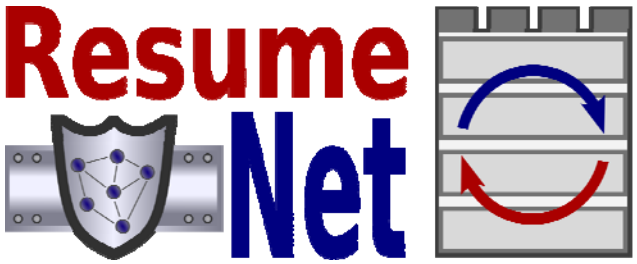
Resilience and Survivability for future networking: framework, mechanisms, and experimental evaluation



Deliverable number:	D4.1a
Deliverable name:	Federation requirements
WP number:	4
Delivery date:	28/02/2009
Date of preparation:	09/04/2009
Editor:	M. Karaliopoulos (ETHZ)
Contributor(s):	M. Karaliopoulos (ETHZ), G. Popa (ETHZ), C. Lac (FT), A. Fessi (TUM), A. Fischer (UP), C. Rohner (UU)
Internal reviewer:	D. Hutchison (ULANC)

Summary

The deliverable is one of the project “light” deliverables requested by European Commission in the context of the project’s commitment to close interaction with the FIREworks Coordination Action. The deliverable aims at providing inputs to FIREworks for the compilation of a deliverable on federation requirements, which will aggregate contributions from all FIRE projects.

The deliverable consists of three main sections: an introduction briefly summarizing the scope of the experimentation within ResumeNet and the current thinking regarding the kind of experimentation facilities to be used. Note that since the project is only six months old and the experimentation activities officially start in M18, some of these decisions will be definitely reiterated and may be modified in the year to come.

The second section outlines the four experimental scenarios in the project and the testbeds that will be used to realize them, whereas the last section brings together some thoughts on the, rather limited, federation requirements coming out of ResumeNet.

Contents

1. Introduction

The experimentation work in ResumeNet is carried out in the project WP4 and its main aim is to evaluate the resilience framework aspects defined in WP1 and the mechanisms realizing this framework in the project WPs 2 and 3.

Four experimental scenarios are defined with the aim to address a wide variety of resilience issues in the context of different types of networks and service provision settings. Therefore, these four experimental scenarios have been selected to be complementary with respect to:

- network type: one scenario is considering wired networks and three scenarios consider wireless networks. The first of them is related to wireless multihop networks, the second evolves around self-organizing, opportunistic networks, and the third addresses smart environments, i.e., spaces filled with sensors and actuators within which devices are moving.
- types of network service faults that will be considered: they range from node misbehavior at different layers (MAC, routing), to software misconfigurations, and DDoS attacks.
- network functions: routing, wireless medium sharing, transport, but also signaling functionality provision will be addressed in the experimental scenarios.

The actual experimentation facilities to be used in these four scenarios are mainly in-house test beds. This choice minimizes interdependencies that might pose high overhead in the experimentation process while still satisfying the objectives of the experimentation tasks in the project. Examples of project in-house testbeds are the Wireless Mesh Network test bed of the ETHZ TIK group and the Haggle test bed maintained by the Uppsala University. Nevertheless, there are scenarios that would definitely benefit from larger-scale experimental facilities that may become available from FIRE or other EU R&D activities. An example of this second alternative regarding experimentation facilities is PlanetLab (Europe), currently under the responsibility of OneLab2 project and/or the test bed that will come out of the EC FP6 ANA (Autonomic Network Architecture) project.

The exact dependence on these external testbeds has been under assessment since the beginning of the project; in the case of OneLab2, for the example, project representatives have attended the kick-off meeting and there have been discussions with the groups involved there regarding the progress of work. The baseline at the moment is to largely rely on the in-house testbed experimental facilities, while constantly assessing the possibility to use larger-scale experimentation facilities from FIRE or other EU R&D projects. In the context of the project experimentation, these test beds will be enhanced to facilitate the four scenarios envisaged in the project.

An additional experimentation facility that is used in the project is the Wireless Multihop Network operating in the village of Wray, situated approximately ten miles from the city of Lancaster in the north-west of England. The facility is being used in the task of assessing challenges and their impact on network resilience [ResD1.1] and may accommodate further experimentation in the course of time.

.

2. Description of ResumeNet experimentation scenarios and related testbeds

2.1. Experimentation scenario: “Wireless Mesh Networking”

2.1.1. Scenario description

Ad-Hoc wireless networks are suitable for a vast array of applications where central control cannot be provided due to various reasons. However, the lack of central control also deprives ad-hoc networks of important properties. For instance, cooperation cannot be taken for granted when this assumption does not hold, that is, in situations each of these nodes is owned by a different entity. Therefore, a number of specific challenges stem from this essential characteristic of ad-hoc networks. Generally speaking, misbehavior can be classified as being either selfishness or maliciousness.

When referring to *selfishness*, one of the essential challenges is cooperation between nodes. The simplest form of cooperation is represented by forwarding data on behalf of other stations. Nevertheless, since cooperation is a global objective that is not recognized by individual nodes, a mechanism should be provided in the network to determine the nodes to collaborate or, in other words, to make them aware of the individual and global advantages that collaboration can provide.

The research effort has focused largely on game theoretic approaches as a viable way of offering incentives to follow the protocol specifications. Examples of such mechanisms are the Ad Hoc VCG [Ande03] and Corsac [Zhong05]. However, there are a number of aspects to be studied in practice (ranging from the time frame within which nodes converge to collaboration to the quantification of overhead that such protocols are involving). One experiment will be devoted to implementing a representative protocol and analyzing its operation on the TiKNet testbed (ref. section 2.1.2). The experiment will be implemented in software and will emulate the behavior at the network layer of the studied protocol. Experimentation should show whether the examined protocol can be used efficiently in practice and, if, so, what are the restrictions. In the end, this should lead to improvements of practical value in the area of game theoretic protocols.

With respect to node *maliciousness*, the goal is to examine how various types of attacks can influence the performance of wireless networks at local and global level. More precisely, the experiments will attempt to quantify the practical impact of common attacks at the network layer (relay nodes that drop packets totally or partially, periodic dropping, Jelly Fish behavior (Aad04)). For the purpose of this experiment, the software installed on the nodes will emulate the described behaviors at the network layer and will collect data related to the achieved throughputs in the studied cases. The comparison between simulation results and experimental results should provide an insight on how various parameters influence the performance of such networks in the presence of a certain percentage of misbehaving nodes and show which are the main factors that lead to resilience in this case.

2.1.2. Testbed description

The experiments will take place on the wireless multihop network testbed of the Computer Engineering and Networks laboratory (TiKNet), at the G floor of the building housing the Department of Information Technology and Electrical Engineering in Swiss Federal Institute of Technology (ETH Zürich). The testbed is readily available, having already been used in prior research work.

TiKNet consists of approximately 20 nodes (PCs) split in two categories:

- Dell PCs with 2GHz processor and 512MB RAM memory;
- PCs with 866MHz processor and 512MB RAM memory.

All the nodes are currently equipped with D-Link DWL-AG530, 108/54Mbit Tri-Mode Dualband WLAN Adapters. The testbed is currently being upgraded through the addition of new nodes.

Each node is running GNU/Linux and is connected also to the local wired network in order to ensure a safe way for the configuration and maintenance operations. Thus, such operations are usually performed out of band (via the wired network). The aforementioned tasks can be performed by using a standard web interface installed on every node [webmin] or by using ssh.

In-house software for emulation of various types of attacks is currently being developed. In addition, readily available traffic generation tools such as tcpdump, mgen, iperf, tudp shall be employed.

The testbed can be accessed by members of ETH Zürich involved in the ResumeNet project and by master students that help with the testbed development.

2.2. Experimentation scenario: “Opportunistic Networking”

Opportunistic networks rely on the mobility of their nodes to achieve to make up for the episodic network connectivity. No end-to-end connectivity is assumed at any given moment; instead a path is constructed by storing the data until connectivity changes and then forwarding the data. By introducing redundancy, chances to reach a destination are increased. At the same time, redundancy makes the network resilient to both node failures and lack of connectivity.

The bottleneck in opportunistic networks is storage capacity. The challenge is therefore to find a balance between increasing redundancy to maximize delivery, and limiting redundancy to avoid storage overflow.

The experiments investigate aspects related to congestion management and make use of the in-house Haggle testbed, which runs on both mobile phones and virtual machines. Details are described below.

2.2.1. Scenario description

In our experimentation we intend to investigate aspects related to congestion management, that is dissemination strategies, resource management (e.g., data ageing), and resilience to attacks. The goal is to investigate the influence of different strategies on system behavior and performance to finally improve resilience. We use the Haggle architecture for implementing the different strategies.

Dissemination strategies: Resolution and forwarding

Resource management: Ageing

Resilience to attacks: Spam, faked routing information, selfishness

2.2.1.1. Benchmark Application

The Haggle architecture provides a search-based data dissemination framework suitable for opportunistic communication, making it easy to share content directly between intermittently connected mobile devices: A Haggle search is a user's expressed interest in content, and these interests (in form of metadata attributes) propagate to other devices, which store them until replaced by newer ones. When a set of received interests match the stored content on a device, the latter tries to push the content to the matching devices. The push-based dissemination occurs among the group of devices with matching interests, or can be enhanced by using a dedicated forwarding algorithm using other nodes as well.

We evaluate the search-based networking architecture of Haggle by its *precision* and *recall* and use them as benchmark to evaluate resilience to the above described scenarios. Precision corresponds to the ratio of received content matching the interests of a node, while we define recall as the ratio of received content over all available matching content in the network.

LuckyMe is an application that periodically generates content with random metadata attributes out of a given attribute pool of integers in the interval [1,N]. At the same time, LuckyMe expresses its interest in content. The interests are expressed in the form of consecutive attributes with binomially distributed weights approximating a normal distribution $N(\mu,\sigma)$.. Content matching the interests therefore have a match value M corresponding to the sum of the weights of the matching attributes. We use LuckyMe to disseminate content into the network (i.e., traffic generation) and trigger traffic. Constant seed values can be used to achieve

repeatability in the random number generators to compare different forwarding or resource management algorithms with exactly the same traffic input.

2.2.2. Testbed Description

The Haggle testbed allows emulating a mobile opportunistic network and conducting repeatable tests in a controlled and easy to manage environment. The Xen virtual machine monitor is at the core of the testbed. Xen supports execution of multiple guest operating systems (or emulated Haggle devices), on a single physical machine, that are monitored by a host system.

Opportunistic network connectivity is emulated over a virtual Ethernet bridge. Network topology changes are performed by controlling connectivity with traffic filtering. This is done by setting rules in the iptable on each node in the testbed and thereby blocking traffic from certain nodes. The order in which the rules are configured is specified in a scenario file. These scenario files can automatically be generated from either real-world traces or statistical models. The Ethernet bridge also allows interconnecting with other networks, for example another testbed (scalability), or a wireless network with real-world Haggle nodes running on mobile phones.

A graphical management console allows starting/stopping nodes, controlling connectivity, and visualizing the internal state of the diverse nodes and their interaction. After an experiment, an collection of analysis scripts automatically generates statistics and graphs from the logfiles of the nodes (e.g., about delay, delivery ratio, dissemination topology, etc.)

The current testbed runs at a dual-core desktop computer (3GHz CPU, 4GB RAM), using Linux for both host and guest operating systems. Up to 30 Haggle devices are supported at the moment. Further scaling is planned although not critical to produce interesting results.

The testbed is available as a software distribution that project partners can install on their own hardware. Federation would be possible through a virtual network but has not been tested yet.

2.3. Experimentation scenario “Service-Level Resilience”

2.3.1. Scenario description

In this experimentation scenario, two concepts for service-level resilience will be evaluated; P2P-based services, and virtualization.

2.3.1.1. Experimentation with P2P overlays for service resilience

Services will be classified in two categories:

- Services that are carried out by a large scale P2P network where all end points are contributing to the service. Examples of these services are distributed rendez-vous points, for example, for establishing a VoIP phone call. This example is depicted in Figure 1.

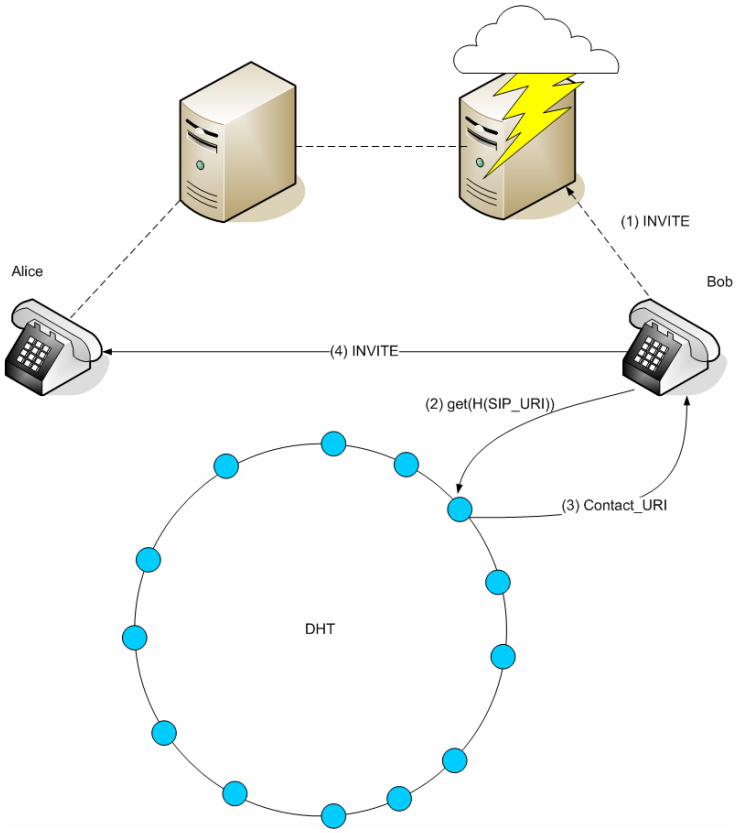


Figure 1. P2P-Session Establishment in case of server failure

- Services provided by a service provider, where the nodes offering the services are part of the provider’s infrastructure. In this case, servers and clients are organized in a P2P network.

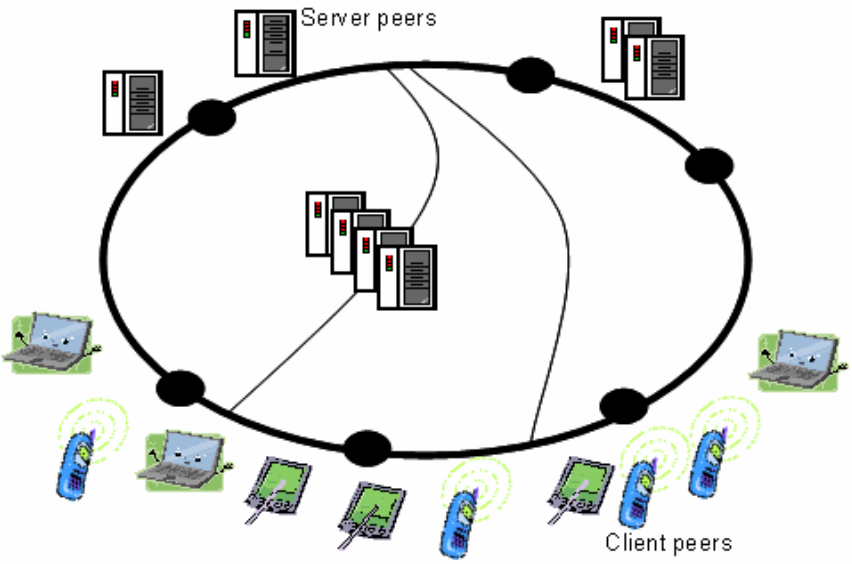


Figure 2. Hybrid P2P Overlay Networks with Clients and Servers Peers

The concepts that will be developed by TUM in WP3 for P2P-based service resilience and overlay-assisted service discovery will be evaluated. The overlay-assisted service discovery will involve the client peers actively in the remediation process when a challenge occurs, so they will be able to find other server peers, and re-establish connectivity in the overlay. Furthermore, this concept will facilitate locating servers when they change their location. This is the case, for example, when virtual servers have been migrated. Migration of services is a remediation mechanism anticipated by UP (see below). Therefore, it is planned that TUM will cooperate with UP for this purpose.

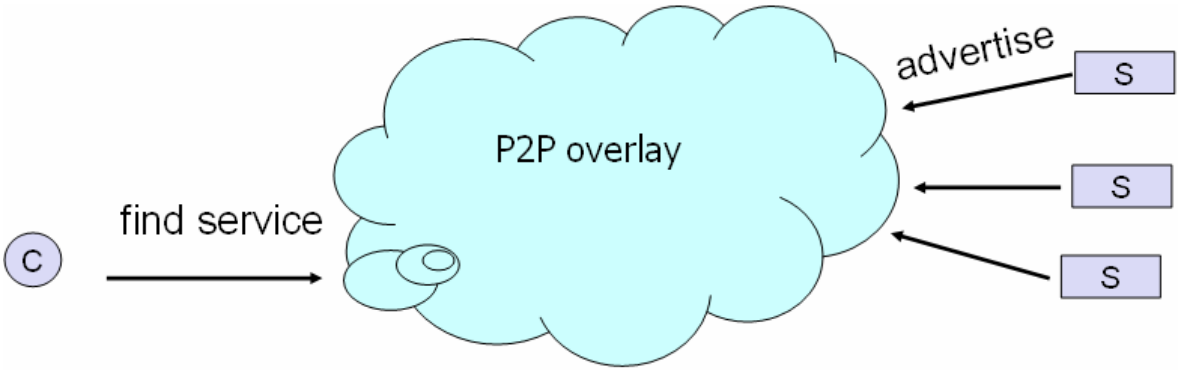


Figure 3 Overlay-based Service Discovery

The experimentation scenario will include the performance degradation up to complete failures of some major components in the service overlay as well as the underlying network. The failure of servers can be emulated by shutting down the server. The signaling for session establishment should remain possible. The failure of routers can be emulated by shutting down a subset of the peers, which are supposed to be connected behind the router. The overlay should be able to recover from this kind of challenge.

Performance degradation could be emulated by activating a higher CPU load at the server. It might be evaluated whether this could introduce additional complexity, or whether session establishment at the peers' side can still be performed smoothly.

It is worth mentioning that all these objectives for experimentation, challenge emulation and remediation mechanisms remain subject for changes, since the project is at an early stage and even the concepts for service resilience are still work-in-progress. Concepts and evaluation methods might be adapted during the project lifetime.

2.3.1.2. Experimentation with virtualization for service resilience

The second major concept for service-level resilience is virtualization. Virtualization provides mechanisms to abstract a service from the underlying hardware, enabling service mobility in order to enhance service resilience. An experimentation scenario will examine the requirements and benefit of dynamic service mobility, taking both cold migration and life migration into account as a remediation mechanism. Restrictions on mobility, both on the service side and on the virtualization side, have to be identified. Additionally, appropriate control mechanisms to trigger a service migration will be defined.

2.3.2. Testbed description

2.3.2.1. Testbed for P2P Services

In order to perform an extensive functional and quantitative evaluation of the service-level resilience with P2P, a large-scale testbed would be required.

Software: TUM will use a self-implemented prototype for a solution for highly-reliable signaling for VoIP, which is called Cooperative SIP (CoSIP) [Fessi07]. The prototype is implemented in Python programming language. It currently makes use of an implementation of the Kademlia DHT algorithm called "Entangled"[1], (also in Python). The implementation will include a control framework for establishing (and re-establishing) a large-scale distributed testbed and collecting measurement data. For example tools such as "parallel ssh" will be used in order to automate the setup procedure of the testbed. A lightweight web server will be running on each P2P node in order to better monitor the system.

Testbed Platform: TUM has been working on initial experiments with PlanetLab. Current experiments involve a number of peers up to 500. Future experiments with several thousands of peers or even more would be useful for better evaluation. The possibility for evaluation with mobile nodes would open new opportunities for experimentation. The concept of "slices" used by PlanetLab has been sufficient since we were able to install Python on it and run our P2P software.

Hardware: Nodes equipped with a minimum of 2 GBytes RAM and an up-to-date CPU power would be useful[2].

Bandwidth: The P2P system is used for signaling only which is a small portion of the traffic. The bandwidth requirements are rather minimal.

Access policies: TUM is currently already running initial experiments on PlanetLab. The access is provided to students at TUM, based on the access policies of PlanetLab. (An account is

[1] http://entangled.sourceforge.net/

[2] Our experiments with PlanetLab up to now showed that we had problems with PlanetLab nodes with 1GBytes of RAM.

needed to login to an assigned PlanetLab “slice”). The experiments involve emulating phone calls with SIP UAs that are running on the PlanetLab nodes.

It might be possible to provide a demo where users can perform an “echo” call to one of these nodes, or use the running testbed on PlanetLab to initiate phone calls peer-to-peer between each other. However, it is not clear yet, how such facilities could contribute to the evaluation of the anticipated resilience goals.

2.3.2.2. Testbed for virtualization

In order to evaluate the requirements and options of Virtual Machine migration, a large-scale testbed consisting of virtualized servers in different subnets is required.

Software: UP research currently focuses on system virtualization, using the XEN hypervisor with additional management primitives developed in the JAVA programming language.

Hardware: The hardware should be able to host several Xen virtual machines per physical host.

Bandwidth: Since Virtual Machine migration includes the movement of large amounts of data (Virtual Machine images), a high amount of available bandwidth is necessary in a testbed.

Testbed Platform: A large-scale testbed with multiple subnets, consisting of machines virtualized with XEN would be useful to evaluate the implications of service mobility. In order to carry out tests, it would be necessary to get full root access to the machines in question, being fully able to create and move any number of Virtual Machines. The current access policies in PlanetLab, for example, would not be sufficient to perform these experiments.

2.4. Experimentation scenario “Smart Environments”

Home Plug-n-Play enriched communications is the theme of this experimentation scenario[3]. The objective is to exploit both IP Multimedia Subsystem (IMS) and Universal Plug and Play protocols in order to establish rich communication involving several audio-video streams between subscribers who have at their disposal a huge heterogeneity of communication devices. The interoperability of these LAN and WAN technologies, resp. UPnP and IMS, ensures the distribution of new services provided by NGN architectures over all user's terminals interconnected in a LAN, with features such as synchronized switching from one terminal to another and end-to-end transmission management between residential terminals and exterior ones.

Today's commercial offers for home terminals are diverse: multimedia cell phones (integrated in the domestic network through WiFi); digital TV decoder (VoD, visiophony) using Ethernet, WiFi or PLC, PCs (media center, VoIP); residential gateways holding a privileged position for shared services (e.g., data storage, peripheral equipments, distributed services); nomadic terminals (pocket computers, audio/video walkman); and others.

The service offers possibilities include:

- subscription to multiple operators and content visualization using the TV set; automatic backup during a service transmission or reception; private/commercial contents shared in home or exterior
- content continuity from one terminal to another (indoor/outdoor or outdoor/indoor);
- service integration in a multi-(service)operator context, e.g., display of the incoming call (phone service of Operator1) on the screen showing a video received from Operator2;
- service distribution over several terminals

2.4.1. Scenario description

Two content-sharing scenarios for interpersonal communications have been studied up to now.

In the first one, the residential user A, while pursuing a conversation with another residential user B, decides to choose some video content and to send it to B. B will have the choice of any of his home terminals with audio-video features to view this content.

In the second scenario, during a voice communication with A, B initiates the start of his camera in order to pursue the communication in a video mode. A will then select any UPnP device at home in response to A's initiative, allowing the non-stop communication flux to go on.

Different resilience challenges can be considered: customer relying on secured means to access to his/her different terminals (no video contents, or personal data, piracy); access to customized audio/video services according to the user's profile (preferences, rights); service, i.e., connection, availability.

As WP4 will only start in 2010, more scenarios, in the framework of service offers sketched previously, will be studied then, and new resilience challenges (operator needing to verify and maintain the identification of user's rights between networks, etc.) will arise. Together with the

[3] The experimentation theme is inspired from Systermin@l, a project of the competitiveness pole "Images & Networks"

challenges described earlier, all these threats will be analyzed and studied in the light of WP1, WP2, and WP3 results and recommendations.

2.4.2. Testbed description

Two home environments are used for testing the usage scenarios. The connection between WAN and LANs is provided by a residential gateway, a modem-router allowing Triple-Play offer through an ADSL line. Note that the testbed targets to maximize the coverage of home application domain and protocol heterogeneity: IP plug & play protocols (e.g., UPnP, Jini), home automation field buses (such as ZigBee), personal area protocol (Bluetooth, USB, IrDA, WiFi, Ethernet), inter-personal communication (SIP-IMS, Jabber/XMPP).

On the user A side, a residential gateway is connected to different communication solutions (phone, laptop, desktop computer). In addition to the same hardware environment, the user B includes a camera for running the second scenario.

As of today, this testbed is accessible only internally. A larger scale testbed has not been considered up to now.

3. With respect to federation

As already mentioned in Section 2, ResumeNet is largely relying on in-house testbeds to carry out its experimentation activities. This is the case with three out of the four experimental scenarios in the project. On the contrary, the third experimental case will need a larger facility in the scale of PlanetLab.

In the longer term, leveraging the other three experimental scenarios to use larger experimental facilities is an opportunity, which is only viewed positively in the project. In general, federation of testbeds seems to make sense when interested in scalability and heterogeneity and these two aspects are relevant when assessing resilience. At this stage of research, however, simple scenarios with a few nodes might still be complex enough to fully understand the interactions and implications of all the involved mechanisms and therefore difficult to project on requirements for federation.

The governance model of the in-house testbeds is rather ad hoc, the responsibility being with the network administrators or nominated staff members in each lab. They are used mainly internally within research groups for research and academic purposes. There are usually two-three levels of access, e.g., {default, user, admin} and remote experiment execution is enabled via password sharing with the remote party.

The testbeds could eventually become made available to the FIRE and the broader research community.

References

(Aad04) Aad, I., Hubaux, J., and Knightly, E. W. 2004. Denial of service resilience in ad hoc networks. In Proceedings of the 10th Annual international Conference on Mobile Computing and Networking (Philadelphia, PA, USA, September 26 - October 01, 2004). MobiCom '04. ACM, New York, NY, pp. 202-215

(Ande03) Anderegg, L. and Eidenbenz, S. 2003. Ad hoc-VCG: a truthful and cost-efficient routing protocol for mobile ad hoc networks with selfish agents. In Proceedings of the 9th Annual international Conference on Mobile Computing and Networking (San Diego, CA, USA, September 14 - 19, 2003). MobiCom '03. ACM, New York, NY, pp. 245-259

(tik) TIK-Net testbed, http://tiknet.ee.ethz.ch/doku.php

(Zhon05) Zhong, S., Li, L., Liu, Y. G., and Yang, Y. 2005. On designing incentive-compatible routing and forwarding protocols in wireless ad-hoc networks: an integrated approach using game theoretical and cryptographic techniques. In Proceedings of the 11th Annual international Conference on Mobile Computing and Networking (Cologne, Germany, August 28 - September 02, 2005). MobiCom '05. ACM, New York, NY, 117-131

[Fessi07] Fessi A. et al., “CoSIP – a hybrid architecture for reliable and secure SIP services”, PIK - Praxis der Informationsverarbeitung und Kommunikation. Volume 30, Issue 4, Pages 206–212, Dezember 2007

[ResD1.1] Understanding challenges and their impact on network resilience. ResumeNet Deliverable D1.1, March 2009

[webmin] http://webmin.com