardware and Sites

Each SURFnet site and network device is assigned an unique ID. The ID of failed hardware, or SURFnet site(s) where failures are occurring, is required to do in-depth geographical analysis of tickets. In particular, this information is necessary for predicting future events or computing the vulnerability of the network. Additionally, information on hardware failures can tell us which specific type/brand of devices performs better, or worse, than others.

Tickets do not contain a failed hardware field, thus the ID of the failed hardware needs to be extracted from the text of the tickets themselves. Unfortunately, there are a number of problems with this approach:

- Sites and devices are typically mentioned when a connection to them fails (or has lost its redundant path).
- As tickets contain communication between employees and between SURFnet and customers and providers, they also contain references to hardware that has not failed. For example, employees can ask if a certain device is affected, mention that a device or point of presence is not affected or compare the problem to a similar issue of another device.

Table 2.5: Causes of "Grote Storingen" (Major Failures)

	Number of Tickets
Node issues/failure	8
Fiber(s) down	4
Power failure/cut	5
Loss of power because of a hardware failure	3
Human error	2
Third party issues	2
Unmentioned	1
Duplicate Ticket	2
Total	27

Table 2.6: Tickets caused by a loss of power (based on an automatic search for keywords)

Impact	**Caused by a loss of power**	**Total Tickets**	**Percentage caused by loss of power**
P3	120	5287	2.27%
P2	119	2786	4.27%
P1	194	1083	17.91%
P1 Internationaal	2	175	1.14%
Grote Storing	8	27	29.63%
All tickets	448	9934	4.51%

- Tickets can be created when no hardware has failed at all, but still refer to a site or device. In particular, all e-mails sent to the SURFnet helpdesk automatically result in a ticket.

Although these problems can be mitigated, for example by filtering out connection IDs (which contain device/PoP IDs), analyses of hardware logs will still be much more accurate. These logs lack the context tickets can provide, but we have already shown that for example the root cause is very difficult to accurately extract from tickets. As such, at the moment we propose using hardware logs instead of tickets for geographical analysis of past network failures. If possible, combined analysis of tickets and logs might provide more detailed results.

2.2 Recommendations

Tickets are used to keep track of, and keep employees up to date about, current issues and problems. Thus certain choices were made that while useful for their purpose, makes the analysis of tickets less accurate and more difficult. We briefly go over some of the problems we encountered while analyzing the tickets dataset.

Many tickets in the dataset are wrongly classified as problems and incidents. One common example are test tickets, created simply to test if the ticket system functions correctly or to learn a new employee how to use the ticket system. However, the biggest source of misclassified tickets are e-mails sent to the helpdesk. The helpdesk converts all incoming e-mails to trouble tickets. This includes questions from customers, mail from providers and even mail sent by SURFnet employees who preferred contacting the helpdesk directly over creating an issue in JIRA.

This leads us to a related problem. When e-mails are converted to tickets, the subject line is used as the ticket summary. The result of this is that many ticket summaries do not accurately summarize the ticket. For example, the dataset contains a ticket simply summarized as "Storing." (meaning Outage). This type of summary gives insufficient information about the issue itself. Additionally, the e-mail itself did not describe an outage at all (only an authentication issue), and should have been sent to SURFconext instead. In contrast to the previous problem, this is also a detriment to the actual functionality of tickets, as employees need to read the ticket descriptions to figure out what the underlying issues are, instead of simply glancing at the ticket summaries.

Tickets themselves contain all communication towards customers, which can quickly add up and make manually reading through tickets much less manageable. These communications include all e-mail traffic. For example, some tickets contain tens of automated e-mails informing SURFnet that persons they are trying to inform of an issue is currently on holiday. Mail is included as is, and are not made more readable

before being added to the ticket. Because a reply can contain the complete conversation, often the entire mailing history is copied multiple times in the same ticket.

A large issue is the usage of free fields. It is difficult to automatically extract useful information from these fields. In addition, there are not many fields which are obligatory to fill in. This has been a deliberate choice by SURFnet, because they found they often needed to add more options to multiple choice fields, or add more custom fields, as there was always a new situation they did not account for beforehand. Evidence of SURFnet's problems with custom fields can be found by looking through the list of all custom fields. Many fields saw (almost) no use. In addition, some fields are extremely similar to other fields.

Finally, it seems as if filling in tickets is deemed much less important after a problem has been solved, as evidenced by the lack of root causes and solution given in tickets. Keeping in mind the purpose of tickets (tracking current problems/incidents), this is to be expected.

2.2.1 Proposal

We propose introducing a separate system for archiving past incidents for analysis, as archiving past incidents for analysis and keeping each other up to date about current incidents are seemingly incompatible. This system should take up as little time as possible from SURFnet employees, to keep them willing to properly update the system and to prevent wasting time.

The system can be included in the current JIRA setup by creating a new type of JIRA issue. This avoids having to introduce completely new software. We propose creating a new issue after resolving a incident or problem with significant impact. By limiting by impact, we avoid wasting time by creating issues for all small problems and incidents, which might quickly tire out employees. Note that these new issues should only be created for actual problems and incidents, and not for misclassified tickets.

The issues should contain the following fields:

- Related ticket ID(s)
- Impact
- Summary
- Brief description
- Timestamp of the start of the problem
- Timestamp of the resolution of the problem
- Root cause (preferable multiple choice)
- Failed Device/PoP ID(s)
- List of affected connections

Most of this is readily available and easily filled in. The summary and description fields might take some time to fill in. However, these fields are not important for automatic analysis, so they could potentially be omitted.

If possible, relevant hardware logs could be attached to these issues to provide additional information. This can potentially be automated based on device ID and timestamps.

2.3 Updated Dataset

Towards the end of our project, SURFnet provided us with an updated ticket dataset, spanning from January 2010 to end-November 2017. In this section we briefly go through some of the changes in the tickets since March 2017.

First, we reconsider the tickets created over time. In section 2.1.1 we saw that there was a steep increase in the number of created tickets since June 2016. It would be interesting to know if this trend has continued or the increase in tickets was only a temporary peak.

As we can see in figures 2.6 and 2.7 the number of created tickets did not go back to the old level in 2017. To make matters worse, towards the end of 2017 there was another large increase in the number of created tickets.

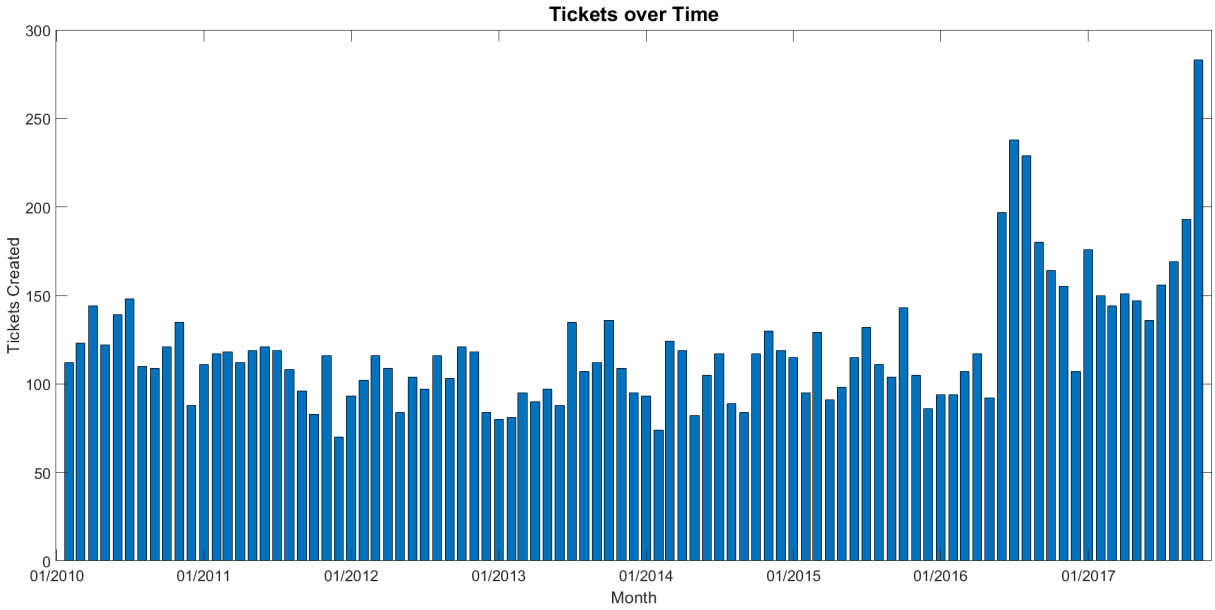

Figure 2.6: Tickets created per month. (November 2017 Ticket Dataset)

Table 2.7: Open Tickets (November 2017 Ticket Dataset)

Creation Date	Open Tickets
2013-2017	64
2013-2015	5
2016	9
01/2017-07/2017	4
08/2017	14
09/2017	5
10/2017	12
11/2017	15

In figure 2.8 we plot the moving average of the created tickets per impact category. Interestingly, the peak in created tickets in 2017 is not only the result of an increase in P3 tickets, but also of P2 and P1 tickets. This points towards the increase being caused by actual problems with the network, instead of simply by an increase in wrongly classified problem and incident tickets, as we would not expect wrongly classified tickets to get assigned an impact category above P3.

If we exclude all P3 tickets (see figures 2.9 and 2.10), we can see there has been a steady decrease in created tickets with an impact higher than P3 since the start of the dataset. However, there are still periods with temporary increases in the number of created tickets, including the summer of 2016.

Next, we take a look at the number of open tickets in the new dataset. Note that as before, open tickets are filtered out of all other statistics we generate.

Table 2.7 shows the number of open tickets in the November 2017 ticket database. Coincidentally, the number of open tickets is still 64. However, while none of the open tickets created in 2013-2015 were closed, most tickets that were open in the March 2017 database have been closed.

Finally, we analyze the number of problems and incidents caused by a loss of power. The number of created tickets caused by a loss of power in the November 2017 database are shown in table 2.8. Perhaps more interestingly in this context are the more recent tickets. In table 2.9 we limit ourselves to the tickets created in 2017.

The amount of problems and incident tickets caused by a loss of power was lower in 2017 than overall. In particular, based on our keyword search, none of the 10 'grote storingen' (large outages) in 2017 seem to be caused by a loss of power. In 2017 the P1 and P2 impact saw a larger amount of loss of power tickets than overall.

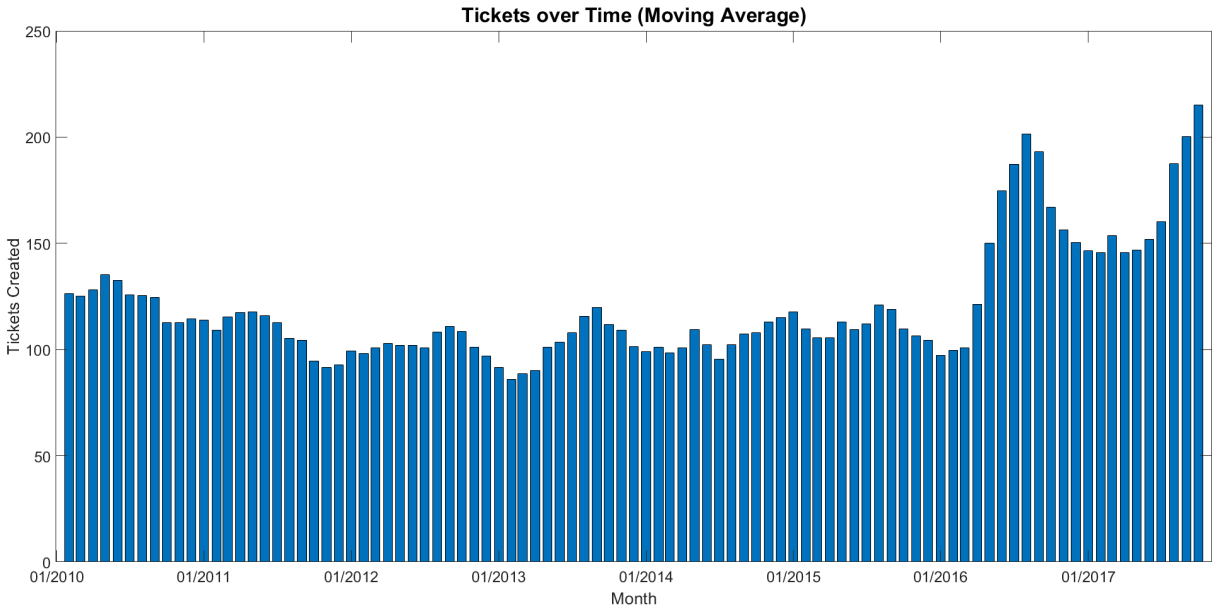

Figure 2.7: Central moving average (with a window size of 5 months) of tickets created per month. (November 2017 Ticket Dataset)

Table 2.8: Tickets caused by a loss of power (based on an automatic search for keywords) (November 2017 Ticket Dataset)

Impact	**Caused by a loss of power**	**Total Tickets**	**Percentage caused by loss of power**
P3	142	6446	2.20%
P2	136	3025	4.50%
P1	213	1147	18.57%
P1 Internationaal	5	214	2.34%
Grote Storing	8	34	23.53%
All tickets	509	11442	4.45%

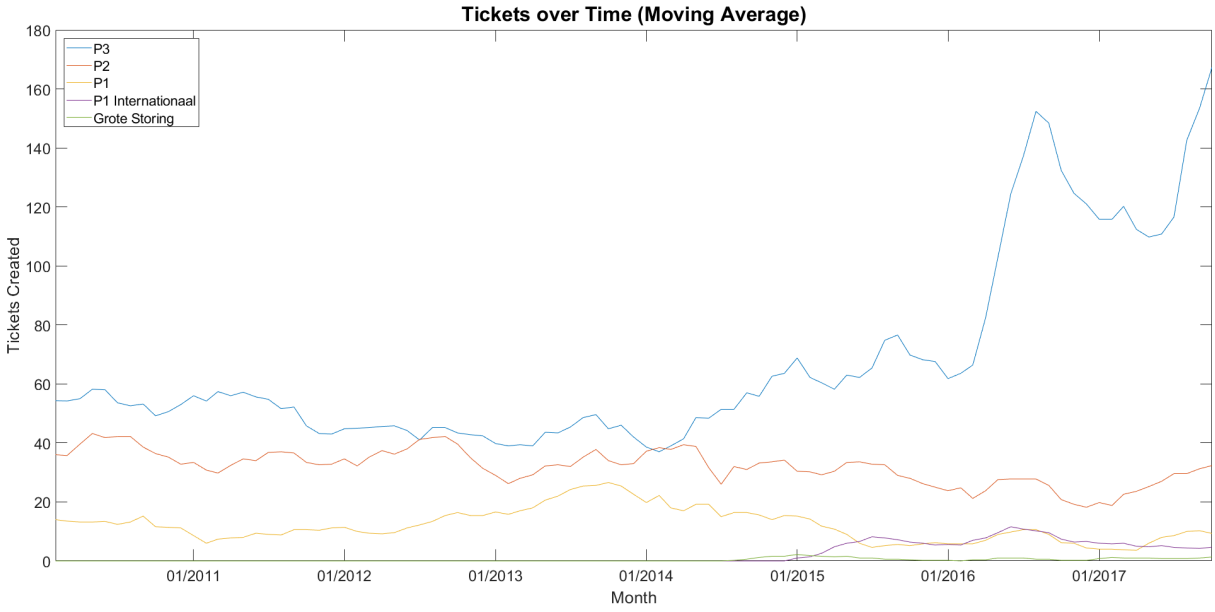

Figure 2.8: Central moving average (with a window size of 5 months) of tickets created per month per impact category. (November 2017 Ticket Dataset)

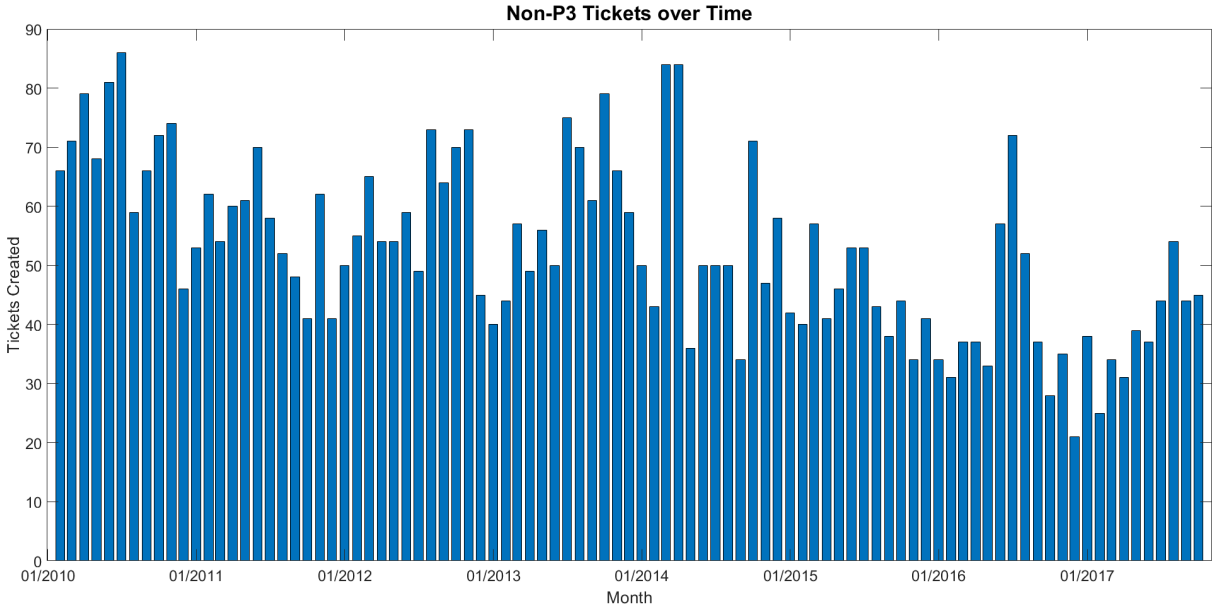

Figure 2.9: Tickets created per month, excluding P3 tickets. (November 2017 Ticket Dataset)

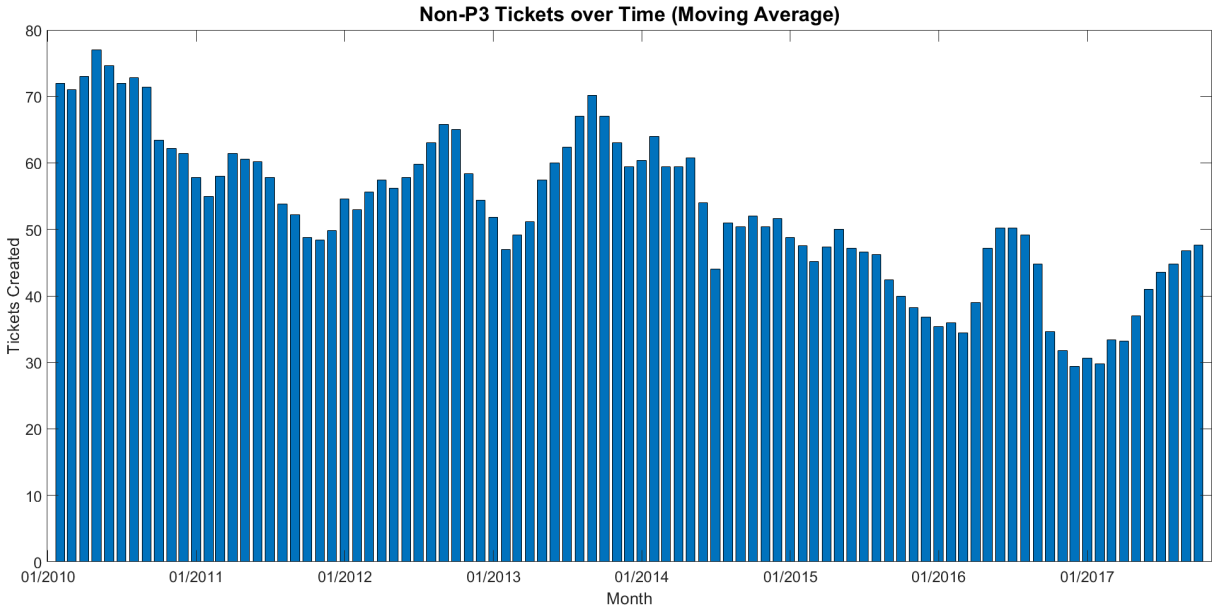

Figure 2.10: Central moving average (with a window size of 5 months) of tickets created per month, excluding P3 tickets. (November 2017 Ticket Dataset)

Table 2.9: Tickets created in 2017 caused by a loss of power(based on an automatic search for keywords) (November 2017 Ticket Dataset)

Impact	**Caused by a loss of power**	**Total Tickets**	**Percentage caused by loss of power**
P3	30	1445	2.08%
P2	20	286	6.99%
P1	23	76	30.26%
P1 Internationaal	4	54	7.41%
Grote Storing	0	10	0.00%
All tickets	77	1871	4.12%

Chapter 3

Port State Logs

3.1 Dataset

SURFnet keeps track of port states. Every 5 minutes, their system checks the state of all ports. This state, either up, down or unknown, then gets logged. SURFnet has provided these logs to us for analysis.

SURFnet has only recently begun storing these states, so the dataset only spans from 01-11-2015 upto 01-06-2017. In addition, when a port is not in use anymore, all its logged data is archived. These archived logs are not in our possession. This reduces the amount and accuracy of the information we can obtain from the dataset.

We assume the logged data is completely accurate. i.e. no down ports were logged as up or vice versa.

The dataset consists of a total of 847597644 states. 91721521, or around 10.82%, of these are 'down' states, while 8820621, around 1.04%, of the states were 'unknown'. A large amount of these down states are caused by the inclusion of disabled ports in the dataset. The state of disabled ports is not typically tracked by SURFnet. Unfortunately, due to a problem with their system, some of these ports are monitored anyway.

The set contains the states of 5702 ports, spread out over 428 devices and 258 locations (Points of Presence). Some of these ports have been monitored the whole period (i.e. from 01-11-2015 to 01-06-2017), while others were added more recently. Some ports are missing one or more months of data. We have filled in these missing months with "unknown".

We consider a device down if all its ports with known state are down. If one of its ports is up, the device is considered to be functioning as well. Similarly, a location is considered down if and only if all its devices with known state are down.

If we filter out all ports without any 'up' states, as these are presumable disabled ports, we are left with 5221 ports. The resulting dataset contains a total of 785516484 states, of which 30938117 (3.94%) are 'down' and 7522865 (0.96%) are 'unknown' states.

In some cases ports are down for very long periods of time. Presumably, these ports have been disabled and enabled again (or were in use at 01-11-2015, but were later disabled), and thus are not filtered out by the above filter operation.

Looking at figure 3.1, we can see that most periods of port downtime are very short. Presumable, these are the actual cases in which a port has gone down unexpectedly. However, we also see many 'peaks' of much longer periods. We believe these are caused by (temporarily) disabled ports.

In figure 3.2 we only show the periods of downtime ≤ 4000 minutes. Here, we can clearly see the distribution of port outage durations. Based on this plot, 24 hours, or 1440 minutes, seems a suitable cut-off point for filtering out disabled ports. We keep 94.41% of total outages.

An alternative, more conservative, cut-off point would be at 136565 minutes, which is almost 95 days. This is the result of 2-means clustering of the outage durations. In this case we would filter out only 1.44% of all outages.

In table 3.1 we give the deciles (10-quantiles) of the port 'down' durations. It should be clear that there are only a relatively small amount of periods of prolonged 'downtime', which together cause most of the 'down' states in the dataset.

If we remove periods of more than 24 hours of downtime from the logs (by changing the states to 'unknown'), we are left with 424120 (0.054%) 'down' states.

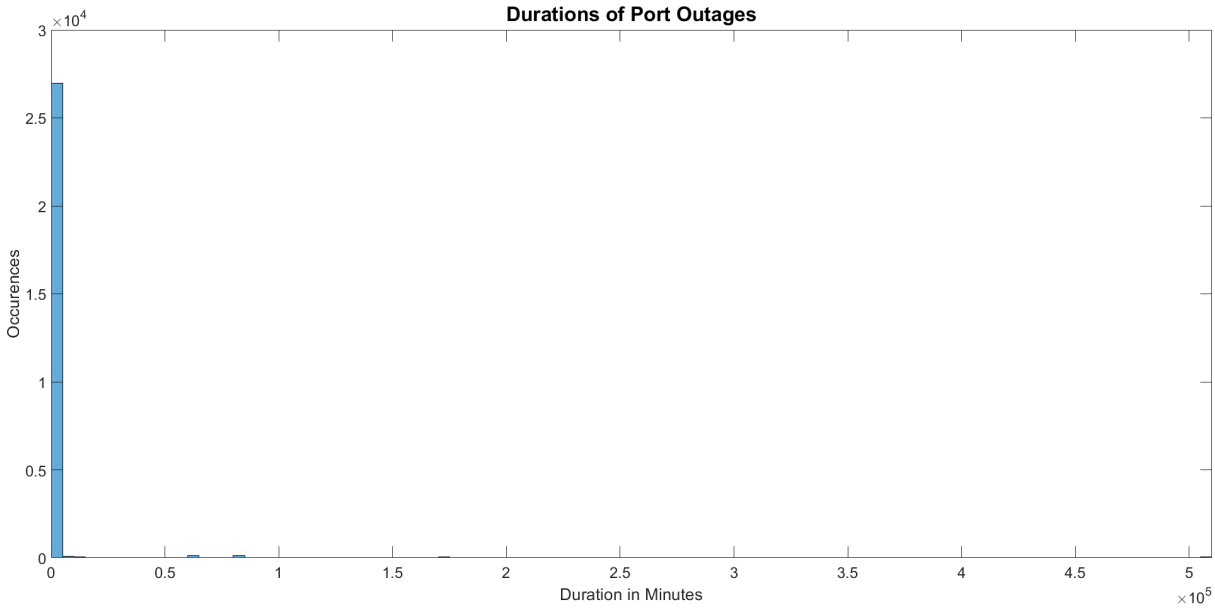

Figure 3.1: Durations of periods of port 'down' states. Port without any 'up' states were filtered out.

Table 3.1: 10-quantiles of the port downtime durations in minutes. Port without any 'up' states were filtered out.

0 (minimum)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 (maximum)**
5	10	10	10	15	25	45	75	150	355	506875

3.2 Percentage of Ports Down

We first consider a simple metric that can give us a broad overview of the state of the network: the percentage of down ports at any given time.

In figure 3.3 one can see both a gradual increase in down ports over time caused by the addition of disabled ports to the dataset, and the temporary increases caused by actual failures.

The two to three downward spikes do not indicate a decrease in down ports, but a temporary problem with the monitoring tool, causing most to all port states to be logged as 'unknown'.

Even after filtering out all ports without any 'up' states, there still seem to be a large amount of disabled ports in the dataset (see figure 3.4). As mentioned in the previous section, these are ports that have been temporarily disabled.

In figure 3.5 these periods of downtime have been filtered out as well. We can clearly see that the maximum amount of down ports is around 5%. Most periods of downtime are very short-lived, but there are a few periods of prolonged port downtime.

Of course, by filtering out prolonged periods of downtime, we might have actually filtered out some actual downtime as well, instead of only disabled ports.

A more useful metric might be the average percentage of time ports are down per month. In essence averaging out figure 3.5. The results have been plotted in figure 3.6.

In section 2.1.1 we mention that starting in June 2016 the number of created trouble and incidence tickets increased sharply. Figure 3.6 offers no explanation for this increase. Based on the available data, we can not conclude that port downtime was significantly higher than normal in the summer of 2016. This is mostly because the dataset only spans a short period of time.

March 2016 saw the most amount of port downtime. The total amount of port downtime in this month was around 3216.67 hours.

3.3 Outages of PoPs and Devices

If a single port is down, this typically does not have a very large effect on the network. In this section we analyze the outages of entire PoPs and devices, which have a much larger impact on the network.

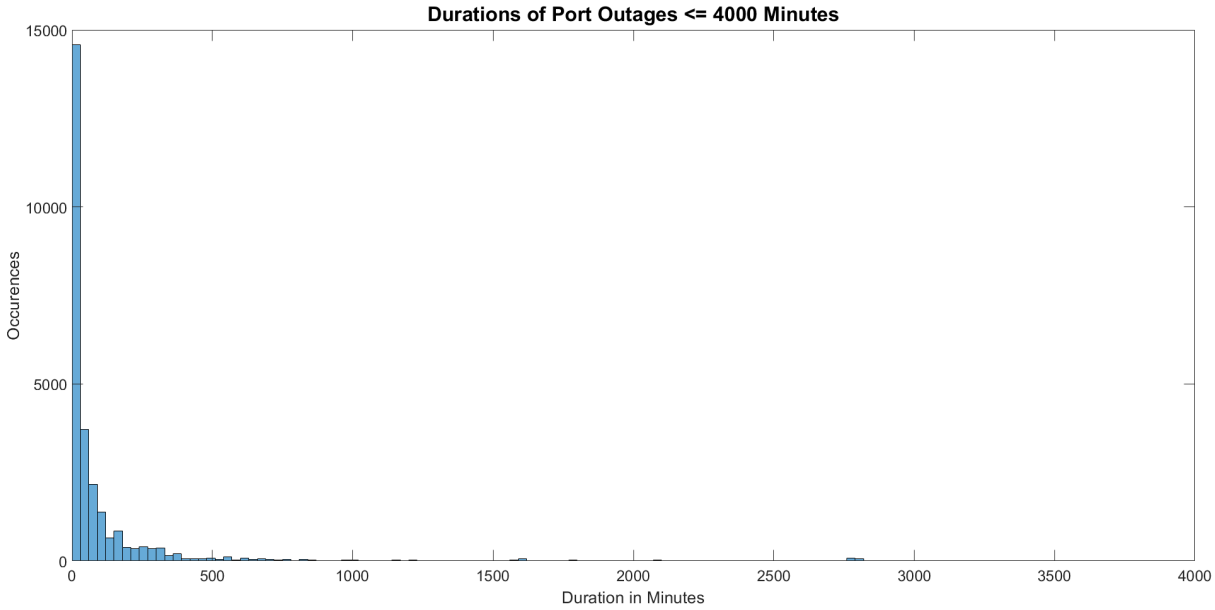

Figure 3.2: Durations of periods of port 'down' states. Port without any 'up' states and durations of more than 4000 minutes were filtered out.

All results in this section are based on the completely filtered dataset. That is, without any periods of prolonged downtime.

3.3.1 PoPs

The filtered dataset contains no occurrences of entire PoPs failing at once. However, based on the ticket dataset we know that in the period between October 2015 and may 2017 there was at least one PoP outage: on 12-07-2016 a PoP in Tilburg experienced a loss of power, causing its nodes to fail and stay down for more than 2 hours.

There are multiple possible reasons why we might miss some outages in this dataset. If the state is unknown instead of down, we do not register this as an outage. In addition, as mentioned in section 3.1, the dataset does not contain the logged data of disabled devices and ports, reducing the completeness of our data. Finally, by filtering out possible disabled port states, we might also be filtering out some actual outages.

In the case of Tilburg, all port states were unknown during the outage.

3.3.2 Devices

We can detect three separate device outage incidents in the dataset. In all three cases the same device had failed: asd001a_5150-05.

- asd001a_5150-05: down from 2016-05-10T06:20 to 2016-05-10T06:30
- asd001a_5150-05: down from 2017-01-24T05:05 to 2017-01-24T05:15
- asd001a_5150-05: down from 2017-04-18T04:55 to 2017-04-18T05:10

None of these outages lasted longer than 15 minutes, although some ports of asd001a_5150-05 may have stayed down for longer.

We would have expected much more device failures. It is very likely that, due to the same reasons as mentioned in the previous section, we can not find all device outages.

3.4 Hardware Type

In this section we compare the port downtime of different equipment. Knowledge about which hardware has more issues than other hardware might help when buying new equipment.

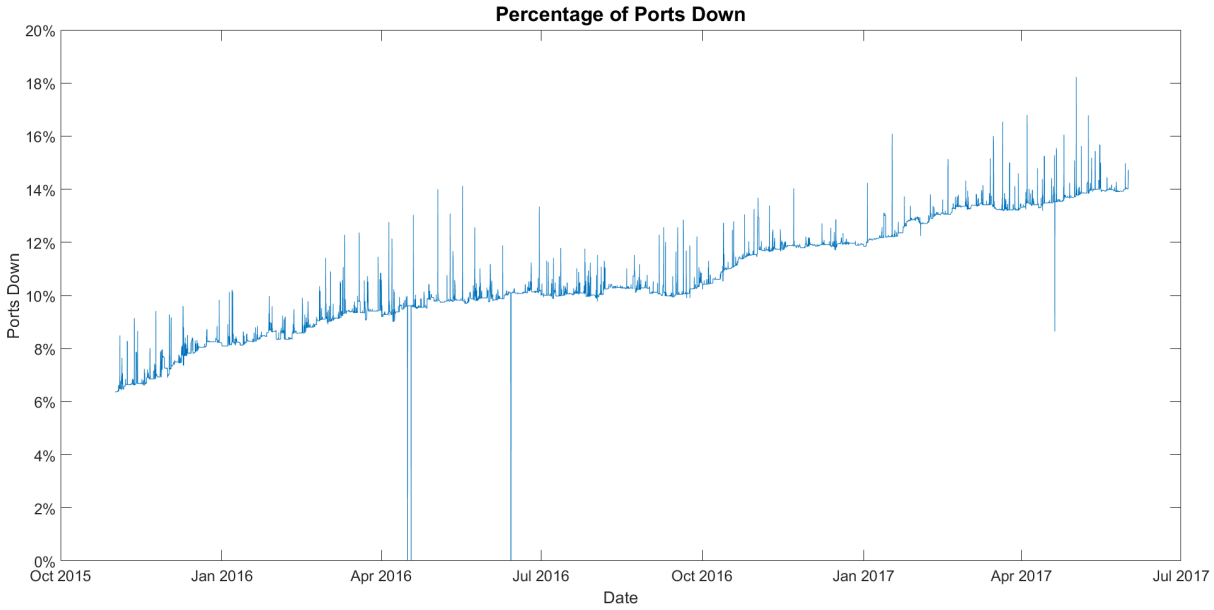

Figure 3.3: Percentage of ports down.

Table 3.2: Percentage port downtime per hardware type. Ports without any 'up' states were removed. Prolonged periods of downtime (> 24 hours) were replaced by 'unknown'.

Equipment Type (Abbreviation)	Percentage Down	Percentage Unknown
5150	0.062%	2.414%
5410	0.057%	7.075%
5142	0.055%	4.393%
All	**0.054%**	**4,842%**
8700	0.049%	9.543%
3930	0.047%	2.693%
5160	0.034%	5.325%

Note that faulty hardware can be expected to be disabled by SURFnet. As most disabled ports are not part of our dataset, this could affect the accuracy of our analysis.

In table 3.2 we give the percentage of 'down' states and 'unknown' states of each equipment type. In essence this is a combination of the frequency of port outages and the duration of these outages.

We have noticed the choice of threshold to filter outage durations on (in our case 24 hours) has a significant effect on these percentages, as well as the ordering of the different types of machines.

In addition, as link, and thus port failures, can be caused by a wide array of reasons not related to the node itself this metric might not be an accurate reflection of the quality or failure rate of devices.

Perhaps an improvement would be to compare the downtime of entire devices, instead of single ports. Unfortunately, these failures are not readily obtainable from the dataset (see sections 3.3 and 3.7)

The dataset contains 6 types of equipment. Of those, half have more port downtime than average. Most active ports in the network were part of the Ciena 5410, followed by the Ciena 3930 and the Ciena 5150.

The hardware logs do not provide any context, e.g. the cause. Some equipment might simply fail more often because of where and how its used. In the next section we analyze the effect of the physical location of equipment on the amount of downtime.

3.5 Physical Location and Failure Rate

Tables 3.3 and 3.4 show the sites with the highest and lowest percentages of port downtime. Based on these results there seems to be a clear difference in downtime based on the physical location of the hardware.

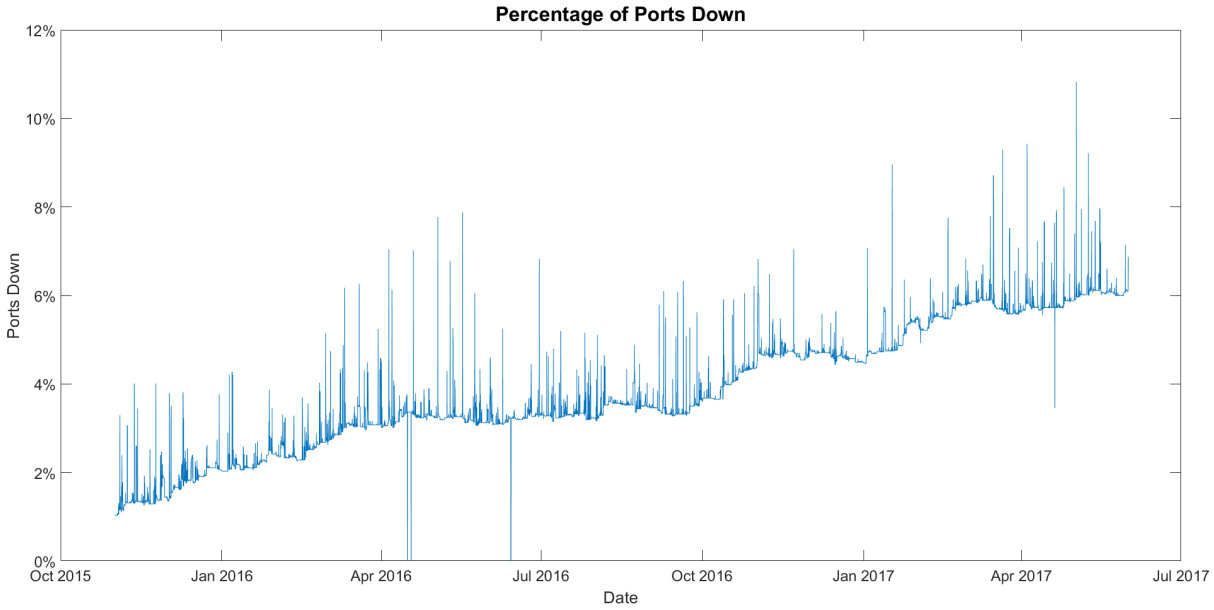

Figure 3.4: Percentage of ports down. Ports without any 'up' states were filtered out.

Table 3.3: 5 sites with the highest percentage of downtime. Ports without any 'up' states were removed. Prolonged periods of downtime ($>$ 24 hours) were replaced by 'unknown'.

Site	Percentage Down	Percentage Unknown
rt002c	0.228%	23.797%
pet001a	0.195%	0.666%
tb005a	0.193%	1.323%
ap001b	0.191%	0.448%
tb005b	0.171%	12.457%

In tables 3.5 and 3.6 we show the places with the highest and lowest percentages of port downtime. While less pronounced, there is still a large difference in port downtime based on the city equipment is placed in.

As in the previous section, the choice of filter threshold had an effect on the both the percentage of downtime and the ordering of PoPs and places.

3.6 Comparing Tickets and Hardware Logs

Table 3.7 shows a quick comparison of the ticket and (port) hardware log dataset.

The biggest advantage of the hardware logs is that it gives very precise data of when ports are down. In comparison, the tickets dataset does not provide this kind of very precise data, but does provide some context. For example the cause and impact of issues with the network.

In addition, a single incident that spans multiple ports is grouped into a single ticket in the ticket dataset. The port dataset does not provide this kind of context. If you want to group the effects of a single incident together, you would have to do some data analysis on the logs or make use of the information in the ticket.

One could say that the hardware log gives the exact state of all the individual parts of the network itself (ignoring 'unknown' states), while the ticket dataset gives an idea of the user experience/performance of the network as a whole.

Both datasets have some problems which reduce the accuracy of any analysis on the data. Many tickets in the ticket dataset are wrongly classified as problems and incidents. However, this presumably only impacts P3 tickets, as we can assume that a wrongly classified ticket is always given the lowest impact value (P3). Thus, the larger issue with the ticket dataset is that most ticket fields are not very machine-readable. This problem does not necessarily impact the accuracy of all analyses, but does limit the amount of information we can (easily) get out the dataset.

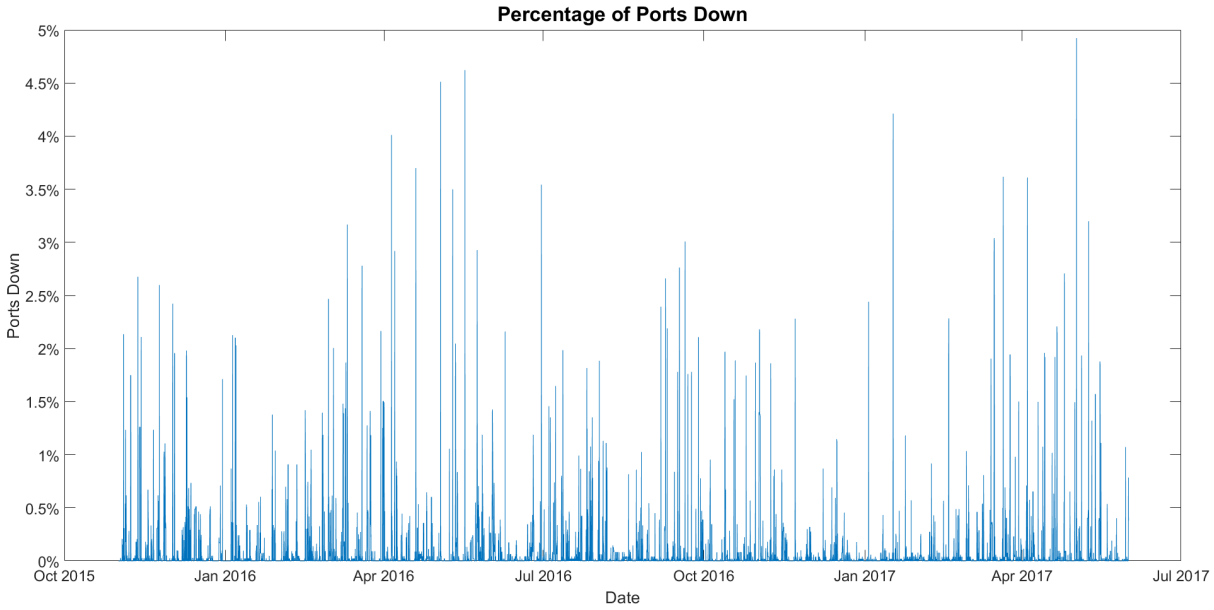

Figure 3.5: Percentage of ports down. Ports without any 'up' states and prolonged periods of downtime ($>$ 24 hours) were filtered out.

Table 3.4: 5 sites with the lowest percentage of downtime. Ports without any 'up' states were removed. Prolonged periods of downtime ($>$ 24 hours) were replaced by 'unknown'.

Site	**Percentage Down**	**Percentage Unknown**
mdmr001a	0.000%	0.550%
ddt003a	0.000%	5.458%
gn016a	0.000%	0.052%
ht006a	0.001%	6.001%
ht007a	0.001%	20.920%

The issues with the hardware logs arguably impact the accuracy of any results much more. Firstly, due to some fault in SURFnet's systems, the logs contain many false 'down' states. The choice of how to filter these states out of the set significantly impacts all results, especially when comparing hardware types and locations. Secondly, the dataset contains a large amount (1.04%) of 'unknown' states, which can be either 'up' or 'down'.

There lies a lot of potential in combining the two datasets. By coupling tickets to specific port outages we can combine the context of any incident with the precise information on which ports (and services) are down at which times. This could be done for example by searching the descriptions of a selected range of tickets for specific locations/hardware/ports based on when which ports are down.

3.7 'Hidden' Device Failures

As mentioned in the previous section, the dataset contains many 'unknown' states. As we could see in section 3.3.1, this can hide large outages in the network. These large outages are exactly the outages that are interesting in the context of geographically correlated failures.

If the device, or even all the hardware in a PoP, is failing, it might be more difficult, if not impossible, to monitor the state of a port. Thus it might be the case that if a device or location goes down, all its states will be 'unknown' and not 'down', effectively hiding the outage from us.

After a quick look through the data, it seems that in many cases a period of 'unknown' states is either very large or very short. This is very similar to what we saw earlier for the 'down' state. Our hypothesis is that these shorter bursts of 'unknown' states are mostly outages as well. In particular, we expect these short periods of 'unknown' to include the location and device failures we were expecting in this dataset.

In figures 3.7 and 3.8 we have plotted the durations of periods of 'unknown' states. Note that there are indeed many short bursts of 'unknown', and some very long periods of 'unknown'. However, there

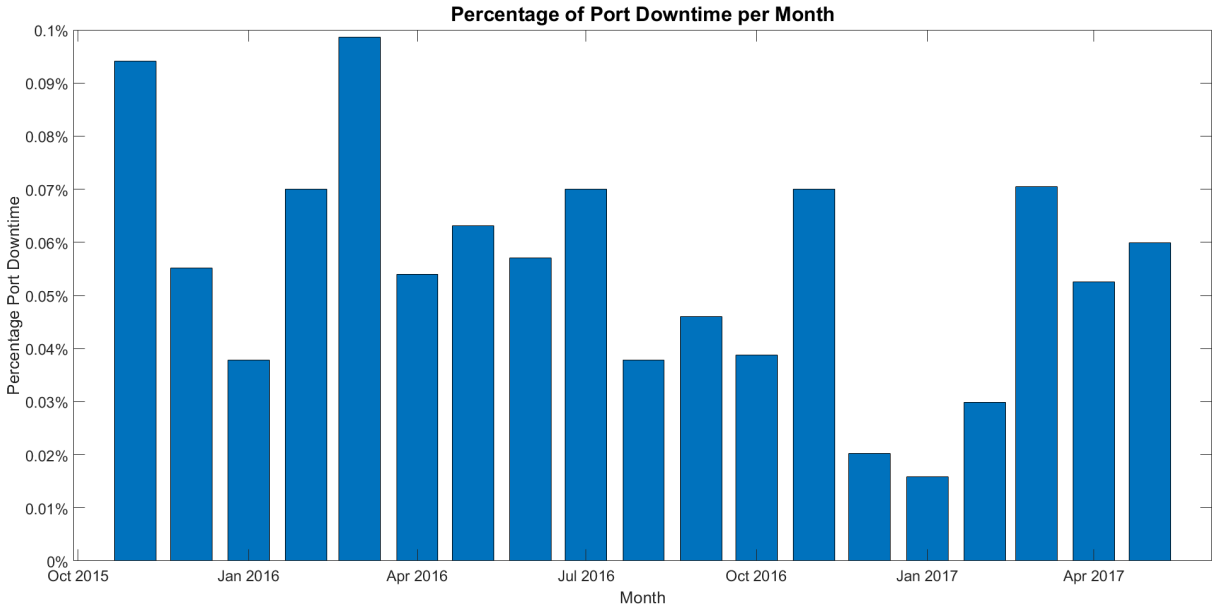

Figure 3.6: Percentage of port downtime per month. Ports without any 'up' states and prolonged periods of downtime ($>$ 24 hours) were filtered out.

Table 3.5: 5 places with the highest percentage of downtime. Ports without any 'up' states were removed. Prolonged periods of downtime ($>$ 24 hours) were replaced by 'unknown'.

Place	**Percentage Down**	**Percentage Unknown**
Harlingen (hl)	0.156%	0.729%
Apeldoorn (ap)	0.138%	0.332%
Drachten (dtn)	0.135%	0.127%
Ossendrecht (osd)	0.125%	2.458%
Petten (pet)	0.117%	0.357%

are some differences from the 'down' periods (see 3.1) as well. Most notably, it seems as if there are two separate clusters of short bursts of 'unknown' states. Maybe these are caused by two different underlying issues, such as device failures and PoP failures. Or perhaps the periods in one of these clusters are caused by failures, and the other by monitoring issues.

As we think that device failures are 'hidden' by the 'unknown' states, it might be more interesting to consider the periods where the state of a device is unknown. That is, the periods during which the state of each of the ports of a device is unknown. Figures 3.9 and 3.10 show the durations of these periods.

In figures 3.11 and 3.12 we do the same for locations (PoPs) instead of devices. These plots rule out the possibility that one of the two clusters was caused by device failures, and the other by PoP failures, as both are still visible in figure 3.12.

We should filter out all cases where the state of every port in the network is unknown, as in these cases the problem probably lies with the monitoring system. In figures 3.13 and 3.14 we have plotted the durations of all periods of location 'unknown' states that do not overlap a period of 'unknown' of all ports.

This filters out most of the 'unknown' periods. It is difficult to split the remaining occurrences into outages and monitoring issues. In the case of the disabled ports, we could simply assume that in most cases a port is disabled for more than 24 hours. The biggest issue was that by filtering out these disabled ports we possibly also filtered out some long outages. Unfortunately, it is not possible to use the same method to tell apart a monitoring problem and an outage.

3.8 Recommendations

Currently, these port statuses are (presumably) only used to monitor the current state of the network.

Table 3.6: 5 places with the lowest percentage of downtime. Ports without any 'up' states were removed. Prolonged periods of downtime (> 24 hours) were replaced by 'unknown'.

Place	Percentage Down	Percentage Unknown
Middenmeer (mdmr)	0.000%	0.550%
Brussels (bru)	0.002%	0.582%
Winschoten (ws)	0.003%	0.053%
Haarlem (hlm)	0.004%	6.939%
Sittard (std)	0.005%	0.707%

Table 3.7: Comparison between the ticket dataset and the hardware log dataset.

	Tickets	Hardware Logs
Machine-Readable	Minimal	Fully
Timespan	January 2010 to mid-March 2017	November 2015 to June 2017
Time precision	Not very precise	Precise to 5 minutes
Cause of incidents	Often included	Never included
Impact of incidents	Almost always included	Can be derived
Problems	Many tickets are wrongly classified as problems and incidents	Falsely classifies some disabled ports as 'down' and contains many 'unknown' states

There is a lot of potential information to gain about past network issues from port state logs, as well as other kind of hardware logs.

However, to be able to gain accurate results from this dataset, the problem of logging disabled ports as 'down' should be resolved first. Disabled equipment can either not be monitored and logged at all or a separate state should be used for disabled ports. As soon as the ports are enabled again, monitoring should resume.

The dataset contains customer-facing ports as well. It does occur that customers keep some of these ports 'down'. These 'disabled ports' should not be included in the dataset as well [1].

Maintenance on the network might require SURFnet to turn of some ports. This is typically not done gracefully, so these events are also included in the logs as periods of downtime. Further analysis might want to take this into account, and filter out these events. However, it could be argued that to accurately reflect the state of the network you should not filter out maintenance work, as these ports are still part of the network.

An interesting possibility lies in coupling the information from the hardware logs to individual tickets. This would add some context to the hardware logs, or precise network state information to the tickets.

In general, coupling different sources of information (e.g. port states, the network topology, tickets, maintenance impact tickets) together could have significant potential. It would also make it easier to filter out unwanted events such as disabled ports and maintenance.

As during a device or PoP failure the state of a port often can not be monitored properly, their statuses might have to be extracted from other types of (preferable hardware) logs.

Of course, data on port states mostly gives information about which links are down at which times. Unfortunately, we had no info on which ports were linked together, so we could not look further into this. We recommend to use this dataset, or a similar one, to look into the performance of different links of the network. Possibly in combination with their geographical location.

[1] Note that by filtering on outage duration we automatically filtered out these events too.

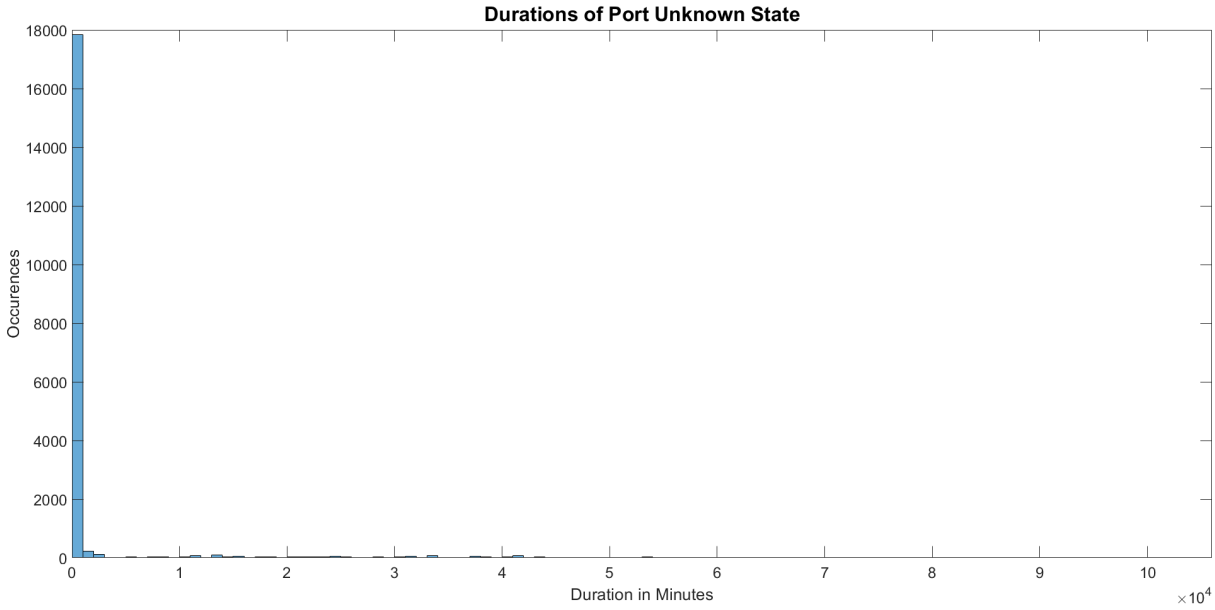

Figure 3.7: Durations of periods of port 'unknown' states. Port without any 'up' states were filtered out.

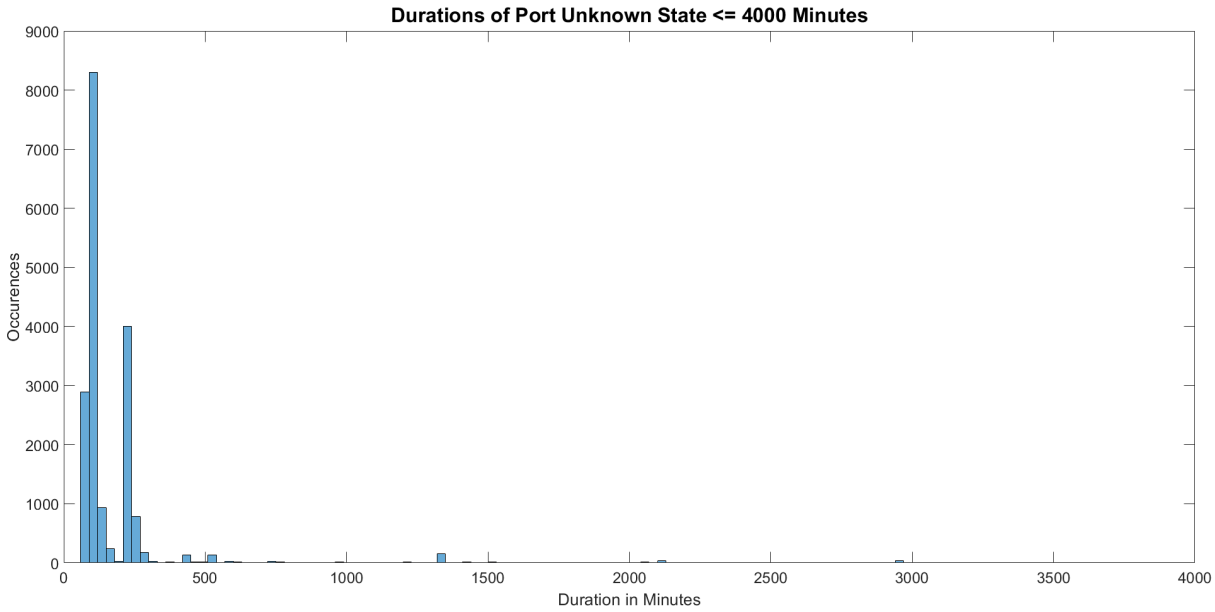

Figure 3.8: Durations of periods of port 'unknown' states. Port without any 'up' states and durations of more than 4000 minutes were filtered out.

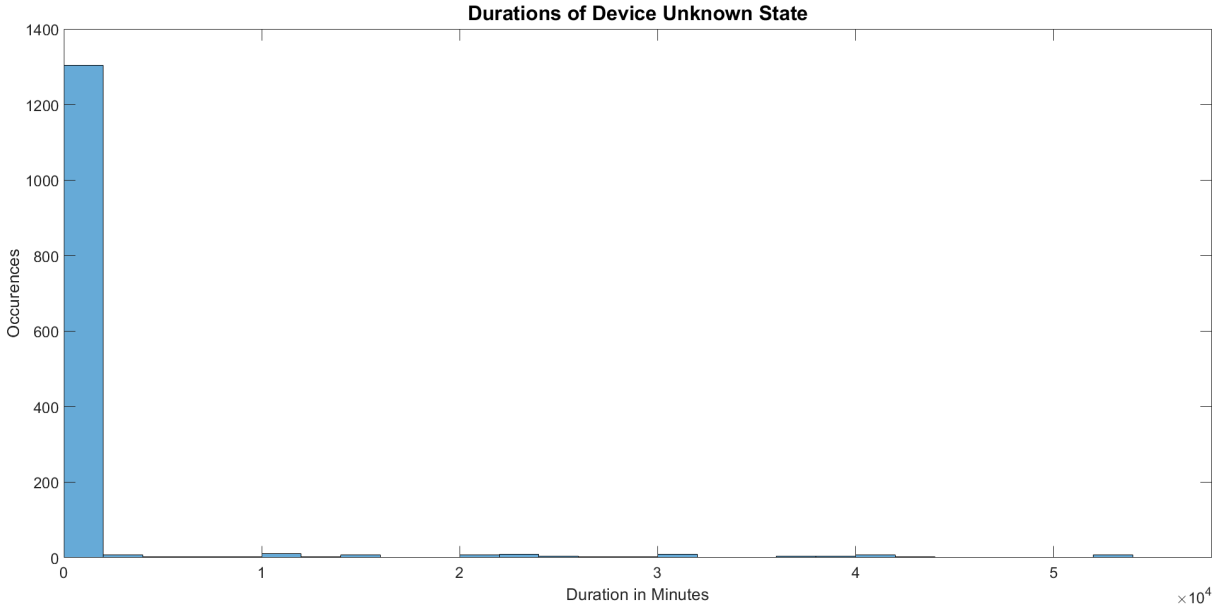

Figure 3.9: Durations of periods of device 'unknown' states. Port without any 'up' states were filtered out.

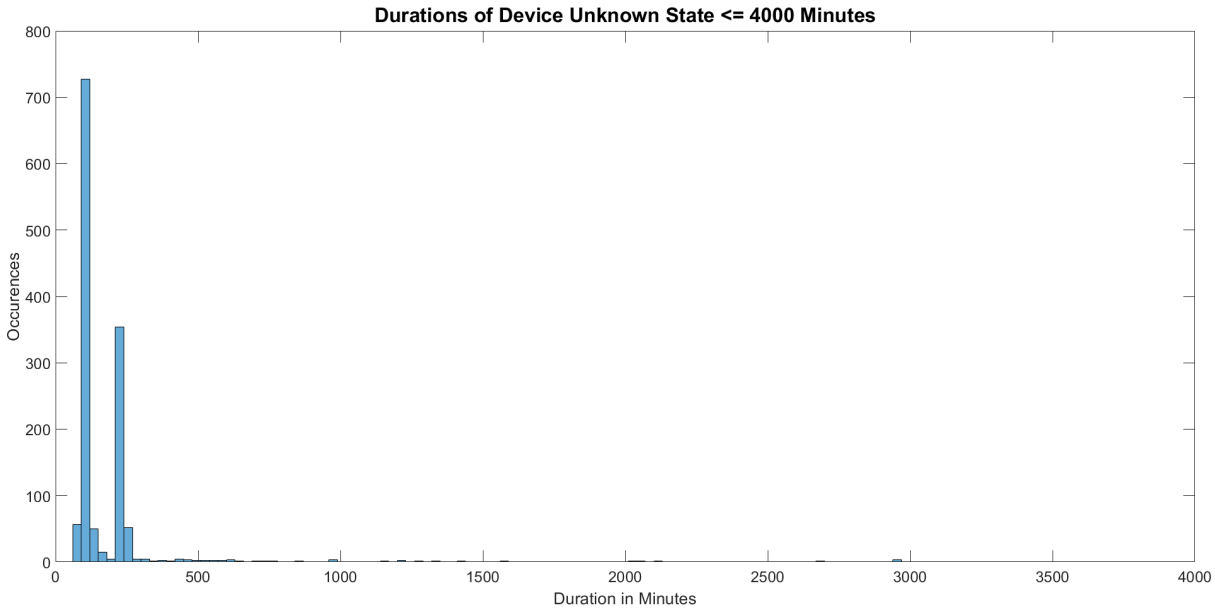

Figure 3.10: Durations of periods of device 'unknown' states. Port without any 'up' states and durations of more than 4000 minutes were filtered out.

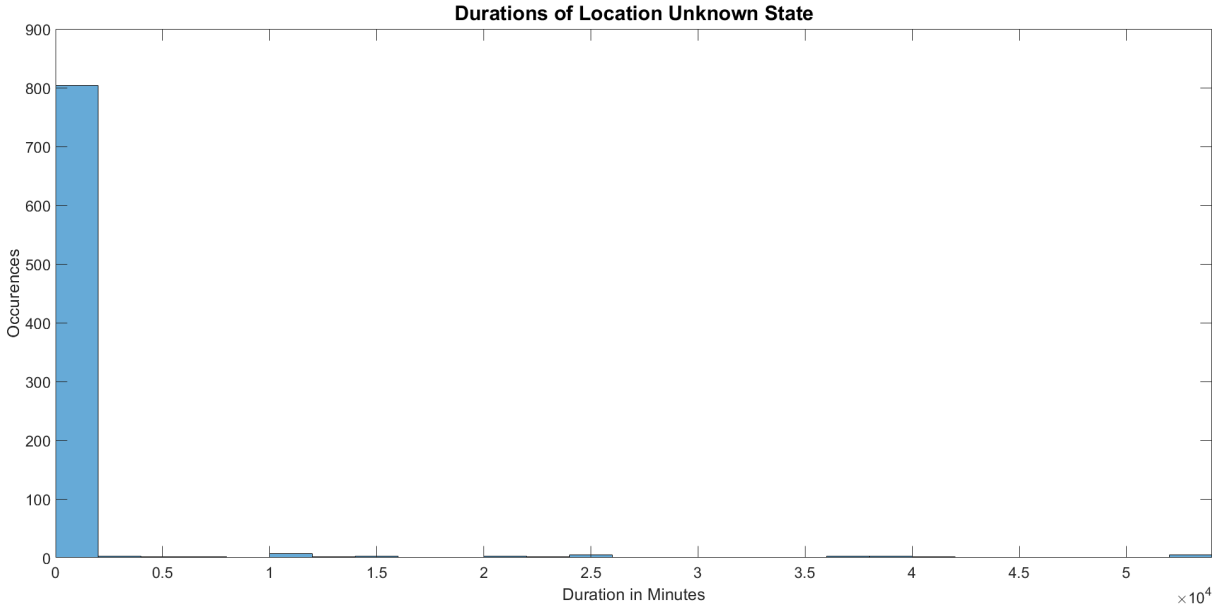

Figure 3.11: Durations of periods of PoP 'unknown' states. Port without any 'up' states were filtered out.

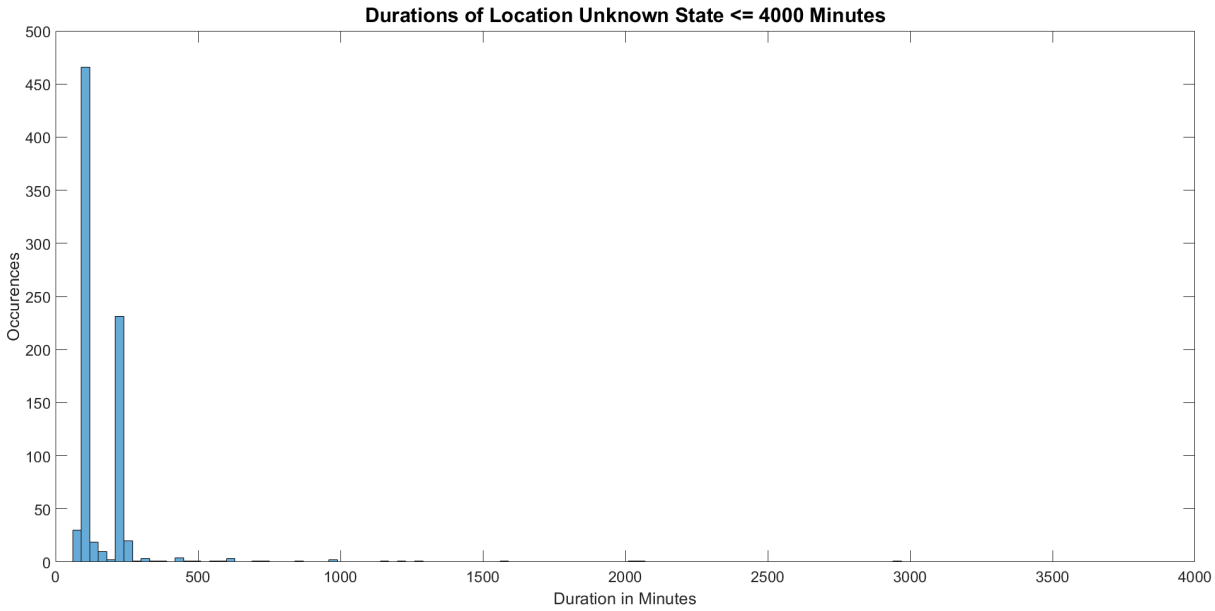

Figure 3.12: Durations of periods of PoP 'unknown' states. Port without any 'up' states and durations of more than 4000 minutes were filtered out.

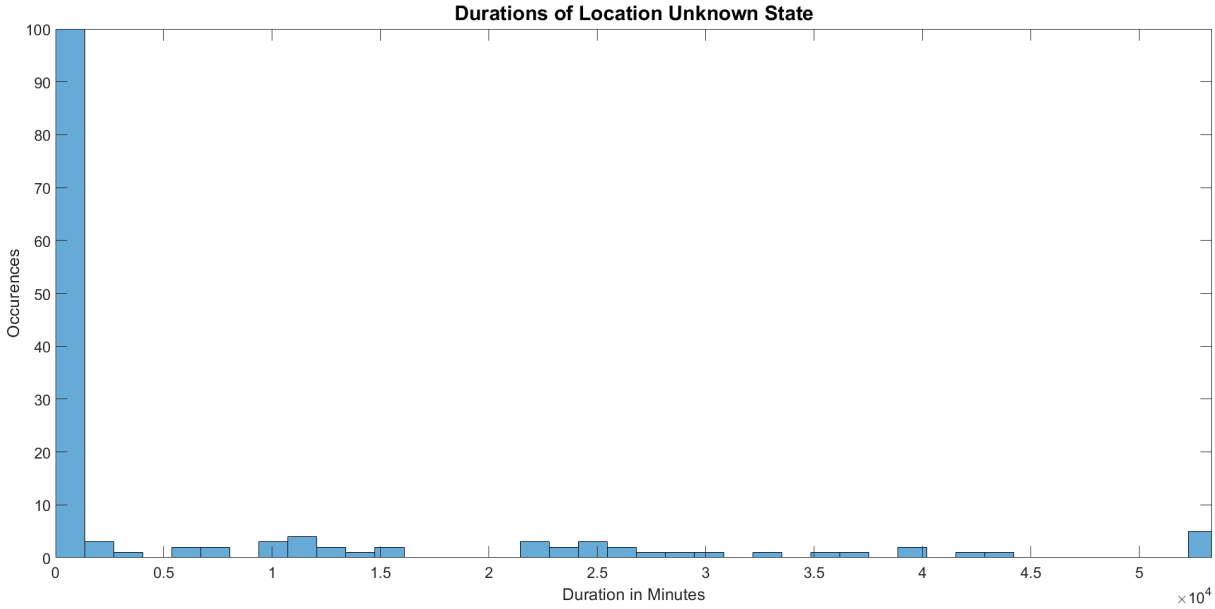

Figure 3.13: Durations of periods of PoP 'unknown' states that do not overlap periods where all port states are unknown. Port without any 'up' states were filtered out.

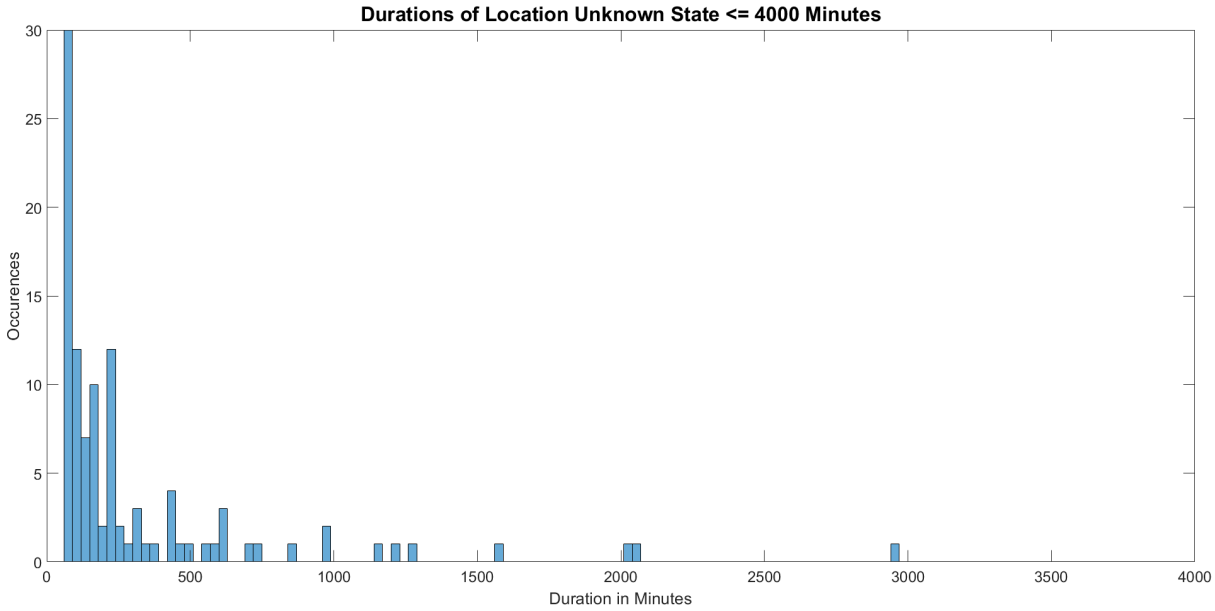

Figure 3.14: Durations of periods of PoP 'unknown' states that do not overlap periods where all port states are unknown. Port without any 'up' states and durations of more than 4000 minutes were filtered out.

Chapter 4

Conclusion

Currently, SURFnet only seems to consider their current situation. However, looking back into the past network performance and issues can give numerous insights about not only the past, but also the present and the future of the network. It can help to put the current network situation in context by comparing it to past situations. For example, at the moment the number of created tickets per month is at an all-time high.

The datasets we were given can give many insights into network. However, these insights are mostly superficial (e.g. the percentage of port downtime per month), as more in-depth information would require changes to SURFnets systems and/or a large time investment.

To be able to gain deeper insights into the network, We propose introducing a separate system for archiving past incidents for analysis, as archiving past incidents for analysis and keeping employees up to date about current incidents are seemingly incompatible.

In addition, we recommend the coupling of hardware logs and incident tickets, as this would combine the context information from tickets with the precise details on the state of network components from the hardware logs.

Currently, the major blockade to gaining accurate results from the port state dataset is that some ports that are not currently in use are logged as 'down'.

As the failure of a device hinders the ability of the monitoring system to read the state of its ports, these ports will typically be given the 'unknown' state during such a failure. To gain more accurate results on device failures from the port state dataset, these failures would need to be properly separated from issues with the monitoring system.

Chapter 5

Appendix

5.1 Categorization of "Grote Storing" Tickets

Cause	Ticket IDs
Node issues/failure	17864, 24339, 24543, 24561, 49812, 53453, 56225, 60275
Fiber(s) down	24761, 51207, 52150, 62581
Power failure/cut	49794, 50127, 51825, 57359, 57739
Loss of power because of a hardware failure	24682, 50706, 57549
Human error	49790, 62132
Third party issues	24738, 50751
Unmentioned	24185
Duplicate Ticket	56377, 62582