All quiet on the Internet front?

Doerr, C; Kuipers, FA

DOI
10.1109/MCOM.2014.6917401

Publication date
2014

Document Version
Accepted author manuscript

Published in
IEEE Communications Magazine

All Quiet on the Internet Front?

Christian Doerr and Fernando A. Kuipers, *Senior Member, IEEE*

***Abstract*—With the proliferation and increasing dependency of many services and applications on the Internet, this network has become a vital societal asset. This creates the need to protect this critical infrastructure, and over the past years a variety of resilience schemes have been proposed. The effectiveness of protection schemes, however, highly depends on the causes and circumstances of Internet failures, but a detailed comprehensive study of this is not yet available to date. This paper provides a high-level summary of an evaluation of Internet failures over the past 6 years, and presents a number of recommendations for future network resilience research.**

***Index Terms*—Internet failures, Network resilience, Multi-layer network, SRLG failures, Mitigation schemes.**

I. Introduction

The Internet constitutes a vital societal network-of-networks infrastructure in which even small "hick-ups" could have detrimental consequences, resulting in significant economic damages to institutions and entire economies.

The importance of the Internet and its services to society make it evident that it should be (made) resilient to failures. This awareness has instigated a large body of research on how to protect networks, although they typically consider a single (simplistic) failure model in which the network is represented by a graph consisting of nodes and links. In order to protect the Internet against failures, we believe that it is essential to understand what kind of failures exist, the impact they have, and the frequency at which they occur; in other words a taxonomy of Internet failures. Even though large-scale Internet incidents have been reported in the media, and some papers include a brief list of several such failures, a taxonomy of key Internet failures showing the cause, duration, range and effect does not yet exist.

In this paper, we will present such an overview and discuss the resulting implications for effective challenge mitigation. Our findings indicate that even failure scenarios for which mitigation strategies exist still pose a major source of outages, indicating that more fine-grained network risk assessment methods and better resilience planning and responses are still needed.

The remainder of this article is organized as follows. Section II offers an overview of our findings on Internet failures and presents our major conclusions. Based on this, Section III discusses the effectiveness of current mitigation strategies and gives recommendations to better avoid Internet failures. We conclude in Section IV.

C. Doerr and F. A. Kuipers are with the faculty of Electrical Engineering, Mathematics and Computer Science at Delft University of Technology, the Netherlands (e-mails: {C.Doerr, F.A.Kuipers}@tudelft.nl).

II. A Timeline of Internet Failures

To arrive at a comprehensive overview of Internet failures, a broad foundation is needed. For the work presented in this article, a variety of sources were consulted. We started by interviewing practitioners and representatives from regional Internet Service Providers (ISP), national research and education network operators (NRENs), national incumbent operators, and multi-national networks about their experiences, incidents and their root causes. Our findings and recommendations [4] were validated in a formative workshop hosted by the European Network Information and Security Agency (ENISA). Subsequently, we augmented our overview with operator reports and literary searches in academic and trade articles, as well as news websites, blogs, forums, and operator mailing lists about Internet incidents. In the following, we will limit our discussion to "Internet" services as commonly referred to by the end user, and will not extend the discussion into IP-based enterprise networks.

From our list of Internet incidents, 54 major and representative Internet failures over the period of June 2007 – December 2013 were chosen, which are displayed in Figure 1. The figure visualizes the time, duration, impact size, and ultimate root cause of each event, denoted by a circle where the size of the circle's area proportionately indicates the approximate number of affected customers and the color the incident duration on a log scale. The markers are centered at the time and ultimate root cause, i.e., if a service failed because of a database replication issue that was due to a defective core router, the event will be marked as a networking issue. In case no accurate number of the affected customer base was available and no meaningful estimate could be derived from operator reports or the literature, the figure only marks the time, root cause, and duration by a square. For details on these and other incidents beyond the space constraints of this paper, we refer the reader to www.internetview.org.

There are a variety of ways to structure the most prevalent types of Internet failures. A first crude classification one could make is into intentional failures, i.e., attacks, and unintentional failures. However, by analyzing the listed incidents and their causes, it becomes apparent that most Internet failures were unintentional and only a few of the incidents were the result of malicious attacks. We, therefore, adopted a slightly different categorization into infrastructure failures, Border Gateway Protocol (BGP)-related failures, and service failures resulting from an attack. Each category is further subdivided as follows:

1. **Infrastructure failures** list all instances where a component necessary to provide a particular service has failed, either directly as part of the operator's service development or outside of the operator's scope but still indirectly having an impact on the assets of the operator. Common failure types comprise network and cable failures, power failures, hardware failures (such as server failures, issues with storage systems, cooling facilities, structural failures, etc.), failures in the service architecture or failures in software components necessary to provide a particular Internet service (ranging from server-side end-user applications to database applications). In addition, we also list service impairments in this category that specifically stem from an accident or natural disaster, such as a hurricane or a fire in a datacenter.
2. The Internet is a network of networks, where each network (called autonomous system) possesses its own range of IP addresses and operates its own routing protocol. The **Border Gateway Protocol (BGP)** facilitates the routing between autonomous systems; it is the necessary "glue" to hold the tens of thousands of networks together into a commonly accessible Internet. Despite this key importance, the BGP is surprisingly susceptible to malfunctions, Internet service impairments and service failures due to the BGP are listed in this category. Most common are the BGP hijacking events, where a network announces some IP address space that it actually does not own. As a result, traffic towards a particular network that is the actual user of that IP prefix is temporarily misdirected. Other previous incidents related to the BGP were hardware- and protocol-based, e.g., unusual but valid BGP messages let key routers in the Internet crash due to software bugs, thereby also effectively cutting off networks from the overall Internet.
3. Finally, **service-related failures** list those Internet service incidents stemming either from failures in some underlying enabling service or direct attacks upon the service itself. The category assembles all incidents on the DNS, which is necessary to translate URLs to their corresponding IP addresses (and without which websites become practically invisible to the end user), as well as impairments and outages of the Secure Socket Layer (SSL) infrastructure that enables encryption between a service and the end user. This category also lists Distributed Denial of Service (DDoS) attacks. These are malicious attacks executed from hundreds or thousands of hijacked computers simultaneously, with the intent to overload a system so that its real end users are denied service. In the classification "Miscellaneous," we collect various events aimed to interrupt a particular service, such as insider attacks, hacks, etc. For an overview of attack types in the Internet and their economic incentives, we refer to Kim et al. [6].

Figure 1: A timeline of Internet failures between June 2007 and December 2013.

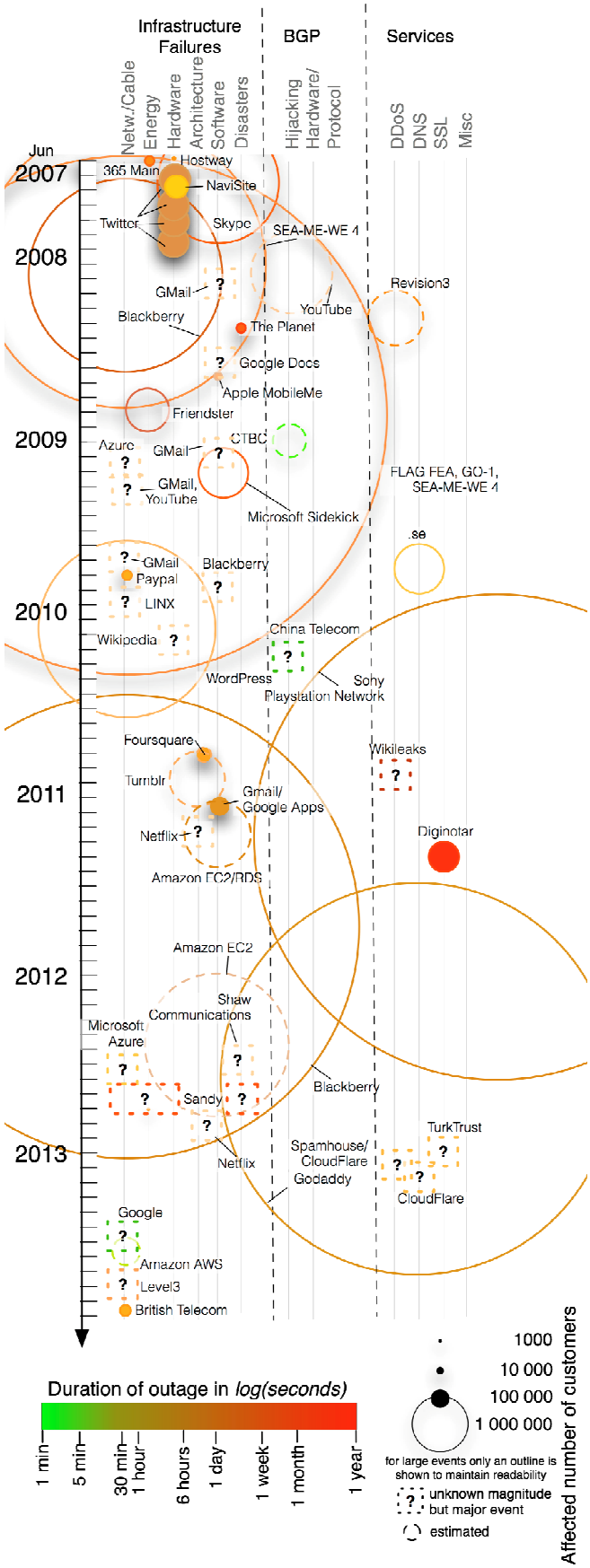

Figure 1 demonstrates that large-scale Internet service failures are occurring with regularity – at least and usually more than once a year – even when the plethora of usually unnoticeable smaller incidents and those events related to national security are not considered. It also becomes evident that the vast majority of events visible to the end user revolve mostly around the failure of the infrastructure and enabling services. This is noteworthy as problems around the notoriously vulnerable BGP protocol (for an excellent survey, see [2]) capture much attention, as it is theoretically possible to generate a large impact on the global interconnection system with comparatively little complexity. While those incidents in practice do occur, their frequency and impact is, however, usually bounded, thanks to the established monitoring infrastructures such as the BGPmon.

Based on the analysis of the incidents and their root causes, we also arrive at several other surprising conclusions: Much of the recent research work on network resilience has focused on the development of algorithmic link/path protection schemes that try to place backup routers and fiber optic cables in such a way in the network that most end-to-end connections are protected while minimizing cost. In the review of Internet failures however, almost no major incidents were identified that were ultimately caused by fiber cuts *and* that could have been prevented by such protection schemes. Major events such as the cuts of the "South East Asia – Middle East – Western Europe" (SEA-ME-WE), the "Fiber-Optic Link Around the Globe" (FLAG FEA), and GO-1 submarine cables in late 2008 in the Mediterranean Sea, or prior events such as the 2006 Taiwan earthquake (during which 8 submarine cables were cut) are usually not in the scope of such protection schemes that typically only plan for a limited number of simultaneous failures. On the other hand, these resilience methods do seem effective against small-scale localized events that according to the conducted ISP interviews probably are not directly visible due to their magnitude, successful mitigation, and routine status.

Network infrastructure failures, however, do not only involve issues such as cable cuts, but also failures of core routers and switches that we found to be a surprisingly common root cause of major outages, especially as it is a common good practice in the ISP community [4] to deploy critical core components at least redundantly or even with entire pools of hot spares. Nevertheless, there were multiple instances where a faulty networking element resulted in a failure of some higher-layer software component, such as a database breakdown that ultimately caused an entire service to fail.

Part of this issue is due to an increasing complexity of Internet services and a tendency to build services by federating lower-level building blocks. While cost effective, this, however, results from an availability standpoint in a tightly coupled system, and with the introduced co-dependency on multiple systems, the frequency and impact of breakdowns increase. This is, on the one hand, true for intra-organizational services that all rely on a common core component so that in case of a failure a variety of services are impaired, e.g., simultaneous failures of Google Apps, Gmail etc. in early 2011. This, on the other hand, is also the case for inter-organizational services and infrastructures, where service from one organization critically depends upon the availability of another one. How services depend on each other, as well as the strength and amount of co-dependencies is less and less known the higher one goes up the stack, so that multiple competing and apparently redundant services in the end are actually relying on the same infrastructure. With the advent of cloud providers, this issue seems to have amplified, as it repeatedly became visible over the past years. A failure in the Amazon Web Services (AWS) cloud infrastructure for example will render dozens of very diverse services unusable at the same time. This issue was on one occasion illustrated in an exemplary manner when several commercial uptime monitoring providers that track and alert website and service providers about an outage all failed simultaneously, as they all procured an underlying, but critical piece of their monitoring solution from the same cloud provider. In these cases, the common good practice to geographically distribute resources frequently does not seem to save the day, as the relatively less impactful connectivity and energy failures are traded against the apparently more frequent failures in the system architecture and software stack. Especially when such diversification is done via the same providers and components, not much is gained. For instance, if an application is hosted in different data centers by the same cloud provider, a service might be more vulnerable as it relies on a centralized system and now has an architectural single point of failure (SPoF).

As can also be seen in figure 1, the actual impact of many Internet failures is not known at all, predominantly because no global measurement and monitoring infrastructure exists as for example in the case of BGP, where monitors distributed worldwide record changes in the global routing table and allow an estimation of which networks are affected by the BGP prefix hijacking and routing issues. While some monitoring providers exist that test the uptime of Internet services, we believe that their deployment sizes (of a few hundred nodes in data centers) are not sufficient to get a good real-time view of the state of the Internet as experienced by the end user and a good localization of failures.

Finally, the results should prompt us to think differently about mitigation strategies currently being used in network resilience engineering. Given that major events have a much longer duration and different root cause (not predominantly network and fiber-driven) than commonly assumed, this suggests that more attention should be directed at resilience engineering of the entire service stack and specifically to the decoupling and challenge containment in tightly coupled systems. Our findings and recommendation for resilience optimization will be further discussed in the next section.

III. Recommendations and Challenges

In this section, we will discuss several commonly used failure mitigation strategies, exemplify under what circumstances they have failed, and provide recommendations and challenges on how to reach a more effective Internet failure mitigation plan. Since the Internet consists of a network of networks, some of our (intra-network) recommendations could be followed or implemented by Internet Service Providers (ISPs) and network operators to strengthen their infrastructures against accidental failures and malicious attacks, while other (inter-network) recommendations may warrant action by policy makers that govern the global Internet to lead to a more resilient Internet ecosystem.

A. *Network risk assessment*

The first step in obtaining a (more) robust network is creating a risk profile of the network that identifies possible network vulnerabilities, as well as a method to measure and assess the resilience of a network. [3] provides a comprehensive overview of various resilience classification approaches in the literature. In addition to a suitable metric, obtaining an accurate risk profile that can serve as a solid foundation for resilience engineering will require a number of aspects.

Going beyond a graph representation: A network typically consists of physical (point-of-presence) locations, the hardware at those locations, and the physical (optical fiber) connections between locations. On top of this network, the operator could run several logical network services, e.g., DWDM, SDH/Carrier Ethernet, and Ethernet, each constituting a layer on top of the previous one.

Regardless of the complexity of a network, they are often modeled as a graph consisting of nodes and links, and as a result, much work on improving network robustness has directed its attention to improving various graph connectivity metrics. In practice, a connection, however, typically does not form a straight line between the locations it connects, and such lines hide a number of underlying dependencies. For example, identifying the location of all single points of failure (SPoFs) in a network[1] based on a graph representation of the network could miss the vulnerabilities of several links being closely together. Geographical SPoFs may exist and should be identified at and across different layers.

Data to determine shared risks: During our study, it became evident that many – especially small – providers do not have sufficient information about their used resources, which are typically leased, to detect shared risk groups, and to correspondingly provision a resilient network. In addition to such geo-localized data, inference tools need to be developed to efficiently determine share risk groups and improve network design even for medium-sized operator networks. A noteworthy example in this direction is [1].

[1] Where a SPoF could be interpreted loosely in the sense that two network components are within close proximity of each other, e.g., 5 meters, or to use other terminology, these elements in close proximity are vulnerable to the same challenge and thus share the same risk group.

Probabilistic embedded risk assessment: Not only is geo-information on the network important, but so is its embedding in a geographical region and the context in which they operate. As network failures could be the result of natural disasters or abound in densely populated areas where fiber cuts are more frequent, the geographic areas in which the network is embedded clearly affects the risk to which the network is exposed. In addition, not all disasters and failures are created equal, and resilience engineering approaches should take the estimated likelihood and projected impact of a challenge into account for a cost- and risk-optimized mitigation strategy.

B. *Business continuity management and mutual aid*

From studying the crises responses that have been published and via several interviews with network operators, it became apparent that a business continuity management (BCM) plan does not always exist or is not up-to-date, leading to many failures being addressed in an ad-hoc approach. When investigating the incidents that were successfully overcome with minimal impact, mutual aid between operators (such as in temporarily lending equipment or routing traffic of another's ISP infrastructure) seemed to be a key factor to challenge containment. This, on the one hand, underlines the importance of BCM. Moreover, the extent of BCM policies and planned responses if they exist at all (typically only at larger operators) tend to greatly differ between network operators. On the other hand, this also highlights that resilience engineering is not and should not be limited to a single network. When addressing global incidents, it is important to have coordinated actions or agreements where one could rely on someone else's network for offloading traffic. Following such approaches similarly for the technical side of network design and resilience optimization would allow that higher resilience levels for a particular deployment could be achieved at a lower overall cost and network complexity, as – it is in insurance – risk and impact are distributed over more shoulders.

C. *Resilience by design*

Depending on the outcome of the risk assessment, the network may need to be augmented, i.e., adding nodes and/or links, to improve its resilience against the identified risks. The art of network augmentation is how to best balance resilience and augmentation costs. An overview of network planning under traffic and risk uncertainty can be found in [7].

Resilience of the entire stack: Despite the importance of communication networks, their security and resilience has long been only marginally addressed, typically as a later add-on, while in other critical systems (like airplanes) resilience has been designed from the get-go from a resilience perspective and tested continuously. As a result, several dependencies have been introduced in communication networks that might cause a ripple-through effect when only a single component fails. For instance, in October 2011, a core switch within the Blackberry network failed. Such hardware failure is in practice usually quickly resolved by proper fail-over schemes, but in this particular case it had caused a malfunctioning of a database that was much harder to resolve

and eventually led to an outage lasting three days. Resilience engineering in networks should, hence, look at the entire networking and application stack as even minor challenges that are remediated within the allowed mitigation margins may amplify and pose a large impact at other layers. Cross-layer resilience engineering – in contrast to, for example, cross-layer performance optimization in wireless networks – has unfortunately received little attention to date.

Spare resources. The end points of interconnections between individual networks take place in data centers that follow a wide variety of practices to increase resilience. There, the level of redundancy and protection against typical failures is described by tiers, with specific guidelines as to what practices must be implemented for a data center to meet these levels and be certifiable as such. The Amsterdam Internet Exchange (AMS-IX), for example, has extended these available standards and further refined them into a list of 141 minimum baseline (technical design, operational, and business continuity) requirements for the data centers providing service to the exchange. While, as stipulated in these standards, it is recommended to overprovision network elements by a factor of two and to create independent availability regions capable of securing network operations, there is currently an ongoing trend where providers are operating their networks at higher and higher loads (as Google for instance is doing with their software-defined wide-area network that is connecting their data centers). The "hotter" the network is operated, the fewer backup resources are available, and the higher the risk in case of failure, since backup paths/resources might not be available. Moreover, running a network at high utilization introduces a risk of overload, as we have seen for instance with popular applications, like Twitter, in their early days. Finally, adopting new technologies, such as software-defined networking (SDN and its protocol OpenFlow) could pose new vulnerabilities, for instance with respect to the robustness of the SDN controller now introducing a new SPoF or the security of the OpenFlow protocol.

Implications of tightly coupled systems, shared infrastructure and unknown SPoFs. In the past few years, the role of cloud computing, in which the infrastructure, the platform, and even the software used by IT operations are outsourced services, has become more prominent. The flexibility of cloud services certainly has its advantages, since they can be used when and only for how long they are needed, and be leased for prices charged in small increments of actual usage. However, these shared infrastructures also pose a risk that failures of a data center could cripple many services. This is supported by some analysts proclaiming that 2012 was the year of cloud (computing) outages. For many Internet services building on such cloud infrastructures this creates un-mitigateable risks, as customers typically do not have much insight into the concrete building blocks of the used infrastructure and potential architectural SPoFs. This general issue, however, greatly extends this particular scenario of cloud computing. In recent years, services have become increasingly coupled and integrated, which has also increased the vulnerability of Internet services due to common shared or cross-dependent infrastructures. Similar to the intensified linkages among actors in the financial market that led to the housing bubble burst in 2008, we might have created similar systemic or hyper risks [5] in Internet services that might explain the comparatively large magnitude of outages. The resilience of such tightly coupled systems is, however, both in general and in specific for the case of computer networks and the Internet as its most prominent example still largely unknown. More research is needed to understand risk and failure trajectories in these tightly coupled systems to develop effective challenge mitigation strategies for Internet services operating under such circumstances.

D. Monitoring of Inter-network resilience

The key to inter-domain routing resilience is the establishment of redundancy at multiple physical end points and if possible, also across multiple levels. The most fundamental inter-domain protection concept is the establishment of multi-homing, i.e., the presence of at least two distinct uplink connections towards non-local destinations. To realize the maximum possible resilience from such a setup, the critical dependencies of the upstream providers should ideally be investigated (such as where the transit providers' fibers run along, from which grid their equipment is powered, or where they interconnect), but obtaining a comprehensive view of this is frequently difficult.

In case a network operator has established several interconnection points with another ISP, the BGP protocol provides additional means to manage and thereby strengthen the interconnection. By tuning the individual BGP configuration at each location and influencing through which points traffic should enter or exit the autonomous system, such as the BGP multi-exit discriminators, local preferences or path attributes, providers can obtain a fine level of control on the traffic flows between networks, privileging or relieving particular hardware over others.

For such setups at network operators and to further deepen the insight in the resilience and reliability of Internet services and their underlying infrastructures in academia, a large monitoring framework should be established. It would be able to build up an assessment and track record of "how good" connections via a particular autonomous system are, what the stability of individual paths is within an autonomous system, and what particular hardware resides at certain geographical locations. Such monitoring systems have contributed a great deal to minimize the impacts of BGP hijacking incidents, as malicious and accident prefix announcement can today be rapidly detected. Establishing a similar system to understand the exact mechanics, location, and impact of Internet failures could promise to generate a similar leap to a more resilient Internet.

IV. Conclusion

Is it "all quiet on the Internet front?" In this article, we investigated a wide scale of Internet failures during the course of the past 6 years – where it became apparent that failures abound – and analyzed their root cause, frequency, duration, and societal impact. Such a study, to date, was missing, yet is vital in establishing proper Internet failure mitigation schemes.

In the second part of this article we scrutinized currently employed mitigation schemes, exemplified in which cases they failed and why, and proposed recommendations and challenges, to be on the road towards reaching fine-grained network risk assessment methods and better resilience planning and responses.

Additional resources: Details about the incidents described in this paper as well as other resources can be found on https://www.internetview.org, a new website dedicated to Internet infrastructure monitoring and resilience.

Acknowledgments

Part of this work has been supported by the EU FP7 EINS project under grant agreement No. 288021.

References

[1] N. Adam, R. Stiles, A. Zimdars, R. Timmons, J. Leung, G. Stachnick, J. Merrick, R. Coop, Va. Slavin, T. Kruglikov, J. Galmiche, and S. Mehrotra, Consequence Analysis of Complex Events on Critical U.S. Infrastructure, Communications of the ACM 56(6): 83-91, 2013.

[2] K. Butler, T. R. Farley, P. McDaniel, and J. Rexford, "A Survey of BGP Security Issues and Solutions," Proceedings of the IEEE, vol. 98, no. 1, January 2010.

[3] P. Cholda, A. Mykkeltveit, B. E. Helvik, O. J. Wittner, A. Jajzczyk, A Survey of Resilience Differentiation Frameworks in Communication Networks, IEEE Communications Surveys, 9(4) 2007.

[4] C. Doerr, R. Gavrila, F.A. Kuipers, and P. Trimintzios, "Good Practices in Resilient Internet Interconnection," European Network Information and Security Agency (ENISA) report, June 2012.

[5] D. Helbing, "Globally networked risks and how to respond," Nature 497: 51-59, 2013.

[6] W. Kim, O.-R. Jeong, C. Kim, J. So, "The dark side of the Internet: Attacks, costs and responses," Information Systems 36: 675-705, 2011.

[7] S. Yang and F.A. Kuipers, "Traffic Uncertainty Models in Network Planning," IEEE Communications Magazine, vol. 52, no. 2, pp. 172-177, February 2014.

Christian Doerr is an assistant professor in the Network Architectures and Services group at Delft University of Technology (TUDelft). He received a M.Sc. degree in Computer Science and a Ph.D. degree in Computer Science and Cognitive Science from the University of Colorado at Boulder. His research interests revolve around critical infrastructure protection, cyber security and resilience engineering.

Fernando A. Kuipers (SM) is an associate professor in the Network Architectures and Services group at Delft University of Technology (TUDelft). He received the M.Sc. degree in Electrical Engineering at TUDelft in June 2000 and subsequently obtained his Ph.D. degree (cum laude) in 2004 at the same faculty. His research interests mainly revolve around network algorithms and cover Routing, Quality of Service, Network Survivability, Optical Networks, and Content Distribution. His work on these subjects includes distinguished papers at IEEE INFOCOM 2003, Chinacom 2006, IFIP Networking 2008, IEEE FMN 2008, IEEE ISM 2008, and ITC 2009.