

Dependability of complex open systems: A unifying concept for understanding Internet-related issues

Nicholas Kyriakopoulos*
Department of Electrical and Computer Engineering
The George Washington University
Washington, DC – USA
kyriak@seas.gwu.edu

Marc Wilikens
Institute for Systems, Informatics and Safety
Joint Research Center
Ispra (VA) – Italy
Marc.Wilikens@jrc.it

Keywords: complex open communications systems, interdependencies, multi-layer dependability requirements, quality of service, survivability, research collaboration

Abstract

Dependability of complex open systems, such as the global information infrastructure, implies end-to-end performance requirements. Survivability and security are viewed as attributes of dependability. In this paper, we present a methodology and a set of quantifiable attributes for specifying end-to-end quality of service, in terms of the quality of service of the constituent parts of the complex system. Areas of potential collaborative research are identified.

1. Motivation and Scope

The convergence of computer and communications technologies has created a global infrastructure for the transmission and processing of information. This new infrastructure is based on the ability of diverse networks to be interconnected and provide global coverage for the transmission of data. The defining characteristic of this global infrastructure is the absence of central monitoring and control in contrast to conventional computer communications systems. By its nature, the resulting system, Internet, has created a new environment for conducting human activities and has given rise to the term Information Society. It has also created the need for new concepts and tools for analysing its behavior and impact on the services based on it.

Reliance on the services provided by and through this new medium is predicated on its ability to deliver them in a manner inspiring confidence to the users of these services. The need for confidence is pervasive not only in a business context but also all social interactions. Confidence may be eroded either by the faulty delivery of desired and promised services, or by the manifestation of unexpected and/or undesirable side effects. For example, heavier than expected use of services available through the Internet could affect the quality of service of other unrelated applications sharing the same resources. Applications relying on the communications infrastructure for the transport of data may deliver faulty services either due to faults within the application or due to faults in the data transport service of the communications infrastructure. On the other hand, the causes of unexpected and/or undesirable side effects may be more complex. They may lie in an application, in the communications infrastructure or in some interdependencies between applications and the communications infrastructure. Because of these interdependencies, one needs to understand not only the behavior of the communications medium and of each application relying on it, but also the effects of the interaction between applications and the communications medium.

Confidence in the information infrastructure, is derived from its dependability, namely, its ability to deliver specified services to the user. Thus, to build confidence one needs to specify requirements and to

* The work presented in this paper was performed while the author was a Visiting Scientist at the Institute for Systems, Informatics and Safety of the Joint Research Center of the European Commission.

monitor the performance of a system in order to ensure that the requirements are satisfied in the presence of threats to the various components of the information infrastructure. For a complex open system such as the Internet, the development of quantitative measures is not a trivial task and presents major research challenges.

A global, or macro, perspective brings forth concepts such as dependability, survivability, fault-tolerance and security and follows the path of system decomposition. Conversely, a micro perspective views a complex system as the result of synthesis based on interconnections of elementary components. The synthesis approach imposes quantitative constraints on the interfaces and interactions among the components. The question then becomes how to relate the qualitative global requirements and the quantitative constraints. Either approach has limitations when applied to complex open systems. A good theoretical foundation for translating the qualitative criteria into quantitative ones through decomposition is not available. Interconnection of components allows for the generation of quantitative requirements at any level of complexity of a system. However, for large complex systems, the number of constraints and variables become too large and, therefore, difficult to manipulate.

This paper will present a conceptual framework for formulating the problem and a methodology for addressing issues arising from the complexity of the problem. Particular attention will be placed on relating survivability and security to dependability and quality of service. The work reported in the paper is indicative of potential transatlantic collaborative projects. It is performed in the frame of the Joint Research Centre's institutional support activities to the European Union Information Society Technologies (IST) programme in the area of dependability. A number of industrial sectors representing utilities, telecommunications and health care contributed to this work. Between December 1999 and March 2000, three meetings were held with these organisations for this purpose.

2. Conceptual Framework and Methodology

The term "information infrastructure" is commonly used to indicate collectively the communications infrastructure that provides data transport services and the various systems, hardware and software, that deliver services requiring and relying upon the transport of data. The main characteristics of this complex system that consists of applications relying on a global communications infrastructure are its global coverage, high connectivity, and data volume and speed. As a result, the "Internet" has become a high level abstraction that provides little, if any, help in analysing the performance of existing systems or designing new ones. Therefore, we start our analysis by defining a framework for decomposing the problem into simpler ones, some of which either have been solved, or are solvable by known methods. The first level decomposition creates two systems, one that provides data transport services and another that delivers services requiring the transport of data. Under such decomposition, distributed computing is an application. So is voice over IP, electronic mail, etc. For other applications, such as financial services, health care, monitoring and control of electric power systems to name a few, the communications infrastructure is one of many systems that are essential to the delivery of services by each of these applications. Each application imposes its own performance requirements for data transport on the communications infrastructure. The collection of these requirements becomes the input for determining the design parameters for the infrastructure.

A service offered to the user by an application relying on internetworking requires data transport services based on IP. Although each application has its own specific requirements, there are some quantifiable attributes that are common to all applications. These are: *availability* of the data transport service (end-to-end connectivity X% of the time averaged over Y time units); *integrity* of the data during transport (X% of application data frames transmitted over Y time units satisfying a user-to-user validity check); *quantity* of data to be transported per unit time (X amount of data transferred between users) and *timeliness* of

transfer from source to sink ($X\%$ of application data frames arriving at the destination user within Y time units from the moment each time unit is transmitted by the source user). To the extent that the performance requirements of a given service depend on the data transport service of the Internet, these four attributes are both necessary and sufficient for specifying the availability, integrity and timeliness of the service. Thus, the dependability attributes of the application, namely, availability, integrity and timeliness can easily be mapped into the dependability attributes of the Internet. These attributes can be used as variables either for analysing existing systems or for designing new ones, because the attributes are quantifiable and can be translated into the *quality of service* attributes of communications systems. From the perspective of the designer of an application, the dependability requirements for data transport is an output derived from the dependability requirements of the application. In turn, this output becomes input for specifying the dependability requirements of the communications infrastructure. The open question for the Internet is whether the dependability requirements imposed on it by a given application are feasible and at what cost.

It has become convenient to refer to the environment created by internetworking as the “cloud”. Although this designation provides an easy visualization of what happens to the data of a user after they enter the server of the Internet service provider (ISP), it is not very helpful for the evaluation of the dependability of the data transport service. Each ISP has control over and is responsible for the operation of a communications network with known characteristics including topology and bandwidth. At the IP level, these networks are connected by internetwork routers, or gateways, under the control of peer ISPs. Thus the topology of the internetwork is known, the gateways being the nodes and the links those connecting peer ISPs. The interconnection between any two peer ISPs is through a relatively small number of gateways. The performance requirements at the IP interface could also be specified in service level agreements between any two peer ISPs, although there are still outstanding issues for implementing such agreements. Thus, a mechanism exists for specifying quality of service requirements among peer ISPs. In view of the preceding discussion about the nature of the IP, the outstanding question is whether dependability requirements for data transport can be specified and met at the interconnections among peer ISPs. If such a goal can be reached, then the proposed methodology allows for the quantitative specification of dependability of the information infrastructure.

3. Security, survivability and dependability

The methodology presented in this paper provides a systematic mechanism for addressing the issues of security and survivability in the context of dependability. The dependability of application services can be compromised by threats to the application subsystem, the communications infrastructure, or both.

In order to understand the nature of the vulnerabilities caused by various types of threats and to assure the dependability of the applications we will follow the established methodology and decompose the problem into a series of simpler problems. For any application, two categories of threats can be identified:

- Threats to the application
- Threats to the data transport service.

The threats to the application could be either related to the transport of data through the communications infrastructure, or caused by other factors. For the purposes of this paper we will limit the discussion to the threats through the data transport service. The effect of a threat to the application through the data transport service can be quantified in terms of the four attributes. Incorrect data or commands, whether generated maliciously or innocuously, do not satisfy the attribute of integrity and can be detected by appropriate integrity checks at the application. Similarly, a threat to an application can be realized through delayed or unavailable data and can be detected at the application by monitoring the received data for timeliness and availability. The dependability requirements for the data could be satisfied at the

application level by providing end-to-end protection such as encryption and redundancy in the data transport paths.

To investigate the threats to the data transport service it becomes easier to divide the problem into threats at the communications network level and threats at the internetworking level. As in the case of applications, the threats can be assessed in terms of their effects on the dependability attributes of availability, integrity, quantity and timeliness. The end-to-end virtual link could become unavailable due to faults that may manifest themselves within one or more communications networks or at the internetworking level. The integrity and timeliness of the data could be compromised within a given network or at one or more internetwork gateways. Faults can appear in either hardware or software. A noisy communications channel, loss of power in a satellite, cutting of a trunk line, malfunctioning computer program, crashing of an operating system, are examples of faults that degrade the performance of the data transport service with respect to the specified dependability attributes.

The partition of threats into the two categories (communications network and internetwork levels) leads to the development of two possible approaches for satisfying the dependability requirements of the applications. One seeks to satisfy the dependability criteria of the data transport service offered by the communications infrastructure as required by the applications. Another seeks to satisfy the dependability criteria of the application in the absence of dependable data transport services. The selection of either approach or some combination of the two becomes an interesting problem in cost-benefit analysis.

Finally, to provide security or/and implement a fault-tolerant design, some knowledge of the nature of the threats is necessary. From the perspective of survivability and security the decomposition of the threats into benign and malicious seems natural. For analysis, however, models for each type of threat are necessary. In order to develop quantitative measures of security and fault-tolerance for a given application and data transport service, the primary issues that need to be addressed are:

- How to identify the external threats for a given application
- How to develop models describing the statistical properties of these threats.

4. Conclusions and Recommendations

This paper has presented a concept and methodology for addressing the problems of security and survivability of the information infrastructure in a systematic and quantitative manner. By specifying a minimum set of necessary attributes for dependability, it has laid the foundation of a unified approach to the design of secure and survivable information systems. It also identifies some relevant topics that might be addressed by collaborative R&D:

1. Relations between qualitative global requirements and quantitative constraints on interfaces and interactions between components.
2. Evaluation of the viability and cost effectiveness of dependability requirements imposed on Internet
3. Specification of dependability requirements for data transport at interconnections between ISP's
4. Feasibility and implications (technical, operational, regulatory, legal) of monitoring performance at the IP level, namely, among peer ISPs
5. Approaches for threat identification and classification
6. Concepts and analytical tools for assessing the relative merits of fault-tolerance, threat mitigation and reduction, and protection from threats.