Network Resilience: A Systematic Approach

Paul Smith, Lancaster University
David Hutchison, Lancaster University
James P. G. Sterbenz, University of Kansas and Lancaster University
Marcus Schöller, NEC Laboratories Europe
Ali Fessi, Technische Universität München
Merkouris Karaliopoulos, NKU Athens
Chidung Lac, France Telecom (Orange Labs)
Bernhard Plattner, ETH Zurich

ABSTRACT

The cost of failures within communication networks is significant and will only increase as their reach further extends into the way our society functions. Some aspects of network resilience, such as the application of fault-tolerant systems techniques to optical switching, have been studied and applied to great effect. However, networks — and the Internet in particular — are still vulnerable to malicious attacks, human mistakes such as misconfigurations, and a range of environmental challenges. We argue that this is, in part, due to a lack of a holistic view of the resilience problem, leading to inappropriate and difficult-to-manage solutions. In this article, we present a systematic approach to building resilient networked systems. We first study fundamental elements at the framework level such as metrics, policies, and information sensing mechanisms. Their understanding drives the design of a distributed multilevel architecture that lets the network defend itself against, detect, and dynamically respond to challenges. We then use a concrete case study to show how the framework and mechanisms we have developed can be applied to enhance resilience.

INTRODUCTION

Data communication networks are serving all kinds of human activities. Whether used for professional or leisure purposes, for safety-critical applications or e-commerce, the Internet in particular has become an integral part of our everyday lives, affecting the way societies operate. However, the Internet was not intended to serve all these roles and, as such, is vulnerable to a wide range of challenges. Malicious attacks, software and hardware faults, human mistakes (e.g., software and hardware misconfigurations), and large-scale natural disasters threaten its normal operation. Resilience, the ability of a network to defend against and maintain an acceptable level of service in the presence of such challenges, is viewed today, more than ever before, as a major requirement and design objective. These concerns are reflected in, among other ways, in the Cyber Storm III exercise carried out in the United States in September 2010, and the “cyber stress tests” conducted in Europe by the European Network and Information Security Agency (ENISA) in November 2010; both aimed precisely at assessing the resilience of the Internet, this “critical infrastructure used by citizens, governments, and businesses.”

Resilience evidently cuts through several thematic areas, such as information and network security, fault tolerance, software dependability, and network survivability. A significant body of research has been carried out around these themes, typically focusing on specific mechanisms for resilience and subsets of the challenge space. We refer the reader to Sterbenz et al. [1] for a discussion on the relation of various resilience disciplines, and to a survey by Cholda et al. [2] on research work for network resilience.

A shortcoming of existing research and deployed systems is the lack of a systematic view of the resilience problem, that is, a view of how to engineer networks that are resilient to challenges that transcend those considered by a single thematic area. A non-systematic approach to understanding resilience targets and challenges (e.g., one that does not cover thematic areas) leads to an impoverished view of resilience objectives, potentially resulting in ill suited solutions. Additionally, a patchwork of resilience mechanisms that are incoherently devised and deployed can result in undesirable behavior and increased management complexity under chal-
The EU-funded ResumeNet project argues for resilience as a critical and integral property of networks. It advances the state of the art by adopting a systematic approach to resilience, which takes into account the wide-variety of challenges that may occur. At the core of our approach is a coherent resilience framework, which includes implementation guidelines, processes, and toolsets that can be used to underpin the design of resilience mechanisms at various levels in the network. In this article, we first describe our framework, which forms the basis of a systematic approach to network resilience. At its core is a control loop comprising a number of conceptual components that realize the real-time aspect of the D²R² + DR strategy, and consequently implement network resilience. Based on the resilience control loop, other necessary elements of our framework are derived, namely resilience metrics, understanding challenges and risks, a distributed information store, and policy-based management. The remainder of this section describes the resilience control loop, then motivates the need for these framework elements.

**RESILIENCE CONTROL LOOP**

Based on the real-time component of the D²R² + DR strategy, we have developed a resilience control loop, depicted in Fig. 1, in which a controller modulates the input to a system under control in order to steer the system and its output towards a desired reference value. The control loop forms the basis of our systematic approach to network resilience — it defines necessary components for network resilience from which the elements of our framework, discussed in this section, are derived. Its operation can be described using the following list; items correspond to the numbers shown in Fig. 1:

1. The reference value we aim to achieve is expressed in terms of a resilience target, which is described using resilience metrics. The resilience target reflects the requirements of end users, network operators, and service providers.

2. Defensive measures need to be put in place proactively to alleviate the impact of challenges on the network, and maintain its ability to realize the resilience target. A process for identifying the challenges that should be considered in this defense step of the strategy (e.g., those happening more frequently and having high impact) is necessary.

3. Despite the defensive measures, some challenges may cause the service delivered to users to deviate from the resilience target. These challenges could include unforeseen attacks or mis-

---

**FRAMEWORK FOR RESILIENCE**

Our resilience framework builds on work by Sterbenz et al. [1], whereby a number of resilience principles are defined, including a resilience strategy, called D²R² + DR: Defend, Detect, Remediate, Recover, and Diagnose and Refine. The strategy describes a real-time control loop to allow dynamic adaptation of network resources in response to challenges, and a non-real time control loop that aims to improve the design of the network, including the real-time loop operation, reflecting on past operational experience.

The framework represents our systematic approach to the engineering of network resilience. At its core is a control loop comprising a number of conceptual components that realize the real-time aspect of the D²R² + DR strategy, and consequently implement network resilience. Based on the resilience control loop, other necessary elements of our framework are derived, namely resilience metrics, understanding challenges and risks, a distributed information store, and policy-based management. The remainder of this section describes the resilience control loop, then motivates the need for these framework elements.

---

**Figure 1. The resilience control loop: derived from the real-time component of the D²R² + DR resilience strategy.**
Defining a resilience target requires appropriate metrics. Ideally, we would like to express the resilience of a network using a single value, \( R \), in the interval \([0,1]\), but this is not a simple problem because of the number of parameters that contribute to and measure resilience, and due to the multilayer aspects in which each level of resilience (e.g., resilient topology) is the foundation for the next level up (e.g., resilient routing). We model resilience as a two-dimensional state space in which the vertical axis \( P \) is a measure of the service provided when the operational state \( N \) is challenged, as shown in Fig. 2. Resilience is then modeled as the trajectory through the state as the network goes from delivering acceptable service under normal operations \( S_0 \) to degraded service \( S_r \). Remediation improves service to \( S_0 \) and recovery returns to the normal state \( S_0 \). We can measure resilience at a particular service level as the area under this trajectory, \( R \).

We have developed a number of tools for evaluating network resilience. For example, we use MATLAB or ns-3 simulation models to measure the service at each level and model results under various challenges and attacks, as in Fig. 2, where each axis is an objective function of the relevant parameters [4]. Furthermore, we have developed the Graph Explorer tool [5] that takes as input a network topology and associated traffic matrix, a description of challenges, and a set of metrics to be evaluated. The result of the analysis is a series of plots that show the metric envelope values \( (m_i(\text{min}), m_i(\text{max})) \) for each specified metric \( m_i \), and topology maps indicating the resilience across network regions.

Figure 3 shows an example of the resilience of the European academic network GÉANT2 to link failures. The set of plots in Fig. 3a show metric envelopes at different protocol levels — the aim is to understand how jitter responds in comparison with metrics at other levels, such as queue length and connectivity. Surprisingly, jitter is not clearly related to queue length, and a monotonic increase in path length does not yield a similar increase in queue length for all scenarios of link failures. In fact, the fourth link failure disconnects a region of the network; whereas up to three failures, the heavy use of a certain path resulted in increasing queue lengths and jitter. The partition increases path length, because route lengths are set to infinity, and decreases connectivity, which is accompanied by a reduction in jitter, shown with the blue arrows in Fig. 3a. The topology map in Fig. 3b highlights the vulnerability of regions of GÉANT2 with a heat map, which can be used by network planners.

Our framework for resilience metrics (i.e., the multilevel two-dimensional state space and the use of metric envelopes) can be used to understand the resilience of networks to a broad range of challenges, such as misconfigurations, failures, and attacks. The ability to evaluate a given network’s resilience to a specific challenge is limited by the capability of the tools to create complex challenge scenarios — this is an area for further work, in which our effort should be focused on modeling pertinent high-impact challenges.
Figure 3. Example output from the Graph Explorer, developed in the ResumeNet project: a) plots showing the relationship between metrics at various layers in response to link failures on the GÉANT2 topology; b) a heat map showing vulnerable regions of the topology with respect to a given set of metrics. Reprinted from [5].
We advocate the use of a policy-based management framework to define the behavior of real-time loop instantiations. Consequently, the implementation of resilience mechanisms can be decoupled from the resilience management strategies, which are expressed in policies.

Understanding Challenges and Risks

Engineering resilience has a monetary cost. To maximize the effectiveness of the resources committed to resilience, a good understanding of the challenges a network may face is mandatory. We have developed a structured risk assessment approach that identifies and ranks challenges in line with their probability of occurrence and their impact on network operation (i.e., how disruptive they are to the provision of its services). The approach should be carried out at the detailed design stage of network design when proactive defensive measures are deployed, and repeated regularly over time as part of the process of network improvements.

Central to determining the impact of a challenge is to identify the critical services the network provides and the cost of their disruption: a measure of impact. Various approaches can be used to identify the critical services, such as discussion groups involving the network’s stakeholders. Networked systems are implemented via a set of dependent subsystems and services (e.g., web and Session Initiation Protocol [SIP] services rely on Domain Name Service [DNS]). To identify whether challenges will cause a degradation of a service, it is necessary to elaborate these dependencies.

The next phase is to identify the occurrence probabilities of challenges (challenge_prob). Some challenges will be unique to a network’s context (e.g., because of the services it provides), while others will not. In relation to these challenges, shortcomings of the system (e.g., in terms of faults) should be identified. The aim is to determine the probability that a challenge will lead to a failure (fail_prob). We can use tools, such as our Graph Explorer, analytical modeling, and previous experience (e.g., in advisories) to help identify these probabilities. Given this information, a measure of exposure can be derived using the following equation:

\[
\text{exposure} = (\text{challenge}_\text{prob} \times \text{fail}_\text{prob}) \times \text{impact}
\]

With the measures of exposure at hand, resilience resources can be targeted at the challenges that are likely to have the highest impact.

Information Sources and Sharing for Resilience

For the most part, network management decisions are made based on information obtained from monitoring systems in the network (e.g., via Simple Network Management Protocol [SNMP]). However, to be able to make autonomous decisions about the nature of a wide range of challenges and how to respond to them — a necessary property of resilient networks — a broader range of information needs to be used. In addition to traditional network monitoring information, context information, which is sometimes “external” to the system can be used. Earlier work has demonstrated how the use of weather information, an example of context, improves the resilience of millimeter-wave wireless mesh networks, which perform poorly in heavy rain [4]. Also, in addition to node-centric monitoring tools, such as NetFlow and SNMP, task-centric tools can be used to determine the root cause of failures. For example, X-trace [6] is a promising task-centric monitoring approach that can be used to associate network and service state (e.g., router queue lengths and DNS records) with service requests (e.g., retrieving a web page). This multilevel information can then be used to determine the root causes of failures.

We are developing a Distributed Store for Challenges and their Outcome (DISco), which uses a publish-subscribe messaging pattern to disseminate information between subsystems that realize the real-time loop. Such information includes actions performed to detect and remediate challenges. Information sources may report more data than we can afford or wish to relay on the network, particularly during challenge occurrences. DISco is able to aggregate information from multiple sources to tackle this problem. Decoupling information sources from components that use them allows adaptation of challenge analysis components without needing to modify information sources. To assist the two phases of the outer loop, DISco employs a distributed peer-to-peer storage system for longer-term persistence of data, which is aware of available storage capacity and demand.

Policies for Resilience

We advocate the use of a policy-based management framework to define the behavior of real-time loop instantiations. Consequently, the implementation of resilience mechanisms can be decoupled from the resilience management strategies, which are expressed in policies. Two immediate benefits: the nature of challenges changes over time — management strategies can be adapted accordingly without the need for network down-time; and policies allow network operators to clearly express when they would like to intervene in the network’s operation (e.g., when a remediation action needs to be invoked).

Research outcomes from the policy-based management field can help address the complexities of resilience management [7]. A difficult task is deriving implementable policies from high-level resilience requirements, say, expressed in service level agreements (SLAs). With appropriate modifications, techniques for policy refinement can be used to build tools to automate aspects of this process. Policy-based learning, which relies on the use of logical rules for knowledge representation and reasoning, is being exploited to assist with the improvement stages of our strategy. Techniques for policy ratification are currently used to ensure that invocation of different resilience strategy sets does not yield undesirable conflicting behavior. Conflicts can occur horizontally between components that realize the resilience control loop, and vertically across protocol levels. For example, a mechanism that replicates a service using virtualization techniques at the service level could conflict with a mechanism that is rate-limiting traffic at the network level. Example policies of this sort are shown in Fig. 4. So that these forms of conflict can be detected, Agrawal et al. [8] provide a theoretical foundation for conflict resolution that needs to be extended with domain-specific knowledge, for example, regarding the nature of resilience mechanisms.
DEFENSE AND DYNAMIC ADAPTATION ARCHITECTURE

In this section, we describe a set of defensive mechanisms and an architecture that realize our systematic approach to resilience, described earlier. The architecture, shown in Fig. 5, consists of several subsystems implementing the various tasks of the communication system as well as the challenge detection components and adaptation capabilities. The behavior of all these subsystems is directed by the resilience manager using policies, which are held in a resilience knowledge base. Central to this architecture is DISco, which acts as a publish-subscribe and persistent storage system, containing information regarding ongoing detection and remediation activities. From an implementation perspective, based on the deployment context, we envisage components of the architecture to be distributed (e.g., in an Internet service provider [ISP] network) or functioning entirely on a single device (e.g., nodes in a delay-tolerant network).

DEFENSIVE MEASURES

As a first step, defensive measures need to be put in place to alleviate the impact of challenges on the network. Since challenges may vary broadly from topology-level link failures to application-level malware, defensive measures against anticipated high-impact challenges need to be applied at different levels and locations. Defensive measures can either prevent a challenge from affecting the system or contain erroneous behavior within a subsystem in such a way that the delivered service still meets its specification. A selection of defensive measures developed in the ResumeNet project is shown in Table 1.

DETECTION SUBSYSTEMS

The second step is to detect challenges affecting the system leading to a deviation in delivered service. We propose an incremental approach to challenge analysis. Thereby, the understanding about the nature of a challenge evolves as more inputs become available from a variety of information sources. There are two apparent advantages of this incremental approach. First, it readily accommodates the varying computational overhead, timescales, and potentially limited accuracy of current detection approaches [9]. Second, relatively lightweight detection mechanisms that are always on can be used to promptly initiate remediation, thus providing the network with a first level of protection, while further mechanisms are invoked to better understand the challenge and improve the network response. Lightweight detection mechanisms can be driven by local measurements carried out in the immediate neighborhood of affected nodes.

For example, consider high-traffic volume challenges, such as a DDoS attack or a flash crowd event. Initially, always-on simple queue monitoring could generate an alarm if queue lengths exceed a threshold for a sustained period. This could trigger the rate limiting of links associated with high traffic volumes. More expensive traffic flow classification could then be used to identify and block malicious flows, consequently not subjecting benign flows to rate limiting. Challenge models, shown in Fig. 5, describe symptoms of challenges and drive the analysis process. They can be used to initially identify broad classes of challenge, and later to refine identification to more specific instances.
REMEDIATION AND RECOVERY SUBSYSTEMS

The challenge detection subsystem interfaces with the remediation and recovery subsystem, the third and final step, by issuing alerts to DISco using the publish(challenge) primitive. These alerts contain information about the challenge and its impact on the network, in terms of the metrics that are falling short of the resilience target. The network resilience manager takes this information as context data, and, based on policies, selects an adaptation strategy. In doing so, the network resilience manager realizes the resilience management functionality in Fig. 1. If further information is required by the network resilience manager that is not contained in the alert, the lookup(alarm) primitive can be used. Furthermore, the network resilience manager can make use of consultants, such as path computation elements, which can compute new topological configurations, such as new channel allocations or new forwarding structures. Resilience mechanisms are deployed by enforcing new configurations on the managed entities (e.g., routers and end hosts) in the network. To implement the resilience estimator, the network resilience manager assesses the success of chosen remedies. The assessment is stored in DISco to aid the diagnosis and refinement steps of the background loop. Carrying out this assessment is not straightforward since it requires spatio-temporal correlation of changes in network state, which is an issue for further work.

RESILIENCE IN SMART ENVIRONMENTS: A CASE STUDY

We are currently evaluating our generic approach to resilience through concrete study cases that cover a range of future networking paradigms: wireless mesh and delay-tolerant networks, peer-to-peer voice conferencing and service provision over heterogeneous smart environments. Herein, we focus our discussion on the last study case. The widespread use of smart mobile devices, together with identifiers such as radio frequency identification (RFID), embedded in objects such as products, enables communication with, and about, these objects. The French national project Infrastructure for the Future Trade (ICOM) has developed an infra- and inter-enterprise infrastructure, depicted in Fig. 6, that allows the connection of objects with enterprise information systems and fixed or mobile terminals. This ICOM platform can be used as a foundation for a number of enterprise applications. The experimentation makes use of three different entities:

• The data acquisition site is the data source — items identified by RFID, for example, are read and their information sent to a processing centre located remotely.
• The data processing site houses different modules of the platform (e.g., data collection, aggregation, and tracking), which will forward the enriched data to the core application.
• The application provision site hosts the platform’s central element — it is also where the data subscriber applications (web services, legal application, etc.) are linked.

Based on outcomes of our risk assessment approach, high-impact challenges to the platform include those that are intentional and accidental: malicious attacks that threaten the confidentiality and integrity of commercially sensitive data, DDoS attacks by extortionists, and, given the immature nature of the platform, software and hardware faults. This understanding ensures that we implement appropriate
defensive measures and dynamic adaptation strategies.

Consequently, defensive measures primarily include secure VPN connections between sites, enabling confidentiality and integrity of the data in transit. Security mechanisms, such as authentication and firewalls, are also implemented. Redundancy of infrastructure and implementation diversity of services are exploited to maintain reliability and availability in the presence of failures caused by software faults.

Incremental challenge analysis is realized using the Chronicle Recognition System (CRS), a temporal reasoning system aimed at alarm-driven automated supervision of data networks [10]. Lightweight detection mechanisms generate alarms based on metrics, such as anomalous application response times and data processing request rates. Finally, policy-based adaptation, implementing remediation and recovery, is achieved through the specification of the platform’s nominal and challenge context behavior (i.e., its configuration in response to anticipated challenges). In our case study, challenge context policies describe configurations in response to alarms indicating a DDoS attack. For example, modified firewall configurations are defined to block traffic deemed to be malicious; service virtualization configurations that make use of redundant infrastructure are also specified to load balance increased resource demands. The transition between behaviors is based on alert messages, generated via challenge analysis, and outcomes from continuous threat level assessment. The case study sketched above illustrates the gain from applying our resilience strategies in a systematic approach: starting from a risk assessment, challenges are derived, allowing defense measures to be deployed. The following step is the specification of chronicles — temporal descriptions of challenges — for detection by the CRS, and policy-driven mechanisms to remediate and recover from unforeseen failures.

**CONCLUSION**

Given the dependence of our society on network infrastructures, and the Internet in particular, we take the position that resilience should be an integral property of future networks. In this article, we have described a systematic approach to network resilience. Aspects of our work represent a longer-term vision of resilience that span multiple dimensions, including infrastructure’s reliability and availability, application-level measures, such as service virtualization, necessitate fewer changes at the network core and lend to easier implementation. Further benefits for network practitioners are anticipated through the use of tools like the Graph Explorer, which can explore correlations among metrics at various levels of network operation.

**ACKNOWLEDGEMENTS**

The work presented in this article is supported by the European Commission under Grant No. FP7-224619 (the ResumeNet project). The authors are grateful to the members of the ResumeNet consortium, whose research has contributed to this article, and in particular to Christian Doerr and his colleagues at TU Delft for the work presented in Fig. 3.
Figure 6. The ICOM platform connecting enterprise sites that perform data processing and application provisioning with objects in a smart environment. Selected resilience mechanisms are shown that can be used to mitigate identified challenges.

REFERENCES

MARCUS SCHÖLLER is a research scientist at NEC Laboratories America, ACM, IET/IEE, and IEICE. He is involved in Low-Latency Communication. He is a member of the NSF FIND program. He received a doctorate in computer science from the University of Karlsruhe, Germany, in 2001 and his doctorate in engineering in 2006 on robustness and survivability. His research currently focuses on resilient, survivable, and disaster-tolerant networking for the future Internet for which he is the principal author of the book High-Speed Networking: A Systematic Approach to High-Bandwidth Low-Latency Communication. He is a member of the ACM, IET, IEEE, and IEICE.


BIographies

PAUL SMITH is a senior research associate at Lancaster University’s School of Computing and Communications. He submitted his Ph.D. thesis in the area of programmable networking resource discovery in September 2003, and graduated in 1999 with an honors degree in computer science from Lancaster. In general, he is interested in the various ways that networked (socio-technical) systems fail to provide a desired service when under duress from various challenges, such as attacks and misconfigurations, and developing approaches to improving their resilience. In particular, his work has focused on the rich set of challenges that face community-driven wireless mesh networks.

DAVID HUTCHISON is director of InfoLab21 and professor of computing at Lancaster University, and has worked in the areas of computer communications and networking for more than 25 years, recently focusing his research efforts on network resilience. He has served as member or chair of numerous TPCs (including the flagship ACM SIGCOMM and IEEE INFOCOM), and is an editor of the renowned Springer Lecture Notes in Computer Science and the Wiley CNDS book series.

JAMES P. G. STERBENZ is director of ResilNets research group at the Information & Telecommunication Technology Center and associate professor of electrical engineering and computer science at The University of Kansas, a visiting professor of computing in InfoLab21 at Lancaster University, and has held senior staff and research management positions at BBN Technologies, GTE Laboratories, and IBM Research. He received a doctorate in computer science from Washington University in St. Louis, Missouri. His research is centered on resilient, survivable, and disruption-tolerant networking for the future Internet for which he is involved in the NSF FIND and GENI programs as well as the EU FIRE program. He is principal author of the book High-Speed Networking: A Systematic Approach to High-Bandwidth Low-Latency Communication. He is a member of the ACM, IET, IEEE, and IEICE.

ALI FESI is a researcher at the Technische Universität München (TUM). He holds a Ph.D. from TUM and a Diplom (Master’s) from the Technische Universität Kaiserslautern. His research currently focuses on the resilience of network services, such as web and SIP, using different techniques (e.g., P2P networking, virtualization, and cryptographic protocols). He is a regular reviewer of several scientific conferences and journals, such as ACM IPTComm, IEEE GLOBECOM, IFIP Networking, and IEEE/ACM Transactions on Networking.

MERKOURIS KARALIPOULOS is a Marie Curie Fellow in the Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Greece, since September 2010. He was a postdoctoral researcher in the University of North Carolina, Chapel Hill, in 2006 and a senior researcher and lecturer at ETH Zurich, Switzerland, from 2007 until 2010. His research interests lie in the general area of wireless networking, currently focusing on network resilience problems related to node selfishness and misbehavior.

CHIDUNG LAC is a senior researcher at France Telecom (Orange Labs). Besides activities linked with network architecture evolution, for which he contributes to the design of scenarios and roadmaps, his research interests are centered on network and services resilience, particularly through his involvement in European projects such as the ReSIST Network of Excellence (2006–2009) and the present STREP ResumeNet (2008–2011). He holds a Doctorat d’Etat-é-s Sciences Physiques (1987) from the University of Paris XI Orsay.

BERNHARD PLATNER is a professor of computer engineering at ETH Zurich, where he leads the communication systems research group. He has a diploma in electrical engineering and a doctoral degree in computer science from ETH Zurich. His research currently focuses on self-organizing networks, systems-oriented aspects of information security, and future Internet research. He is the author or co-author of several books and has published over 160 refereed papers in international journals and conferences.