The dependability of an IP network – how to assess it?

Ilkka Norros, Urho Pulkkinen, Pertti Raatikainen and Eija Myötyri

The final report of the IPLU project

January 9, 2007

Contents:

1 Introduction 3
1.1 About this document and the IPLU project 3
1.2 IPLU's approach to the dependability of a ubiquitous IP infrastructure 3
1.3 The structure of this document 5
2 Definitions 6
2.1 Networking 6
2.2 Actors 6
2.3 Dependability 7
3 A system look at the dependability of IP networks 8
4 Dependability indices and requirements 10
4.1 General remarks 10
4.2 Qualitative requirements 10
4.2.1 Criticality of services 10
4.2.2 Qualitative redundancy requirements 11
4.2.3 Security 11
4.2.4 Personnel qualification 11
4.3 Indices and quantitative requirements 12
4.3.1 Downtime-frequency curves 12
4.3.2 On/off-availability and quality-availability 15
4.3.3 Focused availability indices 15
4.4 Component level reliability 17
4.5 Other network dependability indices 17
4.5.1 Security 17
4.5.2 Robustness 18
4.5.3 Traffic statistics 18
4.5.4 Indices related to operating personnel 18
5 Reliability analysis 19
5.1 Qualitative reliability analyses 19
5.1.1 Failure modes, effects and criticality analysis (FMECA, FMEA) 19
5.1.2 Operating experience analysis 20
5.2 Quantitative reliability analysis 21
5.2.1 Elements of system reliability modelling 21
5.2.2 Challenges of IP-network reliability analysis 22
5.2.3 Prerequisites of quantitative reliability modelling of IP-networks 25
5.2.4 Example 26
6 Dependability case 31
6.1 Background and general principles 31
6.2 Elements of a dependability case 31
6.3 Implementation of dependability case for IP-networks 35
6.4 An example of dependability case 35
7 Conclusions and suggestions 37
References 38

1 Introduction

1.1 About this document and the IPLU project

This document is the final report of the Finnish IPLU project in 2006 (see http://iplu.vtt.fi). IPLU's task was to create a conceptual framework for considering the complex problem "Can one rely on IP technology?" and to identify and develop methods for assessing the dependability of IP networks.

This report presents the conclusions and recommendations of IPLU. Other major deliverables of the project are the following:

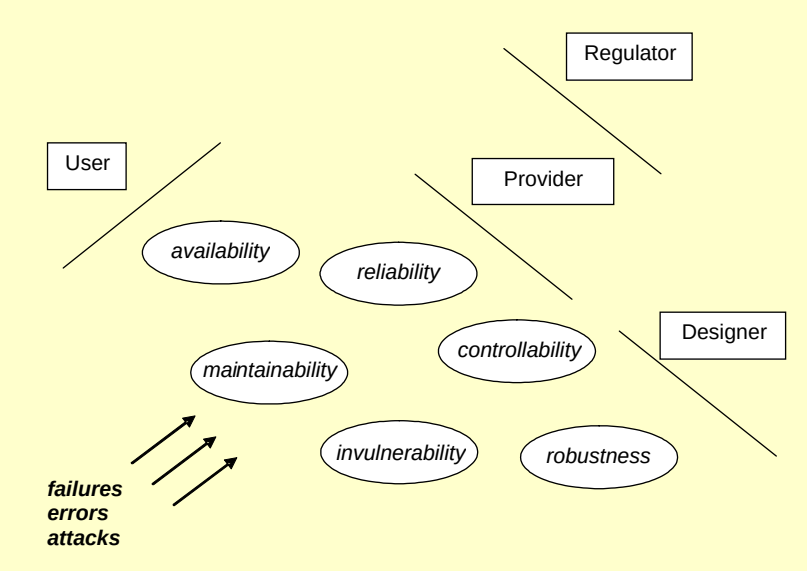
- IPLU's baseline paper "The dependability of an IP network – what is it?"
- IP-verkkojen luotettavuus - mitä se on?
  A public seminar organized by IPLU on May 17, 2006 (in Finnish)
- Dependability of All-IP networks
  A multidisciplinary international workshop organized by IPLU on May 18-19, 2006
- Kari Seppänen: Resiliency in Ethernet Based Transport Networks
- Pertti Raatikainen: Next Generation Network and Reliability

1.2 IPLU's approach to the dependability of a ubiquitous IP infrastructure

A project like IPLU was felt necessary because of the deep changes that are now happening in the telecommunications infrastructure when legacy networks are being replaced by IP networks. Classical telephone networks and television broadcasting are vitally important infrastructures, and with all the grandeur of IP technology, stability and reliability have not belonged to its foremost attributes. The following theses, debatable of course, present our concerns in more detail:

- the significance of IP networking has already grown tremendously, but the convergence towards All-IP is making it a still more critical infrastructure
- in longer run, everybody should have a reliable IP connection at a reasonable price
- the slogan that the network would disappear from the user's perception is deeply misleading - the future user also needs "network literacy"
- dependability is the Achilles' heel of the All-IP vision
- the present situation is not a stable one but one of deep-going changes

A broad, multidisciplinary approach was considered necessary from the start, because the number of relevant aspects and points of view to the problem is large, and we wanted to see more specific problem areas within a holistic map. The following picture from IPLU's baseline paper presents the aspects of dependability (ovals) that were found most relevant in this context, together with the generic actors (rectangles) who are concerned with or taking care of these aspects.



The notion of dependability and the somewhat unusual selection of its main aspects adopted here are discussed in detail in the baseline paper. The following remarks explain the main activities that the picture hints to:

- the **User**'s basic requirement is the **availability** of connectivity to the global Internet
- the **Provider**, who in this picture combines the roles of **Operator** and **Manufacturer**, makes his best to provide this availability by taking care of the **reliability** and **maintainability** of the networks, utilizing also the existing degree of **controllability** of network traffic behaviour
- the **Designer** is the collective acronym for the research and standardization activities that create the network protocols aiming at **robustness** and **invulnerability**
- the **Regulator** imposes rules to the **Provider** that force the latter to maintain the dependability of the networks at a level considered sufficient

These additional remarks clarify our focus further:

- *security* is seen mainly as an enabler for **invulnerability**
- *quality* and **availability** do not have a sharp boundary, sufficiently bad quality can be considered as unavailability
- the reason for *traffic congestion* causing bad *quality* or un**availability** may be the **Provider**'s bad dimensioning or a lack of **controllability** to which the **Designer** is responsible also

Multidisciplinarity was needed to combine the expertises in telecommunications technologies with those in safety and reliability. As its main (though tentative) proposal, IPLU developed the Dependability Case methodology that makes it possible to combine a heterogeneous set of relevant requirements, facts, tools, techniques etc. into an organized whole around a "claim" about a network. The methodology is an adaptation of the Safety Case methodology used in contexts like nuclear power but new in the world of networking. In the next phase, the multidisciplinarity should be further extended to include the expertises related to human activity, the users' (see [8]) as well as the telecom operators', and media theory (e.g. [17], [12]), considering global IP connectivity as a generic medium of future.

1.3 The structure of this document

- Chapter 2 specifies the terminology used in the sequel
- Chapter 3 discusses the character of the IP infrastructure from a system analytic point of view
- Chapter 4 considers different kinds of requirements and computable indices for formulation of quantitative requirements
- Chapter 5 gives an overview of various techniques and procedures of dependability analysis
- Chapter 6 proposes a general method for setting and verifying dependability requirements in the multi-actor scenario
- Chapter 7 summarizes IPLU's conclusions and suggestions

2 Definitions

In this document we use the following terminology. Interpretations of central terms also define the scope of the IPLU project.

2.1 Networking

IP network: a network that uses Internet Protocol to carry data

Internet: the globally connected set of autonomously operated IP networks

Access network: the part of a transport network that connects the end-user devices to the network

IP connectivity: ability to have an end-to-end connection from an IP interface to another IP interface between two terminals: IP packets are transmitted correctly; of one user: the user's access transmits IP packets correctly to and from the Internet

Virtual Private Network: a set of layer 2 and/or layer 3 tunnels that provide connectivity between sites of a user organisation separated from connectivity of other organisations

Transport: in this document all functionality that transfers an IP packet from an IP interface to another without inspecting the IP header; in practice, transport involves protocol layers below the IP layer, i.e. physical signal transport, MAC and link layers as well as possible intermediate layers, such as MPLS

Telephone: (i) POTS, GSM, UMTS; (ii) various VoIP services

Television: public audiovisual broadcast service

2.2 Actors

Individual user: human individual, possibly with several IP access potentials and home network

Corporate user: organization with IP access; may have corporate network and/or VPN

Telecom operator: an organisation providing carrier service for end-users or virtual operators

Internet service provider: an organisation that provides Internet access to end-users

Content provider: an organisation that provides services to end-users or telecom operators; examples of possible services are shopping, web surfing, chat rooms, playing games, accessing data, music or books, TV programs; from the network point of view, content provider is a corporate user

Regulator: authority with law-enforced power to set requirements on networks and their operation practices

Manufacturer: producer of (i) networking equipment, (ii) terminal equipment, (iii) networking software

Designer: research, standardization (generic term for people and institutions who create network algorithms and protocols; non-standard term)

2.3 Dependability

Dependability: An integrating concept encompassing several attributes. In this report we apply the approach developed in [2], see Section 1.2 above. Some variations:
Avizienis et al [4]: availability, reliability, safety, integrity, maintainability
ASC, IEC [1]: reliability performance, maintainability performance and maintainability support performance
Villemeur [30]: reliability, availability, maintainability, safety

Dependence: dependence of system A on system B represents the extent to which System A's dependability is or would be affected by that of System B [4]

Trust: accepted dependence [4]

Availability: readiness for correct service [4]

Reliability: (i) wide sense: about same as dependability; (ii) continuity of correct service [4]

Maintainability: ability to undergo modifications and repairs [4]

Robustness: a property of algorithms and protocols, meaning stability with respect to arbitrary inputs; more generally [4]: dependability with respect to external faults

Vulnerability: possibility of behaviour that leads to serious degradation of performance by a cause that comes from outside of the system's normal constituents

Controllability: in our context, characterized by the possibilities of single agents (mainly the telecom operators) to accept, reject, or route traffic offered to the network, and to open and close individual services or the whole network

Security: *integrity* of information and *confidentiality*, absence of unauthorized disclosure of information, and availability for authorized actions [4]

Redundancy: system's property to possess multiple, more or less independent ways to realise its service

Failure: transition of a component or system to a state where it does not any more deliver its correct service (a failure is rather on the "surface", its cause being somewhere "deeper") [4]

Error: deviation of a state of the system from its correct service state (need not be manifested) [4]

Fault: cause of an error [4]

Resilience: fault tolerance [4]

3 A system look at the dependability of IP networks

This chapter presents some general facts and thoughts about what kind of system we are studying when we want to consider the dependability of IP networks as a whole.

First, the basic service provided by this system is to transfer a data packet from any access point to any other access point. This task is conceptually very simple and natural. Whatever developments we shall meet in future telecommunications, it is hard to think that the global packet transfer service now realised by the Internet would ever become obsolete.

Second, the service of the system is provided by the collaboration of autonomously owned and operated subsystems that are connected in a highly complex way. The topology of the network of autonomous systems was characterized as "soft hierarchy" by Reittu and Norros [23]. The autonomous systems share a few common resources like the Domain Name System and authorities like the Internet Assigned Numbers Authority. The dependability of this supersystem is a complex problem with technical, commercial and political dimensions. From a purely topological point of view however, the soft hierarchy structure is extremely resilient (Reittu [24]).

Third, coming to the national level in Finland, the system consists of a few independently owned and operated transport networks and a larger number of additional Internet Service Providers who mainly hire the transmission links from the previously mentioned ones. The ISPs exchange traffic mostly within a single institution, the national exchange point Ficix. International connections are not provided by Ficix, but all major ISPs maintain their own international links. National regulations are issued by Finnish Communications Regulatory Authority (Ficora).

Fourth, at the level of a major IP telecom operator, the network can be clearly divided into a core network and access networks attached to it (the higher parts of access networks may also be called by more specific terms, for example metropolitan area networks). The core networks have redundant topology, whereas the access networks have typically, perhaps with the exception of their higher, metropolitan level, a tree topology. The IP traffic carried by a mobile telephone network is typically carried within the mobile network infrastructure to one (or a few) router(s) connected with an IP core network.

Fifth, the usage of the IP service requires a variable number of auxiliary services that are implemented in servers and must be available as well as properly configured: DNS, DHCP, authentication.

Sixth, the transmission networks and link layers of user access technologies are and will be very heterogeneous and have further layered structures. They are complicated systems that are configured by the telecom operators' highly qualified personnel. The increasingly popular Ethernet networking seems to be a particularly error-prone and vulnerable technology (Seppänen [26]).

Seventh, the use of IP packet transfer to real-time services like telephony and television requires not only the functioning of the respective application layer systems but also the *availability of quality* at the IP layer in terms of bandwidth and delay. It is important to realise that the basic design of IP networking does not include quality provisioning as an inherent feature. IP routing protocols and

TCP traffic control have built-in robustness in their design, whereas quality provisioning is essentially more complicated and usually limited within autonomous systems.

Eighth, all digital networking requires electricity. Without dedicated efforts, the IP networking vanishes together with electricity, unlike the traditional telephone network which carries the needed power to the terminals.

From this description, we can draw some characterizations of the system: is it a “hard” or “soft” system, is it complex, and what are its control mechanisms like?

1. The system is not a single though huge "machine" in the sense in which such a characterization might still have applied to the world-wide POTS. The dependability concerns of pure telephone networks focus on protection switching, reliability of switches and components, and the availability and securing of electricity supply. IP networking should rather be analysed as a "soft system" in terms of Checkland [9], driven by a large number of organizations and renewing continuously its technological base. New protocols are introduced and existing ones have frequent version updates. Traditional reliability analysis and “hard system” theory are meaningful for various technological subsystems, but for the total system, which also contains several relevant human activity systems:

- the users' activities where universal IP connectivity is used as a generic medium: the dependability requirements depend on the actual usage of different applications that pose different requirements on the IP network, and in this respect the situation is continuously changing in unpredictable ways
- the reflection of users' requirements in legislation and regulations
- the telecom operators' competition and collaboration
- the human operators' work

2. Is our system complex? The answer is not as unconditional “yes” in every respect as one might think at first sight. Despite the huge size of the system, the system seems to have so much hierarchy and modularity that meaningful reliability analyses of abstracted subsystems are possible. As a purely technical system, the TCP/IP layer network is not an extremely complex system, if all components, including software components, are modelled as binary ones (functioning or not). Indeed, the fulfilment of a set of conditions that are all necessary is logically simple and corresponds to a series system. Possibly the transmission networks are in fact more complex than the IP layer. Interworking for connecting services is greatly facilitated by concentrating it on IP layer. On the other hand, routers are sophisticated computers with complicated potential behaviours, and detailed analysis of failure modes of distributed software components of IP networking would be a very complex task. Thus, the question on the complexity of the IP networks depends on the level of detail.

3. As regards control, the total system is obviously not operatively controlled by anybody. In longer timescale, the system develops in complex ways and surprises abound. Main dynamic control tools available for an operator are traffic engineering on one hand and traffic filtering, i.e. the removal of unwanted traffic from the network on the other hand.

4 Dependability indices and requirements

4.1 General remarks

Dependability requirements set for a telecommunication network (or part of it) are requirements set on observable characteristics of various aspects of the dependability of that (part of the) network. Requirements can be set by the User, by the Regulator or by the network Provider itself. They may be specified in regulations or in agreements between User and Provider or between two Providers.

Dependability requirements can be either qualitative or quantitative. We discuss below both types and suggest some for further consideration.

By a *dependability index* for a telecommunication network (or part of it) we mean a *summarizing quantitative* characteristic of some aspect of the dependability of that (part of the) network. Obviously, quantitative dependability requirements should concern only indices that are *measurable*.

The following table shows what actor (column) is expected to set requirements on what aspect of dependability (row) to what actor (table entry).

	User	Operator	Manufacturer	Designer	Regulator
Availability	Operator	Operator 2			Operator
Reliability	Manufacturer (terminal)	self Manufacturer			Operator
Maintainability	Manufacturer (terminal)	self Manufacturer	Manufacturer 2		Operator
Robustness		Designer	Designer	self	
Controllability	Operator	Manufacturer	Designer	self	Operator
Invulnerability	Operator	User Operator Manufacturer Designer		self	Operator Manufacturer Designer

4.2 Qualitative requirements

4.2.1 Criticality of services

It is natural to require that the most critical functions of the society should be best protected and secured. Various critical infrastructures tend to become more and more tightly interlinked with the communication infrastructure (for the status of critical information infrastructure research in Europe, see [10]). The development of the latter is in turn characterized by the growing role of IP networking. This prospect poses serious concerns, because the regulations can only follow the actual developments with the delay of a few years.

IP-based services appear that are in the beginning more like a curiosity or just fun, but because of their cheapness and often good quality they have the potential to grow within a short time to dominant role. “Toys” become serious tools, and legacy tools get antiquated. A big part of the users can well accept services that have inferior dependability in one or several aspects.

Services using common Internet packet transfer are difficult to secure. The common protocol and universal connectivity create unavoidable dependence relations with all other processes utilizing the Internet.

One important example are emergency services. Emergency calls are protected with several special regulations (see Ficora’s regulation 33B/2005 M). It is not clear how they should be extended if a significant part of telephony moves to pure IP networks. However, emergency calls are only one of a large number of critical functions that are moving to the Internet. Commerce, banking and various kinds of logistics all need IP connectivity that can be trusted.

4.2.2 Qualitative redundancy requirements

It is a basic rule that the Regulator shall not tell how the networks are built, because then it would make itself responsible for the outcome. Instead, certain more general characteristics are required. One of them is the qualitative requirement of *redundancy* in important parts of networks. Ficora gives redundancy requirements for example in regulation 27E/2005 M. The general idea behind them is the prevention *single points of failure*. Most of the network’s functionalities should survive the loss of any single link or node. A crucial additional requirement is that two elements protecting each other should be as independent of each other as possible in the sense that the loss of one of them should not much increase the probability of the loss of the other.

In an IP network, however, the principle of no single points of failure is difficult to realise literally, because common protocols and distributed functionalities carried by them are themselves single points of failure, if their functioning becomes for some reason seriously disturbed, for example by a dormant fault in commonly used software.

4.2.3 Security

Information security is the most important enabler for the invulnerability aspect of dependability. By the way, this might be the only aspect of dependability where requirements are directed also toward the User.

4.2.4 Personnel qualification

High qualification of operating personnel is an indispensable feature of any high-dependability socio-technical system. Human errors made in network configuration are generally recognized as one of the major sources of faults in networks. Ethernet seems to be particularly dangerous in this respect. In everyday operation the freshly made configurations go often directly into use without much testing. Faults in configuration get particularly easily created when changes are made to equipment or software.

The change of telecommunications from stable monopoly-governed infrastructure to a rapidly changing multi-system to which companies with very different capacities contribute has made the personnel qualification challenge particularly urgent. Besides error-free configurations, the operating personnel also has to take care for information security in the continuous change of technologies and products. Finally, telecom operators face a hard competition that creates the pressure to cut personnel costs. Particularly dangerous from the dependability point of view is the erosion that may result if several experts leave a company at the same time.

In addition to formal education and courses, requirements should be set for good working practices and culture. These are normal requirements in safety-critical work (nuclear power, aviation etc.), but the special qualifications of work and expertise of operating networks has not been given the attention it deserves.

4.3 Indices and quantitative requirements

4.3.1 Downtime-frequency curves

The traditional way to set quantitative requirements for the availability of telecommunication system is to give a single number, usually consisting of decimals "9", for example 0.99999, a.k.a. "five nines". This characteristic is however rather uninformative, because it speaks nothing about the lengths of the individual downtimes. The five nines are satisfied as well by one five minutes break in a year as by a one second break almost every day. The acceptability of these interpretations may be totally different depending on the user's needs.

We propose instead the use of *downtime-frequency curves* that characterize the frequency of each down-period length separately. They are defined as follows.

Consider first the characterization of the reliability of a system or, similarly, availability of a resource, with binary nature: at each timepoint t, it can be unequivocally stated whether the system is up or down. Thus, its performance is described by a {0,1}-valued stochastic process:

$$I_t = 1_{\{system\,down\,at\,time\,t\}}.$$

The probability of failure, $P(system\,down\,at\,time\,t) = E\,I_t$, is already a characteristic of the reliability of the system. Assuming stationarity and ergodicity, this number is independent of t and obtained almost surely as the limit of the observed relative frequency:

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T I_t \, dt = E\,I_0 \text{ a.s.}$$

Let us define the ongoing down-period length at time t as

$$W_t = \inf\{s \geq t : I_s = 0\} - \sup\{s \leq t : I_s = 0\}.$$

When the system is up, we have $W_t = 0$. The relative share of time spent in down-periods lasting longer than τ during an observation period of length T is then given by the random variable

$$\varphi_T(\tau) = \frac{1}{T} \int_0^T 1_{\{W_t > \tau\}} \, dt.$$

Considered as a random function of τ, $\varphi_T(\tau)$ is non-increasing. Its initial value $\varphi_T(0)$ equals the relative overall downtime of the system in the observation period (for example one year).

If the system is stationary, W_t is a stationary stochastic process, and we find that the expectation of the random function $\varphi_T(\tau)$ equals the tail distribution function of the random variable $W = W_0$:

$$F_T(\tau) = E\varphi_T(\tau) = \frac{1}{T}\int_0^T P(W_t > \tau)\, dt = P(W > \tau).$$

Using this framework, we can now formulate reliability criteria that take into account the down-period lengths also: let us consider the performance of the system acceptable if

$$\varphi_T(\tau) \le \psi(\tau)$$

for some selected function ψ. The function ψ can, for example, be specified in a Service Level Agreement. The telecom operator has to build the system in such a way that the expected curve $F_T(\tau)$ lies sufficiently much below $\psi(\tau)$. Since the relevant values of both the down-periods and the probabilities extend over many orders of magnitude, the curves should be drawn in a log-plot or even log-log plot.

A standard mathematical model of this kind of process is the alternating renewal process, where the up- and down-periods are *independent* random variables with distributions G_{up} and G_{down} and means μ_{up} and μ_{down}, respectively. When I_t is a stationary version of an alternating renewal process, the distribution of W is

$$P(W > \tau) = \frac{1}{\mu_{up} + \mu_{down}} \int_\tau^\infty y\, G_{down}(dy).$$

Note that the distribution G_{up} has an effect only through the expectation μ_{up}.

Here is a formal example of such plots. Assume that the down-periods and up-periods are independent, time unit is one hour, the up- and down-periods have means 3000 and 1, respectively, and the down-period length has a Weibull distribution

$$1 - G_{down}(y) = \exp(-\beta\, y^\alpha),$$

where α and β are parameters. The choice $\beta = \Gamma(1+1/\alpha)^\alpha$ yields the desired mean 1. We can now compute and plot the functions $F_T(\tau)$ for three qualitatively different parameter values $\alpha = 0.5,\ 1$ and 2. This example also illustrates the usability of linear, log-linear and log-log plots for various purposes.

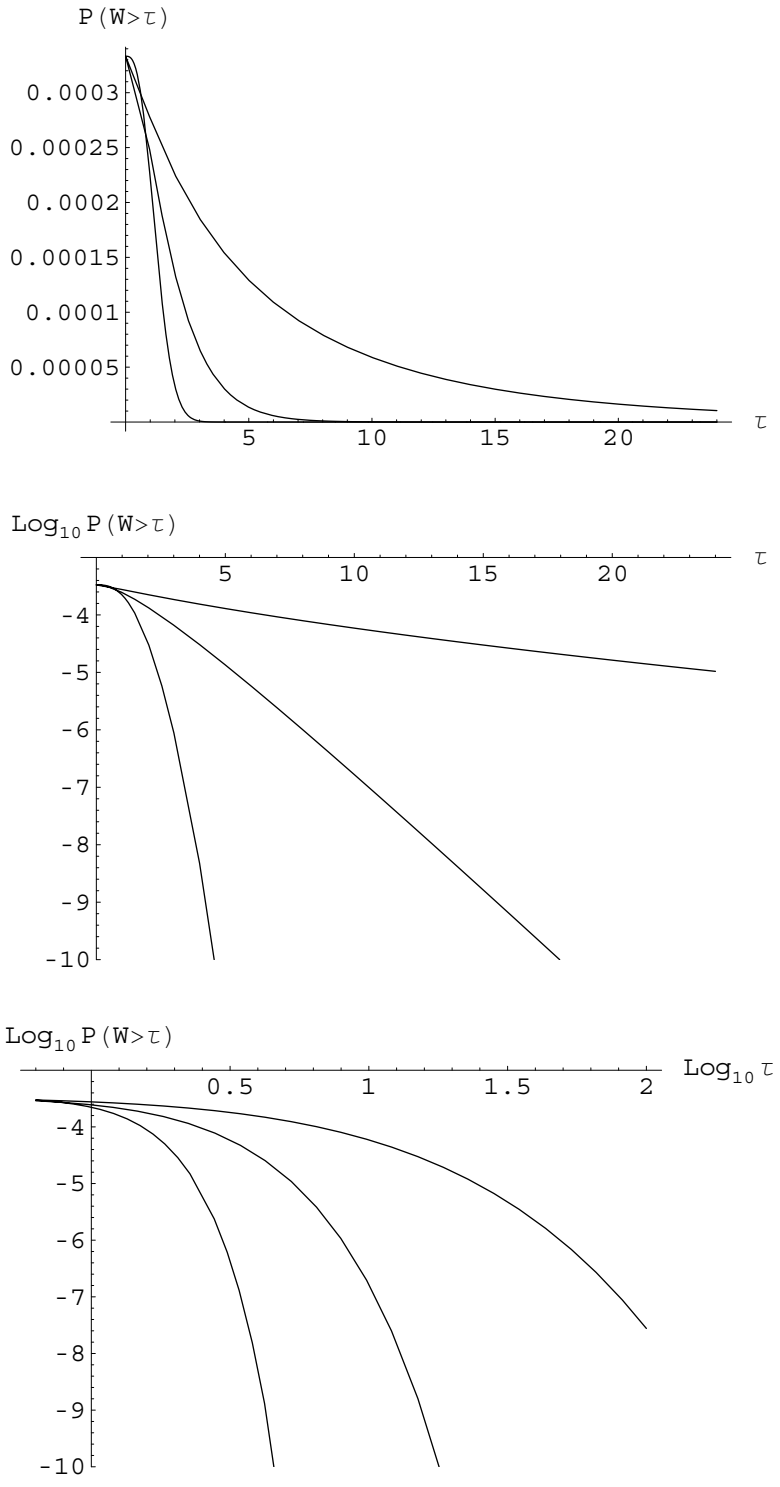


Figure 1. Downtime-frequency curves, when an individual down-period has Weibull distribution with mean 1 and exponent α= 0.5 (highest), 1 (middle), 2 (lowest). The relative mean downtime is 1/3001 in all cases.

As an example how empirical data might look like in this framework, assume that the downtimes of a system within a year consist of intervals with lengths 2, 2, 2, 3, 3, 5, 5, 6, 7, 9, 25, 35, 240 minutes (in ascending order). The empirical tail distribution function of W is then determined by the points

marked as triangles in Figure 2. Note that the few long down-periods have the effect that the whole point set looks almost horizontal. The other three point sets show the corresponding plot when 1, 2 and 3 largest values are removed from the data set, respectively.

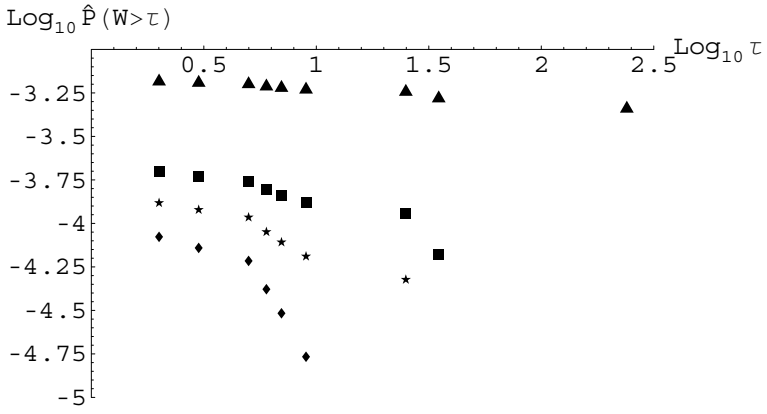


Figure 2. Data plot example in log-log scale.

4.3.2 On/off-availability and quality-availability

The downtime/frequency curves discussed in the previous subsection refer to a service whose availability at a given time is a binary variable. This is often not sufficient when we consider IP systems. The packet transfer may work "reliably" in both directions but proceed with much lower rate and/or higher delays than in normal conditions. From a mathematical point of view, however, this problem can be reduced to the binary case simply by considering the *set* of binary processes

$$1_{\{q(S_t)\le r\}},\ r \in \Re,$$

where S_t is the system state at time t, q is some characteristic of it (rate, delay,...), and R is the set of possible or relevant values of that characteristic.

For example, an SLA may require that the bandwidth of an MPLS path be higher than 50 Mbit/s with an availability of at least 0.999. Then, the set R may contain the value 50 Mbit/s alone. However, since IP-based services are usually quite flexible with respect to bandwidth requirements, it would make sense to require additionally that the availability of 5 Mbit/s be at least 0.99999.

One can also, at least in principle, let R be a whole interval and replace the binary-case criterion that the empirical values should lie below a curve to the two-dimensional criterion that they should lie below a surface. If higher $q(\cdot)$ means better quality, the monotonicities behave similarly in both dimensions if r is replaced by some inverse parameter β by writing, for example, $r = 1/\beta$.

4.3.3 Focused availability indices

The availability indices of the previous section can be applied to different network sections. We divide the discussion into the personal user, corporate user and operator points of view. We note however that although the availability of IP connectivity (with or without quality specification) is in

a sense the only thing that really matters, measuring and sanctioning only the realised availabilities as black boxes is not a fruitful approach for really improving the dependability of IP networks. The Dependability Case approach proposed in Chapter 6 below is a richer framework where the interface of actors discussing dependability matters is not restricted to the targeted and realised availability indices alone. Nevertheless, aiming at quantitative and verifiable target levels and at the computability of probabilistic availability estimates is an indispensable part in developing networking with better and more transparent dependability. Indeed, the computability of availability would be a very useful design requirement for networks!

4.3.3.1 Individual user point of view

Relevant section: The special requirement of a personal user is the availability of the access path from the user's terminal to a core (or core-edge) router.

Verification: It may be impossible to monitor the access availability comprehensively even for fixed access technologies, on the simple reason that users should recommendably save energy and switch their terminals off when they are not used. The verification of a claimed availability level of IP connectivity should however be possible for the path between the first multiplexing element like DSLAM and a core router. The availability of necessary facilities like DNS and DHCP can also be monitored from the DSLAM.

Discussion: Despite of the often declared vision of ubiquitous IP-connectivity, it depends very much on the access technology what availability requirements are reasonable. As regards quality-availability, only probabilistic guarantees can be given for technologies with capacity sharing, which includes public wireless access technologies and cable modems. On the other hand, for an IP television subscriber it is natural to require the same level of quality-availability as that of the usual cable television. This exceeds considerably the availability level that people are accustomed to in their data communications. If the goal of ubiquitous IP connectivity is taken seriously, high availability is required also for lower QoS or speed levels. Users' SLAs should contain realistic target values of availability indices, reflecting the access point's location and technology.

4.3.3.2 Corporate end user point of view

Relevant section: For corporate users demanding continuously working IP connectivity, we identify two different cases:
(i) Internet access alone: the path(s) between one or several customer's hubs or routers and a core router.
(ii) VPN connecting the customer's premises: continuous connectivity between all premises.

Verification: Monitoring is possible in principle because the user's access is always on.

Discussion: Although the user's primary need is quality availability alone, a SLA can contain further requirements, for example, that connections should be redundant.

4.3.3.3 Operator point of view

Relevant section: A full-scale telecom operator should maintain availability indices for several network sections:

(i) access networks and facilities (DHCP, DNS)

(ii) availability of core: all edge-router/edge-router pairs

(iii) external connections: core/FICIX, international links

Verification: Monitoring is possible and done already at present, except perhaps for (i).

Discussion: Summarizing statistics of unavailabilities should include estimates of the number of end users having suffered because of each. Criticality of connections is another dimension that should be recorded. Core networks should aim at continuous availability, even during the update breaks of individual routers. The operators' practical experience tells that failures are much more common in access networks than in the core. Most faults concern single customers. The order of magnitude of the fault frequency of multiplexing devices like DSLAMs might be about one per week in whole Finland.

4.4 Component level reliability

The standard reliability indices for hardware components are *fault probability* for off-the-shelf components and *mean time to failure (MTTF)* for working components and network elements like

- fibres, cables
- linecards
- switches
- router hardware
- server hardware

Manufacturers can be required to provide estimates of these numbers, and network providers should be aware of them. It is however not common among Finnish operators at least to set numeric reliability requirements for elements like routers. One obvious reason for this is that the routers outdate because of lacking capacity faster than physically.

Software reliability is harder to characterize numerically. This is due to the fact that software errors are design faults, which manifest as failures only when a suitable input sequence occurs. The software reliability depends both on the software and its operating environments. In principle, standard reliability indices, like failure rate of *MTTF* are applicable also for software failures. However, data on software errors are not systematically collected.

4.5 Other network dependability indices

4.5.1 Security

Although requirements related to the invulnerability aspect of dependability are qualitative by nature, statistics on observed intrusion attempts and successes are a useful and important tool in the work towards invulnerability. *Barrier thinking*, which is discussed in more general terms later in

this document, can be implemented here by recording how vital parts of the network were endangered.

The network providers are also obliged by Ficora's regulation 13/2005 M, 5§, to follow the amount of nuisance traffic and to filter it out when observed. Statistics about such traffic tell about the controllability aspect of the network's dependability, and to some extent also about its vulnerability.

4.5.2 Robustness

The robustness of the routing protocols' performance and the magnitude of route flapping can be to some extent monitored by recording the frequency of the computing of IGP routing tables or the link state advertisement packets.

This kind of monitoring could deliver valuable feedback to the designers of protocols. Such communication between Provider and Designer is one feature of good "dependability culture".

4.5.3 Traffic statistics

As noticed already, traffic is itself one factor of the dependability in IP networks. Ficora's regulation 50B/2003 M, 5§, obliges the network provider to monitor traffic volumes regularly and partly continuously. Network must be dimensioned so that blocking is rare even in busy hours. There are however no prescribed numbers what traffic loads are considered safe, but this is left to the operator's judgment. The current rule of thumb seems to be that a load of 30-40 % be safe in the core network, whereas 50 % is already critical. The growth of traffic has been such that the routers are typically changed every 3 years to 10 times more powerful ones.

4.5.4 Indices related to operating personnel

Competence requirements are of course mainly qualitative, but the amount of people competent to configure the network together with their location (remoteness – whether in same country for example) may be of some use in the characterization of the human factors of network dependability.

5 Reliability analysis

5.1 Qualitative reliability analyses

The qualitative reliability analyses aim at identifying failure modes and their impact on the system. Basically, they are systematic approaches, which help the designers and reliability analysts in finding the vulnerabilities of the system, that is, in identifying the causes and impact of failures or other disturbances and possibilities to prevent failures. Many of the qualitative reliability analyses have been standardised and they are often required in the agreements between system or component vendors and purchasers.

The qualitative reliability analyses also form the basis of reliability modelling and quantitative reliability analysis. It is not possible to create a reliability model without a careful qualitative analysis. These analyses can also be applied to systematise the collection and analysis of operating experience.

5.1.1 Failure modes, effects and criticality analysis (FMECA, FMEA)

One of the most important and well known qualitative reliability analyses is failure modes, effects and criticality analysis (FMEA, FMECA). It is an easy to use and yet powerful pro-active engineering quality method that helps to identify the weak points of the system in the early conception phase. It is applicable to all kinds products (hardware, software) and processes. It is a structured approach which makes it easy to use even for non-specialist. It is a standardised method (see e.g. [14]).

There are several types of FMEAs, some of which are used much more often than others. FMEAs should always be done whenever failures would mean potential harm or injury to the user of the end item being designed. The types of FMEA are:

- System FMEA- focuses on global system functions
- Design FMEA- focuses on components and subsystems
- Process FMEA- focuses on manufacturing and assembly processes
- Service FMEA- focuses on service functions
- Software FMEA- focuses on software functions

The basic objective of FMEA is the identification of failures and their impacts.

In the FMEA-process, the system under analysis is first divided into functional parts or components, and the functionality of parts is described. The failure modes of each component are identified systematically, i.e. the possible failures of each part are postulated. The possible causes of failures are identified, and, more importantly, the possible impacts of failures both on system and component level are identified. The possibility to detect the failures both before and after their occurrence are analysed, and the ways to prevent the failures and their impacts are identified. Finally, the criticality of the failure is analysed and failure rate is evaluated. The FMEA-analysis is documented carefully by using FMEA-forms. There are several computer-aided tools to perform

FMEA (see http://www.fmeainfocentre.com/tools.htm). FMEA can be made as a teamwork or by an individual analyst.

FMEA is often combined with a maintenance effects and criticality analysis (MEA, MECA), which focuses on identifying the impacts of the corrective or preventive maintenance on the components and system. In MEA, the repair/maintenance actions of each part and its failure mode are described (compared FMEA), and the impact of maintenance is identified both at system and component level. A simple example of a FMEA sheet is given in Table 1.

Table 1. An FMEA sheet.

FMEA								
Part	Function of the part	Failure mode	Failure cause	Consequences of failure		Detection of failure	Failure rate	Other observations , e.g. failure criticality, failure prevention
				Part level	System level			
	...	...	...	...	...	...	...	...

In addition to the above basic versions FMEA and MEA, there are several versions of the methods. They are often focused on a specific type of systems or technology, but they follow the above generic approach. Further, methods have been developed for identifying human errors in maintenance and installation, as well as for identifying common cause failures. In the case of IP-network reliability analysis, it could be advantageous to develop a FMEA-type method for identifying protocol configuration errors and their impacts.

The application of qualitative reliability analyses is highly recommended. They form the necessary basis for the reliability modelling and quantitative analysis and provide information which can be used to develop the maintenance and operation procedures.

5.1.2 Operating experience analysis

The collection and analysis of operating experience is important for many reasons. First, authorities may require the analysis of critical failures. The analysis and feedback of failure data and information to designers is useful in designing new components and systems. Together with FMEA, the analysis of operating experience helps in the identification of failure modes, human errors and common cause failures.

The problem in practice is that operating experience data is not collected or is not documented in a good format. In an ideal case, each failure/disturbance event should be recorded at least with the following information

- exact location of the failure
- exact time of the failure event
- cause of failure (e.g. random, configuration error, external fault)
- analysis of the impact of failure on both system and component level, impact on the network

- description of the maintenance and repair action
- outage time, repair time (at system and component level)

Actually, the above information is a part of a FMEA, and the operating experience data can be documented in FMEA-type sheets.

The modern system and component monitoring techniques help the automatic collection of operating experience data. However, the analysis part cannot be automatised, and it requires that certain amount of personnel and effort is allocated to it.

5.2 Quantitative reliability analysis

5.2.1 Elements of system reliability modelling

Quantitative reliability analyses aim at evaluating the systems reliability by using probabilistic or statistical measures. Depending on the case under analysis, the reliability is evaluated as

- the probability that the system operates without failures over a certain period of time in certain circumstances,
- the probability that system performs its function when demanded, or
- the probability that the system is operating at certain timepoint (= availability).

The above probabilities are usually very near to 1, and it is more convenient to use their complements (probability that at least one failure occurs in certain time period, probability that system fails to perform its function, and the probability that system is not in functioning state at certain time point (unavailability)).

System reliability models are used to determine systems reliability properties on the basis of the reliability properties of its components. To do this, two types of models are applied. First, the systems reliability structure is describe by using suitable models, which explain the systems failure events as a function of the failure events of the components. Usually it is assumed the state of the system and the states of the components are binary (failed or functioning). Thus the systems state is a binary function of the states of its components. This kind of functions can be presented graphically in the form of fault trees. Binary functions can be presented by using the concept of minimal cutsets (minimal set of component failures leading to a systems failure), e.g. in the form

$$S = \phi(X_1, X_2, ..., X_n) = \sum_{i=1}^{M} \prod_{j \in C_1} X_j ,$$

in which S is a Boolean variable representing the systems state, X_j, $j=1,...n$ are the Boolean variables representing the component states, $\varphi(.)$ is the Boolean structure function, C_i, $i=1,...,M$ are the minimal cutsets and the sums and the products are Boolean. Corresponding expressions have also been developed for multistate systems.

The probability that the system is failed is determined by using second type of models, which describe the probability that a component is failed (or the probability that a minimal cutset occurs). These models require information about the components failure rate, failure time distribution, and repair time distribution, or repair policy. The systems failure probability (unavailability etc), can be evaluated by using the above minimal cut set representation:

$$P(S=1)=P(\phi(X_1,X_2,...,X_n)=1)=P\left(\sum_{i=1}^{M}\prod_{j\in C_1}X_j=1\right)\approx\sum_{i=1}^{M}P\left(\prod_{j\in C_i}X_j=1\right),$$

where the approximation is valid when the components' failure probabilities are small.

It is also possible to describe the stochastic behaviour of the system state by using finite state Markov chains.

As it can be seen from the above, the systems reliability structure can be described by a structure function, which can be represented in rather simple form (the minimal cutset representation), and the systems reliability properties can be determined by using suitable component (or sub-system) failure probability models. There is an extensive literature on reliability modelling, here a textbook by Rausand and Høyland [22] can be referred to.

The reliability modelling of IP-network is, however, more demanding, due to features of the system. This will be discussed in the next section.

5.2.2 Challenges of IP-network reliability analysis

Certain features of IP-networks make it difficult to be modelled by using the conventional reliability analysis methods. Basically, the conventional approach discussed in previous section is, at least to some extend, applicable to systems having a network structure.

5.2.2.1 Structure of the IP network

The basic difficulty in the IP-network reliability analysis is that e.g. fault trees are intended for analysis of systems which have a tree structure. In the case of networks, tree-like models cannot be directly applied.

However, it is possible to develop algorithms for finding the minimal cutsets for example for the connection between any two nodes in the network, and the reliability properties of such connection can be determined as discussed in previous section. If the network becomes large, the search for minimal cutsets gets computationally demanding: it is known that the general problem of finding the minimal cutsets is NP-hard. The minimal cutset search can, however, be made rather efficient, if the structural properties of the network are taken into account. Some examples of minimal cutset search algorithms and their computational requirements are given by Tsukiyama et al [29], Suh & Chang [28], and Sharafat & Ma'rouzi [27]. It is possible to conclude that if the networks structure is given, it is possible to find minimal cutsets, and determine the quantitative reliability properties as in the basic case discussed in previous section.

5.2.2.2 IP-specific dynamic features

Another difficulty in the IP-network reliability analysis is the layered structure of the network. There are several logical layers in the network, and the routing of packets through the network is controlled dynamically by various protocols operating in different logical layers. Thus, the logical

(or virtual) network structured differs from the physical structure. If the correct operation of protocols is assumed, the reliability analysis reduces to the analysis of physical network discussed above. However, if the operation of protocols is erroneous, it is possible that the nodes of the network become dependent, and their operation may be lost simultaneously. Similarly, the overload to network may have similar type of impact on the system. There may be some possibilities to take this into account in the reliability modelling. First, the possible failure modes related to the operation of protocols (and other dynamic features) are identified. Then the impact of these failures on the system (i.e. the network structure) is identified. Finally, the reliability of the network (impacted by the failures) is determined conditionally on the occurrence of these failures. This approach reminds the Probabilistic Safety Assessment (PSA), which is the basic tool for evaluating nuclear power plant safety (McCormick [20]). The approach has been applied for reliability analysis of the Finnish power transmission network (Haarla et al [13]) In PSA, initial disturbances (corresponding to the above mentioned failures) or initiating events and the response of the system to initiating events are identified. Depending on the success or failures of the protective actions, the course of the disturbance divides to different disturbance sequences, and finally to different end states. The minimal cutsets corresponding to each end state can be determined, and if the frequencies of the initiating events and the failure probabilities of protective actions are known, it is possible to determine the probabilities of each end state.

The applicability of PSA-type modelling has not been tested in the case of IP-network reliability analysis. The amount and nature of different initiating events may be large, and their impact on the system is not easily identified. The IP-network has not been designed in the “defence-in-depth”-way, and the barriers against various disturbances are not defined in the way which is required in PSA-type modelling.

5.2.2.3 Role of software

The operation of IP-networks is based on software. Almost all components in the network include software. The possibilities of software errors should be taken into account in reliability models. From the reliability point of view, software differs from hardware in many ways. For example, software doesn’t age in the way a hardware component does. Some salient features of software failures are the following:

- software faults are caused by hidden design flaws rather than wear-and-tear or physical failure.

- each software is, at least to an extent, unique. Even minor differences in the program code might mean large differences in the behaviour of the software. Therefore experience with the reliability of other software is of very limited use at best.

- software faults manifest themselves only under particular conditions.

- in practice it has been observed that the mean time to failure (measured in number of runs) of a software system in large systems is inversely proportional to program size; this would indicate that the number of faults per line is roughly constant.

- when used in redundant components, similar software creates dependence between the failure behaviour of components, and the impact of software erro0r on the systems structure is not always easily identified

There are methods and tools for quantitative software reliability analysis (see [18]). The methods are focused on the determining the probability distribution of the number of residual errors or faults in the software, or evaluating the failure rate of the software. The predictions of software reliability models cannot always be applied in systems reliability models as such. However, they can be used as useful background information for expert judgement.

Software errors are one cause of components' functional failure. As such they increase the components' failure probability, and they can be taken into account in sensitivity analyses included in quantitative reliability assessment. In the cases, where software errors have impact over several components or subsystems, the quantitative assessment becomes difficult.

5.2.2.4 Human errors

Human errors in operation, installation and maintenance of system are often one often most important reasons for systems unreliability. Human errors made in maintenance and installation of systems are often left hidden into the systems, and they manifest themselves only in certain occasions (like software errors). Further, they may introduce dependencies between systems and components, and thus cause common cause failures, which make several redundant systems unavailable simultaneously.

From the quantitative reliability analysis point of view human errors may only one random cause of failure, which only increase the failure rate of the system, or cause systematic failure, which have impact on the systems reliability structure. The latter type of human errors are the most difficult ones to be modelled.

5.2.2.5 Reliability of services

The Internet is used to obtain various services, such as e-mail, banking services, etc. The number of different services is growing, and the nature of services is becoming more and more complex. In order to get a service, several functionalities of the network are needed.

One way to look at the reliability of the network is to evaluate the availability of services, or sets of services. The reliability of a service could then be measured as the long term part of time during which the service is available for a defined group of users. In order to get a more general view, this measure can be averaged over a set of services. In principle, the reliability of the network functionalities needed to provide a service for a user can be modelled using the methods described above and, consequently, the availability of a service can be quantified. However, it seems very difficult to analyse the simultaneous availability of a large set of services.

5.2.3 Prerequisites of quantitative reliability modelling of IP-networks

As discussed above, quantitative reliability modelling of certain parts and features of IP-networks seems possible. However, the models and their results are tightly connected to the level of abstraction behind the models. Certain features are included in the models by making suitable assumptions, which should be made explicit in order to understand the results and to make valid conclusions. The modelling should be transparent, and follow a good reliability modelling process.

The starting point of reliability modelling is the understanding of the systems functionality: what are the systems main functions, what are the components and subsystems which perform the systems functions, how the functions depend on each other and what are the support functions needed to perform the main functions. There are several ways to describe the systems functionality. Holistic meta-models applied in describing critical infrastructures are one possibility [5]. Another possibility is to apply models based on the IDEF (IDEF = Integrated Computed Aided Manufacturing DEFinition language) family of models [15], [16]. The above mentioned models are basically static; if dynamic functionalities need to be described the modelling techniques get more complex.

After the functional description of the system, qualitative reliability analysis methods must be applied in order to identify failure modes and possible disturbances which can prevent the intended functionalities of the system. This is a necessary step in reliability analysis, without it the structural reliability models cannot be made (see section 5.1). Qualitative reliability analyses are made both on system and component level. The component reliability models are also based on qualitative reliability analyses (e.g. FMEA).

The structural reliability model is built on the basis of functional description and qualitative reliability analysis. In this phase of modelling, the degree of detail and the abstraction level of the models are decided. The models applied can be fault-trees or network reliability models.

Figure 3 summarises the prerequisites of reliability modelling.

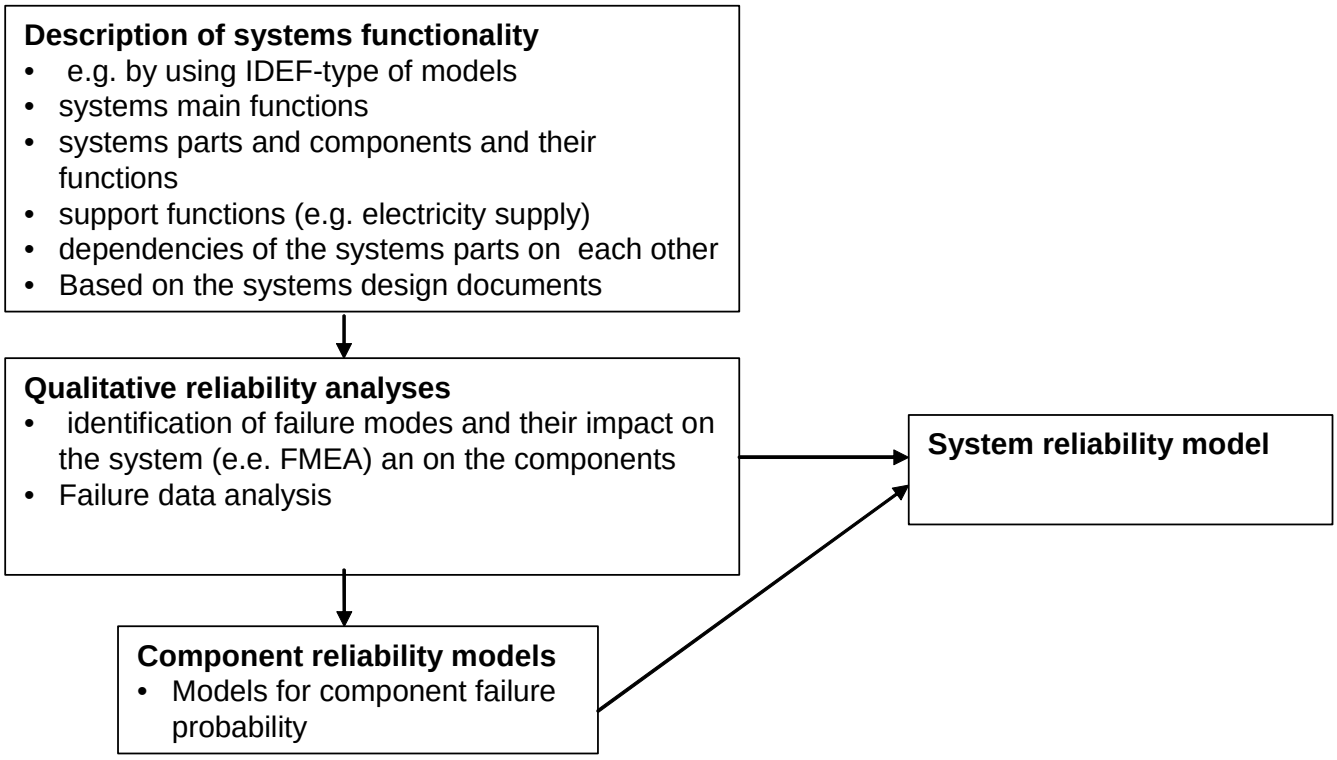


Figure 3. Reliability modelling.

5.2.4 Example

As an example of reliability modelling, a simple network is considered (Figure 4). The network consists of a set of terminal users and the reliability of their access to the core network is modelled. The access requires the failure free operation of set of links (denoted by L-··-u·· in Figure 4), DSLAMs (Ci), routers (Ri) and name severs (DNSi). The structure of the core network is not modelled, and it is assumed to be one unit.

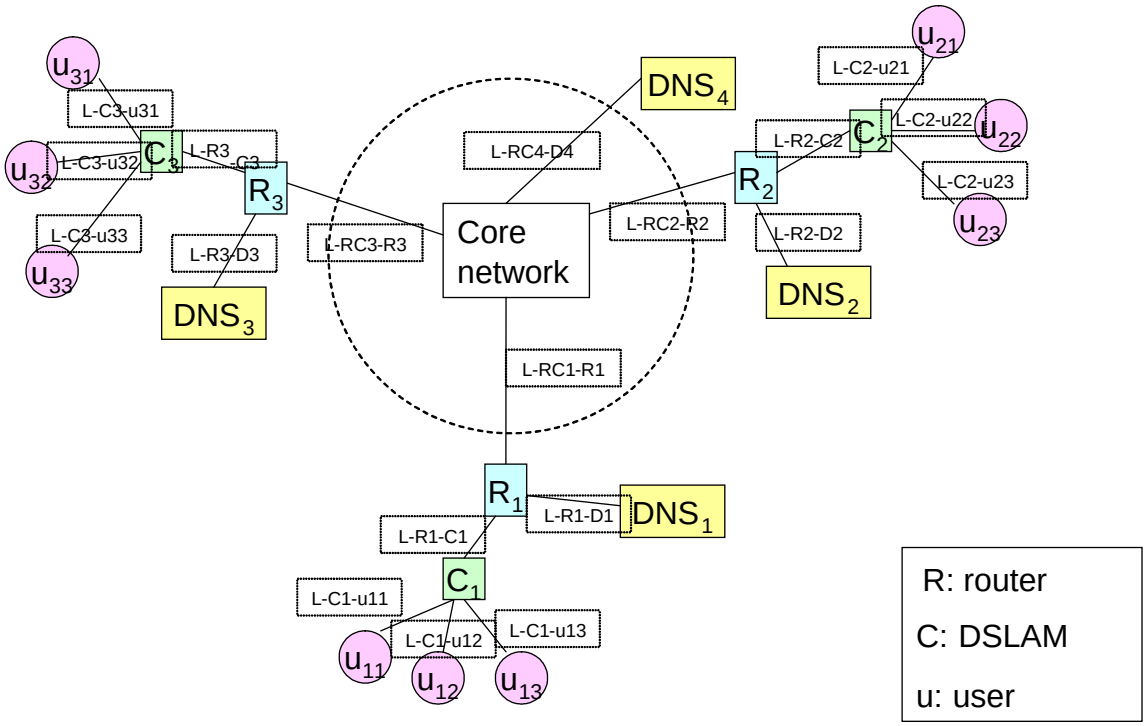


Figure 4. Simple example network.

The fault-tree model for the example network is developed for the TOP-event (i.e. the system failure) that none of the users u_{11}, u_{12},…,u_{33} is connected to the core network. The fault-tree is in Figure 5 - Figure 11.

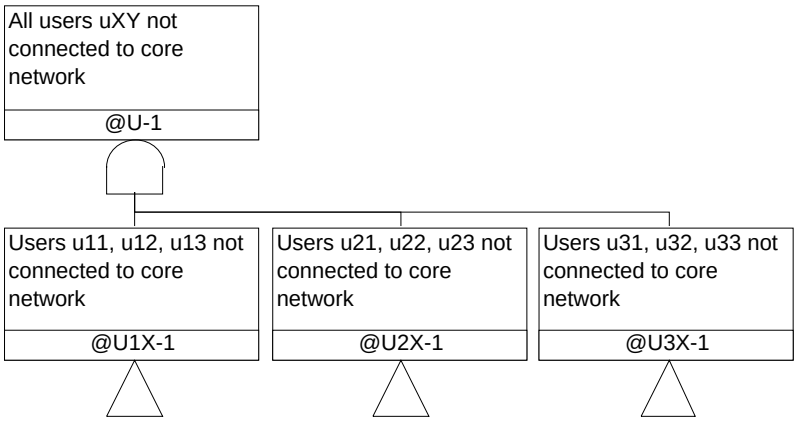


Figure 5. Example fault tree, part 1.

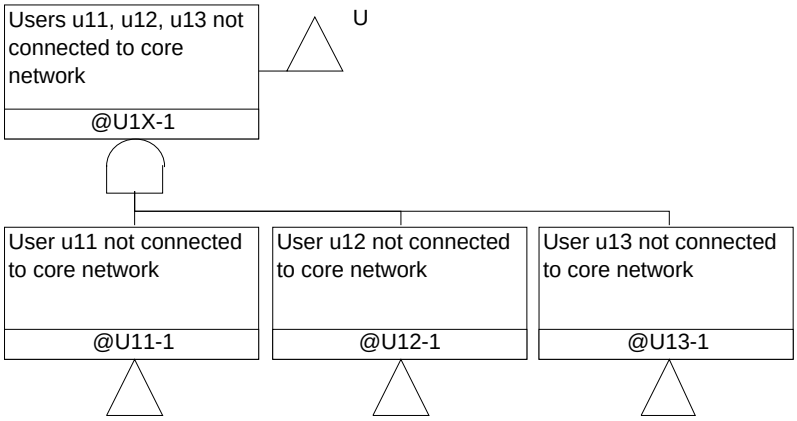


Figure 6. Example fault tree, part 2.

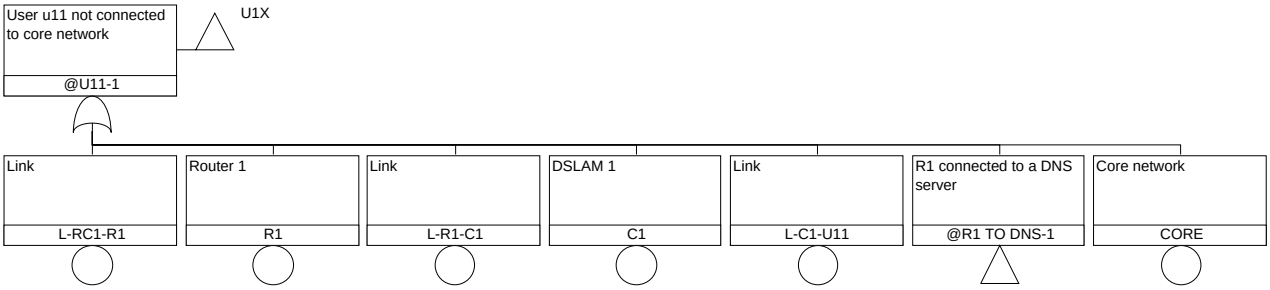


Figure 7. Example fault tree, part 3.

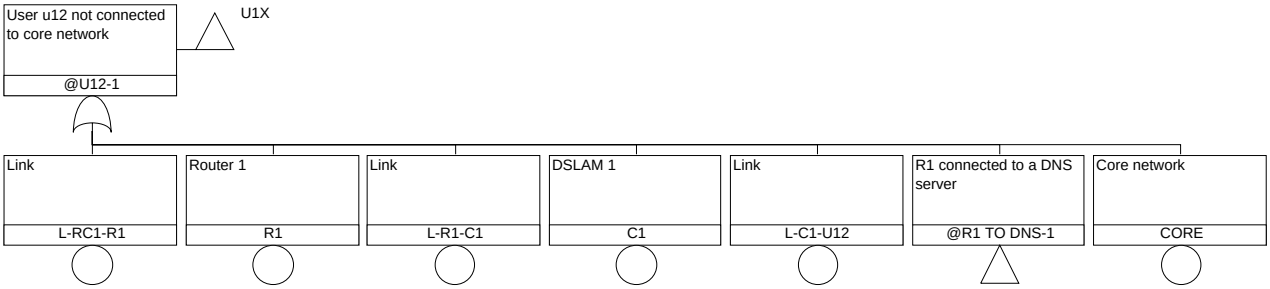


Figure 8. Example fault tree, part 4.

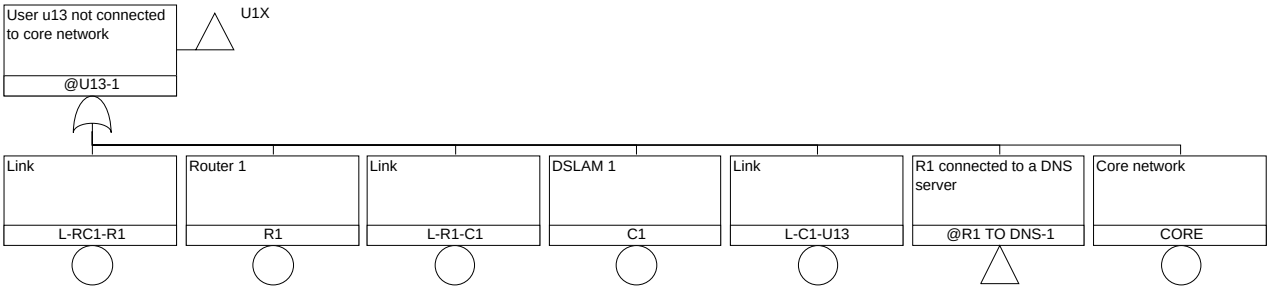


Figure 9. Example fault tree, part 5.

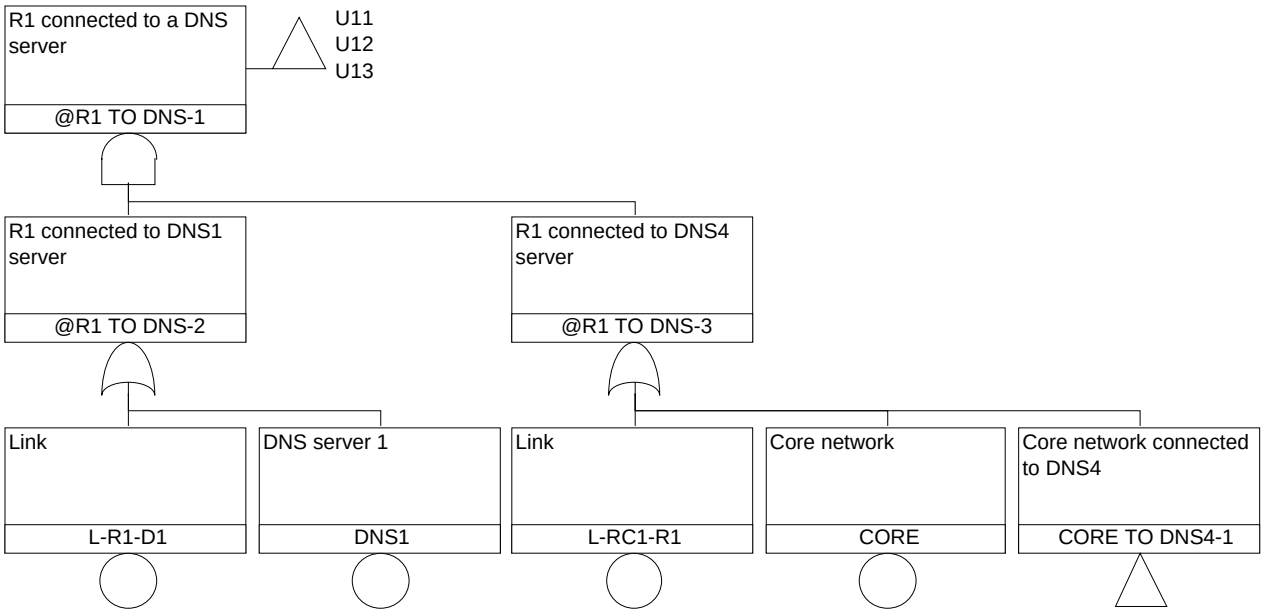


Figure 10. Example fault tree, part 6.

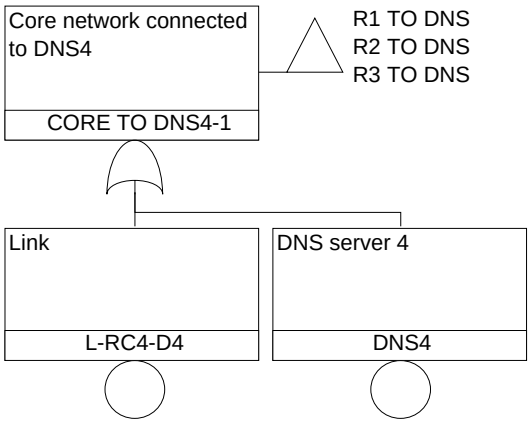


Figure 11. Example fault tree, part 7.

The fault tree model is used to determine the minimal cutsets of the TOP-event "None of the users u_{11}, u_{12},…,u_{33} is connected to the core network". The total number of minimal cutsets is 562. The 66 shortest minimal cutsets (of order 3) consist of the combinations failures of routers, d-slams and their connecting links. The number of minimal cutsets consisting of four failures is 303. The list of minimal cutsets reveals, for example, that the number of low order minimal cutsets (which represent the highest vulnerabilities) is rather high. This is due to long non-redundant sequences of components from the periphery to the core network. Some of the most important minimal cutsets are in Table 2.

Table 2. Minimal cutsets in the fault tree example.

No.	Event1	Event2	Event3
1	L-R3-C3	L-RC1-R1	L-RC2-R2
2	L-R1-C1	L-RC2-R2	L-RC3-R3
3	L-R1-C1	L-R2-C2	L-R3-C3
4	L-R1-C1	L-R3-C3	L-RC2-R2
5	L-R1-C1	L-R2-C2	L-RC3-R3
6	L-R2-C2	L-RC1-R1	L-RC3-R3
7	L-R2-C2	L-R3-C3	L-RC1-R1
8	L-RC1-R1	L-RC2-R2	L-RC3-R3
9	C1	L-R3-C3	L-RC2-R2
10	C2	L-RC1-R1	L-RC3-R3
11	C3	L-RC1-R1	L-RC2-R2
12	C1	L-RC2-R2	L-RC3-R3
13	C3	L-R2-C2	L-RC1-R1
14	C2	L-R3-C3	L-RC1-R1
15	C1	L-R2-C2	L-R3-C3
16	C1	L-R2-C2	L-RC3-R3
17	C3	L-R1-C1	L-R2-C2
18	C2	L-R1-C1	L-R3-C3
19	C3	L-R1-C1	L-RC2-R2
20	C2	L-R1-C1	L-RC3-R3
21	C1	C2	L-R3-C3
22	C1	C2	L-RC3-R3
23	C1	C3	L-R2-C2

6 Dependability case

6.1 Background and general principles

IP-networks are complex systems, the dependability of which cannot be described with a single numerical index. This is due to both the complexity of the network system and the concept of dependability. As discussed earlier, dependability has several aspects (e.g. availability, safety, maintainability, reliability, robustness, invulnerability and controllability), which should be taken into account when measuring or characterising the systems dependability properties. Furthermore, different requirements concerning the above mentioned aspects are set upon the IP-network system, Thus, it seems that instead of set of indices, the dependability of a IP-network system should be characterised by systemic and holistic view upon the system.

In the safety critical industry (nuclear power plants, aerospace), the need to characterise the safety of the system has been solved by introducing a goal oriented approach, which is called Safety Case (see e.g. Bishop and Bloomfield [6]). The idea of Safety Case approach is that it identifies the goals or requirements se on the system, and collects evidence for the demonstration of the goals. Bishop and Bloomfield [6] define the concept of safety case as:

A documented body of evidence that provides a convincing and valid argument that a system is adequately safe for a given application in a given environment.

The safety case approach has been applied e.g. in safety assessment of smart devices (see Pulkkinen and Bloomfield [21], Bishop et al [7]).

For the case of IP-network dependability, we modify the concept of safety case and introduce the concept of dependability case. Basically, it does not differ much from safety case, despite its focus on dependability feature. The definition of dependability case is:

A documented body of evidence that provides a convincing and valid argument that a system is adequately dependable for a given application in a given environment.

The implementation of a dependability case requires an explicit set of claims about the dependability properties of the system, the evidence supporting the claim, set of dependability arguments that link the claims to the evidence and clear assumptions and judgements underlying the dependability arguments. In addition to these, different viewpoints and levels of detail on the system are needed.

6.2 Elements of a dependability case

The basic elements of a dependability case with their definitions are collected in Table 3. The claims are based on the dependability requirements or goals set on the system. They originate from the regulatory, user, design, or operator requirements. The evidence is any set of information on the

system, to be used as a support for the truth of the claim. The arguments and inference provide a link between evidence and claim, indicating for example the strength of evidence.

Table 3. Elements of a dependability case.

Element of dependability case	Definition
Claim	a statement about a property of the system or subsystem
Evidence	A basis of safety argument; facts (e.g. based on scientific principles and research), assumptions, subclaims derived from lower level sub-arguments
Argument	Link between evidence and claim, can be deterministic, probabilistic or qualitative
Inference	The mechanism that provides the transformational rules for the argument

The general structure of a dependability case is given in Figure 12.

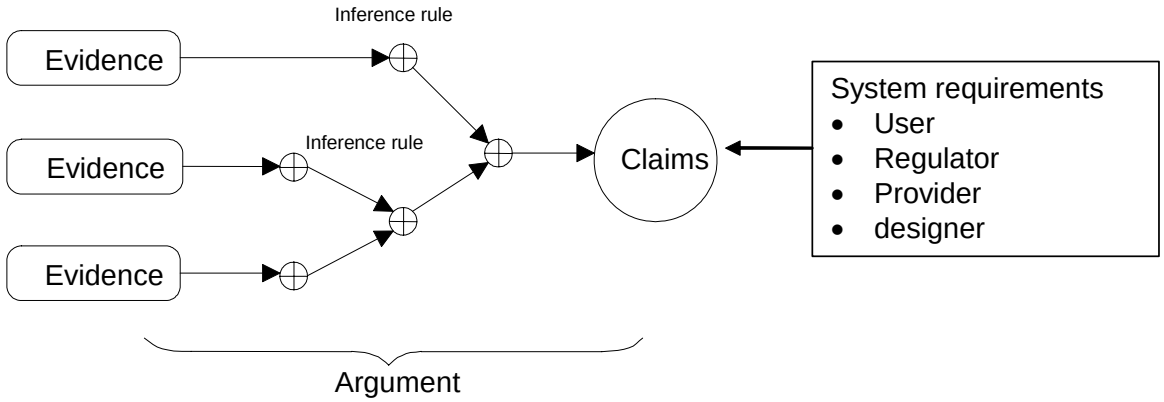


Figure 12. Structure of a dependability case.

The type of claims depends on the system under analysis. For IP-network, typical claims could concern:

- reliability and availability (qualitative, quantitative)
- length of unavailability time
- usability
- functional correctness
- time response
- maintainability
- fail-safety
- accuracy
- robustness to overload

- modifiability
- structural integrity
- economical trustworthiness of provider
- etc.

It should be noted that the claim structure depends also on the goals of the dependability analysis: e.g. the analysis may focus on certain type of failures and the set claims corresponds to the requirements on the co0nsidered failure types.

The arguments may be deterministic, probabilistic or qualitative. The deterministic arguments involve application of predetermined rules to derive a true/false claim; e.g. a formal proof of compliance to a specification or demonstration of dependability requirement

The probabilistic arguments involve quantitative statistical reasoning to establish a numerical level of some statistical (dependability) property of the system. Examples of probabilistic arguments are the determination of statistical estimates for failure rate, mean repair time and availability of the system.

Qualitative argument can be interpreted as compliance with rules that have an indirect link to the desired attributes of the system. Examples on these are the compliance with standards, evaluation of the staff skills and experience, as well as analysis of the economical status of the the company providing critical services.

An example of arguments is given in Table 4.

Table 4. Argument in a dependability case.

Attribute	**Design etc. features**	**Evidence/ assumptions**	**Subsystem requirements**	**Claim**
Fail-safety	Use of functional diversity Fail-safe architectures	System hazard analysis Fault tree analysis	Fail safety requirements for subsystems; response to failure conditions	Claim that dependability is maintained under stated conditions, assuming the subsystems are correctly implemented
Reliability/availability	Architecture levels of redundancy, segregation Fault tolerant architectures Design simplicity	Reliability of components, common cause failure assumptions Failure rate, diagnostic coverage, test and maintenance intervals, repair time Prior field reliability in similar applications Operational experience	Hardware component reliability Software integrity level Component segregation requirements Fault detection and diagnostic requirements Maintenance requirements	Reliability claim based on modelling and CCF assumptions, together with fault detection and repair assumptions Reliability claim based on experience with similar systems
Response time	Design ensures overall response time is bounded	Assumes subsystem time budgets can be met	Time budgets for hardware interfaces and software	Claim that overall system design can meet time response
Functional correctness	Partitioning of system according to criticality Design simplicity	Assumption that segregated functions cannot affect each other	Subsystem integrity level Functional segregation requirements	Claim that response behaviour of the critical functions implements the overall function
Maintenance quality	Design simplicity Good procedures for corrective and preventive maintenance The maintenance agreements clearly stated	Observed number of maintenance errors Maintenance effects analysis (MEA)		Claim that the maintenance of the system can sufficiently ensure the operability of the system
Configuration correctness	Good procedures configuration Experienced staff	FMEA on configuration errors Observed number of configuration errors		Claim that configuration errors are not too frequent
Trustworthiness of service provider	The system is divided into several criticality levels	Economical record has been good Operating experience from similar applications		Claim that the service provider can keep up the critical operations

6.3 Implementation of dependability case for IP-networks

The implementation of a dependability case starts with the analysis of requirements set upon the system. In this phase, the requirements from various sources (regulator, user, provider, designer) are collected. The requirements may cover several attributes of the system properties (see the list in previous section). It is important that all relevant types of requirements are covered.

The second phase of dependability case implementation is the establishment of the claim structure: the claims are classified according to the aspect of dependability, and they are divided into sub-claims if necessary. The top level claims may concern the system properties, which are then divided into sub-claims on sub-system, or on component level.

Next, the evidence is collected. The evidence include e.g. information the structural properties and design features of the system, analyses of the operating experience, results available quantitative and qualitative reliability analyses, analysis of maintenance procedures and agreements, vulnerability assessments, system hazards analyses etc. Basically any facts that are related to dependability properties of the system can be interpreted as evidence. In addition to this, expert judgements about the systems dependability features may be taken into account. However, it is important that the evidence is transparent.

Finally, the argumentation and inference linking the evidence with the claims is made. The argumentation should be transparent and based on clearly stated assumptions. The argumentation evaluates the pieces of evidence and provides justification on whether the evidence supports the claims. Furthermore, the argumentation identifies alternative pieces of evidence, and provides an evaluation of the strength of the evidence.

The dependability case may consist of a large set of claims and sometimes on rather complex argumentation structure. The management of large dependability cases may be difficult without computer based tools. Recently, such tools have been developed (see e.g. ASCE [3]).

6.4 An example of dependability case

In the following, a simple demonstration of a dependability case is given (Figure 13). The aim is to provide evidence on the sufficient availability of a connection of terminal user to the core network. The example network has been discussed in Section 5.2.4, where a fault-tree model of the availability of the connection to the core network is presented.

The requirement that availability of the connection is sufficient is first divided into two claims concerning the failure rate and the unavailability time of the connection. In Figure 13, the claim that the duration of unavailability of the connection is further analysed. In this simple case, the evidence is assumed to consist of operational experience, information on the architecture of the system and the result of a quantitative reliability analysis.

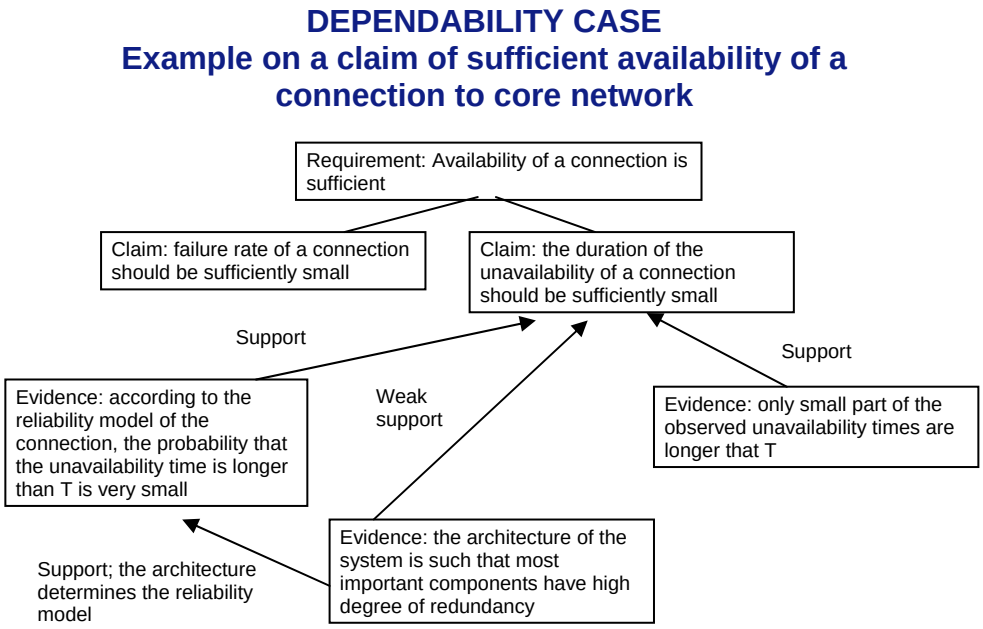


Figure 13. Example of a dependability case.

7 Conclusions and suggestions

Taking care of the dependability of the IP infrastructure is a common cause of all the actors: User, Provider, Regulator and Designer. Improving and maintaining a high level of dependability of the IP infrastructure is a continuing process that should find appropriate institutional forms and practices. The User's role should not be underestimated. General awareness about the dependability of IP networking should be raised so that the users would better understand to require adequate, differentiated reliability for different communication networks and services.

Availability monitoring and methods for collecting dependability-related data should be developed towards common standard practices. This kind of systematic data production is an indispensable part of the "dependability process". Summarizing reports on such data could in some phase be required by the Regulator. Qualitative operation experience should be utilized systematically (see Section 5.1.2).

In the design of future networks, dependability should be one of the built-in principles. "Defense-in-depth" thinking should be applied consequently. This facilitates also modelling and assessing networks from the dependability point of view. The current networks have several kinds of dependability problems. For example, long unprotected tree structures are common in network periphery, Ethernet networking is often applied uncautiously (see [26]), and BGP has its problems (see [2]).

Existing quantitative reliability modelling and analysis techniques are useful in many specific contexts, and they should be introduced into this field more extensively. Various meaningful indices can be computed to measure and characterize the availability and other aspects of network dependability. However, indices are not sufficient for creating and maintaining the whole picture of dependability situation. The Dependability Case method, proposed in this report, collects all kinds of indices together with other types of relevant arguments.

The access techniques to the IP infrastructure are diversifying and differ much in dependability also. The availability of different access networks increases the total dependability of the IP infrastructure from the user's point of view. In particular, alternative accesses (e.g. wired and wireless) improve the overall availability in network periphery.

The emergence of a "dependability process" as outlined above would help in the development of "best practices" in various contexts. Such practices also provide some protection against a deterioration of "dependability culture" by economic pressures. The operating personnel's high qualification work deserves similar studies as those common in safety critical work.

A criticality classification of telecommunication services should be created, including the counterpart of emergency calls in the IP era. Regulator's requirements should be formulated in a technology independent way, with the criticality classification as a starting point. Regulations could also be structured in the form of generic dependability cases.

References

[1] Reliability Division of the American Society for Quality, http://standardsgroup.asq.org/dependability/tc56/faq.html.

[2] K. Ahola, E. Myötyri, I. Norros, L. Norros, U. Pulkkinen, P. Raatikainen and T. Suihko. The dependability of an IP network – what is it? Baseline paper of the IPLU project. http://iplu.vtt.fi/iplu_baseline_2006.pdf.

[3] Assurance and Safety Case Environment (ASCE). http://www.adelard.co.uk/software/asce.

[4] A. Avizienis, J.-C. Laprie, B. Randell and C. Landwehr. Basic concepts and taxonomy of dependable and secure computing. IEEE Trans. Dependable and Secure Computing **1**(1), 11-33, 2004

[5] Beyer, U, Flentge, F. Towards Holistic Metamodel for Systems of Critical Infrastructures. ECN European CIIP Newsletter. Volume 2, Number 3. 2006.

[6] Bishop, P.G., Bloomfield, R.E. "A methodology for safety case development", Safety-Critical Systems Symposium (SSS'98), Birmingham, UK, Feb 1998.

[7] Bishop,P.G. Bloomfield, R E , Guerra, A S L , Tourlas, K. "Justification of smart sensors for nuclear applications". In Proceedings of the 5th International Conference on Control and Instrumentation in Nuclear Installations, 21-23 September 2004.

[8] Design of All-IP Mobile Networks (DAIMON). Final report of project. To appear within a forthcoming VTT publication.

[9] P. Checkland. Systems Thinking, Systems Practice. London, John Wiley & Sons, 1981.

[10] CI^{2}RCO project homepage: http://www.ci2rco.org

[11] J.C. Doyle, J. Carlson, S.H. Low, F. Paganini, G. Vinnicombe, W. Willinger, J. Hickey, P. Parrilo and L. Vandenberghe. Robustness and the Internet: Theoretical Foundations. Draft. March 2002. http://netlab.caltech.edu/internet.

[12] V. Flusser. Writings. University of Minnesota Press, 2002.

[13] Haarla, L, Pulkkinen, U, Koskinen, M, Jyrinsalo, J. "A method for analysing the reliability of a transmission network. Submitted for publication in Reliability Engineering and System Safety, 2006.

[14] IEC 60812. Analysis techniques for system reliability – Procedure for failure mode and effects analysis (FMEA). International Standard.

[15] IEEE Std 1320.1-1998. IEEE Standard for Functional Modeling Language—Syntax and Semantics for IDEF0. New York: IEEE, 1998.

[16] IEEE Std 1320.2-1998. IEEE Standard for Conceptual Modeling Language Syntax and Semantics for IDEF1X. New York: IEEE, 1998

[17] M. McLuhan. Understanding Media - The Extensions of Man. McGraw-Hill, New York, 1964.

[18] Musa, J D,. Iannino, A, Okumoto, K. Software reliability – measurement, prediction, application. McGraw-Hill, 1987.

[19] D. Pei, D. Massey and Lixia Zhang. A framework for resilient Internet routing protocols. IEEE Network **18**(2), 5-12.

[20] McCormick, N J. Reliability and risk analysis. Methods and nuclear power applications. Academic press, 1981.

[21] Pulkkinen, U, Bloomfield R.E. "Assessment of smart device software, ASDeS". SAFIR-seminar. To appear in 2007.

[22] Rausand, M, Høyland, A. "System reliability theory. Models, statistical methods and applications." Second edition. Wiley Interscience, 2004.

[23] Reittu, H. & Norros, I.: On the effect of very large nodes in Internet graphs. Globecom 2002, Taipei, Taiwan.

[24] Reittu, H.: On the robustness of power-law random graphs. ICMS Workshop "New Directions in Applied Probability: Stochastic Networks and Beyond". Edinburgh, July 2006. http://icms.org.uk/archive/meetings/2006/stochnet.

[25] J. Salzer, D. Reed, and D. D. Clark, "End-to-end arguments in system design," ACM Transactions on Computer Systems, 2, no. 4, Nov. 1984, pp. 277-288.

[26] Seppänen, K.: Resiliency in Ethernet Based Transport Networks. An IPLU project report, 2006. http://iplu.vtt.fi/ethernet-transport.pdf

[27] Sharafat, A R, Ma'rouzi, O R."A novel and efficient algorithm for scanning all minimal cutsets of a graph.". eprint arXiv:math/0211436, 2002.

[28] Suh. H, Chang, C K. "Algorithms for the minimal cutsets enumeration fo networks by graph search and branch addition. 25th Annual IEEE International Conference on Local Computer Networks (LCN'00). p100., 2000.

[29] Tsukiyama, S, Shirakawa, I, Ozaki, H, Ariyoshi, H. "An algorithm to enumerate all cutsets of a graph in linear time per cutset". J. Assoc. foir Comp. Mach. vol 27, 1980, pp.619-632

[30] A. Villemeur. Reliability, Availability, Maintainability and Safety. John Wiley & Sons, 1991. Two volumes.