Nicola Blefari-Melazzi
Giuseppe Bianchi
Luca Salgarelli *Editors*

Trustworthy Internet

Springer

Trustworthy Internet

Nicola Blefari-Melazzi
Giuseppe Bianchi · Luca Salgarelli
Editors

Trustworthy Internet

Springer

Editors
Nicola Blefari-Melazzi
Department of Electronic Engineering
University of Rome "Tor Vergata"
Via del Politecnico 1
00133 Rome
Italy
e-mail: blefari@uniroma2.it

Giuseppe Bianchi
Department of Electronic Engineering
University of Rome "Tor Vergata"
Via del Politecnico 1
00133 Rome
Italy
e-mail: giuseppe.bianchi@uniroma2.it

Luca Salgarelli
Department of Information Engineering
University of Brescia
Via Branze 38
25123 Brescia
Italy
e-mail: luca.salgarelli@ing.unibs.it

Selected contributions from the 2010 International Tyrrhenian Workshop on Digital Communication

ISBN 978-88-470-1817-4
e-ISBN 978-88-470-1818-1

DOI 10.1007/978-88-470-1818-1

Springer Milan Heidelberg Dordrecht London New York

Library of Congress Control Number: 2010937688

Cover design: eStudio Calamar S.L.

Printed on acid-free paper, June 2011

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

This book aims at providing a snapshot of the various facets (and underlying challenges) that contribute to make the Internet trustworthy. It originated from the 21st International Tyrrhenian Workshop on Digital Communications, an event traditionally organized by CNIT, the Italian inter-university consortium for telecommunication research. The workshop is devoted each year to a specific theme in the area of digital communications and its applications, and the theme selected for this year's edition was "Trustworthy Internet" (tyrr2010.cnit.it), which the European Union's research program defines as a network characterized by the following properties: secure, reliable and resilient to attacks and operational failures; guaranteeing quality of service; protecting user data; ensuring privacy and providing usable and trusted tools to support the users in their security management.

The workshop program comprised both peer-reviewed papers and invited contributions from leading experts in the field. It provided a lively and exciting discussion on the technical challenges and issues involved in the trustworthy rethinking of the Future Internet. This book aims at offering the findings and discussions carried out during the workshop days to a broader audience. For this purpose, the book contains a selection of the works presented at the 21st International Tyrrhenian Workshop on Digital Communications. Each contribution has been extended with background material on the specific topic in question. Moreover, the sections account for the supplementary insights gathered from the workshop discussions, and, when appropriate, also include additional technical results.

We thank the persons who have been instrumental in making our workshop a success, and who have permitted this book to exist: the contributors for their high quality works and their supplementary effort spent in making their findings accessible to the broader audience expected for a book, the reviewers of the papers for their thorough work; the speakers, the session chairs and the audience for attending and making the sessions a lively and fruitful environment; the organizers of the invited sessions, who also contributed sustantially to the Introduction and organisation of this book: Dr. Thorsten Strufe, Dr. Sonja Buchegger, Prof. Fabio

Massacci, Dr. Federica Paci, Dr. Saverio Niccolini, Dr. Sandra Tartarelli, Dr. Leonardo Chiariglione; a special thank to Leonardo who gave also the keynote speech: "*Catering to the sustainable evolution of the digital media foundations*".

We are also grateful to our main sponsor Confcommercio, and its President, Dr. Carlo Sangalli, for the generous support in such a difficult time. Indeed, we are proud that such an highly recognized Italian institution has tangibly demonstrated a strong interest in sustaining research efforts targeting a more secure, efficient, and trusted Internet environment.

We thank the research projects belonging to the 7th Framework Programme of the European Union in which CNIT is involved, and working in areas relevant to the workshop theme, for providing technical support: CONVERGENCE, DEMONS, FLAVIA, MOMENT, PERIMETER, PRISM, with a special thank to the CONVERGENCE project that supported the workshop also with additional funding.

And we finally thank the CNIT president, Prof. Giancarlo Prati, the CNIT Director, Prof. Silvano Pupolin, and the CNIT executive board, for having given us the opportunity to organize this workshop on the very timely and important topic of trustworthy Internet.

N. Blefari Melazzi
G. Bianchi
L. Salgarelli

Introduction

The term "trustworthy" is explicitly defined by the European Community's FP7 research program as: secure, reliable and resilient to attacks and operational failures; guaranteeing quality of service; protecting user data; ensuring privacy and providing usable and trusted tools to support the user in his security management. As such, the Trustworthy Internet not only has to include mechanisms, architectures and networking infrastructures that intrinsically provide basic security guarantees, but it also has to ensure users, service providers and application providers alike that their requirements in terms of Quality of Experience, manageability and efficiency are fully met. Providing such combined guarantees in a rapidly evolving, complex infrastructure such as the Internet requires solving challenging issues that encompass many fields of theoretical and applied information engineering, spanning all levels of the protocol stack, ranging from finding new intrinsically secure transmission systems, to radically novel routing models, to new architectures for data dissemination and for interconnecting an unprecedented number of devices and appliances.

This book aims at representing a view of the state of the "Trustworthy Internet" as we enter the second decade of our century. The material included in this book originates from a workshop, organized in September 2010, and specifically dedicated to the several aspects which contribute to make the today's and tomorrow's Internet trustworthy. The workshop comprised either invited contributions from renowned researchers with complementary expertise, as well as independent, peer-reviewed contributions stimulated through an open call for papers. The book includes a selected subset of the workshop papers. Each contribution has been edited and extended after the workshop, taking into account the discussions carried out during the event, incorporating when appropriate additional technical material, and, perhaps most importantly, complementing the specific technical aspects presented with background material devised to more comprehensively introduce the reader to the specific topic of trustworthiness tackled.

For the reader's convenience, the selected scientific contributions are further grouped in homogeneous chapters that present scholarly visions of many of the aspects defined by the term *trustworthy*. As such, it is our hope that this material can serve scholars, practitioners and students alike as a guide not only to help understanding the state of current research in this topics, but can also represent a guided look at its medium term future.

The material in *Part I* focuses on the research that aims at imagining how *the future Internet will support trustworthiness*. Although predicting the future is always risky business, it is clear that the Internet is changing quickly, starting from the communication paradigms that are shifting rapidly from the client-server model to more advanced ones, such as publish-subscribe, overlay and community-oriented networking. These new communication models not only pose interesting and novel challenges to operational aspects of networking, such as efficiently supporting scalable QoS requirements and dealing with fast routing in heterogeneous networks, but also stimulate the research of new mechanisms to support the security and trust of such new infrastructures. Contributions in this chapter revolve around the last aspect.

Chapter 1 discusses the security and privacy properties of a new prospective Internet architecture based on the publish-subscribe paradigm. Indeed, the information oriented nature of publish-subscribe, the decoupling it offers between information providers and information consumers, its location-identity split properties, and the clean-slate ambition of replacing the actual Internet protocol stack rather than complementing it, calls for specifically tailored security and privacy features, and opens up a number of key challenges which are duly presented and discussed in this section. Chapter 2 looks at security aspects in programmable, high-performance routers, which are basic building blocks for network resource virtualization and dynamic service configuration in future Internet scenarios. Router programmability requires a careful scrutiny of the security aspects which arise when either customers as well as service providers access the routing nodes. This chapter specifically discusses promising solutions to the support of secure services through the separation of control and forwarding elements.

Identity management, i.e. governing the fine relationship between privacy and trust, in overlay networks is explored by the last two chapters of this Part. Chapter 3 looks at the problem of how to provide secure pseudonymous access from the perspective of an overlay data dissemination network. Chapter 4 attacks the problem of building a wide-area, application-layer identity management for a Semantic Web architecture aimed at supporting seamless data interoperability to facilitate sharing and exploitation of distributed and heterogeneous user profile information.

Part II overviews specific trustworthiness issues, tools, methodologies and their applications to different aspects of the current Internet architecture. Chapter 5 surveys context privacy protection approaches and solutions for the Internet of Things, which represents a challenging issue given the scalability requirements of the sensory world. This chapter specifically deals with the various facets of privacy issues, here categorized as Application privacy, concerning the protection of the network application being run, Correspondent privacy, concerning identity protection of the other peer in a communication, and Location privacy, concerning the protection of the user location information. Chapter 6 first discusses emerging and novel ways for providing security directly at the physical layer, and then provides a thorough overview of a new modulation technique, able to encrypt the radio signal without any a priori common secret between the two nodes. Here, the information is modulated, at physical layer, by the thermal noise experienced by the link between two terminals, and a loop scheme is designed for the recovery of mutual information. Chapter 7 introduces the reader to the important area of secure multiparty computation. While showing a special case of privacy-preserving computation, the chapter permits the reader to get acquainted with the application of homomorphic encryption to the computation of vector multiplications, scalar products, renormalizations, etc. The tools and methodology here described have a wide applicability in many fields, and show that non elementary operations can be performed without impairing the privacy of the data. Chapter 8 presents a fully distributed security framework for Peer-to-peer Voice over IP, which does not rely on a centralized PKI, but leverages and adapts brilliant approaches, such as ZRTP from Zimmermann, which are deemed to significantly influence future Internet community/distributed services.

Finally, the last two contributions tackle, for different contexts (4G femto cells and ZigBee sensor networks), the issue of spontaneous, autonomic configuration. Specifically, Chapter 9 takes a look at novel ways of dynamically sharing a scarce resource, i.e., the spectrum, so as to provide QoS guarantees in wireless environments with high density of low coverage wireless femto-cells, such as the ones typical of home or community networks. The solutions presented do not require coordination, and may be applied also to the futuristic (and here shown to bring performance advantages) scenario of *different* operators sharing a same subset of frequency bands. Finally, Chapter 10 uses network monitoring information for automatically re-configuring autonomic, self-organizing wireless personal area networks in an energy-aware and robust manner.

Part III deals with the topics related to Online Social Networks, which are increasingly popular with Internet users. They attract ever-growing groups of participants, who in addition spend increasing amounts of time on these sites. Facebook, the currently largest Online Social Network, alone claims an active user base of over 550 million individuals. Both with respect to the number of available pages and served bandwidth it constantly ranks among the top three web sites worldwide (see for example http://www.hitwise.com). Considering the time the users spend on a web brand's pages, it has surpassed the former top competitor, Google and its affiliated services, with users spending almost three times as long on Facebook's, compared to Google and its affiliates' pages (over seven hours per month on Facebook compared to just under 2.5 hours on Google's sites in January 2010, see "Top web sites and Brands" at The Nielsen Company). In addition to user time spent on online social networks, there is an increase in personal information revealed online. Users of Online Social Networks upon registration create a profile describing themselves and they henceforth can (and are expected to) connect their profile with links to the profiles of their friends, family, and acquaintances. Online Social Networks thus represent a subset of the relations their users entertain in the real world. They additionally offer rich functionality to publish and share various content, to communicate by different means, and to collaborate and play games together. Online Social Networks by nature contain a wealth of personally identifiable information, since each published item and each message inherently is linked to the personal profile identifying the publisher or sender.

Collaborative games in Online Social Networks frequently leverage a player's social environment: players do not play alone or in predetermined groups, but the users added in the contact list of the player automatically act as partners or supporters in the game. This causes a change in the perception of friendship and trust, since contacts are made for the sole reason of two strangers wanting to support each other in one of the games. Chapter 11 deals with this issue and the consequences for trust and privacy from the security perspective. The highly accurate and authentic identification of users comes with both beneficial and adverse consequences. Recognition and reputation now can be attributed to individuals and have a direct effect on real life, as opposed to the reputation of anonymous clients, devices, or personae before. Profiling and targeting users is performed directly on individuals as well. This conflict is the topic of Chapter 12, where the field of participatory sensing is described, and new methods are proposed to leverage the positive social effects, while using abstraction to mitigate the consequences for the privacy of the participants. The voluntarily self-maintained and easily exploitable database of personally identifiable information that Online Social Networks represent is of high value to third parties. The possibility to create highly detailed and perfectly identified behavioral profiles of individuals is not only very attractive to the advertisement industry, but also to miscreants with various adverse motivations. This situation is the focus of Chapter 13, where several attacks are described that are possible and additionally exceptionally successful due to the social and seemingly trustable environment that Online Social Networks provide. Finally, Chapter 14 motivates the decentralization of Online Social Networks for the purpose of load balancing and, more importantly, for the protection of the users' privacy. This chapter subsequently presents a survey of the current proposals for decentralized Online Social Networks and classifies them by characteristic properties.

New technologies and tools including Web services, blogs, social network sites, mash ups, wikis have dramatically changed the way users communicate, collaborate and share personal information on the Internet.

Part IV groups contributions that present the various facets of trustworthyness in Web 2.0 platforms. The main focus here is on finding mechanisms able to guarantee that the composition of services satisfies certain security properties, while maintaining the system flexible and scalable. Web service technology allows users to pull together content from different sources and services to build a new service. Such technology has also facilitated the collaboration intra- and inter-organizations by making accessible organizational business processes through the Web. Blogs, wikis and social networking sites allow users to share new content and to collaborate and communicate with others. The power of Web 2.0 brings up a number of serious security issues ranging from identity management and reputation, privacy protection and anonymous usage, access control to content and services to integrity protection of composite services. Securing Web 2.0 applications is challenging because such applications are targeted towards making people, information, and resources available to all who need it, and to adapt swiftly to changes and evolutions in software platform and usage models. In contrast, information security seeks to restrict access only to those with proper authorization.

Chapter 15 presents a broad survey of the many challenges faced by the mechanisms used to protect social networks, where content is dynamically updated and accessed by millions of users on a daily basis. In this scenario we must be able to enforce access control policies on social network users' profiles that meet the privacy preferences of the users. This chapter discusses the various alternatives based on risk, trust and other metrics and the related trade-off. The other three chapters discuss more in details the impact of the evolution of data types, user roles and usage models on the security solutions for Web 2.0 applications. Chapter 16 deals with the problem of managing the temporal evolution of authorizations in which users' attributes can change over time and in particular can change after access has been granted. The presented model is an extension of the (now) classical usage control model by Park and Sandhu: it guarantees that access control policies are enforced not only when a resource is accessed but also during its usage. Chapter 17 investigates the challenges of selecting services that meet specific privacy and security properties to build complex composite applications such as business processes. In particular, the problem of dynamic certification of composite services is analyzed, starting from the composition of the certificates of the composite services. Chapter 18 analyzes the same problem but from the perspective of usability. While it is true that users can change attributes as time goes by, the goals of an organization or the objectives of the individuals in a social or collaborative environments are more stable. We would like to be able to achieve them even in presence of changed security circumstances. This chapter discusses the problem of dynamic resiliency of business processes, that is how to guarantee that business processes can still be executed when users authorizations change and that overall business goals can be achieved by suitably re-distributing the authorizations to the appropriate users.

Part V focuses on various facets of network monitoring with a particular attention to trustworthiness applications. Indeed, network monitoring is an area which is expected to face a number of radical changes in the near future. Security threats, which once represented mere "hacking" or exploitation of hosts for little more than curiosity or vanity, have given way to sophisticated criminal operations that exploit vulnerabilities in network devices and end systems to take over large numbers of nodes, arranging them into botnets, for spamming, phishing, extortion via distributed denial of service attacks, and personal information theft, threatening end-user privacy and the importance of "information as an asset". To make matters more challenging, there is an ongoing trend towards bringing Internet technologies into every end device. A very high number of nodes (e.g., every phone on a network, every device in a household, etc.) are now becoming IP-enabled, more intelligent and more complex, providing new hosts for botnets and bridging legacy and IP-based systems. This poses serious challenges to the operators, as the problems multiply while the requirement on trustworthiness remains unchanged. In addition, monitoring systems must be scalable. Internet traffic growth is reaching volumes previously unimagined. Annual global IP traffic volume nearly doubles every two years, and will exceed half a zettabyte (5 × 1020 bytes) by 2012. This growth poses severe challenges to the Internet scalability, and calls for a decentralized and scalable monitoring infrastructure. Furthermore, monitoring infrastructures must take privacy into account. Indeed, traffic monitoring activities, especially at higher layers of the network stack, pose a serious risk to the privacy of the individual, since they may result in tracking the personal activities of the end users without their knowledge. Monitoring activities undertaken without transparency or accountability with respect to data processing, i.e. without privacy-awareness, lead to a loss of trust in the network as a whole. As a result, care must be taken that privacy concerns are addressed, and that privacy rights and data protection laws are not violated. And, finally, network monitoring applications are moving up the layers. Traditional monitoring applications are no more focused on the analysis of IP-level traffic but instead tend to include more and more application-specific information and semantics in order to reach their objectives.

Even if the six contributions that comprise this Part do not pretend to completely cover the several challenges above discussed, nevertheless they provide a valid picture of the trends and solutions that may characterize future-generation network monitoring infrastructures and approaches. Chapter 19 presents the challenges and solutions for building extensible, programmable and high performance network monitoring solutions inspecting application layers. Especially programmability (which is also closely considered and advocated in Chapter 22) appears to be a key requirement to permit future monitoring infrastructures to rapidly and flexibly accommodate the continuously evolving needs posed by the emergence of new threats and new application-layer monitoring requirements. Chapter 20 clearly exemplifies how monitoring tasks are no more confined to the network layer, by showing techniques for detecting frauds and misuse of telephone services and combat spam over Internet telephony. Chapter 21 uses application layer semantics for monitoring and reducing overload situations in IP-based telephone networks using the SIP protocol. Chapter 22 brings about, again, the need for programmability from a different perspective, i.e., by specifically targeting the design issues and challenges behind programmable monitoring probes, and the possibility to support processing and filtering means directly on the probe itself, thus moving from a traditional centralized vision to a distributed and highly scalable modern “in-network” monitoring vision, where monitoring tasks are directly supported inside the network itself. Chapter 23 can be considered as a concrete example of such an in-network processing and filtering vision, as it discusses the rationale for offloading central intrusion detection systems by implementing approximate intrusion detection rules directly on probes, and presents a proof-of-concept hardware implementation of a probe capable of supporting rules from the topmost known SNORT intrusion detection system. Finally, Chapter 24 addresses the very important and timely issue of how to protect network customers' privacy without compromising information usability for monitoring purposes. This chapter specifically investigates the fundamental principles and requirements for a privacy-aware ontological model in the semantic domain of monitoring-data management and exchange. It proposes a rule-based approach for specifying and automatically enforcing the appropriate privacy policies, and advocates a clean separation between data models and security semantics.

Part VI analyzes the issues and tradeoffs of bringing trustworthiness to digital content and its distribution. The Internet was originally conceived as an "Internet of Hosts", whose underlying protocols were designed to support exchange of simple unstructured information between well-identified nodes. Today, by contrast, it is becoming an Internet of Things (devices and appliances associated with their own IP address), an Internet of Services (in which users in different localities access different functionalities on different hosts), an Internet of Media (shared and managed across different networks) and an Internet of People (boosted by the explosion of social networking and the emergence of the Web 2.0 paradigm). In these "new Internets", the key elements are no longer "hosts" but data and services (or content). As one author put it, "People value the Internet for what content it contains, but communication is still in terms of *where*". In other words, what we are observing is a shift from "host-centric networking" to "content-centric" or "data-centric" networking.

This shift imposes new requirements on middleware, on the underlying networking functionality and on the way content is codified, formatted, described and exchanged in the network. Regarding the organization of content, several of these needs are addressed e.g. by existing MPEG standards. For instance MPEG-21 already defines standard ways of providing meta-information and standard ways of describing the content and structure of complex "Digital Items". However, there is the need of extending the ability to manage and trade "classical" media (e.g. video, music) digital objects to a broader range of digital objects, including descriptors for Real World Objects (RWO), services and people, meeting at the same time new requirements coming from such extended environment. Regarding middleware, there is the need of providing APIs to dynamically define and encapsulate new classes of content, and related meta-information, to create packages of different classes of information resource, to guarantee their security and privacy and integrity, to name them, to support semantic interpretation of metadata and tags, to search for them, filter them, read and write their attributes and content, adapt them for use on different machines, copy them, test their validity and efficiently synchronize them across multiple machines. The MPEG-M emerging standard, an extension of the former MXM MPEG platform, is addressing a significant part of these issues. As regards networking functionality, content-centric architectures are being proposed, where the network layer directly provides users with contents, instead of providing communication channels between hosts.

Chapter 25 presents an outlook on the definition and implementation of distributed architectures that enable the development of distributed multi-media applications on top of them, while offering Digital Rights Management (DRM) features. Chapter 26 develops from the consideration that innovative networks must be aware of which content is actually transported and introduces Scalable Video Coding as an important tool for such networks. Finally, Chapter 27 identifies the main functionality of a content-centric network, discusses pros and cons of literature proposals for an innovative, content-centric network layer and draws conclusions stating some general requirements, which a content-centric network layer should satisfy.

Contents

Part I New Visions for a Trustworthy Internet

1 **Publish–Subscribe Internetworking Security Aspects** 3
Nikos Fotiou, Giannis F. Marias and George C. Polyzos

2 **Security Issues in Programmable Routers for Future Internet** . . . 17
Raul Cafini, Walter Cerroni, Carla Raffaelli and Michele Savi

3 **Secure Pseudonymous Access to Overlay Data Dissemination Network** 31
Anna Del Grosso, Marco Listanti, Andrea Baiocchi and Matteo D'Ambrosio

4 **An Overlay Infrastructural Approach for a Web-Wide Trustworthy Identity and Profile Management** 43
Maria Chiara Pettenati, Lucia Ciofi, David Parlanti, Franco Pirri and Dino Giuli

Part II Security, Energy Efficiency, Resilience and Privacy

5 **Context Privacy in the Internet of Things** 61
Laura Galluccio, Alessandro Leonardi, Giacomo Morabito and Sergio Palazzo

6 **Physical Layer Cryptography in Wireless Networks** 75
Lorenzo Mucchi, Luca Simone Ronga and Enrico Del Re

7 **Gram-Schmidt Orthogonalization on Encrypted Vectors** 93
Pierluigi Failla and Mauro Barni

8 **A Peer-to-Peer Secure VoIP Architecture** 105
Simone Cirani, Riccardo Pecori and Luca Veltri

9 **Improving QoS of Femtocells in Multi-operator Environments**. . . 117
Franco Mazzenga, Marco Petracca, Remo Pomposini and Francesco Vatalaro

10 **Autonomic Network Configuration in IEEE 802.15.4: A Standard-Compliant Solution** 129
Francesca Cuomo, Anna Abbagnale and Emanuele Cipollone

Part III Security in Online Social Networks

11 **On the Concept of Trust in Online Social Networks** 143
Henric Johnson, Niklas Lavesson, Haifeng Zhao and Shyhtsun Felix Wu

12 **Participatory Sensing: The Tension Between Social Translucence and Privacy**. 159
Ioannis Krontiris and Nicolas Maisonneuve

13 **A Summary of Two Practical Attacks Against Social Networks** 171
Leyla Bilge, Marco Balduzzi, Davide Balzarotti and Engin Kirda

14 **Decentralized Social Networking Services** 187
Thomas Paul, Sonja Buchegger and Thorsten Strufe

Part IV Secure Collaborative Systems

15 **Access Control, Privacy and Trust in On-line Social Networks: Issues and Solutions** 203
Elena Ferrari

16 **Dynamic Resiliency to Changes**. 213
Fabio Massacci, Federica Paci and Olga Gadyatskaya

17 **Certifying Security and Privacy Properties in the Internet of Services** 221
Marco Anisetti, Claudio A. Ardagna and Ernesto Damiani

18 Time-Continuous Authorization of Network Resources Based on Usage Control 235
Barbara Martini, Paolo Mori, Fabio Martinelli, Aliaksandr Lazouski and Piero Castoldi

Part V Network Monitoring

19 Towards Monitoring Programmability in Future Internet: Challenges and Solutions 249
Luca Deri, Francesco Fusco and Joseph Gasparakis

20 Analyzing Telemarketer Behavior in Massive Telecom Data Records 261
Nico d'Heureuse, Sandra Tartarelli and Saverio Niccolini

21 SIP Overload Control: Where are We Today? 273
Dorgham Sisalem

22 Towards Smarter Probes: In-Network Traffic Capturing and Processing 289
Nicola Bonelli, Andrea Di Pietro, Stefano Giordano, Gregorio Procissi and Fabio Vitucci

23 IDS Rules Adaptation for Packets Pre-filtering in Gbps Line Rates 303
Simone Teofili, Enrico Nobile, Salvatore Pontarelli and Giuseppe Bianchi

24 Introducing Privacy Awareness in Network Monitoring Ontologies 317
Giuseppe Tropea, Georgios V. Lioudakis, Nicola Blefari-Melazzi, Dimitra I. Kaklamani and Iakovos S. Venieris

Part VI Content

25 Rights Management in Architectures for Distributed Multimedia Content Applications 335
Jaime Delgado, Víctor Torres, Silvia Llorente and Eva Rodríguez

26 Scalable Video Coding in Content-Aware Networks: Research Challenges and Open Issues 349
Michael Grafl, Christian Timmerer, Hermann Hellwagner, Daniel Negru, Eugen Borcoci, Daniele Renzi, Anne-Lore Mevel and Alex Chernilov

27 Network Layer Solutions for a Content-Centric Internet 359
Andrea Detti and Nicola Blefari-Melazzi

Part I
New Visions for a Trustworthy Internet

Chapter 1
Publish–Subscribe Internetworking Security Aspects

Nikos Fotiou, Giannis F. Marias and George C. Polyzos

Abstract Publish–Subscribe is a paradigm that is recently receiving increasing attention by the research community, mainly due to its information oriented nature. Although the publish–subscribe paradigm yields significant security advantages over the traditional send–receive one, various security and privacy challenges are raised when it comes to the design of an internetworking architecture that is solely based on this paradigm, such as the Publish Subscribe Internet (Ψ) architecture. Ψ is the main outcome of the Publish–Subscribe Internet Routing Paradigm (PSIRP) project, which was launched with the ambition to develop and evaluate a clean-slate architecture for the future Internet based on the publish–subscribe paradigm. Availability, security, privacy and mobility support are considered as core properties for this new form of internetworking, instead of being provided as add-ons, as in the current Internet. This paper discusses the security and privacy properties of and challenges for publish–subscribe internetworking architectures and specific techniques and solutions developed in PSIRP for Ψ.

Keywords Clean slate · Future Internet · PSIRP · Publish/subscribe · Security

N. Fotiou (✉) · G. F. Marias · G. C. Polyzos
Mobile Multimedia Laboratory, Athens University of Economics and Business,
Patision 76, 104 34 Athens, Greece
e-mail: fotiou@aueb.gr

G. F. Marias
e-mail: marias@aueb.gr

G. C. Polyzos
e-mail: polyzos@aueb.gr

N. Blefari-Melazzi et al. (eds.), *Trustworthy Internet*,
DOI: 10.1007/978-88-470-1818-1_1, © Springer-Verlag Italia Srl 2011

1.1 Introduction

The Publish–Subscribe paradigm has been in the spotlight of recent research efforts. Its information oriented nature, the decoupling it offers between information providers and information consumers as well as its location-identity split properties, have inspired a variety of—mainly overlay—architectures that focus on multicast [6], mobility [15], indirection [29] as well as on caching [16].

Publish–Subscribe architectures are composed of three main components; publishers, subscribers and a network of brokers [8]. Publishers are information providers that 'publish' information (advertisements). Subscribers on the other hand are information consumers that express their interest in specific pieces of information by issuing subscriptions. Brokers are responsible for matching publications with subscriptions and initiate the (information) forwarding process from information providers towards information consumers. The broker, responsible for the publication–subscription matching, is often referenced to as the rendezvous point and, therefore, the network of brokers is usually referred to as the rendezvous network. Publication and subscription operations are decoupled in time and space allowing for the support of mobility as well as anonymization mechanisms. Moreover a publication can be provided by multiple nodes and similar subscriptions can be aggregated, creating opportunities for multicasting and multihoming. Inherently, the publish–subscribe paradigm has many security advantages compared to the commonly used end-to-end, send–receive oriented paradigm.

PSIRP (Publish–Subscribe Internet Routing Paradigm),[1] an EU FP7 funded research effort, has designed, implemented in prototypes, and initially evaluated a clean-slate, information oriented future Internet architecture; we call it the *Publish–Subscribe Internet* (PSI) architecture, Ψ for short. This architecture aims at overcoming most limitations of the current Internet and at emphasizing the role of information as the main building block of the (future) Internet. This new architecture is based on a paradigm completely different from the current one. Ψ is based on pure, through-the-stack application of the Publish–Subscribe paradigm. Moreover by abiding to the Trust-to-Trust (T2T) principle [4], i.e., all functions take place only in trusted points, the Ψ architecture considers security as a building block of its architecture rather than as an 'add-on'. Ψ harvests the security advantages the publish–subscribe paradigm offers, whilst Ψ-specific security mechanisms are also incorporated.

The purpose of this paper is twofold: to give an overview of the security features of and challenges for the publish–subscribe paradigm, as well as to show the additional techniques and mechanisms developed in PSIRP in order to secure the Ψ architecture. The remainder of this paper is organized as follows. Section 1.2 highlights some of the security and privacy challenges that exist in publish–subscribe architectures. Section 1.3 presents the security advantages of the publish–subscribe paradigm. Section 1.4 overviews the Ψ architecture and its specific

[1] http://www.psirp.org

security solutions. Section 1.5 investigates how other, related architectures handle security requirements. Finally, our conclusions as well as ideas for future work are presented in Sect. 1.6.

1.2 Security and Privacy Challenges in Publish–Subscribe Architectures

As previously mentioned, in the publish–subscribe model, producers publish event notifications to announce information availability and consumers subscribe to specific information items to explicitly declare their interest. Matching is achieved through the rendezvous network, which is envisioned as a distributed service that spans over a large number of providers and administrative domains. In the case where one or more matches are provided by the rendezvous service, then a particular sub-graph over the network topology is determined and activated to support a multihomed and multicasted communication service that transports information elements from publisher(s) to subscriber(s).

Security issues and requirements that arise in a global-scale publish–subscribe system have already been extensively addressed. Wang et al. [31] as well as Lagutin et al. [21] have specified security requirements for a publish–subscribe architecture, whereas Wun et al. [33] have identified and classified possible DoS attacks in content-based publish–subscribe systems. Various mechanisms have been developed in order to secure publish–subscribe systems—such as Eventguard [28]—and most of them base their operation on traditional security mechanisms, adapted to the concept of the publish–subscribe paradigm. In this paper we are focusing on security, trust, and privacy requirements focusing on a different level of abstraction and trying to enrich the existing work with recent results for the publish–subscribe paradigm.

In the information level, integrity, authenticity and validity of information are required. Integrity protection methods will ensure that any violation or fabrication of information elements' content will be detectable. Authenticity means that the information that is received by the subscriber is identical with the subscriber's initial request, and it is not forged. Validity means that the information items announced by the publisher, matched with the subscriber's request, and then forwarded to the subscriber are identical. Detecting integrity violation is a task that mainly is based on public key certificates and signatures, and, thus, it requires trusted third parties or bilateral trust (e.g., symmetric secrets, or HMAC key-based approaches). On the other hand, publication and subscription operations might be decoupled in time. Thus, subscribers might never recognize the publishers' identities, or even their certificates. Thus, information integrity verification should be assisted by the rendezvous-network. In order to avoid bottlenecks due to processing or signing every information element, rendezvous nodes might produce sequences of integrity evidences, such as TESLA seeds if a TESLA approach [25] has been adopted between publication end-points and consumers. Verifying authenticity and validity of the information requires a different, reaction-oriented approach, which is based on subscriber's evaluation on the received information. Such an approach will rank published information elements, and recommend the accurate ones, avoiding DoS attacks [11] or spamming [12].

At the application layer, a main security challenge is the design of a mechanism that grants to subscribers the appropriate access privileges to publication announcements. This is akin to making confidential the existence, and not the content, of publications. Assuming that publishers are always privileged to submit events and announcements, the rendezvous network should enforce an access framework that makes the notifications reachable to preferred subsets of subscribers. For application-level access control, such subsets are formed using scopes [9], role-based access control [3, 26], as well as identification and authentication schemes [24]. On the other hand, publication content confidentiality is achieved mainly through encryption. Finally, when a forwarding topology will be deployed to transport information to subscribers, then there is a potentially strong anonymity requirement to unlink the information and the publisher and subscribers among themselves and from the networking attachment and relay points.

From the subscriber's privacy point of view, a central objective is to unlink his identity from his subscription interests, e.g., by supporting anonymous subscriptions. Subscription privacy might rely on an anonymity framework related to trusted proxies (anonymizers) that receive and process the original request, change its time reference, hide the subscriber's identity and obfuscate his network attachment point. This approach might introduce significant delays, but fulfills the demand for strong anonymity support at the network layer. Additionally, such a system should be designed and deployed appropriately to avoid attacks that have been reported on mix-based privacy enhancement approaches, such as traffic analysis, blending and trickle attacks [32].

1.3 Publish–Subscribe Security Features

The publish–subscribe paradigm can be seen as a remedy to the imbalance of power between senders and receivers in the traditional send–receive paradigm. With the original Internet architecture, the network will make a best effort attempt to deliver whatever any sender sends, irrespective of the interest of and no matter the cost for the receiver and the network(s). This imbalance is often accused for the increasing number of (Distributed) Denial of Service (DDoS) attacks, as well as for the emergence of spamming. In publish–subscribe systems there is no information flow as long as the receiver has not expressed interest on a particular piece of information, i.e, the receiver in a publish–subscribe architecture is able to instruct the network which pieces of information shall be delivered to it. Moreover, and even though the model is so powerful so that there can be subscriptions before the corresponding publications have been published, no information is requested

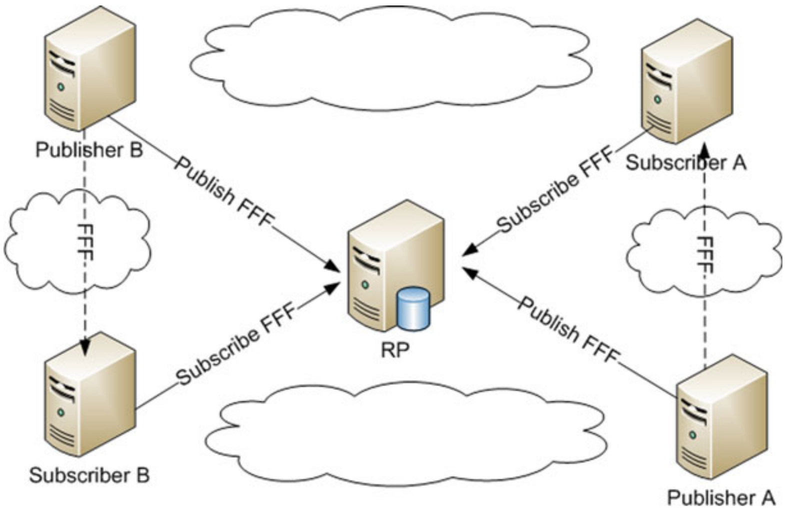

Fig. 1.1 Example of multihoming in a publish–subscribe architecture

from a publisher, unless the publisher has explicitly denoted the availability of that information, i.e., not before the publisher has issued a publication message (for this particular piece of information).

Publication and subscription operations are decoupled in time and space, i.e., they do not have to be synchronized neither do they block each other. Moreover publishers and subscribers do not communicate directly and they can hide their identity as—in general—subscribers are only interested for the information itself rather than on who provides it, and publishers—usually—disseminate publications using multicast so they cannot (and usually should not) be fully aware of the publication's recipients. Therefore, anonymity can be easily achieved in publish–subscribe architectures. Moreover having a point in the network where subscription and publications are matched, effective deployment of access control mechanisms is enabled.

Publish–Subscribe architectures offer great availability. The rendezvous network of a publish–subscribe architecture is usually implemented using a DHT. DHTs provide significant load balancing—usually at the cost of some communication stretch. Moreover in a publish–subscribe architecture multihoming can be easily achieved, as multiple publishers may advertise the same publication to a Rendezvous Point (RP), therefore a RP has a number of options with which it can satisfy a subscription. Figure 1.1 shows an example of multihoming in a publish–subscribe architecture. Publishers A and B, both publish publication FFF. Subscribers A and B subscribe to this publication. For each subscription message the RP knows two publishers that can provide the publication matched, therefore for each subscription message it could choose the publisher that is closer (in any sense) to the respective subscriber, e.g. here, it chooses publisher A to serve subscriber A and publisher B to serve subscriber B.

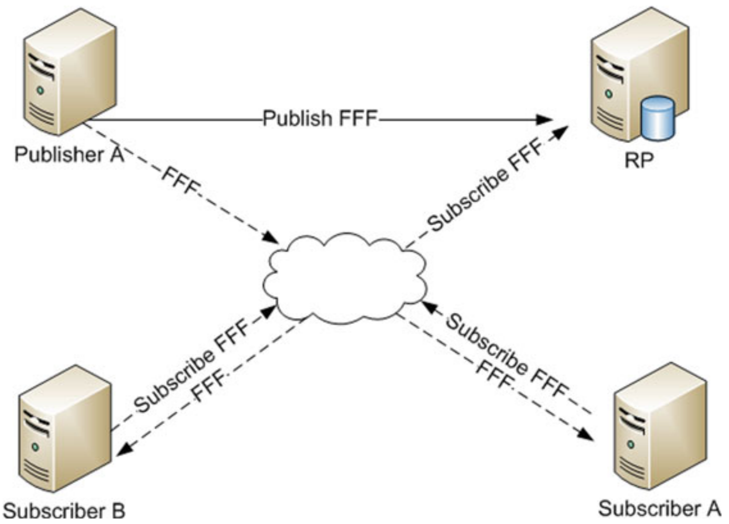

Fig. 1.2 Resource sharing in a publish–subscribe architecture using subscriptions aggregation and multicast

Publish–subscribe architectures allow for subscription aggregation and they create opportunities for multicast to be useful, therefore in these architectures resource sharing can be achieved, leading to greater availability. In Fig. 1.2 both subscribers A and B subscribe to publication FFF. The subscription messages are aggregated within the networks, when following the same path towards the RP. Moreover publisher A forwards a single data flow, which is copied (bifurcated) in an appropriate place in the network in order to serve both subscribers.

1.4 The Ψ Architecture

The core element of the Ψ architecture is information; information is everything and everything is information [30]. In Ψ every piece of information is identified by a unique, flat, self-certified identifier, known as the *rendezvous identifier* (RId). Information is organized in *scopes*. Scopes are physical or logical structures that facilitate the finding as well as access control over a piece or collection of information. A physical scope can be for example a corporate network, whereas a logical scope can be a group of friends in a social network. Scopes can be included within each other, creating a flexible structure. Scopes are identified by a flat identifier known as the *scope identifier* (SId). Each SId is managed by a rendezvous point (RP) which can be a single *rendezvous node* or a large *rendezvous network*.

The publication operation in Ψ involves three steps [10]; initially the SId of the publication scope is identified, then the RId of the publication is created and,

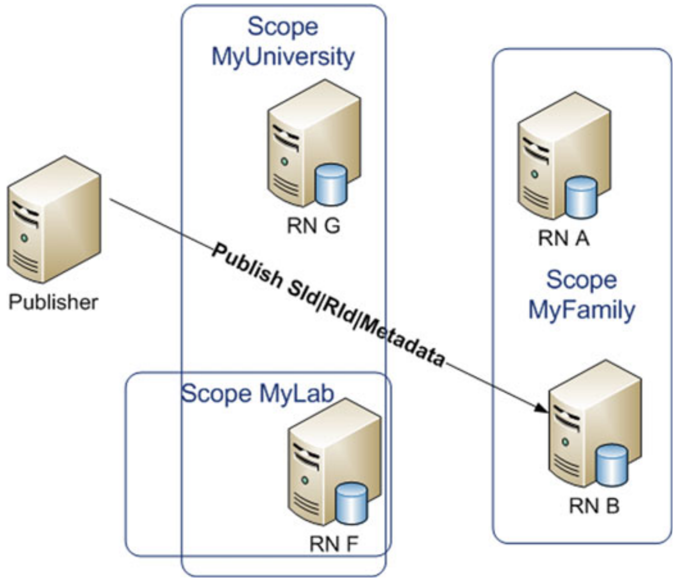

Fig. 1.3 Publication in a Ψ network

finally, the publication is published in i.e. the publication message, including the Rid and Sid, is sent to the RP responsible for handling this SId. The publication message may also contain *metadata*—such as size of the data, encoding and other general information about this publication. Figure 1.3 shows the publication operation in a Ψ network with three scopes; the scope MyUniversity and its sub-scope MyLab and the scope MyFamily. As it can be seen in this figure, a publisher issues a publication to the scope MyFamily. The publication message should contain a scope-unique publication identifier (RId), the MyFamily scope identifier (SId) as well as metadata that describe this publication. The publication message reaches the rendezvous node RN B, which is part of the MyFamily rendezvous network.

The subscription operation involves the identification of the SId and RId of a publication—which can be done, for instance, with the help of a search engine—and the sending of a subscription message. Initially the subscription message will be forwarded to the appropriate scope as all the other scopes are not aware of the publication in question. When the subscription reaches the appropriate scope it will be forwarded to the publication RP. The network is responsible for routing publication and subscription messages towards the RP as well as for forwarding publications from publishers towards subscribers. Figure 1.4 shows the subscription operation. A subscriber subscribes to an already published publication. When the subscription message reaches the appropriate RP, and as long as there is a publication that matches this subscription message, the RP creates a forwarding path, from the publisher towards the subscriber, and instructs the publisher to send

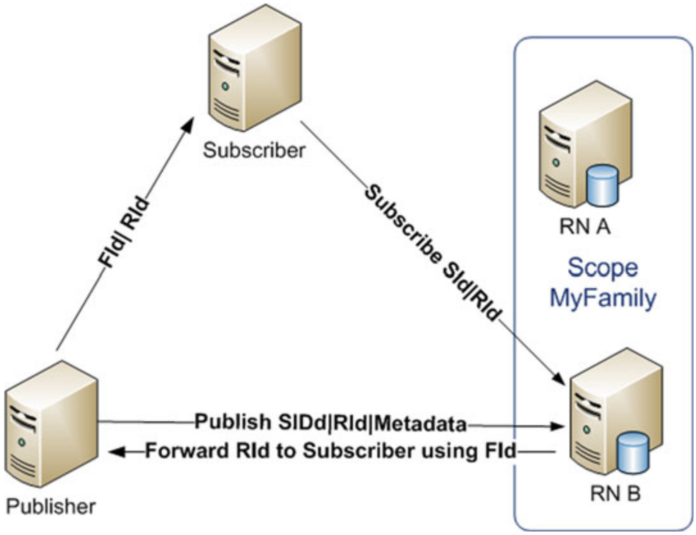

Fig. 1.4 Ψ subscription: initially (e.g.) the publisher issues a publication, then a subscriber, subscribes to this publication and the rendezvous point instructs the publisher to forward this publication to the subscriber

the publication using a specifically created identifier (FId) for this path. A forwarding path is realized through zFilters [14], a Bloom filter based structure that contains the link identifiers that a data packet must traverse in order to reach its destination(s). Ψ uses a slow path for signaling, i.e., publication and subscription messages, and a fast path for data forwarding. Moreover multicast is the preferred delivery method.

1.4.1 Ψ-Specific Security Mechanisms

Security in Ψ plays an important role and trust is at the center of a Ψ declared principle. Security mechanisms are considered at all levels of the architecture. Information in Ψ is transmitted in encrypted packets using the Packet Level Authentication (PLA) technique [19]. PLA is a novel mechanism, applied in Ψ, for protecting the network based on the assumption that per packet public key cryptographic operations are possible at wire speed in high speed networks with the help of new cryptographic algorithms and advances in semiconductor technology. Moreover when applied in wireless environments PLA has been proven to offer significant energy efficiency [20].

As already described Ψ's forwarding mechanism is based on the formation of a Bloom filter—called zFilter—that describes the path that a data packet should follow [14]. The computation of the zFilter is based on the identifiers of the links that compose the data path. These identifiers are dynamically generated every time a zFilter is created, making this way almost impossible for an attacker to create crafted zFilters or link identifiers that will lead to DoS attacks or to information leakage. Forwarding using zFilters is achieved at line speed, leading to excellent performance and scalability. Network attachment in Ψ [17] assures proper user authentication protecting both users from improper configuration as well as the network from (D)DoS attacks that can be caused by malicious users who repeatedly try to attach themselves to a Ψ network.

At the higher layers of the architecture, existing security mechanisms can be used. Nikander and Marias [22] studied the application of existing work on cryptographic protocol analysis in a pure publish–subscribe architecture and found out that, even if networking protocols are revised very drastically, current cryptographic protocol analysis can be applied to a certain extent, with only minor modifications, mostly on the notation side. Moreover, novel trust mechanisms should be considered applied to information ranking [12] rather than ranking end-users.

Ψ security is going to be primarily based on the notion of scopes. Although not yet fully designed and implemented, scopes are expected to control information dissemination as well as to play a significant role in applying access control policies, as well as accounting mechanisms. Scopes are expected to be Ψ's information firewalls.

1.5 Security Aspects of Comparable Internetworking Architectures

CCNx [7] (Content-Centric Networking, now termed Named Data Networking: NDN) is an ongoing research project that investigates the potential of an information-oriented Internet architecture. In contrast to Ψ, CCNx proposes an architecture organized using hierarchical naming [13]. Moreover CCNx uses a broadcast-based mechanism for information location, rather than a rendezvous driven one. CCNx does not rely on flat self-certified identities, it rather uses a scheme that assures the relationship between publications and their identities and it provides validity, provenance, and relevance [27]. In this scheme every publisher is allowed to generate a user-friendly tag label for their publication, which in a next step is incorporated into the body of the publication as a digital signature. This digital signature is generated by applying the publisher's public key over the publication's data and the publication label. When a subscriber receives the publication, and provided that the publisher is reliable, he is able to verify that the publication he received matches its label. On the other hand in case of a malicious publisher that uses forged labels, this publisher can be held accountable for his behavior, as its public key has been used in order to generate the publication's digital signature.

The Data-Oriented Network Architecture (DONA) [18] and Routing on Flat Labels (ROFL) [5] are two pioneering architectures that introduced flat identifiers. DONA aims at replacing DNS with flat self-identifying labels that will enable data location and retrieval. In contrast to Ψ, DONA uses the same path, for information location and forwarding. DONA's main security mechanism is its self-certified naming. DONA names are organized around principals and they are of the form P:L, where P is the cryptographic hash of the principal's public key and L is a label chosen by the principal, who ensures that these names are unique. Every publication is accompanied by a metadata file that includes the principal's public key as well as her digital signature over the publication data. Users in DONA are expected to learn a publications' name using external, reliable mechanisms. In order to defend against DoS attacks, DONA relies on IP-level mechanisms, as well as on the limits that providers will pose on users' publications and subscriptions. Finally DONA assumes the existence of third trusted parties for public key status retrieval and revocation.

ROFL creates an internetworking architecture in which routing takes place solely based on the data—flat—identifiers. In ROFL there is no information hierarchy, as there is in Ψ (with the usage of scopes) and DONA. ROFL security is also based on self certified identities. In ROFL, in every network node, i.e., router or host, a unique ID is assigned, which is tied to a public–private key pair. This key pair is used to sign-verify every packet that traverses the system. ROFL secures its routing infrastructure by using the so-called *filtering* and *capabilities* techniques. With *filtering*, every host can control its reachability and therefore filter out malicious hosts. With *capabilities* the architecture is able to perform fine-grained access control. Whenever a (legitimate) host requests the creation of a network path, a *capability* token is provided, which proves that the host has the proper access control credentials for this path. *Capability* is a cryptographic token designating that a particular source (with its own unique object identifier) is allowed to contact the destination.

The Internet Indirection Infrastructure (i3) [29] and the Host Identity Protocol (HIP) [2] are two rendezvous-based overlay solutions that aim at supporting mobility, multicast and multihoming. Ψ's rendezvous and topology processes use similar concepts, at all levels of the architecture.

i3 implements an IP overlay network that replaces the point-to-point communication model with a rendezvous-based paradigm. In i3 *sources* (akin to Ψ publishers) send packets to a logical identifier, whereas *receivers* (akin to Ψ subscribers) express interest in packets by inserting a trigger into the network. A distributed lookup service is responsible for matching triggers with packets and an overlay network of i3 nodes is responsible for forwarding packets. An i3's extension, known as the *Secure-i*3 [1], further enhances the security of the proposed architecture by allowing hosts to hide their IP address as well as to defend against DoS attacks without introducing new vulnerabilities. IP address hiding is accomplished with the usage of the so-called *private IDs*; when an end-host issues a new trigger, instead of using its real IP address, it uses the public ID of an i3 (reliable) node that acts as the end-host's representative. The public ID of this i3 node is the private ID of the end-host. Even if the representative node removes its public ID it will not affect the already established end-host's connections. Every node in i3 may have multiple public IDs. In case of DoS attacks a node may remove all of its public IDs to eliminate the attack, or remove some of them in order to mitigate the attack. Moreover, puzzles can be used as a countermeasure against DoS attacks; before a suspicious host is allowed to send a new packet, it is requested to solve a cryptographic puzzle. Finally, hosts in i3 can manipulate the path that a packet should follow in order to reach them, this way they are able to circumvent parts of the network that are under attack.

HIP introduces a new layer that decouples host identity from location identity in the internetwork stack, between the IP layer and the transport layer. When HIP is used, the applications no longer connect to IP addresses, but to separate *Host Identifiers*. A Host Identifier is a cryptographic hash of the host's public key, which in turn, is used for securing communication between hosts. The resolution from a Host Identifier to an IP address can be achieved either by using a DNS-like mechanisms or a DHT. *Host Identity Indirection Infrastructure* (Hi3) [23] is the secured version of the HIP protocol, which utilizes Secure-i3's rendezvous principles. Secure-i3 is used in order to perform Host Identifier to IP address resolution, whereas IPSec is used for the rest of the communication between hosts.

1.6 Conclusions and Future Work

The Publish–Subscribe paradigm achieves a significant shift from the current end-host driven internetworking towards an information oriented Internet architecture. This paradigm offers significant security advantages, including greater availability and enhanced privacy. The opportunities for multicast, mobility support and caching, as well as, the decoupling it offers between the communicating parties, make the publish–subscribe paradigm a strong candidate for a future internetworking architecture. Nevertheless various security and privacy challenges remain and further research is needed in order to identify and tackle them. Towards this direction the PSIRP project has created the so-called Ψ architecture; a clean slate Internet architecture that is based on the publish–subscribe paradigm. The Ψ architecture demonstrates the significant capabilities of this paradigm and through the development of Ψ-specific security mechanisms shows the road towards a secure future internetworking architecture.

The research in this field is a very active ongoing effort. Various research projects around the world investigate the potential of new internetworking architectures based on the publish–subscribe paradigm—or other similar ones. Security remains in the spotlight of all these research efforts. As far as the Ψ architecture is concerned, its research and development continues during the EU FP7 PURSUIT[2]

[2] http://www.fp7-pursuit.eu/