

# Analysis of Dependencies Between Failures in the UNINETT IP Backbone Network <sup>\*†</sup>

Andres J. Gonzalez (1) , Bjarne E. Helvik (1) , Jon K. Hellan (2), Pirkko Kuusela (3)

(1) Centre for Quantifiable Quality of Service in Communication Systems <sup>‡</sup>

O.S Bragstads plass 2E, N-7491 Trondheim, Norway

Email: andresgm@q2s.ntnu.no , bjarne@q2s.ntnu.no

(2) UNINETT. The Norwegian Research Network

Abels gate 5, NO-7465 Trondheim

Email: jon.kare.hellan@uninett.no

(3) VTT, Technical Research Center of Finland

P.O.Box 1000, FI-020044 VTT, Finland

Email: pirkko.kuusela@vtt.fi

## Abstract

Dependencies between failures in operational networks may have a huge impact on their reliability and availability. In this paper we analyze failure logs to identify simultaneous and potentially correlated failures in routers and links of an IP backbone network. We show that the actual behavior of failure processes does not support the independence assumption commonly used in theoretical studies. Scatter plots are presented to visualize the failure processes, and it is seen that geographical adjacency has a pronounced effect. The existence of high correlation coefficients and high autocorrelation in some failure processes was observed. A formal analysis confirms this. The consequences of these dependencies on the provisioning of guaranteed availability are briefly discussed.

---

\*Regular Paper Submission

†Keywords: Failure correlation, network dependability, failure analysis, real data modelling, events dependence.

‡"Centre for Quantifiable Quality of Service in Communication Systems, Centre of Excellence" appointed by The Research Council of Norway, funded by the Research Council, NTNU and UNINETT. <http://www.q2s.ntnu.no>

# 1 INTRODUCTION

Availability is a significant element for provisioning a good QoS and is one of the most important parameters in setting up a service level agreement (SLA) between providers and users of a network service. Analysis of real failure processes in networks are mandatory in order to get the appropriate information for availability dimensioning and to deal with the risks associated with SLA agreements. In spite of this, for a number of reasons, among them that failures of their network are not what operators like to have exposed in a competitive commercial marketplace, the access to such failure log information is very limited, and therefore few studies based on data from operational networks are performed. In [11] a study of spatial and temporal failures and outages in an access network was performed to assess the availability. Another study in [6] estimates the time between failures and times to repair for elements in a large wireless access network, finding that they are not consistent with exponential distributions, but they may better be described by Weibull or two-stage hyper-exponential distributions. A study of the failure behavior in an operational backbone network is reported by Iannaccone et al. [2]. They examine the frequency and duration of failure events and discuss various statistics, for instance the distribution of inter-failure times and distribution of link failure durations. This work was continued by Markopoulou et al in [5], where failures and repairs in the Sprint IP backbone Network are classified and analyzed. They perform a characterization of the different classes of failures found. In [4] Kuusela and Norros analyze router failure logs from the Finnish academic network, FUNET and in [3] they describe a method that can be used to assess downtimes due to joint failures.

This paper is based on real operational data and focuses on correlations between link failures and between router failures. The investigation is based on 9 years of logged failures made available by the Norwegian academic network operator UNINETT [10].

In the analysis of network reliability, independence between failure events is commonly assumed, in part for mathematical convenience, in part due to lack of a better failure model and in some cases due to ignorance.

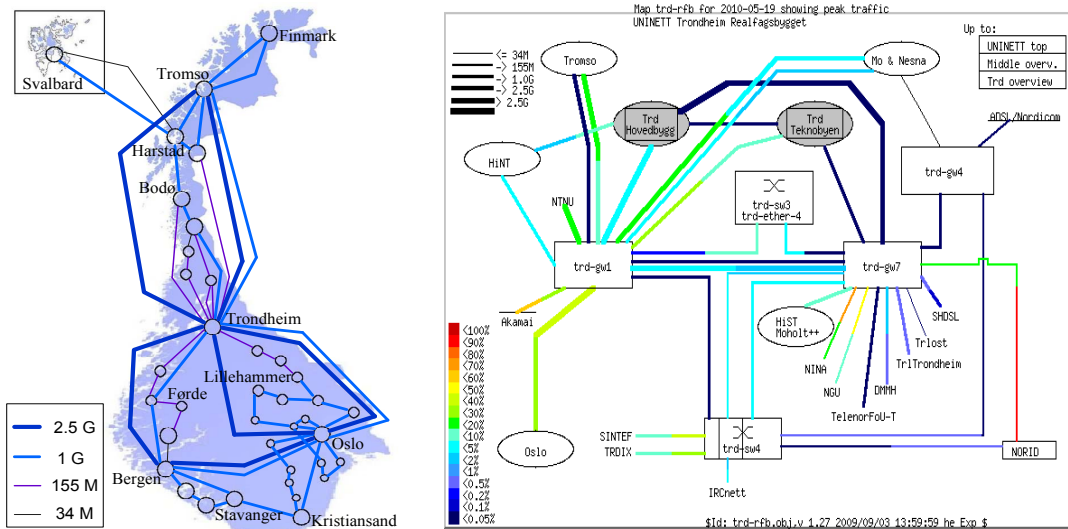
An objective in this paper is to show that independence assumption is incorrect, and in this respect, our findings confirm the initial observations of Markopoulou et al. [5]. The main contribution of this study is to show that the correlation of the failures of network elements is strongly related to their geographical distance. This is demonstrated by visualization and by formal analysis.

This paper is organized as follows. First, the UNINETT's IP backbone network, the information collection method and the data used for the analysis are presented. In Section III, the dissimilarity between the observed failure process and the process that would result if network elements failed according to independent renewal processes, is demonstrated. Section IV describes three different methods used for the study of failure dependence and the respective obtained results are illustrated and analyzed. Finally Section V describes some potential future works and concludes the paper.

## 2 UNINETT NETWORK DESCRIPTION

UNINETT is the network that connects universities, colleges and research institutions in Norway. The core of the network interconnects the main norwegian cities through optical fiber connections of 10 and 2.5 gigabit per second (Gbps) forming rings to ensure that the loss of a single link does not cause any loss of connectivity. In addition there are at least two disjoint paths between the major universities. A large and increasing part of the network is connected to this core through 1 Gbps links and for locations with both a smaller number of users and relatively high costs of establishing links the capacity is typically 155 Mbps and some few with 34Mbps. Figure 1(a) shows a global overview of the UNINETT topology. Each "node" in this map represents a geographical area that may contain several interconnected buildings and smaller locations. For instance Figure 1(b) illustrates more explicitly the network details of the node Trondheim (see Fig. 1(a)) in order to illustrate that each node in the main overview may have several router and links.

The failure logs were obtained through a centralized network management controlled



(a) Global Overview of the UNINETT Network

(b) Detailed Configuration of the Trondheim Node Network

Figure 1: UNINETT Network Topology

form the UNINETT NOC (Network Operations Center) in Trondheim since January 2001 until October 2009. During this period the global design of the network was conserved, but many individual changes were performed e.g. router replacements, new optical fiber installation, etc. Therefore in the performed studies we select relatively network stable intervals. Router’s and link’s failures are registered with a precision in the order of seconds. The data collection method follows SNMP standards, where for each new component installed SNMP agents enable the detection of changes in the network operation. Those changes may be identified in two ways: Either by periodical polling, using GetRequest/GetResponse messages generated from the central server, or by trapping techniques generated on the remote agent that uses notification messages able to capture every change on line. The router state is identified by *no-response/reachable* messages and the link state is reported as failures on specific router interfaces associated with a link using *linkUp/linkDown* messages.

In some cases trap messages are generated by network agents when a device return to an operational state, reporting to the central server the local log of the failure. This information may help to correct and define more precisely a down-time obtained by polling.

For this reason the use of an intelligent mechanism able to perform this correction is necessary. It is also important to clarify that for the case of links, the failures are identified in the SNMP management system as failures in specific routers interfaces associated with a link name. A logical link contains two physical fibers that in most of the cases report failures simultaneously from each side. However, our study avoids trivial conclusions by not considering as simultaneous two interface-events that are reported very close in time in different places, if they have the same link name. Finally, the information obtained from UNINETT contains summaries of events based on SNMP data without any previous processing, therefore a very important phase in our work was the implementation of PERL scripts in order to obtain a clean UP/DOWN state in time for each network component.

The studied backbone is operated using WDM technology, routers form several brands and using IS-IS as routing protocol. A more detailed information of the UNINETT network can be found in [10].

### 3 EMPIRICAL BEHAVIOR OF AGGREGATE FAILURE PROCESSES

A first objective in this paper is to compare the real stochastic behavior of failure events with the Poisson assumption that has been used during many years in the field of network dependability.

The following notation and considerations are used. Routers and links will be analyzed separately. Failure events within a fixed observation period  $T$  will be considered, where  $N$  network components are regarded. Each device  $i$  ( $i = 1, 2, \dots, N$ ) has an operational state that may be described by an UP/DOWN signal as illustrates Figure 2. A failure  $j$  occurs at time  $t_{i,j}$  and the downtime duration is denoted by  $d_{i,j}$  where  $n_i$  is the number of failures of device  $i$  during  $T$  and  $j = 1, 2, \dots, n_i$ .

For more than four decades mathematical models based on Poisson assumptions have

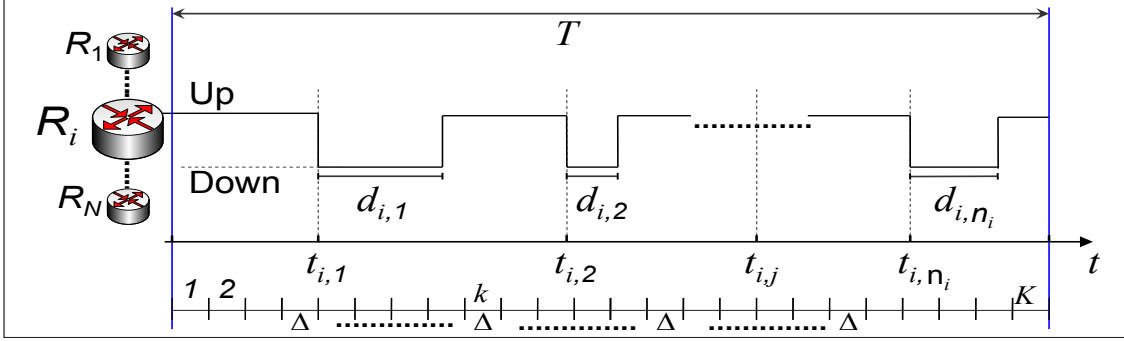


Figure 2: Behavior of a Network Component

been developed. An argument for this is the Palm-Khintchine theorem which states that the aggregation of a large number of processes tends to be Poisson distributed if: each individual process is renewal, no process is *dominant* (failure intensity very high compared to the others) and they are independent from each other [1]. Under these assumptions the aggregation of the  $N$  failure processes will become a Poisson process with parameter  $\lambda$  where

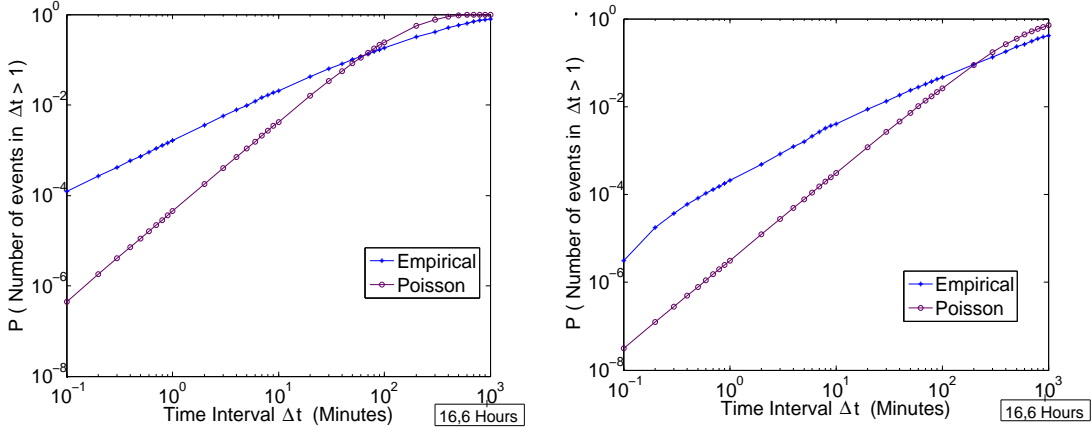
$$\lambda = \sum_{i=1}^N \frac{n_i}{T}. \quad (1)$$

We want to find the coincidence of  $M$  failures from different devices, therefore the occurrence of two or more failure events  $P(M > 1)$  within a given short time interval  $\Delta$  is of interest. From a theoretical point of view, using the Poisson parameter obtained in (1) this value may be written as:

$$P_{Pois}(M > 1) = e^{-\lambda\Delta} + \lambda\Delta e^{-\lambda\Delta} \quad (2)$$

On the other hand we can obtain empirical values for  $P(M > 1)$  using the information from the failure logs. First we split  $T$  in  $K$  small slots  $\Delta$  as is shown in Figure 2. Then we evaluate the number of failures  $m_k$  on each slot  $k$  ( $k = 1, 2, \dots, K$ ) as follows:

$$m_k = \sum_{i=1}^N \sum_{j=1}^{n_i} I\left((k-1)\Delta < t_{i,j} \leq k\Delta\right), \quad (3)$$



(a) Probability of two or more Link-Down events in a time interval (b) Probability of two or more Router-Down events in a time interval

Figure 3: Coincident events based on empirical logs and Poisson assumptions during 2007

where  $I(x)$  is the Indicator function.<sup>1</sup>

Considering all the network components  $(1, 2, \dots, N)$  the empirical  $P(M > 1)$  is obtained as follows:

$$P_{emp}(M > 1) = \frac{\sum_{k=1}^K I(m_k > 1)}{K}, \quad (4)$$

where  $K = T/\Delta$  is the number of slots in the evaluation period.

Figure 3 illustrates  $P(M > 1)$  under different  $\Delta$  on links and routers separately during the year 2007. It is seen that the empirical probability is larger for almost all  $\Delta$  smaller than the inverse of the average failure intensity ( $1/\lambda$ ) which is 95 minutes for links and 304 minutes for routers. The difference is two orders of magnitude for small  $\Delta$ 's, clearly showing that failures in the real life tend to occur more coincidently than what is expected under the Poisson assumption.

Given that no process is found to dominate, according to the Palm-Khintchine theo-

<sup>1</sup> $I(x)$  takes value 1 if condition  $x$  is fulfilled and 0 if not.

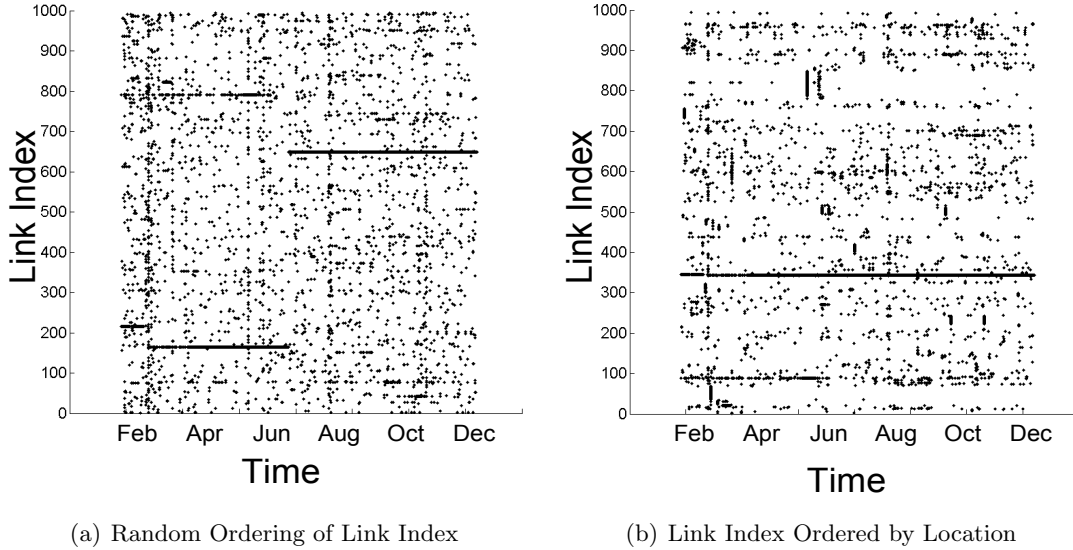


Figure 4: Links Down Events During 2007

rem this difference may be due to the failure processes not being renewal or independent or (most likely) both. This fact suggests a dependency between failure events that has to be verified through more specific methods in the next sections.

## 4 DEPENDENCE ANALYSIS

Three different methods are used to identify and analyze dependencies among failure events. The first is visual, the next identifies coincidence depending on geographical distance and the last regards the evaluation of auto-correlation and correlation coefficient in the failure processes.

### 4.1 SCATTER PLOT ANALYSIS

Scatter plots enable us visually to identify patterns in routers and links failures. In this method the data is displayed as a 2-dimensional collection of points where each of them represents a  $t_{i,j}$  on the horizontal axis and a device index  $i$  on the vertical axis.

The scatter plots in Figure 4 contain link failures during year 2007. They show corresponding results to the observed in [5] where horizontal and vertical patterns may

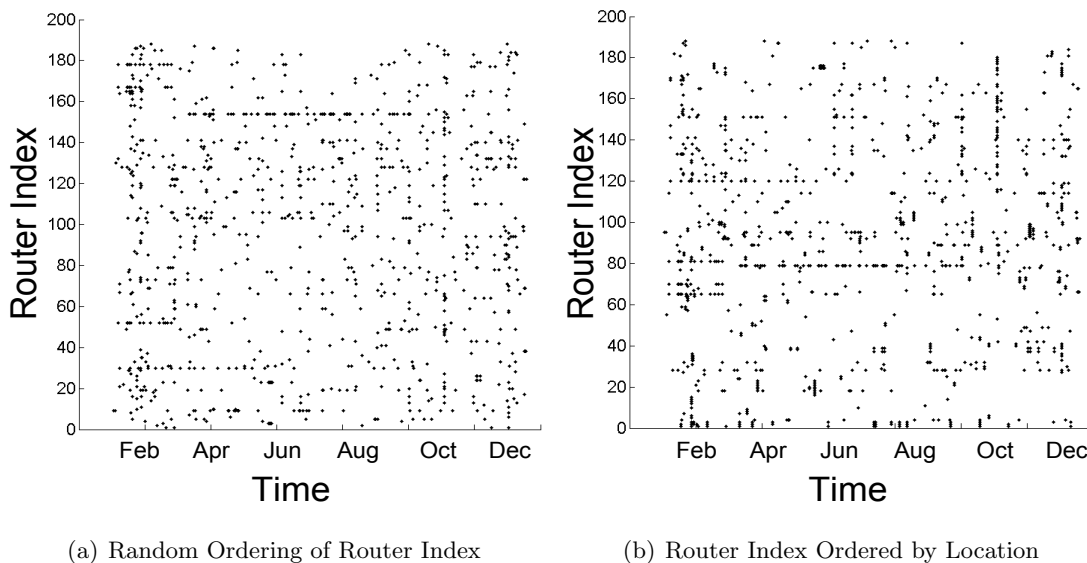


Figure 5: Routers Down Events During 2007

be identified, suggesting the presence of dependencies in space and time. For instance, the broken horizontal stripes indicate burst of failures from a component  $i$  over the duration of the stripe. Therefore these processes are unlikely to be renewal and autocorrelation may exist. On the other hand, the formation of vertical patterns is due to the existence of simultaneous failures and potentially correlated events of different components. This suggests that a failure in one device may have influence on others.

An advantage of the scatter plots is that we may observe changes in the overall pattern by changing the ordering of the  $y$  axis, i.e. the indexing of devices. An interesting result is obtained when the links are indexed according to their geographical location as can be observed in Figure 4 where in 4(a) the links are indexed randomly while in 4(b) the links are indexed according to geographical location from the north to the south of Norway<sup>2</sup>.

An initial observation is that the dots exhibit a more pronounced clustering in Figure 4(b), where some areas are characterized by a low or a high dot density, representing geographical zones and time periods with links relatively stable or with high amount of failures respectively. An important observation that Figure 4(b) offers is the presence of more compact and clearly defined vertical stripes, suggesting not only that failures may

<sup>2</sup>The number of dots is the same on both figures (4(a) and 4(b)).

occur simultaneously but also that geographical adjacency has a high relevance. Similarly, horizontal stripes are identified, indicating that this behavior also tend to coincide in the same geographical area. Note for instance the horizontal stripe across the entire year in Figure 4(b) which is seen to stem from at least three different links by comparison with Figure 4(a).

A similar scatter plot for routers is presented in Figure 5. It shows the presence of horizontal and vertical patterns as well. The comparison of Figures 5(a) and 5(b) yields similar observations where geographical location is an important actor.

We may conclude that failure processes are highly unlikely to be renewal or independent and that the coincidence in failures is significant in the UNINETT network. The dependence is stronger between devices that are geographically close.<sup>3</sup>

## 4.2 SIMULTANEOUS EVENTS ANALYSIS

In this section we introduce a method to identify simultaneous and potentially correlated failure events. First the notion of simultaneous event has to be clarified. Ideally a simultaneous event occurs when two down-events are reported exactly at the same time. Nevertheless, due to clock differences, delays in the system and also due to the propagation time of one failure over the others, this requirement has to be relaxed. Hence we introduce a flexibility gap  $\Delta$  where two failures are considered simultaneous if they occur within  $\Delta$  i.e.  $t_{i,j}$  and  $t_{x,y}$  are simultaneous if  $|t_{i,j} - t_{x,y}| \leq \Delta/2$ . The magnitude of this gap will be determined using the information in Figure 3, selecting a  $\Delta$  where  $P_{Poiiss}(M > 1)$  is very small. More specifically a value of  $\Delta$  equal to 30 seconds was used for routers since the ratio  $P_{emp}(M > 1)/P_{Poiiss}(M > 1)$  has its approximate maximum at this gap value. Choosing a similar ratio value, for links was chosen  $\Delta = 10$  seconds.

Due to the results obtained from the scatter plot analysis in section 4.1, we want to evaluate the existence of simultaneous failures in two different network components and the relation with their geographical distance.

---

<sup>3</sup>In the scatter plots the limited printing resolution makes impossible to discriminate closely located points. Our study will use complementary methods to verify the results obtained.

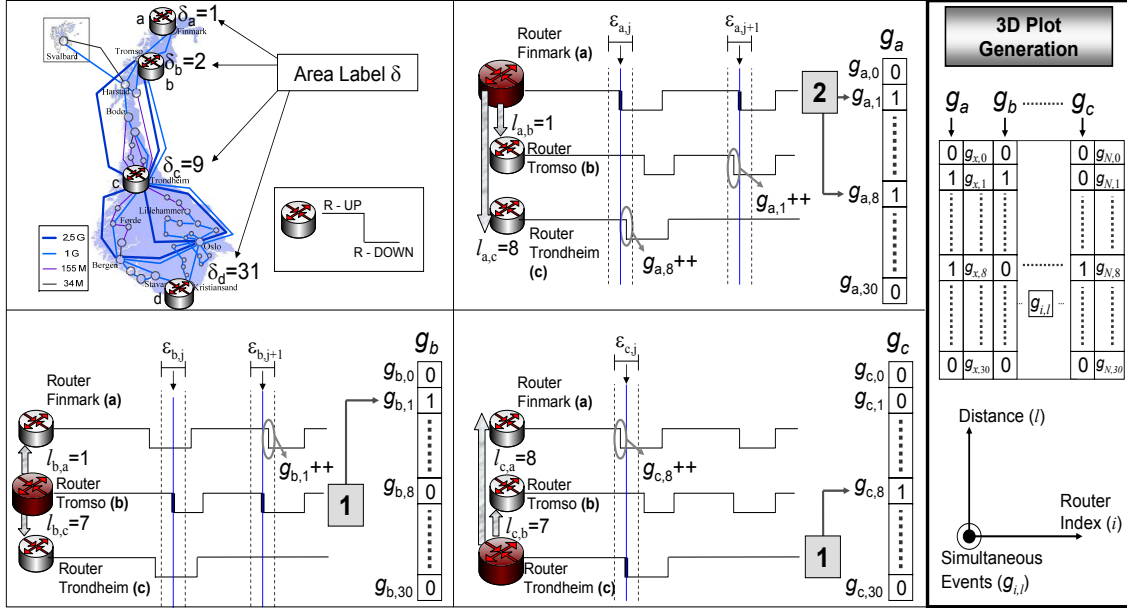


Figure 6: Methodology for the Construction of the Simultaneous 3D Graph

Figure 1(a) illustrates that network components are grouped in different geographical areas clearly defined and well delimited, therefore we can easily group all the components inside them to have an initial idea about the distance between elements, in a practical and organized way. For this reason each area will be labeled with a value  $\delta$ , using a geographic sweep from the north to south of Norway as is shown in Figure 6. The concept of *distance*  $l$  used in this paper will be the difference between the label of components  $i$  and  $x$  denoted by  $l_{i,x} = |\delta_i - \delta_x|$ . In this way for instance a router  $i$  located in Finmark will get a label  $\delta_i = 1$  and will have a distance  $l_{i,x} = 8$  from a router  $x$  in Trondheim which has label  $\delta_x = 9$ .

An appropriate way to illustrate the number of simultaneous failures of network elements and the *distance* between them, is by a three-dimensional plot.

The procedure to generate the mentioned 3D plot is illustrated in Figure 6 and works as follows: The occurrence of simultaneous failures will be evaluated on each network element  $i$  separately. When a down event  $j$  occurs at  $t_{i,j}$  a gap  $\varepsilon_{i,j} = [t_{i,j} - \Delta/2, t_{i,j} + \Delta/2]$  will be used to search which other component  $x$  suffered a failure  $f$  within this interval, taking into account the distance  $l_{i,x}$  between the two affected devices.

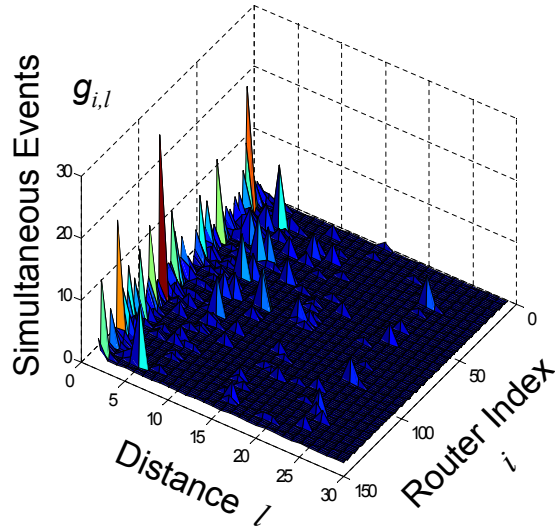


Figure 7: Simultaneous Down Events in Routers During 2007

When all the failures  $n_i$  in a network element  $i$  are checked, a vector  $g_i$  with the number of simultaneous events that occur per distance  $l$  will be obtained. If all these vectors are organized and plotted together, a 3D graph that shows the distribution of simultaneous down events will be obtained where each cell will have a value  $g_{i,l}$  given by the next equation.

$$g_{i,l} = \sum_{j=1}^{n_i} \sum_{\forall x \neq i} \sum_{f=1}^{n_x} I\left(t_{x,f} \in [\varepsilon_{i,j}] \wedge l = l_{i,x}\right)^4 \quad (5)$$

Figure 7 shows the results regarding routers where the  $x$  axis represents the *distance*  $l$  between network elements, the  $y$  axis contains the router index  $i$  and in  $z$  is located the number of simultaneous events  $g_{i,l}$ . A similar 3D-plot for links is shown in Figure 8 where the same kind of patterns are observed. In this case the procedure described in Figure 6 was used as well, but taking into account the link considerations described in section 2.

Figures 7 and 8 for routers and links respectively show a concentration of events on short distances, specially for the value 0, which means that the components are located

<sup>4</sup> $\wedge$  represents logical conjunction (AND operator).

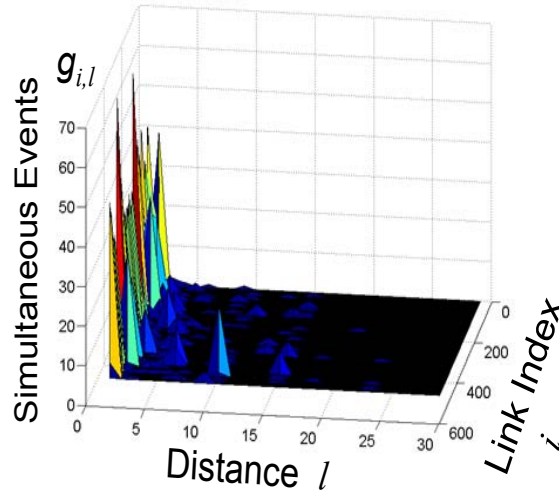


Figure 8: Simultaneous Down Events in Links During 2007

in the same node.

There is a clear dominant presence of simultaneous events within components located in the same geographical area. However, this is more pronounced for the case of links. Finally, when simultaneous events occur in links, the number of elements involved is much larger.

### 4.3 AUTOCORRELATION AND CORRELATION COEFFICIENT

The results obtained with the second method are very illustrative. They explain better the vertical patterns observed in the scatter plots. Nevertheless we will apply a third method in order to verify and observe closely the dependence in failure events.

We analyze the vertical patterns of Figure 4 and 5 evaluating the correlation between failure processes. Here the Pearson's correlation coefficient  $\rho$  will be calculated using two methods as is illustrated in Figure 9. In the first method the state of a network element  $i$  will be modeled as a working/failed signal that is divided in slots of size  $\Delta_1$  during a defined period  $T$ , obtaining a binary vector  $\mathbf{X}_{i,k}$  that takes value 1 if the element is up in the  $k$  interval ( $k = 1, 2, \dots, \frac{T}{\Delta_1}$ ) or 0 otherwise .

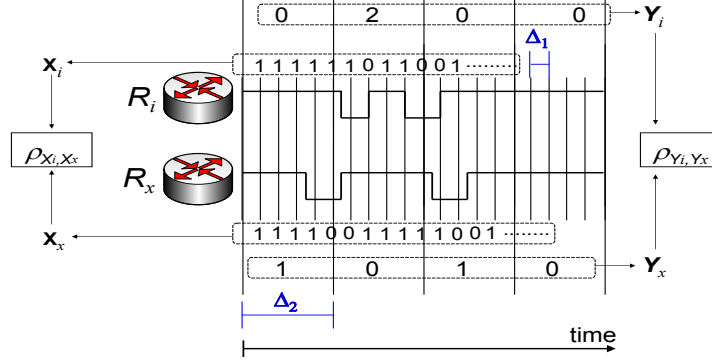


Figure 9: Methodology used to find  $\rho$

$$X_{i,k} = \begin{cases} 0 & t_{i,j} \leq (\Delta_1 \cdot k) < t_{i,j} + d_{i,j} \quad \exists j : 1, \dots, n_i \\ 1 & \text{otherwise.} \end{cases} \quad (6)$$

The value of  $\Delta_1$  is critical for this method, therefore we run several calculations with values between 1 and 100 seconds obtaining coherent and similar correlation values. Nevertheless, the huge size of the vectors obtained with this method is a considerable drawback for its application. On the other hand, bigger slots may be used if a second method is applied in order to obtain a vector  $\mathbf{Y}_{i,k}$  that count the number of down-events of component  $i$  within a given interval  $k$  ( $k = 1, 2, \dots, \frac{T}{\Delta_2}$ ) of size  $\Delta_2$ .

$$Y_{i,k} = \sum_{j=1}^{n_i} I\left(t_{i,j} \in [\Delta_2 \cdot k, (\Delta_2 + 1) \cdot k]\right) \quad (7)$$

In this study we choose  $\Delta_2$  values between 5 and 60 minutes, reducing the size of the vectors considerably.

The correlation should be obtained for every pair of elements  $(i, x)$  using the Pearson's formula:

$$\rho_{Y_i, Y_x} = \frac{E[(Y_i - \mu_{Y_i})(Y_x - \mu_{Y_x})]}{\sigma_{Y_i} \sigma_{Y_x}} \quad (8)$$

The obtained correlation values in approximately 96% of the cases were very close to 0, nevertheless in our study we put special interest on the values that were larger than

0.1.

Each evaluated pair has two features, first the distance  $l_{i,x}$  and second the corresponding correlation coefficient  $\rho_{Y_i, Y_x}$  between the respective vectors obtained according to the procedure explained in Figure 9. Therefore the use of a 3D plot is also appropriate for this case.

For this elaboration, the values of correlation will be grouped in nine discrete groups, defined by gaps  $\rho_{Y_i, Y_x}(\kappa)$  as is described in the next equation.

$$\rho_{Y_i, Y_x}(\kappa) = \kappa \cdot 0.1 \leq \rho_{Y_i, Y_x} < (\kappa + 1) \cdot 0.1 \quad \kappa = 1, \dots, 9 \quad (9)$$

The correlation is evaluated on every vector-pair with distance  $l$ , and assigned to the respective  $\rho_{Y_i, Y_x}(\kappa)$ . A vector  $h_\kappa$  with the number of correlation values that belongs to  $\rho_{Y_i, Y_x}(\kappa)$  per distance  $l$  is obtained. If all these vectors are organized and plotted together, a 3D graph that shows the distribution of  $\rho_{Y_i, Y_x}$  will be generated, where each cell will have a value  $h_{\kappa, l}$  given by the next equation.

$$h_{\kappa, l} = \sum_{i=1}^N \sum_{\forall x \neq i} I\left(\rho_{Y_i, Y_x} \in [\rho_{Y_i, Y_x}(\kappa)] \wedge l = l_{i,x}\right) \quad (10)$$

Figure 10 shows the results for  $\rho_{Y_i, Y_x}(\kappa)$ , using  $\Delta_2$  equal to 60 minutes for failures in routers during year 2007.

Additionally the 2-dimensional plots in Figure 10(b) help to visualize better how are distributed the different correlation coefficients calculated. The values illustrated in these marginal plots are calculated as follows:

$$\rho_\kappa = \sum_{\forall l} h_{\kappa, l} \quad l = 0, \dots, 30 \quad (11)$$

and

$$\rho_l = \sum_{\forall \kappa} h_{\kappa, l} \quad \kappa = 1, \dots, 9 \quad (12)$$

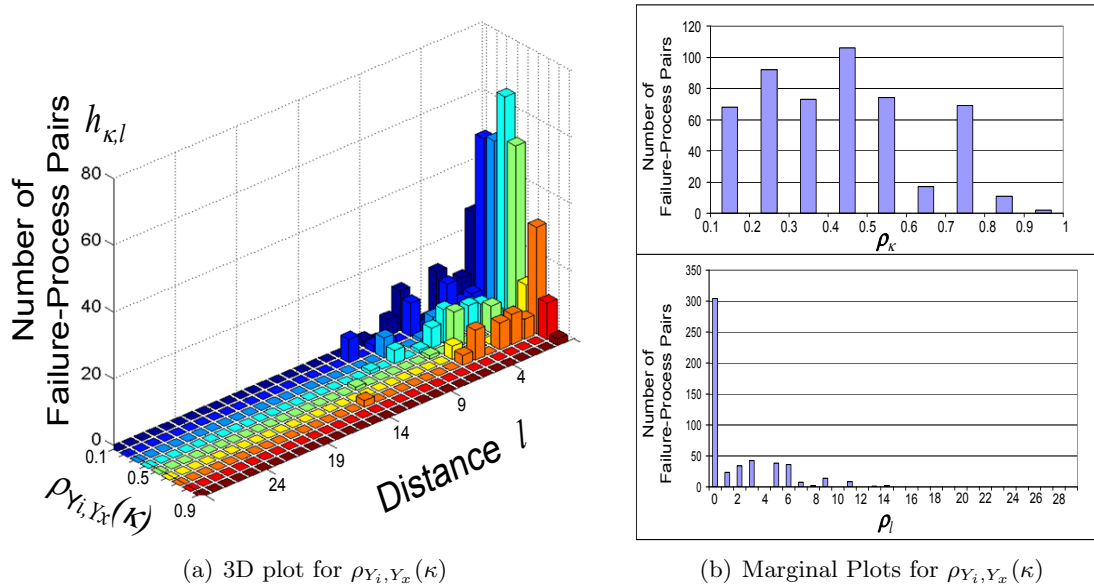


Figure 10: Number of router-failure process pairs with  $\rho_{Y_i, Y_x}(\kappa)$  larger than 0.1

The results obtained with this method show again a clear dominant presence of correlated events within components located in the same geographical area.

On the other hand, we also studied the horizontal patterns observed in Figures 4 and 5 through the formal evaluation of autocorrelation in individual components failures. This is made using a procedure where for each component  $i$  is obtained a vector  $\theta_i$  that contains the time between failures during  $T$  as is described in Equation (13).

$$\theta_{i,c} = t_{i,c+1} - t_{i,c} \quad c = 1, 2, \dots, (n_i - 1) \quad (13)$$

Using  $\theta_i$  we applied equation (14) to estimate the autocorrelation  $\rho_{\theta_i, \theta_i}(\tau)$ .

$$\rho_{\theta_i, \theta_i}(\tau) = \frac{\text{Cov}(\theta_{i,c}, \theta_{i,c-\tau})}{\sigma_{\theta_i}^2} \quad (14)$$

In some cases were found a high autocorrelation as is shown in Figure 11(a), given that for most of the lags the obtained values exceed the 95% confidence bounds that indicate acceptable values if the failure events were independent. These cases may be easily associated with the very pronounced horizontal patterns observed in the scatter

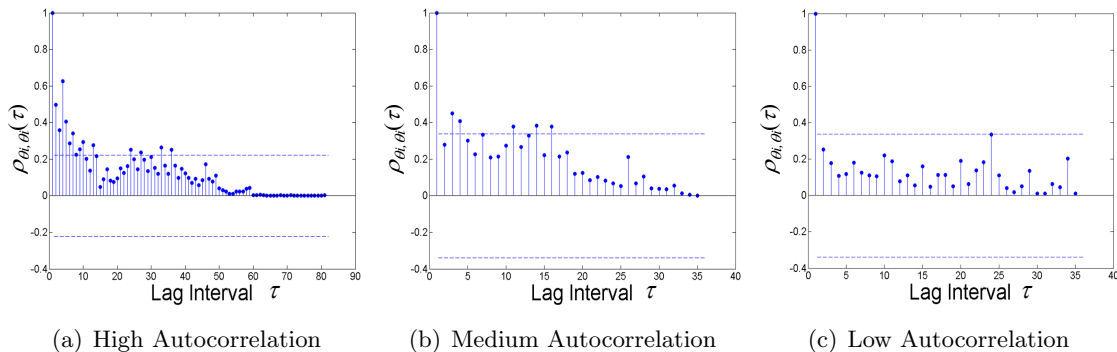


Figure 11: Autocorrelation in Failure Processes

plots. There are also some cases where medium levels of autocorrelation are observed in the sense that there are few lags that exceed the confidence bounds (Fig.11(b)). Finally there are many failure processes that seem to be independent as may be observed in Figure 11(c). The important conclusion is that the presence of high and medium levels of autocorrelation is not a negligible case.

To conclude this section we may say that after these three different methods the obtained results show clearly that the distance among elements has a huge impact in the failure correlation. Nevertheless, there is a small portion from the possible correlated events that occur when they are geographically far. Although this portion is small this finding is very interesting given that it was not expected.

Finally it is important to clarify that the same patterns and behaviors were observed in all chosen periods among the available downtime logs (2001-2009) where many different network configurations may be present due to the dynamics of the network caused by the installations and removing of components, giving to our findings a wider validity.

## 5 CONCLUSIONS AND FUTURE WORKS

The paper yields an improved insight into the failure processes at a real network. Correlation between failures are pronounced in both time and space. A main result obtained is that geographical distance has a significant impact, therefore this effect should be

considered in dependability studies and network design.

The independence assumption commonly used to model network dependability is incorrect, at least for the case of the UNINETT backbone network, nevertheless we believe that this conclusion may be easily extended to a wider amount of networks.

We found that there is a small portion of coincident failures that do not fit with the geographical location explanation. Therefore a deeper study of this kind of events may be analyzed in future works.

The results shown in this paper have direct implications in network design and backup assignments techniques like Shared Risk Group (SRG) used to have robustness under single link failures which means that the connections affected by one failure can not share any backup resource [8]. For instance the UNINETT topology shown in Figure 1 is vulnerable since Trondheim is an area that forms a "bridge" between southern and northern Norway. The "bridge" is well designed and has no single point of failure, but the geographical closeness increases the risk of simultaneous failures. Another use of the results is to complement the theoretical models that have been developed for assess network reliability in presence of interdependence between the component failures e.g. [9] and [7]. These studies assume the existence of some correlation in failure events, but empirical information that allows a proper parametrization is missing.

We have not looked into the potential causes of correlated failures, e.g. failures of the power grid, but we have a clear understanding that best practices are applied to avoid common cause failures. On the other hand, the *distance* used was based on geographical areas given that this makes easier the organization of information. Nevertheless this is rather crude and an improved insight may be obtained with a finer measure.

In this paper there were described two initial methodologies to analyze the failure events correlation from two different network elements, nevertheless there were found some limits given the size of the vectors obtained and the computational efficiency to use them, therefore new and innovative ways to obtain correlation values may be defined.

This work is an initial research that shows specific results in the analysis of depend-

ability logs in the UNINETT network.

## 6 ACKNOWLEDGEMENTS

UNETT provided the log of failure and reparation events in its backbone network for the period 2001-2009. Special help in the understanding of the logs was received from Jon Kåre Hellan (UNETT Scientist).

## References

- [1] D.R. Cox. *Renewal Theory*. Methuen, 1967.
- [2] Gianluca Iannaccone, Chen nee Chuah, Richard Mortier, Supratik Bhattacharyya, and Christophe Diot. Analysis of link failures in an IP backbone. *In Proc. of the Internet Measurement Workshop*, pages 237–242, 2002.
- [3] P Kuusela and I Norros. On/off process modeling of IP network failures. In *Proceedings of The 40th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2010)*, 2010.
- [4] P. Kuusela, I. Norros, and P. Raatikainen. Report on modelling the reliability of an ip-network and strategies for improving the reliability. Technical report, A report of the IPLU-II project, June 2009. Available at <http://iplu.vtt.fi>.
- [5] A. Markopoulou, G. Iannaccone, S. Bhattacharyya, C. N. Chuah, Y. Ganjali, and C. Diot. Characterization of failures in an operational ip backbone network. *IEEE/ACM Transactions on Networking*, 16(4):749–762, Aug. 2008.
- [6] Steven M. Matz, Lawrence G. Votta, and Mohammad Malkawi. Analysis of failure and recovery rates in a wireless telecommunications system. *Proceedings of the 2002 International Conference on Dependable Systems and Networks.*, pages 687 – 693, 2002.

- [7] M. Naldi and G. D'Acquisto. A normal copula model for the economic risk analysis of correlated failures in communications networks. *Journal of Universal Computer Science*, 14(5):786–799, 2008.
- [8] R. Ramamurthy, Z. Bogdanowicz, S. Samieian, D. Saha, B. Rajagopalan, S. Sen-gupta, S. Chaudhuri, and K. Bala. Capacity performance of dynamic provisioning in optical networks. *Journal of lightwave technology*, 19(1):40–48, Jan 2001.
- [9] Nozer D. Singpurwalla and Chung-Wai Kong. Specifying interdependence in net-worked systems. *IEEE Transactions on Reliability*, 53(3):401–405, 2004.
- [10] The Norwegian Research Network UNINETT. Downtime statistics. Available at: <http://drift.uninett.no/downloads/>.
- [11] Baek young Choi, Sejun Song, George Koffler, and Deep Medhi. Outage analysis of a university campus network. *Computer Communications and Networks, 2007. ICCCN 2007. Proceedings of 16th International Conference on*, pages 675 – 680, 13-16 Aug. 2007.