17-03-2015

Open Call Deliverable OCM-DS1.1 Final report on IRINA and software prototype (IRINA)

Open Call Deliverable OCM-DS1.1

Grant Agreement No.:	605243
Activity:	NA1
Task Item:	10
Nature of Deliverable:	R (Report)
Dissemination Level:	PU (Public)
Lead Partner:	iMinds
Document Code:	GN3PLUS14-1294-45
Authors:	Dimitri Staessens

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7 2007–2013) under Grant Agreement No. 605243 (GN3plus).

Abstract

This document provides a summary of the activities and results obtained by the IRINA project in the GN3plus open calls. IRINA has investigated clean-slate network architectures, more particularly the Recursive Internetwork Architecture (RINA). A state-of-the-art and analysis of network architecture proposals and associated research projects in a global context was performed to motivate the need for the project. A survey was conducted to align the scope of the project with the requirements and objectives of the NREN community. In order to evaluate the use case, the rina-tgen tool was developed as part of the project and released as Open Source software. Experimental results from the project show that RINA has a lot of potential benefits in deploying services from a cloud environment.

Table of Contents

Table of Figures

Table of Tables

Executive Summary

The current Internet architecture has been a resounding success since its inception almost 40 years ago; however, it is not without its limitations. On the one hand IP has been an unexpected incomparable success in terms of deployment due to a number of factors. On the other hand, IP was designed to deliver packets on a "best effort" basis, meaning that it is acceptable to discard packets. It was originally designed in an era before VoIP and other streaming media services, prior to mobile devices, and pre-dates modern cloud based infrastructures, making the integration of multi-homing, mobility, security and providing QoS guarantees particularly cumbersome.

RINA is an emerging clean-slate programmable networking approach, centring on the Inter-Process Communication (IPC) paradigm, which aims to support high scalability, multi-homing, built-in security, seamless access to real-time information and operation in dynamic environments. The principles behind RINA were first presented by John Day in his book "Patterns in Network Architecture: A return to Fundamentals"[1]. RINA takes as a starting point the basic premise that "networking is IPC and only IPC" [2] . Networking provides the means by which processes on separate computer systems communicate, generalising the model of local inter-process communications. A Distributed IPC Facility (DIF) is an organising structure, grouping together application processes that provide IPC services. A DIF can be seen as what is generally referred to as a "layer". According to this view, networking is not a layered set of different functions but rather a single layer of distributed IPC that repeats over different scopes - i.e. providing the same functions/mechanisms - that can be tuned with different policies to operate over different ranges of the performance space (e.g. capacity, delay, loss).

IRINA performed an analysis of the Strengths, Weaknesses, Opportunities and Threats for a number of proposed architectures (XIA, SDN, LISP…) and approaches taken in international research projects,(Mobility First, SAIL, TRILOGY,…). The Opportunities and Threats are the same for all Future Internet technologies, the most important being business looking for solutions to speed up service development and deployment, lower CapEx and OpEx and ways to improve network reliability and mobility. The main threats to possible new technologies are resistance from incumbent technologies and the general risks associated with deploying a relatively new technology.

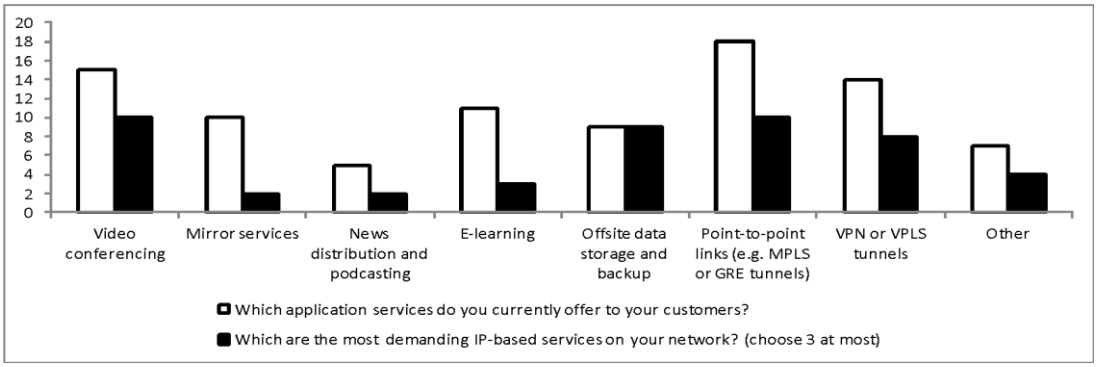
One of the strengths between most current approaches is that they focus on immediately implementable solutions not requiring drastic changes in the underlying infrastructure. However, these solutions are usually point solutions, meaning that they only solve particular use cases and may even introduce new issues.

Internet architectures in general have a slow rate of change and obsolescence. This is due to their inherent heterogeneity and the length of time required standardising new technological

[1] Day, John. Patterns in network architecture: A return to fundamentals. Pearson Education, 2007.

[2] Day, John, Ibrahim Matta, and Karim Mattar. "Networking is IPC: a guiding principle to a better internet." *Proceedings of the 2008 ACM CoNEXT Conference* 9 Dec. 2008: 67.

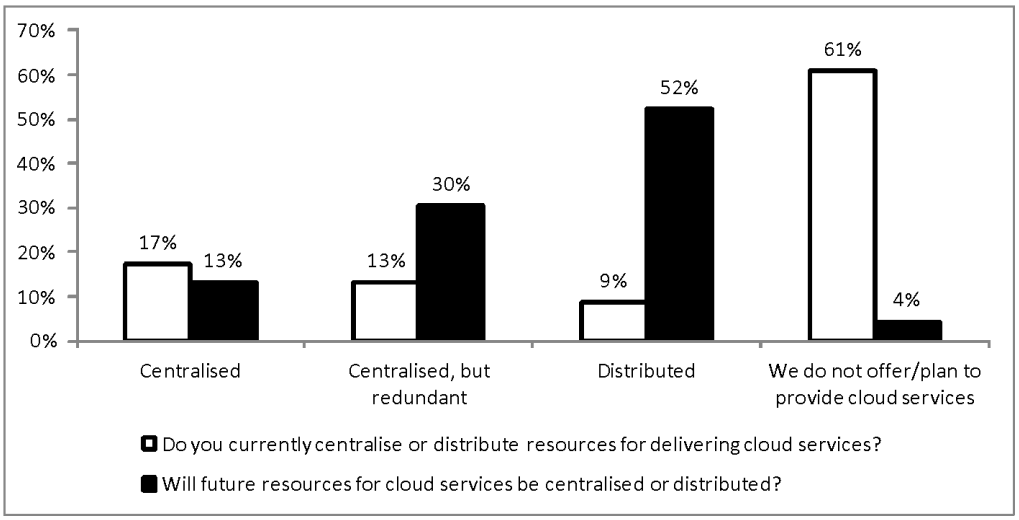
advancements. The current Internet architecture has been in place for the past 40 years, but may require a radical change in order to support required services into the future.
We made a brief summary of current and future NREN requirements, subdivided into two sections, namely service requirements and technical requirements. Some typical NREN service expectations include Network as a Service (NaaS), security, authentication, collaboration tools, multimedia content repositories and eLearning activities. While Technical requirements includes QoS, network virtualization, mobility, multi-homing, scalability, security and network management.
In the GN3+ OC IRINA project, a survey was held to gather the necessary information from the NRENs in order to assess the general requirements, the most demanding applications and key service parameters. Based on the survey results an accurate use case tailored to the NREN environment was derived. 24 NRENs responded to the survey. Offsite data storage and backup, video conferencing, point-to-point links and VPN or VPLS tunnels are considered as the most demanding services, shown below.



Overview of offered application services and the technological difficulty to offer each IP-based service.

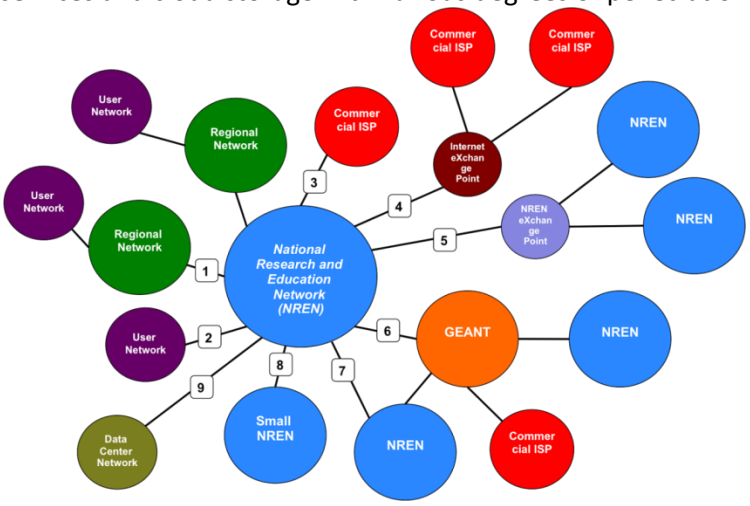
The most stringent requirements for these IP-based services are bandwidth, availability, latency and provisioning time. Mobility is not considered as a stringent requirement. Ten of the 24 respondents currently offer cloud services to their customers. 7 offer the cloud service via centralised resources (of which 3 have redundant resources) and 2 via distributed resources. For the future, the respondents identified security management as the most important operational service (14 responses), closely followed by cloud computing nodes (12) and equipment management (11). Failure management (7), Provisioning services (7) and mobility management (6) are also considered important services for the future.

While only ten of the 24 respondents to the survey currently offer cloud services, all but one is planning to roll out a cloud service in the future. There is also a move from centralised services to distributed cloud services.



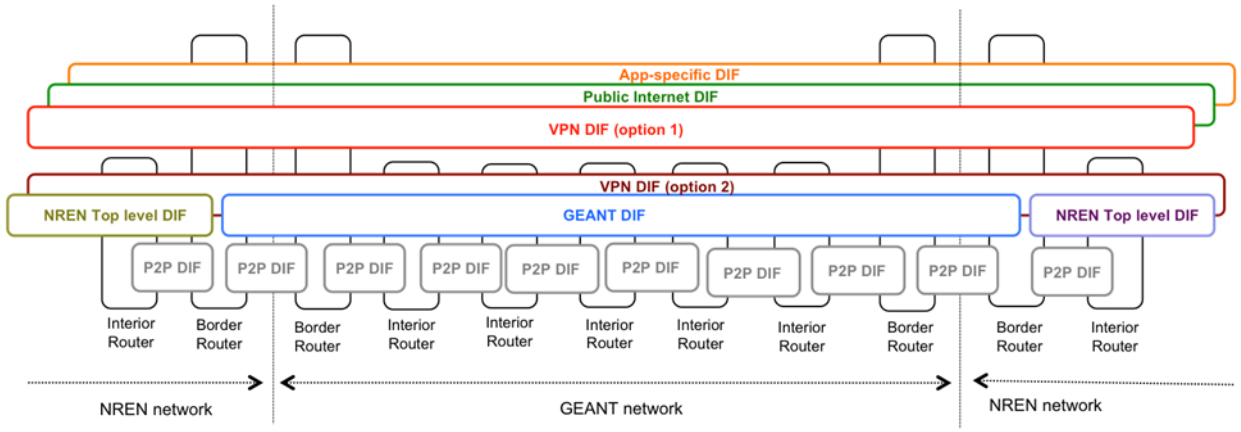
Comparison of current situation and future situation for cloud services

Based on these results, a use case was developed, focusing on three classes on NRENs (Large, Medium and Small) interconnected through GÉANT. These NRENs deploy three key services: video conferencing, VPN services and cloud storage with various degrees of penetration.

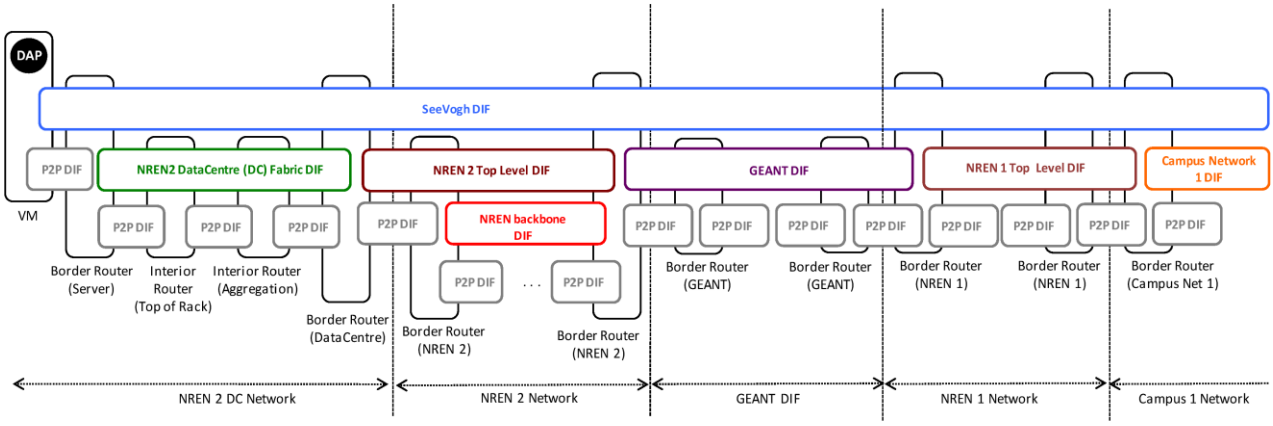


Overarching scenario

The key service that is analysed is the SeeVogh distributed video conferencing application. RINA was applied to this scenario; investigating various interconnections between NRENs and Regional Networks, User Networks, Commercial ISPs, IXPs, GÉANT and peering with neighbouring NRENs.



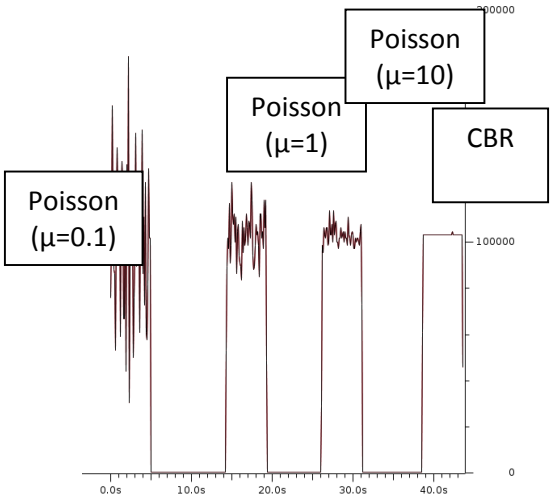
The Figure above shows an example of the GÉANT network modelled with the RINA architecture. A single DIF (the GÉANT DIF) spans all the routers in the GÉANT network and makes it a single resource allocation domain. This DIF can support a number of DIFs on top, such as the Public Internet DIF, application-specific DIFs or a number of VPN DIFs. These VPN DIFs will usually have a greater scope than just the GÉANT network, and will typically involve a number of NRENs and even regional and/or campus networks if they span to end users in labs, for example. Therefore the setup of these VPN DIFs will require the collaboration of a number of management domains. As in the case with NREN networks, VPN DIFs can also be implemented closer to the physical medium if overlaying them over the GÉANT DIF is not enough. The VPN DIF option 2 provides an example of this situation.



Application-specific DIFs can support the operation of a distributed collaboration application such as SeeVogh, which is currently setup as a peer to peer network of USHI instances (roughly speaking, SeeVogh servers) overlaid over the Internet. The USHI instances themselves are responsible for routing the traffic in the USHI peer to peer network. This is a scenario that can be used with the RINA architecture as well, but another interesting design is also possible: create an application-specific DIF that supports the operation of the distributed collaboration application. This "*SeeVogh DIF*" allows all USHI instances to rely on the IPC capabilities provided by the network, such as security, multi-homing, mobility or multicast; as well as to leverage dedicated hardware capable of processing large amounts of traffic if required.

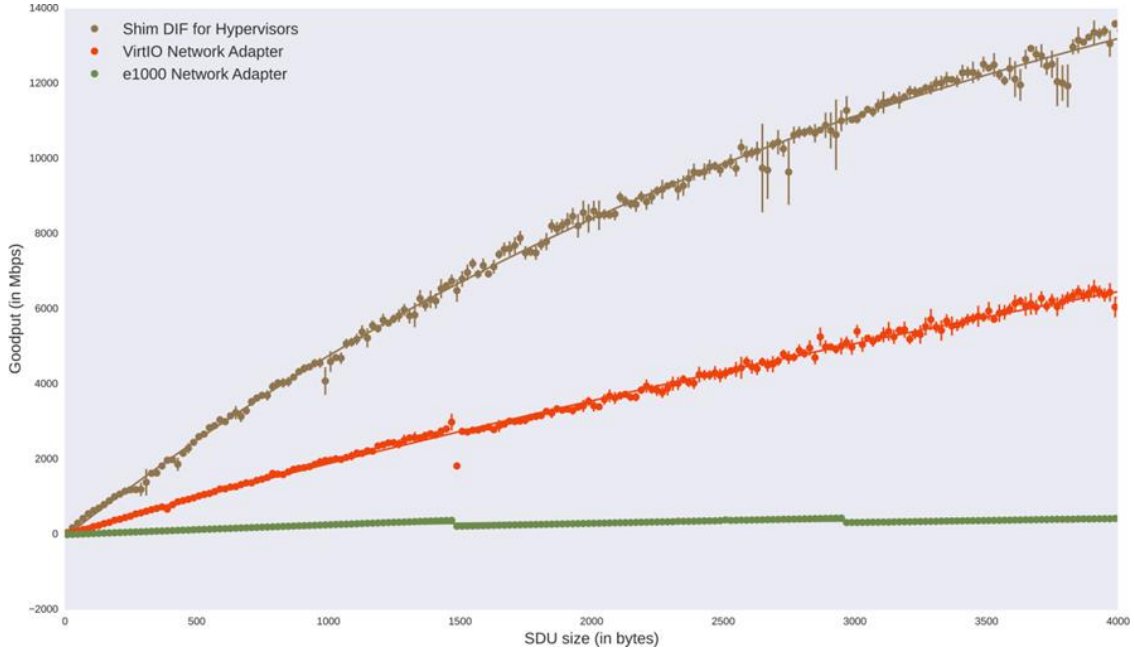
The IRATI stack on which IRINA builds lacks of a traffic generation tool complying with the requirements of the IRINA experimentation objectives since it only provides a RINA-based tool - the `rina-echo-time` application - which only performs basic ping functionality and provides

rudimentary bandwidth testing capabilities. We analysed some available traffic generation tools (most notably netperf, D-ITG and Ostinato) and, after weighing in the development work needed to convert these tools from POSIX sockets to the IPC API, decided to build a lightweight packet generator for RINA. The `rina-tgen` tool is currently supporting constant bit rate and random traffic with poisson distributed interarrival times. It is released as open source on github (http://github.com/irati/traffic-generator) under the GÉANT outward license.



Wireshark trace of traffic generated using `rina-tgen`

A minimal deployment was made in the iLab.t testbed at iMinds. The `rina-tgen` tool was validated to generate constant and poisson distributed traffic.



The performance of (single-host) Inter-VM communication was evaluated in-depth. In RINA, networking is IPC between application processes. As a consequence, there is no need to emulate a

NIC to connect the VM stack to Hypervisor stack. Unlike traditional VM networking, the shim DIF for Hypervisor is not restricted by the limitations of the Ethernet technology.
When performing maximum achievable bandwidth tests, the shim DIF for Hypervisors prototype outperforms an emulated e1000 NIC by more than an order of magnitude, and the virtio-net NIC by a factor of 3, showing that a simpler and cleaner architecture such as RINA also allows for better performance.

1 Introduction

This document provides an overview of the work performed in the GN3+ Open Call 5 project "IRINA", "Investigating RINA as the next generation GÉANT and NREN network architecture".

IRINA sets out to apply the Recursive Internet Architecture (RINA) to the NREN environment. To achieve this, it builds heavily on the results from the FP7 IRATI[3] project (01/2013-12/2014), which has developed a RINA prototype for Linux/OS. IRINA built a use case, analysed the requirements and built a traffic generation tool for experimentation, released as Open Source[4].

In the first section a comparative analysis of the Recursive InterNetwork Architecture (RINA) against the current networking state-of-the-art and the most relevant clean-slate network architectures under research in the context of GÉANT and the National Research and Education Networks (NRENs) environment is provided. The section specifically focuses on the actual and future requirements of the research and education networking communities. In particular, it provides insights into current network architectures designs that might be better suited to a broader range of emerging NREN and GÉANT services than current network architectures. Finally, a comparison is made among these architectures and incremental improvements based on the problem scopes they address.
In the second section the assessment analysis of the relevant network architectures identified in the first section is provided. The following methodologies have been used in the analysis:

- SWOT analysis: A set of objective internal (i.e. strengths and weakness) and external (i.e. opportunities and threats) factors are be defined. For the relevant architectures, the identified factors are quantitatively specified.
- PEST Analysis: This gives consideration to the political, economic, socio-cultural and technological (PEST) environment. Our PEST analysis will help in viewing the bigger picture. It includes an assessment of all four components (political, economic, socio-cultural and technological) as they apply to network architectures with an emphasis on considering the overall environment, and how the current architecture of GÉANT will respond to external variables such as the introduction of a new architectural approach.
- Risk Analysis: The risk assessment presents views on service risk, market risk, people risk, financial risk and competitive risk. In this analysis estimate how likely the threat is and how damaging it could be to GÉANT and the NRENs, and we will outline ways to manage or minimize the risk.

[3] http://irati.eu

[4] http://github.com/IRATI/traffic-generator

In the third section, we analyse the NREN requirements related to security, e-Learning platforms, etc. and a summary of results from a survey held among the NRENs is presented.

From this survey, a use case is refined in Section 4, taking account the overall topology of the interconnected NREN network and the most demanding services that are deployed now and will be deployed in the future.

The fifth section provides a further refinement of the final use case, with the application of the RINA architecture in Section 6.

The scenario takes into account internal NREN design, interconnection to other networks (GÉANT, customers, other NRENs, commercial ISPs), data centre design, network management and a selected service that highlights its potential benefits (SeeVogh).

Section 7 provides a rationale for the development of a software tool, `rina-tgen` which allows the scenario to be validated on the test bed, described in Section 8.

Section 9 summarises the conclusions drawn from the IRINA project.

2 State of the Art Review

2.1 Open challenges in today's Internet

The current Internet architecture has been a resounding success since its inception almost 40 years ago; however, the Internet architecture is not without its limitations. When it was first designed, the types of services it would provide was never envisaged, starting primarily as a data transfer only network to becoming a network providing near real time voice and streaming services. Also the range of devices used to access these services has also dramatically increased from being static nodes to mobile devices. While there have been major innovations in the upper (application) and lower (data link) layers of the Internet's protocols, the middle layers (transport/Internet) have remained relatively unchanged and have lead to many of the current challenges for content delivery specifically streaming services that the research community have attempted to resolve through incremental changes and ad-hoc patches.

Over the last number of decades the Internet Protocol (IP) has emerged as the unifying inter-network protocol facilitating communication between millions of interconnected devices, computers and data-centres (DC) worldwide. On the one hand IP has been an unexpected incomparable success in terms of deployment due to a number of factors. These factors include its technical features such as routing friendly design and relative scalability, its historical role as the protocol suite of the Internet and its open standards and development process which reduced barriers to acceptance of TCP/IP protocols. On the other hand, IP was designed to deliver packets on a "best effort" basis, meaning that it is acceptable to discard packets. It was originally designed in an era before VoIP and other streaming media services, prior to mobile devices, and pre-dates modern cloud based infrastructures, lacking the ability to provide QoS.

Currently deployed network architectures were not designed with built-in security measures, but rather security mechanisms were appended on in the form of additional layer functionalities or separate specific security protocols and as contended by Clark, D. et al., "experience has shown that it is difficult to add security to a protocol suite unless it is built into the architecture from the beginning"[5] which results in unnecessary complexity brought about by an ambiguous approach to network design. This approach to network architecture design - compounded with the exposure of IP addresses and the use of well-known port numbers that are visible to any user - introduce further security risks into the fragile network architecture. Current solutions rely on application layer security - e.g. HTTPS - or lower layer solutions that do not work with Network Address Translation (NAT) - e.g.

[5] "RFC 1287 - "Towards the Future Internet Architecture", 2009, 25 Mar. 2014. <https://www.ietf.org/rfc/rfc1287.txt>

IPsec. NAT boxes provide a means for partitioning and reusing parts of a single IP address space[6]. There has been a proliferation of these so-called middleboxes (firewalls, load balancers, etc.) in an attempt to alleviate current networking problems. However, their deployment further compounds current networking problems by hindering the ability to deploy end-to-end solutions. They represent point-solutions to optimise the current architecture at the expense of a global architectural solution. Network Function Virtualisation (NFV) is currently attempting to provide an answer to the proliferation of expensive middleboxes.

One of the most notable shortcomings of current networks is their lack of support for redundant physical interconnections (i.e. multi-homing - where a single host or router has a number of different network interfaces that are connected to the same or multiple networks simultaneously), making load balancing and reliable switchovers hard to achieve in a timely manner (i.e. they require higher level knowledge and processing, to figure out that the two addresses actually belong to the same device). Current network architectures do not provide separate names for the basic entities in the architecture: nodes, interfaces or points of attachment (PoA) to the network and applications. The only names provided are host PoA names (i.e. MAC and IP addresses). Although these names are commonly referred to as "host addresses", they are in fact interface addresses. The end result of this incomplete naming scheme is that the network has no way of identifying whether two or more IP addresses belong to the same node, making multi-homing impossible to achieve. Both in IPv4 and in IPv6 is an address attached to the network interface, not the host. Some protocols have been designed to support multi-homing (e.g. SCTP), and altering the point of attachment dynamically (Mobile IP). However, they have limitations as they rely on pushing the "intelligence" into the core network in addition to using customised network stacks.

The proliferation of smartphones, tablets, laptops and other mobile devices makes mobility the most important requirement for any future network architecture. Mobility can be viewed as a dynamic version of multi-homing with expected failures, which indicates that any network architecture capable of providing multi-homing will also be capable of providing mobility without requiring any additional protocols. Currently, application mobility is only supported for specific handoff scenarios (e.g. a mobile node connects to a new access point that is attached to the same domain as the old access point), while in most scenarios it is not possible to facilitate seamless mobility; low performance, long delay and frequent disconnections lead to degradation of the service for the end-user. Mobile IP delivers mobility, albeit with additional cost and overheads.

Another shortcoming of current network architectures is the lack of a built-in mechanism that allows the network to provide specific QoS levels as required. Most modern applications such as real-time streaming multimedia, Voice over IP (VoIP), safety-critical applications and other applications have specific requirements for bandwidth, delay, loss, jitter and error probability. It is a vital design aspect for today's network architectures to inherently provide mechanisms to support various QoS levels and particular service guarantees. Current network architectures do not provide mechanisms to support QoS on a large scale, but only over small, dedicated networks as there is no inherent mechanism within the architecture that can be leveraged to provide differing service levels maintained under specific resource guarantees throughout the network. In particular, applications have no means of requesting certain QoS parameters, since the sockets API only permits specification of either a reliable (TCP) or unreliable (UDP) transport service. It should be stated that some of these solutions above do not work well together as they are only point solutions to solve specific issues, and could not be combined into a functionally complete solution that addresses all shortcomings - as one might expect from the next generation network architectures - due to integration constraints.

[6] "RFC 1631 - "The IP Network Address Translator", 1994, 26 Mar. 2014. <http://www.ietf.org/rfc/rfc1631.txt>

2.2 State of the art on network architectures

This section focuses on making a comparative study of RINA against the current state of the art in network architectures. The network architectures have been designed with a view to solving many of the inherent problems in the current Internet, such as QoS, network virtualization, mobility, multi-homing, scalability, security, and network management.

2.2.1 The Recursive InterNet Architecture

RINA is an emerging clean-slate programmable networking approach, centring on the Inter-Process Communication (IPC) paradigm, which aims to support high scalability, multi-homing, built-in security, seamless access to real-time information and operation in dynamic environments. The principles behind RINA were first presented by John Day in his book "Patterns in Network Architecture: A return to Fundamentals"[7]. Since the book was published in 2008, several organizations have stated their interest in further researching RINA, as well as into turning the theory into practice by deploying RINA in the real world. The Pouzin Society (PSOC)[8] was formed in 2009 to coordinate all the international activities around RINA research and development.

RINA takes as a starting point the basic premise that "networking is IPC and only IPC"[9]. Networking provides the means by which processes on separate computer systems communicate, generalising the model of local inter-process communications. A Distributed IPC Facility (DIF) is an organising structure, grouping together application processes that provide IPC services. A DIF can be seen as what is generally referred to as a "layer". According to this view, networking is not a layered set of different functions but rather a single layer of distributed IPC that repeats over different scopes - i.e. providing the same functions/mechanisms - that can be tuned with different policies to operate over different ranges of the performance space (e.g. capacity, delay, loss).

The following figure provides more details of the RINA architecture. The structural blocks (i.e. the DIFs), the interfaces between them (i.e. between the DIFs) and the components within them are identified. The instantiation of a DIF within a system (e.g. a computer) is an IPC Process, an application that provides distributed IPC Services.

RINA is composed of an undefined number of layers on each node. This architectural property implies many benefits. Moreover, the address namespace is controlled by policies, so it is possible to use current or future tools designed to decrease the converging time of the routing algorithms, like a hierarchical addressing space. It is also possible to set up a maximum of nodes per DIF. When this maximum is reached, The DIF can be divided, and those divisions can be connected by creating an upper layer between them. This process can be done indefinitely providing RINA with a native scalability.

[7] Day, John. Patterns in network architecture: A return to fundamentals. Pearson Education, 2007.

[8] "The Pouzin Society - Building a Better Network." 2007. 21 Feb. 2014 <http://pouzinsociety.org/>

[9] Day, John, Ibrahim Matta, and Karim Mattar. "Networking is IPC: a guiding principle to a better internet." *Proceedings of the 2008 ACM CoNEXT Conference* 9 Dec. 2008: 67.

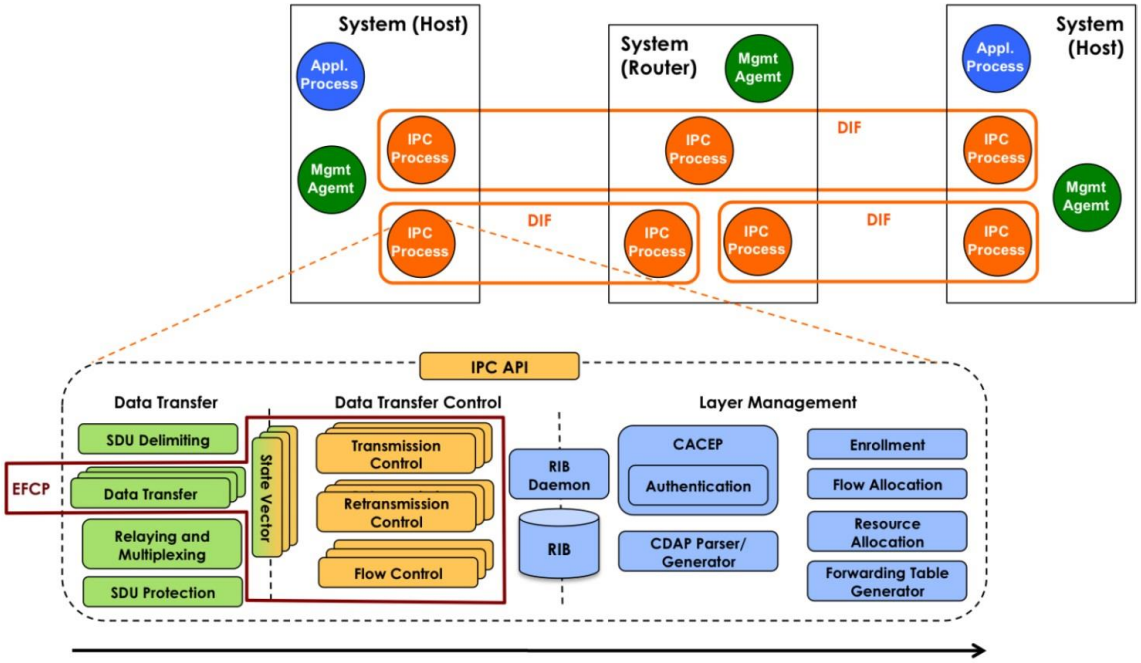


Figure 1: The Recursive InterNet Architecture

2.2.2 Content Delivery Networks

The current Internet based on the TCP/IP architecture is inefficient at delivering time-sensitive multimedia traffic. Hence the content oriented networking paradigm is aiming to alleviate the inherent problems with TCP/IP by redesigning the current Internet architecture to facilitate content-oriented applications and services in an efficient and scalable manner. The main premise behind content oriented networks involves the decoupling of multimedia content from hosts, not at the application layer, but at the network layer[10].

2.2.3 SDN

SDN[11] is changing the way networks are designed and managed. It has two defining characteristics. First, SDN separates the control plane (which decides how to handle traffic) from the data plane (which forwards traffic according to decisions made by the control plane). Second, SDN consolidates the control plane, so that a single control program controls multiple data-plane elements. The SDN control plane exercises direct control over the state in the network's data-plane elements (e.g. routers, switches, and other middleboxes) via a well-defined API.

[10]Choi, Jaeyoung et al. "A survey on content-oriented networking for efficient content delivery." *Communications Magazine, IEEE* 49.3 (2011): 121-127.

[11] Feamster, Nick, Jennifer Rexford, and Ellen Zegura. "The Road to SDN." *Queue* 11.12 (2013): 20.

2.2.4 FORCES

The ForCES IETF working group has created a framework, requirements, a solution protocol, a logical function block library, and other associated documents in support of Forwarding and Control Element Separation (ForCES)[12]. Drawing on the experience gained from developing the standards and from many efforts using this architecture, the ForCES working group is now working on a set of additions to the model, the protocol, and the libraries. The following 5 work items are the chartered tasks of this working group:

1. Extensions to Model and Protocol: This work is to address a set of extensions to the base model and protocol resulting in updates to RFCs 5810 and 5812. This effort will produce 2 Standards Track documents (one for the model and another for the protocol).
2. Inter-FE Connectivity: ForCES processing is often spread across multiple Forwarding Elements (FEs). The original framework identified the interface between FEs as the "Fi" reference point. Protocol and Logical Function Block (LFB) mechanisms to carry metadata across the Fi interface are needed. This effort will produce a standards track document defining the protocol on the wire to address this need, and the LFBs used to represent the Interfaces for sending and receiving such information. It is expected that this work will draw heavily on existing protocol and LFB definitions.
3. Parallelization: An FE can implement an LFB chain with parallelization, but the currently-defined mechanism has no means to represent when synchronization is needed, or to allow the Control Element (CE) to specify where it believes such parallelism is useful. This work item will produce a single standards track document to improve the handling of this case.
4. Subsidiary Management: Deployment experience has demonstrated the value of using ForCES to control the Forwarding Element Manager (FEM) by creating an LFB to represent its function using the same encoding rules as for any other LFB. This allows it to be controlled by the same Control Element (CE). This work item assumes the presence of an initially booted FE whose configuration could then be updated at runtime via an FEM LFB for runtime config purposes (e.g., by adding a new CE and its associated IP address). This work item can also be useful in addressing control of virtual FEs where individual FEM Managers can be addressed to control the creation, configuration, and resource assignment of such virtual FEs within a physical FE. This work would result in a standards track LFB FEM library RFC.
5. In addition to the specific work items listed above, the working group will allow discussions and review work of how to use ForCES to model topics of interest to Network Function Virtualization, I2RS, or OpenFlow. It is understood that the primary responsibility for such documents lives with other working groups, individual contributions or other standards bodies.

2.2.5 4D

While the Internet Protocol (IP) has been a runaway success, today's IP networks are difficult to manage efficiently. The 4D architecture takes a clean slate approach for redesigning different aspects of network control and management, guided by the following three principles:

[12] "ForCES - IETF Datatracker." 2010. 23 Feb. 2014 <http://datatracker.ietf.org/wg/forces/charter/>

Network-level objectives: Running a robust data network depends on satisfying objectives for performance, reliability, and policy that can (and should) be expressed as goals for the entire network, separately from the low-level network elements.
Network-wide views: Timely, accurate, network-wide views of topology, traffic, and events are crucial for running a robust network.
Direct control: The decision logic should provide network operators with a direct interface to configure network elements; this logic should not be implicitly or explicitly hardwired in protocols distributed among switches.
Despite the early design goal of minimising the state in network elements, tremendous amounts of state are distributed across routers and management platforms in IP networks. Many loosely-coordinated actors that create and manipulate the distributed state introduce substantial complexity make both backbone and enterprise networks increasingly fragile and difficult to manage. The 4D architecture[13][14] decompose the functions of network control into 4 planes:

1. A **decision** plane that is responsible for creating a network configuration (e.g. computing FIBs for each router in the network);
2. A **dissemination** plane that gathers information about network state (e.g. link up/down information) to the decision plane, and distributes decision plane output to routers;
3. A **discovery** plane that enables devices to discover their directly connected neighbours;
4. A **data** plane that forwards network traffic

Today, a large body of research exists on the correctness of existing routing protocols. However, analytical frameworks for studying routing dynamics have mostly focused on one single routing protocol instance at a time. In reality, the Internet is composed of, not one (e.g., BGP) but, a multitude of protocol instances that need to interact. For example, routes must be exchanged between BGP and OSPF. The interactions between these protocol instances are governed by the “routing glue” component. However, despite its wide usage and essential role, there has been no formal investigation into how safe its usage is. 4D develop analytical models to rigorously analyze the interactions between multiple routing protocol instances, and its impacts on a network-wide level. Making routing protocols safe alone is not sufficient to ensure the correctness of Internet routing but the routing glue plays an equally important part: Its usage can result in a wide range of routing anomalies including persistent forwarding loops and permanent route oscillations. This routing glue deserves further attention from the networking community.
Flow monitoring supports several critical network management tasks such as traffic engineering, accounting, anomaly detection, identifying and understanding end-user applications, understanding traffic structure at various granularities, detecting worms, scans, and botnet activities, and forensic analysis. These require high-fidelity estimates of traffic metrics relevant to each application. The set of network management and security applications is a moving target, and new applications arise as the nature of both normal and anomalous traffic patterns changes over time. The 4D makes the case for a "RISC" approach for flow monitoring which employs simple collection primitives on each monitoring device and manages them in an intelligent network-wide fashion, to ensure that the collected data will support computation of metrics of interest to various applications. A RISC architecture dramatically reduces the implementation complexity of monitoring elements; enables router vendors and researchers to focus their energies on building efficiently implementing a small

[13] "The 4D Project - School of Computer Science." 2005. 23 Feb. 2014 <http://www.cs.cmu.edu/~4D/>

[14] Greenberg, A. "A Clean Slate 4D Approach to Network Control and Management *." 2005. <http://www.cs.cmu.edu/~4D/papers/greenberg-ccr05.pdf>

number of primitives; and allows late binding to what traffic metrics are important, thus insulating router implementations from the changing needs of flow monitoring applications.

2.2.6 FARA

The NewArch [15][16] project explored the use of top-down architectural design by developing and prototyping a new architectural model that is called FARA. The name "FARA" is an abbreviation for three fundamental elements of the model: "Forwarding directive, Association, and Rendezvous Architecture". In the FARA model, the abstraction of host-to-host communication is replaced by packet exchange between entities. Intuitively, an entity is an application or similar thing. Structurally, an entity is an abstract concept, not linked to any particular implementation approach. An entity could be a process, a thread in a process, several processes, a whole machine, a cluster and so on. The use of the (IP address, port number) pair as a definition of destination identity is replaced in FARA by the notion of an association, which allows sequences of packets to access common state within an entity and synchronizes the communications between entities. Each packet carries an association ID (AId) that enables the receiving entity to demultiplex the message to a particular association. However, an association and its AId are strictly local to the containing entity; FARA does not assume global name spaces either for associations or for entities.
FARA replaces the use of the IP address for packet routing by the more general notion of a Forwarding Directive (FD) which routes the packet through the network and may be used for de-multiplexing within an end system. Each packet carries a destination FD, which provides enough information to permit the forwarding and delivery of the packet to the correct entity. The packet may also carry a source FD, which will permit a return packet to get back to the source. The FD drives all forwarding actions to reach the destination entity, and then the entity uses the AId to locate the association state. An FD may be a generalized source route, it may use topological information that may or may not be globally unique, and it may be rewritten in route. The current IP address plays both the FD role and the AId role, while FARA specifically separates these roles. Note that FARA does not require a single global address space.

2.2.7 LISP

The Locator/ID Separation Protocol (LISP)[17] [18] is a network architecture and set of protocols that implement a new semantic for IP addressing. In a nutshell, LISP separates the 'where' and the 'who' in networking and uses a mapping system to couple the location and identifier. LISP attempts to solve the multi-homing problem by changing the semantics of the IP address, replacing the IP address with two number parts, a Routing Locator (RLOC) part that are topologically bound to a network point of attachment (this is used by routers for forwarding packets through the network) and an Endpoint Identifier (EID) that is assigned independent of the network topology and aggregated by

[15] NewArch Whitepaper, ISI, USC. Available online at: <http://www.isi.edu/newarch/DOCUMENTS/WhitePaper.pdf>

[16] NewArchFinalReport, ISI, USC. Available online at: <http://www.isi.edu/newarch/iDOCS/final.finalreport.pdf>

[17] Farinacci, D. "RFC6830 - IETF Tools." 2013. <http://tools.ietf.org/html/rfc6830>

[18] "Locator/ID Separation Protocol." 2008. 25 Feb. 2014 <http://www.lisp4.net/>

administrative boundaries for identifying devices. Despite that LISP has been adopted by the Internet Engineering Task Force (IETF) as its most promising seamless adoption to solve the multi-homing problem of the current TCP/IP architecture, some novel research has declared it as a non-scalable solution[19].

2.2.8 ITU-T G.803

The ITU-T G.803 [20] is an architecture for transport networks based on the Synchronous Digital Hierarchy (SDH). SDH was originally designed to transport circuit mode communications from a variety of different sources, but it was primarily designed to support real-time, uncompressed, circuit-switched voice encoded in PCM format. The primary difficulty in doing this prior to SDH was that the synchronisation sources of these various circuits were different. This meant that each circuit was actually operating at a slightly different rate and with different phase. SDH allowed for the simultaneous transport of many different circuits of differing origin within a single framing protocol. SDH is not itself a communications protocol *per se*, but a transport protocol.

2.3 State of the art on research projects

This section presents the state-of-the-art in research projects[21] aimed at realising the future Internet network architectures highlighted in the previous section. In many cases the identified projects only tackle parts of the overall problem and as such do not provide a holistic solution to the identified problems.

2.3.1 eXpressive Internet Architecture

The eXpressive Internet Architecture (XIA)[22] [23]project, maintained some features of the Internet, such as providing support for the narrow waist technologies that networks are required to support, and packet switching, but it differed in a number of aspects. The XIA project aimed to build an architecture with native support for what it termed multiple principals (i.e. content, services, or users) and provide the ability to evolve its core functionality to accommodate new principals in the future. XIA also aimed to provide intrinsic security where communicating entities validate that their underlying intent was satisfied correctly without relying on external databases or configuration. The architecture proposed to provide support for a number of communication types such as content-centric networking, service-

[19] Meyer, D and Lewis D - "Architectural Implications of Locator/ID Separation." Dec. 2008 <http://tools.ietf.org/html/draft-meyer-loc-id-implications-01>

[20] ITU-T, "G.803: Architecture of transport networks based on the synchronous digital hierarchy (SDH)." <https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-G.803-200003-I!!PDF-E&type=items>

[21] Available online at: <http://cordis.europa.eu/fp7/ict/future-networks/projects_en.html>

[22]"XIA - eXpressive Internet Architecture - Carnegie Mellon University." 2010. 25 Feb. 2014. <http://www.cs.cmu.edu/~xia/>

[23] Anand, Ashok et al. "XIA: An architecture for an evolvable and trustworthy Internet." *Proceedings of the 10th ACM Workshop on Hot Topics in Networks* 14 Nov. 2011.

based communication, multicast, and mobility. A constraint imposed on the architecture was that it should have been possible to enable or disable support for these services as required. The main contribution that XIA adds to foster evolvability is an architectural element called "*intent*" where a XIA address space represents not only a new type of address semantic, but also backward compatible address pathways to reach the new address space[24].

2.3.2 Mobility First

The Mobility First project[25] was an NSF-FIA project founded on the premise that the Internet was originally designed for interconnecting fixed end points and as such failed to cope with the increasing demands of mobile devices and services. The Mobility First network architecture proposed to extend the narrow waist of internet protocols to include a global name resolution service that separated naming and addressing semantics, a storage-aware routing protocol, transport segmented hop-by-hop, and management and service application programming interfaces. The storage-aware routing (STAR) protocol was implemented using self-certifying public key addresses for increased trustworthiness. The Mobility First architecture could easily cater for context and location aware services required by mobile devices and attempted specifically to address issues around security and trust requirements caused by the open nature of wireless access networks, dynamic association, privacy concerns and the increased chance of network failures.

2.3.3 Nebula

The NEBULA project[26] focused on building a cloud-computing-centric network architecture based on an extensible core network that was built to interconnect data centres with trustworthy transit and access networks that enables provisioning of distributed communication and computing utilities. The NEBULA project was designed with the following constraints; parallel paths provided between data centres and core routers, a policy-based path selection mechanism, and authentication enforced during connection establishment. This enables roaming mobile users to connect to the closest data centre over a variety of access mechanisms such as wireless or wired links. The NEBULA architecture includes the NEBULA data plane that establishes policy-compliant paths combined with flexible access control mechanisms to defend against availability attacks. A NEBULA virtual and extensible networking techniques control plane provides access to application-selectable service and network abstractions such as redundancy, consistency, and policy routing. A NEBULA core redundantly interconnects data centres containing replicated data with high-availability core routers.

2.3.4 Named Data Network

The basic premise behind the Named Data Network (NDN) project[27] is that the current Internet transport mode has transitioned from an end-to-end packet delivery model to a content-centric

[24] Han, Dongsu et al. "XIA: Efficient support for evolvable internetworking." *Proc. 9th USEnIX nSDI* (2012).

[25] "MobilityFirst FIA Overview." 2010. 23 Feb. 2014 <http://mobilityfirst.winlab.rutgers.edu/>

[26] "NEBULA Future Internet Architecture." 2013. 22 Feb. 2014 <http://nebula-fia.org/>

[27] "Named Data Networking, A Future Internet Architecture." 2010. 22 Feb. 2014 <http://named-data.net/>

model. The present day Internet provides support for the client-server model and finds it extremely difficult to provide support for secure content-oriented functionality. NDN retains most of the current Internet's layering architecture, but modifies the middle layers to facilitate content distribution networks. The main novelty of this approach is removing the restriction that packets can only name location endpoints (IP addresses) and instead an NDN packet can be anything (e.g. an endpoint, movie chunk, book, etc). Rather than attempting to secure the transmission medium or data path using encryption, NDN attempts to secure the content by naming the data via a security-enhanced mechanism. The approach taken permits the separation of trust between data and hosts/servers which facilitates optimisation by caching content on the network side. NDN proposed to name distributed content hierarchically in a name tree structure to facilitate scalability due to the fact that NDN names are longer that IP addresses and to make content easier to retrieve. NDN also proposed to secure the data content specifically using public key cryptography methods instead of securing the transport containers that held the data while in transit across a network.

2.3.5 4WARD

The aim of the 4WARD project[28] was to facilitate the design of inter-operable future Internet network architectures. One of the design goals of the 4WARD project was to create a future network paradigm around the concept of networks of information where information objects maintain their own identity and are not required to be bound to location hosts (this is similar in spirit to the NDN project). Another design point for the 4WARD project was for the network path to be an active unit capable of controlling itself while also providing resilience, failover, mobility, and even secure data transmission. The 4WARD project proposed a management capability called "default-on" that was to be an essential part of the overall network and to facilitate the reliable instantiation and interoperation of heterogeneous networks over a single unifying infrastructure. One of the main outcomes of the 4WARD project was the definition of a stratum concept[29] which is the main component if its architecture. This is similar to a DIF in RINA; however, the interface and functionality of a stratum is not clearly specified and there can be many stratums carrying out many functions. The proposed 4WARD solution was targeted towards the completed range of network technologies from fibre backbones to wireless and sensor networks.

2.3.6 ANA

The Autonomic Network Architecture (ANA) project[30] focused on the autonomic behaviour of network architectures with the goal of designing and developing a framework that enabled flexible, dynamic, and fully autonomic formation of atomic functional blocks into network nodes, services, and entire networks. The proposed architecture was constrained to reacting to changes in networking conditions, dynamic adaptation, and re-organisation of network elements adhering to higher-level specifications. One of the mains goals of the ANA framework was to provide a set of generic abstractions that were capable of modelling networking concepts such as compartment, information channel, functional block, and information dispatch point which could then be easily implemented in

[28] "The FP7 4WARD Project." 2008. 23 Feb. 2014 <http://www.4ward-project.eu/>

[29] Johnsson, Martin et al. "Towards a new architectural framework–the Nth stratum concept." (2008).

[30] "Autonomic Network Architecture" 2006. 25 Feb. 2014 <http://www.ana-project.org/>

executable code. The project also aimed to identify and specify generic communication methods that were capable of interacting with the abstractions defined to model a communications network as opposed to defining stringent specifications like protocols and packet header fields. A common communication core was required to be defined and implemented by all components of a network instance that was called upon when needed. The ANA framework has an architectural construct called the compartment[31] that is similar to a RINA DIF, but unfortunately does not concisely define the functionality of a compartment as it capable of performing many different functions. The project outlines some abstract-level concepts regarding the internal workings of a compartment, but fails to highlight the relationship between the abstract and concrete functional elements of a compartment.

2.3.7 SAIL

The Scalable and Adaptive Internet Solutions (SAIL) project[32] aimed to integrate the concepts of Network of Information (NetInf), Cloud Networking (CloNe), and Open Connectivity Service (OConS). The NetInf was developed as an information centric networking architecture that focused on information objects as opposed to network nodes. The NetInf architecture was targeted towards communication and information distribution and included such features as secure naming, name-based routing, in-network caching, and optimised distribution that were provided as general services to all applications. The basic tenet behind OConS was to increase the efficiency of fixed and mobile networks through harnessing multi-path, multi-protocol, and multi-layer networking. While the main idea behind CloNe was to combine both cloud computing and network virtualisation. SAIL's architecture only targets management and control functions in the control plane and leverages the current Internet architecture and associated technologies in the data plane[33] with its inherent limitations and complexity.

2.3.8 TRILOGY

Trilogy[34] proposed a control architecture for the Internet that could adapt in a scalable, dynamic, autonomous and robust manner to local operational and business requirements. The project focused on reachability through inter-domain routing, including policy control and integrating filtering at trust boundaries. It also focused on how to deliver effective and efficient control of resource sharing under the constraints of social and commercial control where the architecture must permit conflicting outcomes to exist and evolve.

The four design principles stated to extend the original Internet's design principles and aptly named 'design for tussle' that the Trilogy project aimed to solve included exposure of information, separation of policy and mechanism, fuzzy ends, and resource pooling. Trilogy was designed based on three architectural aspects: End-to-end transport protocols that could utilise multiple paths through the network in order to achieve multi-homing at endpoints in a bid to improve reliability and utilisation in the network. Multipath routing that permitted routers to select from multiple routes to a specific

[31] Keller, Ariane et al. "A system architecture for evolving protocol stacks." Computer Communications and Networks, 2008. ICCCN'08. Proceedings of 17th International Conference on 3 Aug. 2008: 1-7.

[32] "SAIL." 2010. 23 Feb. 2014 <http://www.sail-project.eu/>

[33] Zhao, Liang et al. "Open Connectivity Services for future networks." Emerging Technologies for a Smarter World (CEWIT), 2011 8th International Conference & Expo on 2 Nov. 2011: 1-4.

[34] "Trilogy: Home." 2007. 22 Feb. 2014 <http://www.trilogy-project.org/>

destination in a bid to improve reliability and resource pooling with the network. An accountability framework to track resources usage designed to ensure that end-users and network operators were held accountable for the impact that their actions may have on other users of the network. The main tenet behind Trilogy was to build a framework to support improved multipath transport, multipath routing and an accountability. The overall Trilogy architecture is divided into three parts: a reachability plane that is responsible for hop-by-hop outgoing link selection which enables network-wide reachability; a resource plane that is responsible for determining how the transmission resource on each link is partitioned between packets and finally transport services that oversee the functions of reliability, flow control, and message framing and that are not visible to the packet framing service. When combined the reachability plane and the resource plane form the packet delivery service.

2.3.9 TRILOGY 2

Trilogy 2 [35] aims to provide a converged architectural framework capable of orchestrating, provisioning, and controlling the usage of heterogeneous resource pools as demanded by emerging highly distributed applications. The basic components of this architecture are the mechanisms and techniques that create liquidity at the different resource domains: bandwidth, storage and processing. A liquid system should ideally allow these heterogeneous resources to be used by any application, whether they are contributed by network operators, data centre operators or end systems. Some of these mechanisms exist in the Internet today and several others are investigated and proposed in Trilogy 2. In addition to creating new liquidity tools, Trilogy 2 provides the integration of these newly proposed mechanisms with the aim of improving the interactions among the heterogeneous resource pools. The interaction between these new mechanisms, along with the interaction between the liquidity tools and the other components of the Internet architecture is also explored. On top of these cross-liquidity tools a uniform resource information model acts as the gluing description language for a converged and seamless control of these heterogeneous resource pools across the Internet. In addition, in order to allow the different stakeholders to be willing to create such liquid pools of resources, Trilogy 2 will also provide the means to control the created liquidity though the means of incentives, information exchange and enforcement tools. Finally, Trilogy 2 will use the novel liquidity mechanisms to enable a set of compelling use cases targeting mobile devices and ISPs network infrastructure.

2.3.10 COSIGN

The Combining Optics and SDN In next Generation data centre Networks (COSIGN)[36] project aims to define and implement a flat, scalable Data-Centre-Network (DCN) architecture facilitated by optical technologies and SDN based network control. In the DCN data plane, the COSIGN project proposes to achieve a fully optical interconnection path from server to server within racks and between racks. COSIGN plans to extend SDN capabilities to leverage the added value of emerging optical technologies. COSIGN also plans to implement the concept of 'DC Infrastructure as a Service' by implementing the required mechanisms for composition and operation of multiple isolated and concurrent virtual DCs (VDCs) sharing the same DC infrastructure.

[35] "Trilogy 2." 2012. 23 Feb. 2014 <http://www.trilogy2.it.uc3m.es/>

[36] "Cosign." 2014. 23 Feb. 2014 <http://www.fp7-cosign.eu/>

2.3.11 T-NOVA

The T-Nova project [37] proposes the introduction of a framework to allow operators to deploy virtualized Network Functions (NFs) for their own and their customers' requirements. T-NOVA aims to design and implement a management/orchestration platform for the automated provision, configuration, monitoring, and optimisation of Network Functions-as-a-Service (NFaaS). T-NOVA plans to leverage and enhance cloud management architectures for the elastic provision and (re-) allocation of IT resources assigned to the hosting of NFs. It also plans to exploit and extend SDN platforms for efficient management of the network infrastructure. The T-NOVA project will establish an "NVF Marketplace", in which network services and functions by several developers can be published and brokered/traded.

2.3.12 UNIFY

The UNIFY project[38] aims to investigate , develop and evaluate means to orchestrate, validate and verify end-to-end service delivery from home and enterprise through aggregation and core networks to data centres. The focus of UNIFY is on a service abstraction model and an associated domain-specific creation language and programming interfaces to automate and optimise the deployment of service chains. This includes management and operation schemes to cope with increased network/service agility and handle network services end-to-end. The proposed framework will be built on a universal node architecture based on standard x86 components and accelerators for network functions virtualization (NFV).

2.3.13 GREENICN

The GreenICN[39] project aims to focus on information delivery, instead of the traditional host-to-host connectivity in IP, this allows the user to obtain content from anywhere in the network. Information Centric Networking (ICN) is aimed at the future internet where the network provides users with named content, instead of communication channel among hosts. The project proposes to design and implement an ICN architecture to facilitate operation of a low-energy information-centric internet.

2.3.14 STRAUSS

The STRAUSS project[40] [41] aims to define a global (multi-domain) optical infrastructure for Ethernet transport, covering heterogeneous transport and network control plane technologies, enabling an

[37] "T-NOVA Project Website." 2007. 23 Feb. 2014 <http://www.t-nova.eu/>

[38] " Home – UNIFY project". 2013. 23 Feb. 2014 <http://www.fp7-unify.eu/>

[39] "GreenICN". 2013. 23 Feb. 2014 <http://www.greenicn.org/>

[40] "STRAUSS Overview." 2013. 23 Feb. 2014. Available online at: <http://www.ict-strauss.eu/>

[41] STRAUSS report. Available online at:
<http://cordis.europa.eu/fp7/ict/future-networks/documents/eu-japan-projects/strauss-final.pdf>

Ethernet ecosystem. The project architecture proposes to leverage software defined networking principles, on optical network virtualisation as well as on flexible optical circuit and packet switching technologies beyond 100 Gb/s.

2.3.15 IRATI

The IRATI project[42] proposes to advance the state of the art of RINA towards an architecture reference model and specifications that are closer to enable implementations deployable in production scenarios. The design and implementation of a RINA prototype on top of Ethernet will permit the experimentation and evaluation of RINA in comparison to TCP/IP.
IRATI has the following main objectives:

- Enhancement of the RINA architecture reference model and specifications, focusing on DIFs over Ethernet. The enhancement of the RINA specifications carried out within IRATI will be driven by three main forces: i) the specification of a DIF over Ethernet as the underlying physical media; ii) the completion of the specifications that enable RINA to provide a level of service similar to the current Internet (low security, best-effort) and iii) the project use cases targeting ambitious scenarios that are challenging for current TCP/IP networks (targeting features like multi-homing, security or quality of service).
- RINA open source prototype over Ethernet for a UNIX-like OS. This is the goal that can better contribute to IRATI's impact and the dissemination of RINA. Besides being the main experimentation vehicle of the project, the prototype will provide a solid baseline for further RINA work after the project. By the end of the project the IRATI partners plan to setup an open source community in order to attract external interest and involve other organizations in RINA R&D.
- Experimental validation of RINA and comparison against TCP/IP. This objective is enabled due to the availability of the FIRE facilities, which provide the experimentation environment for a meaningful comparison between RINA and TCP/IP. IRATI will follow iterative cycles of research, design, implementation and experimentation, with the experimental results retrofitting the research of the next phase. Experiments will collect and analyse data to compare RINA and TCP/IP in various aspects like: application API, programmability, cost of supporting multi-homing, simplicity, vulnerability against attacks, hardware resource utilization (proportional to energy consumption).
- RINA prototype over Ethernet for JunOS. The RINA implementation within the JunOS operating system, using the JunOS SDK, will allow IRATI to increase the impact and realism of the experimentation. JunOS is a FreeBSD based OS, therefore there is no need to start a RINA implementation from scratch: the UNIX-like OS prototype will be adapted to JunOS. IRATI project members will learn to what degree the current router platform architectures can be reused for non-IP based technologies.
- Interoperability with the Pouzin Society RINA prototype over UDP/IP. The achievement of interoperability between independent implementations is a good sign that a specification is well done and complete. Therefore, achieving interoperable RINA implementations is both a necessity and a validation of the RINA specifications; even more taking into account that PSOC and IRATI prototypes target different programming platforms (middleware vs. OS kernel) and work over different underlying technologies (UDP/IP vs. Ethernet).

[42] "IRATI - Investigating RINA as an Alternative to TCP/IP." 2007. 23 Feb. 2014. Available online at: <http://irati.eu/>

- Provide feedback to OFELIA in regards to the prototyping of a clean slate architecture. Experimentation with a non-IP based solution is an interesting use case for the OFELIA facility, since IRATI will be the first to conduct these type of experiments in the OFELIA test bed.

2.3.16 PRISTINE

The PRISTINE project[43] intends to design and implement the innovative internals of the RINA clean-slate architecture. This includes the programmable functions for: supporting congestion control, providing protection / resilience, facilitating more efficient topological routing, and multi-layer management for handling configuration, performance, and security. The project aims to demonstrate the applicability and benefits of the approach and its built-in functions in use-cases driven by the end-users, service providers, and equipment vendors participating in the project with the aim of ensuring that the applications and tools developed will be deployable by providers, and have a greater potential for future exploitation.
PRISTINE has the following main objectives:

- RINA Software Development Kit: making the network programmable. Design and develop a Software Development Kit for the PRISTINE implementation of RINA, that enables programmers to effectively exploit in reality all the theoretical customization capabilities provided by RINA. The SDK will define a set of APIs into each of the components of an IPC Process, allowing developers to modify the behaviour of the DIFs in terms of data transfer, forwarding, authentication, access control, resource allocation and so on. PRISTINE will modify the IRATI project implementation to allow extension modules to be plugged in and out of the prototype.
- Programmable congestion control for effective data transfer. Research, design and implement the mechanisms and algorithms that allow each DIF to explicitly detect congestion generated within the DIF, and take the appropriate measures to quickly react against it. Control loops with different characteristics will be designed, tailored to the requirements of PRISTINE's use cases. The interactions between the control loops at different DIFs will be analyzed. The different congestion control solutions will be incorporated into the prototype through the use of the SDK.
- Distributed resource allocation strategies to support multiple levels of service. Investigate and program a set of distributed resource allocation techniques that enable a DIF to provide different levels of service to honour the requirements of different applications. These techniques will leverage the capabilities that RINA provides in terms of allowing applications to express their desired level of service and the theory unifying connection oriented and connectionless resource allocation. The interaction between distributed resource allocation and congestion control techniques within a DIF will be investigated. The SDK will be used to plug the extensions into the prototype.
- Topological addressing as an enabler of efficient routing. Research and develop topological addressing schemes and its associated routing mechanisms, in order to minimize the size of the forwarding tables within DIFs. Topological address spaces reflect an abstraction of a connectivity graph within a layer, therefore the forwarding decision can be taken by

[43] "Welcome to PRISTINE! Exploring programmability, advanced ... - IRATI." 2014. 23 Feb. 2014. Available online at: <http://irati.eu/welcome-to-pristine-exploring-programmability-advanced-policies-and-dif-management-systems-in-rina/>

examining the destination address and the addresses of the directly attached routers. PRISTINE will investigate what topologies for address spaces make sense, are easily maintained, and scale for the three use cases of the PRISTINE project. Development activities will be carried out through the SDK.

- Authentication, access control and encryption for secure DIFs. Investigate, design and implement different strategies to perform authentication, access control and encryption as required by the three PRISTINE scenarios. Security is an integral part of an IPC Process and does not need to be handled in separate subsystems such as firewalls. Application access control, symmetric/asymmetric key-based authentication protocols and encryption mechanisms will be investigated and adapted to RINA through the use of the SDK.
- Security coordination within a DIF: self-management, attack identification and mitigation. Research and program techniques that enable a DIF to coordinate its internal security mechanisms in a distributed an autonomous way. Management and distribution of credentials, as well as logging and analyzing the key events related to security are the most important issues that will be addresses by this objective. The analyzed information will be used to decide if a DIF is being attacked, and to take measures to protect from the attack. The developed extensions will be incorporated into the prototype through the SDK.
- Multi-homing and self-healing as the basis for resilient networks. Investigate and develop routing algorithms and routing information dissemination strategies that optimally exploit RINA's support of multi-homing for load-balancing and rapid recovery of failures. Distributed resource allocation techniques will also be used in order to re-create the connectivity graph of the DIF, effectively recovering from malfunctioning links or IPC Processes. All the extensions will be plugged in the prototype through the use of the SDK.
- Multi-layer DIF Management System (DMS) for integrated network management. Design and develop a DMS capable of managing multiple DIFs (layers) at once. The commonality provided by RINA allows multi-layer management to be vastly simplified; thus opening the door to more robust, dynamic, responsive and cheaper network management operations. The DMS developed within PRISTINE will take care of configuration, performance and security management.
- Trials of the project use cases: deploying PRISTINE's solutions in the real world. Demonstrate the benefits of the RINA architecture and PRISTINE's solutions by trialling the project use cases in realistic conditions. PRISTINE will bundle the different solutions into three packages, one for each use cases (detailed in the next section), and showcase the technical and business impact of the project results through different trials over a rich infrastructure composed by partner's resources and relevant FIRE facilities.
- RINA Simulator to understand the behaviour of extensions at scale. Design and develop an OMNeT++ based RINA simulator, utilizing part of the IRATI implementation source code as an input. The simulator is a secondary objective, but a useful tool for RINA research. Within PRISTINE it will enable researchers to understand how the solutions for the different problem areas behave at scale.

2.4 SWOT assessment analysis

In this section, the SWOT analysis of the architectures described in the SOTA review (ref. section 0) is presented. A SWOT analysis is a structured planning method to evaluate the **s**trengths, **w**eaknesses, **o**pportunities, and **t**hreats of a business or project:

- Strengths: characteristics of the project/architecture that give an advantage.
- Weaknesses: characteristics that place the project/architecture at a disadvantage.
- Opportunities: elements that the project/architecture could exploit to its advantage.
- Threats: elements in the environment that could cause troubles.

In the context of this document, these terms assume the following meanings respect to a) the RINA architecture - if compared against an architecture - or b) a complete RINA prototype - if compared against a project providing a running prototype as its final outcome. RINA is taken as the reference comparison term since it provides a complete architecture addressing all the problems and issues of the current Internet, as described in section 2.2.1.
Please note that the following analysis does not take into consideration in-progress works, as depicted in the SOTA review. By taking as a reference point RINA, the RINA based projects (i.e. FP7 IRATI and FP7 PRISTINE) are also excluded from this analysis.

Since the analysis presented compares future network architectures, opportunities and threats factors are not reported as in classic SWOT tables but summarized only once as follows:

Opportunities	Threats
Businesses looking for • ways to speed up service development and deployment • ways to lower Capital and Operational expenditures • new business models • ways to improve network reliability and mobility	• Resistance from incumbent technologies • Deployment risks

Table 1: Opportunities and Threats for Future Internet Architectures

2.4.1 XIA

Strengths	Weaknesses
• The architecture proposed to provide support for a number of communication types such as content-centric networking, service-based communication, multicast and mobility. • The project aimed to provide intrinsic security where communicating entities validate that their underlying intent was satisfied correctly without relying on external databases or configuration. • The architecture takes into account scalability problems. • The project aims to build an architecture with native support for what it termed multiple principals (i.e. content, services, or users) and provide the ability to evolve its core functionality to accommodate new principals in the future. • A prototype is available for deployment.	• The architecture does not take into account network virtualization solutions. Therefore, it does not address the associated problems. • It does not take into account network management problems. • Mobility is provided with a mechanism similar to mobile IP, inheriting its complexities. • It does not support QoS. • It does not provide specifications and/or documentation.

2.4.2 Content Oriented Networks

Strengths	Weaknesses

• Given its limited scope due to its major focus on contents:
  ○ It takes into account scalability.
  ○ It takes into account security.
  ○ It takes into account network management problems.
  ○ It takes into account mobility related problems.
• Prototypes and solutions are available for deployment.
• Supports retrieving data from multiple locations
	• The architecture does not take into account network virtualization solutions. Therefore, it does not address the associated problems.
	• It does not provide the means for full-blown QoS (as in RINA). QoS is handled on end-point base, relying on the mechanisms offered by the underlying network.
	• It does not support multi-homing.
	• It does not provide a single specification and/or documentation.
	• There is no reference prototype or solution but a plethora of different solutions and prototypes.
	• The architecture is targeted to contents (and to content-oriented applications and services). Therefore, it addresses only a part of what the Internet is used for.

2.4.3 SDN

Strengths	Weaknesses
• Takes into account network management related problems. • Separates Control and Data planes, providing the means for better network management. • Takes into account scalability problems. An SDN controller can manage multiple data-plane nodes (e.g. OF switches). • Has the possibility to support network virtualization. • SDN solutions are available as OTS software components. Therefore, the solution is deployable. • Targets to provide a well-defined set of APIs. • Mobility issues are being addressed[44].	• It does not take into account QoS (in general). QoS can be offered by prioritizing certain flows based on the requested characteristics (e.g. send best effort through a different path rather than the path used for higher priority ones). • It does not provide multihoming support. • It does not provide additional security mechanisms than the standard ones used in IP networks. • It does not uniquely define an architecture but the principles in order to separate data- and control-planes. Therefore it does not provide specifications or clear documentation but well-known practices and patterns. • It is now considered OTS. A standardization process is not foreseen in the short/mid term (and it might not happen).

2.4.4 4D

Strengths	Weaknesses
• Takes into account scalability problems. • Takes into account security problems. • Takes into account network management related problems. • Analyzes the interactions between multiple routing protocol instances, and its impacts on a network-wide level.	• The architecture does not take into account network virtualization solutions. Therefore, it does not address the associated problems. • It does not support QoS • It does not support mobility • It does not support multihoming • The approach is only theoretical. The project does not provide a prototype that could be used for experimentation purposes.

2.4.5 LISP

Strengths	Weaknesses
• Attempts to solve the multi-homing problem by changing the semantics of the IP address • Takes into account scalability problems. • Solutions are available for either production or experimentation deployments: ○ Cisco has IOS and NX-OS images which support LISP. • There are various open-source implementations available (e.g. UCL, UPCM, LISPmob and AVM GmbH ones)	• It mainly focuses on scalable routing and addressing, therefore the architecture does not address mobility, security, network management and QoS. • The architecture does not take into account network virtualization solutions. Therefore, it does not address the associated problems. • The approach, even if taking into account scalability problems, presents drawbacks[45], for instance testing the Locator Path Liveness in the data plane does not scale in the worst case

2.4.6 4WARD

Strengths	Weaknesses
• The architecture does not take into account network virtualization solutions. Therefore, it does not address the associated problems. • It supports QoS. • It supports mobility. • It provide support to security. • It supports network management. • A prototype is available for experimentation. • Its “Information objects” are not bound to hosts.	• It does not address multihoming problems. • It does not address scalability problems. • The interface and functionality of the common building block, the stratum element, is not clearly defined.

2.4.7 Mobility First

Strengths	Weaknesses
• It takes into account mobility. • It takes into account security (in its storage-aware routing protocol). • Its “global name resolution service” concept separates naming and addressing semantics. • The architecture specifically addresses issues around security and trust requirements caused by the open nature of wireless access networks, dynamic association, privacy concerns and the increased chance of network failures. • A prototype is available for experimentation	• The architecture does not take into account network virtualization solutions. Therefore, it does not address the associated problems. • It does not address QoS • It does not address multihoming issues. • It does not take into account scalability problems. • It does not take into account network management problems. • Only targets mobility and the security mechanisms associated with it. • It primarily aims at defining a network architecture.

2.4.8 Nebula

Strengths	Weaknesses

• The architecture supports mobility (roaming users should be able to connect to the closest DC either via wireless or wired links).
• The architecture targets security concerns.
• The architecture aims at providing redundancy.
• The architecture separates Control- and Data-planes, aiming to enhance the network management aspects.
• A prototype is available for experimentation.
	• The architecture does not take into account network virtualization solutions. Therefore, it does not address the associated problems.
	• It does not provide support to QoS.
	• It does not provide support to multihoming.
	• It does not directly take into account scalability concerns and management problems although some documentation mentions these aspects. However, no clear plans on addressing these properties is available.
	• Is an architecture that is very limited since it is only targeted at interconnecting data centres with trustworthy transit and access networks.
	• The architecture is mainly a cloud-computing-centric network architecture and mostly focus on Data Centres. It was built to interconnect DCs with trustworthy transit and access networks.
	• The architecture imposes constraints such as:
	  ○ Parallel paths between DCs and core routers
	  ○ A policy-based path selection mechanism
	• Authentication enforced during connection establishment

2.4.9 ANA

Strengths	Weaknesses
• Focuses on autonomic behaviours (that would ease network management) such as reacting to changes in networking conditions, dynamic adaptation, and re-organisation of network elements adhering to higher-level specifications. • Focuses on the autonomic behaviour of network architectures. • The framework provides a set of generic abstractions (capable of modelling networking concepts such as compartment, information channel, functional block, and information dispatch point).	• The architecture does not take into account network virtualization solutions. Therefore, it does not address the associated problems. • The functionality of a compartment is not defined, so no guarantees on QoS, mobility, multihoming, scalability, security, network management. • The architecture does not concisely define the functionality of a compartment as it capable of performing many different functions. • (Bad) Autonomic behaviours might harm performances as well as resources utilization.

2.4.10 SAIL

Strengths	Weaknesses
• Takes into account scalability problems. • Takes into account security problems. • Takes into account network management problems. • A prototype is available for experimentation.	• Does not take into account QoS • The architecture does not take into account network virtualization solutions. Therefore, it does not address the associated problems. • Does not take into account mobility • Does not take into account multihoming • Only targets management and control functions in the control plane and leverages the current Internet architecture and associated technologies in the data plane with its inherent limitations and complexity.

2.4.11 TRILOGY

Strengths	Weaknesses
• Takes into account scalability problems. • Takes into account network management problems. • It provides early prototypes to experiment with. The prototypes provide scattered solutions for demonstration purposes. • One of the major contributions (i.e. MTCP) is under "standardization" process.	• The architecture does not take into account network virtualization solutions. Therefore, it does not address the associated problems. • Does not take into account mobility. • Does not take into account multihoming. • Does not take into account security and QoS.[46] • The solution presented is mainly theoretical. • The architecture outlined in the project seems mainly aiming at supporting MTCP research intents instead of providing the general means for a survivable architecture that could last in the middle/long terms.

2.5 PEST Analysis for RINA in the Context of NRENs and GÉANT

A PEST (Political, Economic, Social & Technological) analysis is an analysis technique used to assess and evaluate factors that may impact a decision, a market or a potential new business[47]. A PEST analysis can help identify business opportunities and provides advanced warning of significant threats and helps develop an objective view of a particular operating environment. This section performs a PEST analysis to assess and evaluate the impact of deploying a novel network architecture and in this particular case the RINA architecture within the NRENs and GÉANT scenario in the context of the E.U. market place. The PEST analysis is performed in relation to the deployment of the RINA architecture itself as opposed to performing a PEST analysis on each individual architecture identified in the SoTA document as it better depicts the overall impacts of adopting RINA by NRENs instead of the other partial architectural solutions that may quite possibly fail to meet expectations. This permits placing a greater emphasis on consideration of the overall environment, and how the current architecture of GÉANT will respond to external variables such as the introduction of a new architectural approach. This final section provides an analysis of certain factors (political, economic, social & technological) associated with RINA in a wider deployment context, that of the networking and research community as this is envisaged as the next stepping stone on the path to full RINA deployment. This also permits the testing on the delivery of critical NREN services over RINA on a greater and more realistic scale.

2.5.1 Political factors

There are a number of political factors that can have an impact on the deployment of a clean-slate network architecture such as RINA by GÉANT for NRENs. These political factors include level of governance, network dependability, interoperability and standards, network neutrality, EU privacy directives, commercial legislation and lawful interception.

Governance; Internet Governance is the development and application of shared principles, norms, rules, decision-making procedures, and programs that shape the evolution and use of the Internet. This has to be done by Governments, the private sector and civil society, in their respective roles. The Future Internet calls for an 'open and transparent' governance model based on international agreements and regulations. This should be based on a stable channel of communication and discussion between all stakeholders such as governments, the private sector (ICT industry), academics and civil society, each in their respective roles. For the existing Internet, Internet names and addresses, and corresponding issues are dealt by the Internet Corporation for Assigned Names and Numbers (ICANN). Application focused improvements or alterations are a mostly open process, and standardised by IETF working groups. RINA governance is currently administered under the umbrella of the PSOC. It offers a common forum where to expose, discuss and concentrate the RINA specifications in order to avoid losing contributions and divergence of the topic.

Network Dependability; As the Internet is becoming an increasingly critical infrastructure for the governments, citizens and businesses in their day-to-day activities, the robustness of the network should be taken into consideration when designing any potential replacements. The networks can be impaired for several reasons (e.g. physical or cyber attacks, etc.). RINA offers improvements in provisioning network connections that support required QoS, and security facilities. In particular, the

[47] Aguilar, Francis Joseph. *Scanning the business environment*. New York: Macmillan, 1967.

enrolment process is very useful in identifying and authenticating network endpoints. When combined with the available resilience policies, this allows RINA to provide better guarantees for the dependability of the network, by controlling admission and optional authentication of communicating parties.

Interoperability and Standards; The Internet is split into many legal and administrative domains with many of the Internet players partly rivals in business, yet partly cooperative in their offer of connectivity services to customers. Both aspects demand widely accepted standards for interfaces and rules for interoperability. These should be robust in terms of completeness (complete technical disclosure of APIs, protocols and formats), control (fair and transparent multilateral governance), cost (fair, reasonable and non-discriminatory licensing of essential IPR) and compliance (adherence to relevant standards and industry specifications). RINA relies on open standards as a route to interoperability. The reference documents will be freely available once sufficiently stable (expected in the short term). People will also be free to contribute to the on-going development of the RINA reference set. There are no licensing fees required to use the specifications. Thus RINA standardisation is as open as IETF standardisation process. The PSOC and the RINA community are looking for alternatives in order to build or join well-known standardization groups. One possibility being studied at the moment sets the ISO/IEC JTC 1/SC 6 "Telecommunications and information exchange between systems" group as a candidate to hold the authority in standardising the architecture.

Network Neutrality; The debate on Internet neutrality focuses on the question of access to the physical (transport) layer of the Internet. The discussion could be summarised as follows: Service providers would like to have neutral transport of their (service) data in operator networks. Network operators would like to differentiate on a technical level the transport of the data packets. In principle, it is a) a business related discussion about charging schemes and b) a technical issue concerning the need for traffic shaping related to potential network congestion. Transparent and competitive broadband markets should ensure that service providers remain honest. RINA offers a standard API for specifying QoS, and thus standardizes the interface between "layers" and thus service providers. In practice using RINA, the interconnections between ISPs can be captured as one or more DIFs used to exchange data.

EU privacy directives; The EU privacy directives require that "personal identifying information" is handled appropriately[48]. This includes the secure transport of such information across the network to designated points of attachment of the communicating parties. The listening, tapping, storage or any other kind interception of communications data (network traffic) by persons other than the involved users, without the consent of the users concerned is prohibited. The new EU 'right to be forgotten' law[82] holds network operators responsible for removing personal data stored on any EU citizen. RINA can indirectly facilitate the pseudo- anonymisation of network access for users or applications due to RINA's addressing scheme where flows are aggregated on N-1 DIFs, so discarding a users/applications personal data would merely be a matter of removing the records from the topmost DIF in the RINA stack.

Commercial legislation; EU and national laws also require that commercial transaction information, for example, transmission of credit card information is protected by minimum standards to prevent criminal activity. This is required to prevent fraud, theft, or other unauthorized access to resources

[48] Data protection factsheet, http://ec.europa.eu/justice/data-protection/files/eujls08b-1002_-_protection_of_personnal_data_a4_en.pdf

and services that may harm the commercial interests of either party. Legislation[49] requires that at a minimum, such traffic is protected by SSL or equivalent encryption techniques when in transit. For both of the above aspects, RINA allows policies to be configured to encrypt network data on a "per-layer" basis. Thus communication can be protected without affecting lower layers of the network stack. Indeed, it is possible to have different levels of protection applied on different "layers" or DIF"s which group peer endpoints.

Lawful Interception; The confidentiality of communications and related traffic data by means of a public communication network and publicly available electronic communications services has to be ensured. However, most EU states support the lawful interception of communications traffic in different domains[50]. This is governed by national law, where the responsibilities and requirements for "legal tapping" are made clear. RINA has no explicit support for the lawful interception of traffic. Policies may be configured to provide a very high level of security on virtual connections in a "layer". These policies are provided by the users of the network, so therefore may employ encryption schemes that may be contrary to the lawful interception of traffic.

2.5.2 Economic factors

Economic factors can have an impact on the level of investment by enterprises and governments in new technologies. These economic factors include GDP, inflation, interest rates, unemployment figures, and deployment costs.

GDP; Gross domestic product (GDP) is the market value of all accountable goods and services produced within a country over a given time period, usually one year. The E.U. economy generates over 12 trillion Euros per year[76] which makes it one of the largest economies in the world. In the fourth quarter of 2013, the seasonally adjusted general government deficit to GDP ratio stood at 2.6% in the euro area (EA18), down from 3.1% in the third quarter of 2013. In the EU283 the deficit to GDP ratio also decreased from 3.5% of GDP in the third quarter of 2013 to 3.1% of GDP in the fourth quarter of 2013[51]. The overall cost of provisioning networks is decreasing which is unlikely to have any major impact on the GDP rate. However, overall use of the services built and deployed to run over these networks is on the increase (e.g. mobile use (see

Figure 2), cloud services, etc.) so it is expected that these types of services will contribute more towards GDP in the future. Existing network architectures will fail to meet the requirements of future mobile and cloud service usage. RINA is inherently capable of supporting this rapid increase in mobile and cloud services.

[49] E-Commerce Directive (2000/31/EC) adopted in 2000.

[50] OJ C 329 adopted on 4.11.1996

[82] European Commission, "Factsheet on ECJ's ruling on the 'right to be forgotten', May 2014, Available at: http://ec.europa.eu/justice/data-protection/index_en.htm [31/05/14]

[51] "Eurostat Home." 2006. 29 Apr. 2014 <http://epp.eurostat.ec.europa.eu/>

[81] Meeker, M., "Internet Trends 2014 – Code Conference", May 2014, Available at: http://www.kpcb.com/internet-trends ,[30/05/14].

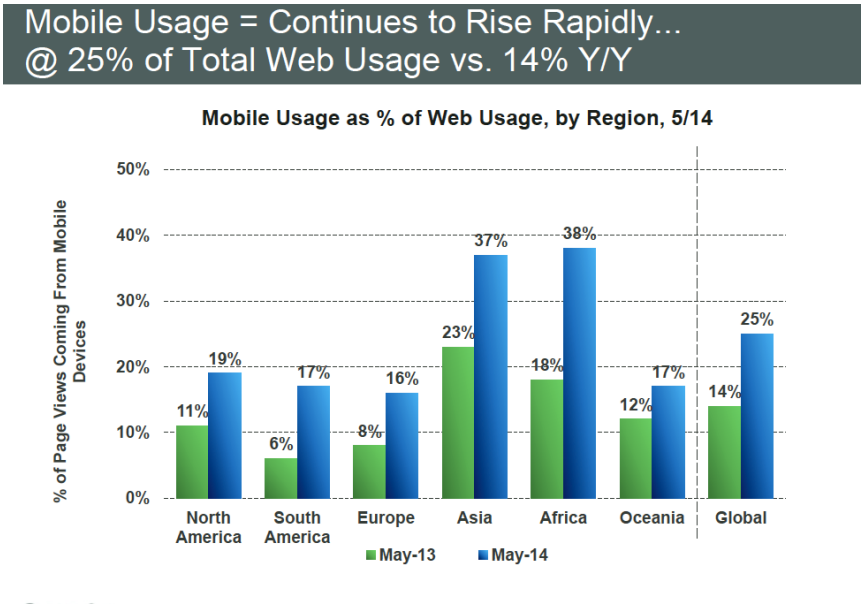


Figure 2: Mobile Usage

Inflation; Inflation refers to a general increase in consumer prices over a period of time and is usually measured by an index. In the E.U., this index has been harmonised across all E.U. member states and is aptly named Harmonised Index of Consumer Prices (HICP)[74]. The HICP is used to measure inflation which is used to define and assess price stability in the euro area as a whole in quantitative terms. The inflation rate in the E.U. was recorded at a low of just 0.50% in March of this year as reported by the Eurostat[73]. This is the lowest rate of inflation seen since November 2009. The inflation rate in the EU has averaged 2.21% from 1991 until 2014, reaching an all time high of 5.00% in July of 1991 and a record low of -0.70% in July of 2009. These figures indicate that the EU economy as a whole is still in recession, which means that individual consumers have a vested interest in obtaining greater value for money by asserting more pressure on network operators to curb their costs. This can have a negative impact on operators' revenues and essentially their bottom-line. For operators, RINA offers more efficient use of their network resources in maximising revenue and ultimately reduce operational expenditure.

Interest rates; The ECB is charged with setting interest rates for the Euro. The current European interest rate is set at a base rate of 0.25% by the ECB[74]. Interest rate have an impact on the level of borrowing by organisation that consider investing in new technologies. However, these low interest rates are only available to large organisations and governmental bodies such as NRENS. Ultimately, this may translate into a competitive advantage for those larger organisations or operators that are capable of providing investment into new network technologies such as adopting RINA due to the fact that their overall borrowing costs are lower and they can upgrade to newer technologies quicker than SMEs.

Unemployment figures; The unemployment rate represents unemployed persons as a percentage of the labour force. The labour force is the total number of people employed and unemployed. At present the unemployment rate for the E.U. as a whole stands at 11.9% with the lowest unemployment rates recorded in Austria at 4.9%, Germany at 5% and Luxembourg 6.1%; while the highest unemployment rates were recorded in Greece at 28% and Spain at 25.8%[73]. A large-scale deployment of RINA will develop expertise in both RINA, and RINA network management aspects that will not only provide

employment opportunities within the EU, but also holds export potential in the form of RINA-based products and consultancy services provided from within the EU to countries outside of the EU with large technology markets such as China, Japan & the U.S.

Deployment costs; It is anticipated that there are some initial adoption costs that will be incurred in deployment of a clean-slate network architecture such as RINA. These adoption costs are varied and may impact on some or all the stakeholders involved. The deployment of RINA may require hardware changes in the form of firmware upgrades to routers and switches (or replacement of older hardware that cannot be firmware upgraded) and would most definitely require some form of software upgrade to the routers and switches. The most significant portion of the labour costs is the cost of training in the use of the RINA architecture. This can be regarded as a once off cost or as an investment in a future technology. The NRENs have a lower relative cost than ISPs as NRENs deal with tech-aware organisations rather than individual customers, who may not be technology aware and need more support.

2.5.3 Social factors

There are a number of social factors that may inadvertently impact on the adoption of RINA. These social factors include levels of education, demand for improved security and inbuilt QoS, chicken and egg problem, network effect, and mobility.

Levels of education; The level of awareness of a nascent network technology within the network operators, operating system providers, service/content providers and device or equipment providers will be a contributing factor. For a complete success, all stakeholders in a network need to be aware of RINA and its relative advantages over the existing TCP/IP stack. The RINA community through the PSOC organised open workshops, where interested parties can come and see the progress of RINA as a revolutionary, clean-slate network implementation. In addition, academic publications are made on aspects of RINA and its approach, to educate the stakeholders of the TCP/IP alternatives and relative benefits of RINA, as the leading contender.

Demand for improved security and inbuilt QoS; It is to be expected that once the consumer, regulatory bodies and industrial fora, become aware of RINA's technical benefits of improved security and inbuilt QoS, they may demand these features. There is a large demand for improvements in network technologies with regard to the security of the systems using it. RINA offers an explicit enrolment stage on initial attachment, so allows for explicit access controls on who can connect to or gain access to the network. In addition it adds configurable packet protection on flows of information making it highly flexible for offering appropriately secured solutions for today's services. The inbuilt QoS support means that the delivery of these services is more consistent and reliable and it is expected consumers will like this certainty.

Chicken and egg problem; One of the key social problems with the uptake of any new networking technology is the "chicken and egg" problem. Equipment providers are not going to include RINA unless there is a consumer demand for it. Consumers through the devices they use cannot use RINA without the assistance of equipment and network providers. Finally, content providers would like to use RINA improved QoS, security and mobility aspects; however, cannot use RINA until the equipment providers provide RINA to consumers, and network operators. Unfortunately, this is one of the biggest causes for inertia within the networking domain.

Network effect; In the simplest form, the value of a product or service is dependent on the number of others using it. This principle is known as Metcalfes law[52] and applies to networking technologies. Metcalfe (working for 3Com) used this principle to describe the uptake of Ethernet. RINA has a very similar position within the networking ecosystem. While it does not require a hardware upgrade to achieve, it will require a software upgrade on most network connected devices. Thus RINA needs to achieve a critical mass, to become a global success.

Mobility; Mobility, gives consumers the ability to roam, and still access the services they demand. There has been a huge growth in the demand for services that can be consumed on mobile devices[50] (see Figure 3). Existing networking technologies use a variety of bespoke techniques (SIP, HIP, Loc/ID split, MIPv6) to cope with the provision of data transport in mobile environments. The management of these protocols adds to the overall complexity of managing mobile networks and therefore adds cost. The complex mix of technologies also adds delays in the perceived response time by the consumer. Within RINA, changing the point of attachment of a mobile device as it roams, is as simple as adjusting the routing table between two "layers" of the network stack. The consistency and simplicity of the approach, will lead to faster switching between points of attachment, and a better overall experience for the end user.

Industry trends: network usage; In general terms the proportion of network data has been increasing year on year. However, Ericsson[53] has been monitoring the proportion of traffic on its networks due to mobile data as opposed to voice. It is interesting to note that the proportion of voice data has only been slightly increasing in the last four years. However, mobile data, i.e. data used from smartphones has seen an exponential growth. This is reflected in other statistics and gives the overview that network traffic is increasing, it is becoming more mobile focussed, and the characteristics of the data are changing. RINA is in a good position with its inbuilt QoS and mobility support to take advantage of these trends.

[52] "It's All In Your Head". Forbes. 2007-05-07. Retrieved 2010-12-10.

[53] Ericsson, "Ericsson Mobility Report", November 2013, Available at: http://www.ericsson.com/res/docs/2013/ericsson-mobility-report-november-2013.pdf ,[05/05/14].

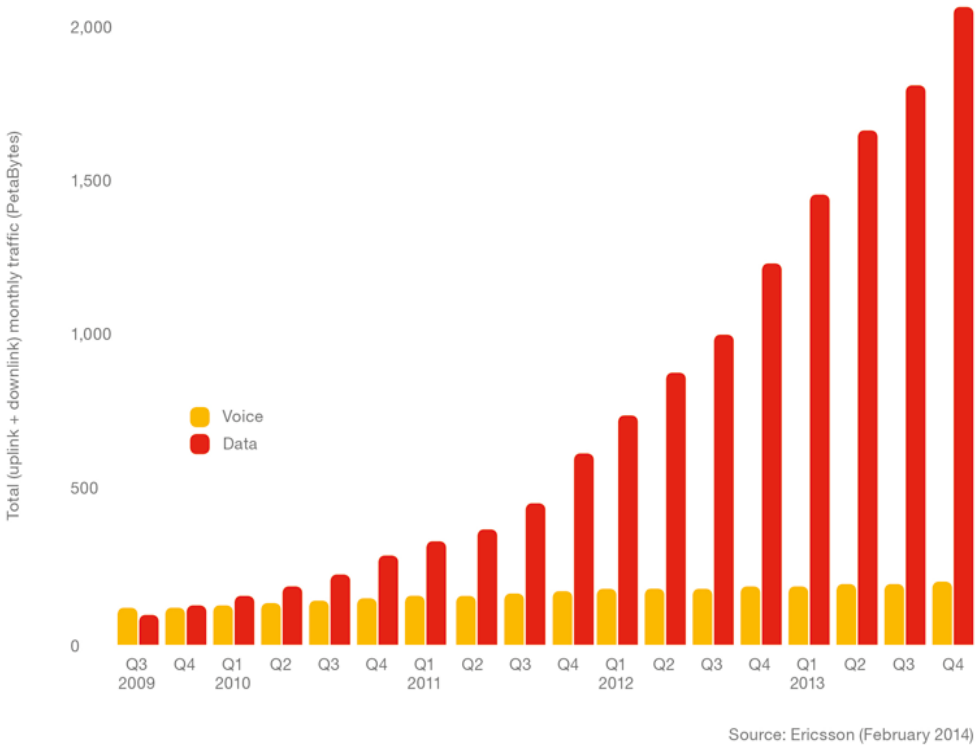


Figure 3: Monthly Traffic

2.5.4 Technological factors

There are a number of technological factors that could affect the deployment of RINA in the NREN scenario. These technological factors include investment in research and development, new technologies and inventions, internet and e-commerce developments, developments in production technology and rates of obsolescence.

Investment in research and development; The E.U. funds a number of related research project in the area: TRILOGY2, COSIGN, T-NOVA, UNIFY, GreenICN and STRAUSS projects. At the time of writing these projects are currently funded to address some of the same network challenges as RINA, but do not offer a complete networking solution to the inherent problems, (e.g. QoS, mobility, multi-homing, scalability, security). The IRATI project will provide an architecture reference model and specifications of RINA that will enable deployable implementations in production scenarios. Currently, the PRISTINE project is investigating the innovative internals of RINA. In regards, to future investment in research and development the EU is currently providing a new round of research and innovation funding called

H2020. The aim of H2020 is to achieve global competitiveness for Europe[54]. As part of this initiative, H2020 recently closed a call for project proposals on smart networks and novel internet architectures.

New technologies and inventions; There are a number of alternative network architectures (e.g. ICN for content networking) that could be used by current NRENs and GÉANT. However, these alternative network architectures provide specific solutions to only particular aspects of the overall inherent networking problems. These alternative network architectures have been previously highlighted in the IRINA SoTA document and summarised in the SWOT analysis. NRENs are investing in IPv6 technologies at present that may present some resistance for NRENs to switch to a clean-slate network architecture. However, the transition period has taken a long time and IPv6 still suffers from the underlying problems associated with TCP/IP based networks. The transition period for RINA is expected to be relatively short due to the fact it can operate alongside TCP/IP network architectures during the transition period.

Internet and e-commerce developments; The Internet is becoming more pervasive with advanced mobile devices and services being added at a rate set to increase exponentially over the coming years. This rate of growth can have a severe impact on the SLAs of service offerings provided by NRENs if the underlying network architecture cannot support the level and types of services required. Network as a Service (NaaS) gives cloud users the capacity to reserve bandwidth on the network dynamically. The main advantage is to give clients access to virtual routers that are physically located at a central location, but could be used by clients as if they were located on their own premises. RINA is a well suited solution for the resource requirements of these distributed applications due to its inherent support for mobility, security and QoS.

Developments in production technology; There are a number of network service providers that could potentially provide similar aspects of required services to NRENs, but not a complete architectural solution. Lately, there has been an increase in the number of NRENs utilising testbeds for experimentation with Software-Defined-Networking (SDN) technologies that enables the creation of virtual networks using the underlying infrastructure with varying characteristics and topologies, adapted to user-needs. RINA can support the utilisation of these testbeds by the various NRENs under very different use case scenarios that are not possible when accessing these testbeds over standard TCP/IP networks.

Rates of obsolescence; Internet architectures in general have a slow rate of change and obsolescence. This is due to their inherent heterogeneity and the length of time required to reach agreement by standards bodies when introducing new technological advancements in the area. The current Internet architecture has been in place for the past 40 years, but requires a radical change in order to support required services into the future. It is expected that if the RINA architecture were adopted by the NRENs, it would facilitate the delivery of current and future NREN services well into the future.

[54] "What is Horizon 2020? - European Commission." 2013. 29 Apr. 2014 <http://ec.europa.eu/programmes/horizon2020/en/what-horizon-2020>

3 NREN requirements

This section outlines general NREN requirements that have been subdivided into two sections, namely service requirements and technical requirements. Some typical NREN service expectations include Network as a Service (NaaS), security, authentication, collaboration tools, multimedia content repositories and eLearning activities. While Technical requirements includes QoS, network virtualization, mobility, multi-homing, scalability, security and network management.

3.1 Services requirements

This section examines the types of services currently offered by NRENs facilitating requirements that many research projects within the NREN community have described which are broad-ranging. These requirements can all be accommodated using current and predicted technology. Some requirements are

- Multiple access of databases from a diverse population of users
- Concentrated flows between key project locations
- Real-time performance requirements

Many research projects have concerns about the quality of service, service level agreement and user support available from the research networking community (in contrast to the level of service provider by a commercial provider). The Trans-European Research and Education Networking Association (TERENA)[55] offers a forum to collaborate, innovate and share knowledge in order to foster the development of Internet technology, infrastructure and services to be used by the research and education community. As part of its work TERENA organises conferences and other networking events for the NRENs. This section bases its observations on the services captured by TERENA in its survey[56] and augments them with the results of our own NREN environment survey[57].

3.1.1 Security Services

Within the scope of security, the NRENs offer multiple key services. The following sub-sections describe the most important ones:

[55] "TERENA." 2003. 24 Feb. 2014 <http://www.terena.org/>

[56] "TERENA." 2013 28 Jan. 2014 <http://www.terena.org/>

[57] NREN Environment Survey, IRINA deliverable 2.1, 2014.

Incident Reporting

This is a mechanism to report Cyber-attacks and vulnerabilities found in existing software. One of the key problems in this area is the lack of an agreed standard for exchanging computer incident information. From the survey, X-ARF is the most commonly used, but not the only format. This service would still be required in a RINA network. Its primary purpose is to report issues in applications making use of the network.

Honey pots

Honey pots are provided as a means to detect unwanted or unusual network traffic. Specifically, they assist in the detection and identification of malware or certain forms of attacks. They do so by "pretending" to be a legitimate network application service, but instead log all access attempts made to them. A secondary purpose is to drive intelligence gathering and security research, i.e. to collect potentially interesting information to identify what kinds of attacks are being attempted. This service would still be required in a RINA network. Its primary purpose is to detect hacking attempts, and malware in applications making use of the network. However, tracking hacking attacks back to the original source would be more difficult within a RINA network due to the use of DIFs.

Firewalls

Firewalls are used at both a network level and also at a deeper application level. The primary purpose is to restrict access to application services running on an internal network, from a wider area network, normally the Internet. Some of the opportunities in an IP based network for attacks, are mitigated by RINA's explicit resource allocation step. Flows can be rejected, as RINA has an explicit flow setup request and response phase in the flow setup.

Anti Spam measures

Some of the services deployed to enforce some anti-spam measures within the network include:

- DNS blacklisting, e.g. known spamming servers are prevented access.
- Whitelisting, e.g. SPF records, used to identify servers that legitimately should be relaying email, as opposed to IP addresses that should not be relaying email.
- Filtering and Antivirus scanning. The most common service used is SpamAssassin, or some other spam filtering software.

Some of the vulnerabilities in IP networks are mitigated by RINA's explicit resource allocation step. RINA DIFs can be configured to require strong authentication between application processes, before flows are accepted. Blacklisting and whitelisting services can still be employed within the scope of RINA to augment RINA, thereby providing greater confidence that legitimate source traffic is identified correctly. Unfortunately, security breaches within the applications themselves can still occur regardless of the underlying architecture, hence the need to still employ application-level security measures (i.e. antivirus).

Authentication

The sensitive nature of the data to be stored leads to a need for a fine grained, distributed Authentication and Authorization (AA) system, it also is virtually imperative that such a system provides Single Sign On (SSO) functionality. There has been some involvement with the related work on AA Infrastructure in the GÉANT project. The scale is thousands of data sets in many countries. In

some projects, such as the surveys of SHARE, the security and confidentiality of the data is vital— for example, when it contains blood samples, DNA, wealth and social security data[58].

DNSSEC

To address weaknesses in the IPv4 and IPv6 DNS address resolution mechanisms, NRENs are currently deploying DNSSEC. DNS covers the resolution of a URL to an IP address; however there is little protection from malicious users running competitive or contradicting DNS services. This allows users to insert incorrect "service" records into the DNS registry (redirecting the user to another site or server). However, in RINA the structure naturally forms a securable container where applications can be required to authenticate with each other before being able to exchange information.

3.1.2 Network collaboration tools

NREN users need access to a range of network collaboration tools to support their teaching, learning and research activities. Over the last decade, network collaboration tools and their associated services have become the mainstay of collaboration among European researchers and providers of higher education. Even though collaboration hardware and software has not advanced significantly in recent years, there has been a significant increase in quality combined with a marked reduction in prices that has made services such as network-based virtual meetings more efficient and cost effective. In key areas of research and education, network collaboration tools are playing a pivotal role in making project, research and administration work more effective, by facilitating the connection of remotely located personnel. Facilitating remote collaboration optimises work time, reduces travel costs, and lowers the environmental impact of travel. Four facilitators of NREN network collaboration include:

1. Numbering schemes and Voice over IP (VoIP) to connect institutional IP telephony deployments or, to a lesser extent, individual end-users. For VoIP to be a realistic replacement for standard public switched telephone network (PSTN) telephony services, customers need to receive the same quality of voice transmission they receive with basic telephone services—meaning consistently high-quality voice transmissions. Like other real-time applications, VoIP is extremely bandwidth- and delay-sensitive. For VoIP transmissions to be intelligible to the receiver, voice packets should not be dropped, excessively delayed, or suffer varying delay (otherwise known as jitter). For example, the following standards must be met[59]:
    a. The default G.729 codec requires packet loss far less than 1 percent to avoid audible errors. Ideally, there should be no packet loss for VoIP.
    b. The ITU G.114 specification recommends less than 150 millisecond (ms) one-way end-to-end delay for high-quality real-time traffic such as voice. (For international calls, one-way delay up to 300 ms is acceptable, especially for satellite transmission. This one-way delay takes propagation delay into consideration—the time required for the signal to travel the distance.)

[58] Commands, SR. "Deliverables - GN3 - GÉANT." 2013.
<http://GÉANT.archive.GÉANT.net/Media_Centre/Media_Library/Pages/Deliverables.aspx>

[59] "Quality of Service Design Overview > QoS ... - Cisco Press." 2004. 23 Feb. 2014.
<http://www.ciscopress.com/articles/article.asp?p=357102>

c. Jitter buffers (used to compensate for varying delay) further add to the end-to-end delay, and are usually only effective on delay variations less than 100 ms. Jitter must therefore be minimized.

VoIP can guarantee high-quality voice transmission only if the voice packets, for both the signalling and audio channel, are given priority over other kinds of network traffic. For VoIP to be deployed so that users receive an acceptable level of voice quality, VoIP traffic must be guaranteed certain compensating bandwidth, latency, and jitter requirements. QoS ensures that VoIP voice packets receive the preferential treatment they require. In general, QoS provides better (and more predictable) network service by providing the following features[60]:

a. Supporting dedicated bandwidth
b. Improving loss characteristics
c. Avoiding and managing network congestion
d. Shaping network traffic
e. Setting traffic priorities across the network

2. Video- and web-conferencing tools should provide a high-quality audio/video-based collaboration environment, often enhanced by other tools enabling joint work. NRENs provide a centrally managed video-conferencing service which indicates the strategic importance of video- and web-conferencing. In order to interact effectively, participants in a video/web conferencing system need to be able to communicate in soft real-time, or as close to real-time as possible. Significant processing is required within the video/web conferencing endpoints, to reduce the latency of processes that compress the raw audio and video into a data stream which is then sent across the network. However, there is a limit as to how much the endpoints can deal with traffic/packet loss or delays across the intervening networks – hence the requirement to provide some protection to the traffic flowing between the endpoints. The metrics used to determine how well a network is performing are:

a. latency (end-to-end delay)
b. packet loss
c. IPDV (Inter Packet Delay Variation) / Jitter
d. bandwidth

- The European Maritime Safety Agency (EMSA)[61] developed the CleanSeaNet[62] service, a satellite-based monitoring system for marine oil spill surveillance and detection in European waters. The service provides rapid delivery of oil spill alert information and images, using radar satellite imagery acquired by SAR satellites, to Member States. Identification of potential spills in near real-time is essential and EMSA has contracted a network of ground stations throughout Europe, able to downlink, process and analyse satellite data within a maximum of 30 minutes after satellite overpass. These images of the sea and coastline around Europe are sent in real-time from the ground stations to EMSA premises in Lisbon, Portugal using the GÉANT-NREN network. The transmission of one image (~500 Mbytes) from a ground station to EMSA should not last longer than one minute. This indicates a required bandwidth of 66

[60] Goode, Bur. "Voice over internet protocol (VoIP)." *Proceedings of the IEEE* 90.9 (2002): 1495-1517.

[61] "European Maritime Safety Agency (EMSA) - Quality Shipping, Safer ..." 2006. 25 Mar. 2014 <http://www.emsa.europa.eu/>

[62] "Satellite oil spill monitoring (CleanSeaNet) - EMSA - European ..." 2012. 25 Mar. 2014 <http://emsa.europa.eu/operations/cleanseanet.html>

Mb/s. EMSA has a 1 Gigabit connection to FCCN the NREN in Portugal. The protocols required to achieve this include VPN, sFTP and a two-way secure SSL - HTTP protocol[63].

- EUMETSAT[64] operate a number of satellites collecting earth observation data and images. The data from the satellites is received by a number of ground stations and sent to EUMETSAT for processing; the resulting data is then disseminated as various data products to users and the research community using commercial telecommunication satellite links, in a similar way to satellite television. There are a wide range of applications for the data products including numerical weather predictions, climate modelling, land surface analysis, Ozone and atmospheric chemistry monitoring, hydrology and water management, and active fire monitoring. There is a continuous flow of data from acquisition, processing to dissemination. The timing is critical as the information must be with the consumers within five minutes of the images being taken. The current satellite dissemination system is based on standard Digital Video Broadcast (DVB) technology with a total bandwidth of about 15 Mb/s. However with new image systems and more detailed products, EUMETSAT expect to require between 100 and 400 Mb/s. Hence they are interested in complementary use of satellite and the NREN network[65].
- To predict the weather, modern meteorology depends upon near instantaneous exchange of weather information across the entire globe; this is coordinated by the World Meteorological Organization (WMO)[66]. The Meteorological offices and associated institutes in each country collect local data and run simulation models to predict the weather patterns on various timescales. Some run climate models. The output, data, together with the corresponding metadata, is made available to the local regional Global Information System Centre (GISC), which is linked via the WMO Weather Information System (WIS) to all the GISCs worldwide. Metadata (and data) updates are synchronised between all the GISCs so the weather information is available throughout the world. High priority messages such as typhoon warnings are also exchanged between the GISCs. Many of the computations must be performed in a timely manner, such as the hourly weather synopses. The current bandwidths used on the commercial networks are a few Mb/s, which is sufficient for the current small text-based data /metadata exchanges. However, with new and more detailed data and images about 100Mb/s or up to 1 Gb/s might be required into the future[67].

3. Group collaboration services; i.e. the bundling of services that allow collaborative groups to form and work together easily, independent of their location. Collaborative groups, sometimes referred to as virtual organisations, can serve individuals from more than one home institution, so the group is not bound to a single institution. There has been considerable growth in this area since 2011. NRENs currently offer a platform of bundled services for collaborative groups of users. These services are federated, allowing access to

[63] Commands, SR. "Deliverables - GN3 - GÉANT." 2013. <http://GÉANT.archive.GÉANT.net/Media_Centre/Media_Library/Pages/Deliverables.aspx>

[64] "Welcome to EUMETSAT — EUMETSAT." 2010. 25 Mar. 2014 <https://www.eumetsat.int/>

[65] Commands, SR. "Deliverables - GN3 - GÉANT." 2013. <http://GÉANT.archive.GÉANT.net/Media_Centre/Media_Library/Pages/Deliverables.aspx>

[66] "WMO: World Meteorological Organization Homepage." 2003. 25 Mar. 2014 <http://www.wmo.int/>

[67] Commands, SR. "Deliverables - GN3 - GÉANT." 2013. <http://GÉANT.archive.GÉANT.net/Media_Centre/Media_Library/Pages/Deliverables.aspx>

them through a web-based authentication scheme. The most common bundled services include mailing lists, a wiki, a document store and calendar/appointment planning.

4. Multimedia content repositories for online presentation of materials recorded by higher education and research organisations to complement remote teaching/learning and science dissemination. The use of multimedia content repositories (i.e. audio/video archives) and the streaming services they offer is increasing. The repository sizes vary enormously, from a few Gigabytes to 250 Terabytes (Finland). A number of NRENs provide video-sharing functionality, which enables the user community to publish and manage the content they themselves have created. A fewer number of repository providers are able to exchange metadata with other content aggregators, but many of them plan to implement this capability in the near future. Similarly, user-initiated live streaming support is not yet a common functionality, but this is being planned by some NRENs. Almost half of the NRENs already offer a web-based multimedia content repository for storage and retrieval of audio/video recordings created by research and higher-education communities. Many of them also feature or plan to introduce video-sharing functionality that enables direct content management by the end-user.

3.1.3 Network e- Science resources

Across the various NREN research groups several common service requirements can be identified that include storage, grids, networks and general computing. The current data scales are petabytes moving to exabytes, the data growth has been exponential for many years, and has exceeded improvements in CPU and storage. There is a need to ensure that the major research centres have network connections with sufficient bandwidth, or point-to-point circuits where appropriate, to allow the required exchange of data to occur in a timely manner. In many cases, the NRENs provide the networking infrastructure for such services and are expanding into offering additional services to the Grid community. The activities of the environmental sciences research groups has quite varied networking requirements with data moving needs varying from bulk data transfer, through to collecting data from intelligent networks of autonomous sensors and observatories to quasi real-time requirements such as the operation of remote and/or mobile instruments and observatories. Significant investment is proposed in both fixed and mobile server systems, collecting data from land, sea and air measurements, using fixed and mobile data collection. The ICT challenges associated with the sector include data capture, particularly from sensor networks, and the combining, processing and storage of large and complex data sets. There are also some significant real-time requirements in terms of collecting and processing data[68].

1. Grid middleware: Grid services have become an important area for NRENs. Projects and organisations such as the new European Grid Infrastructure[69] aim to introduce a production Grid service for scientific research purposes, using distributed computing services. In many cases, the NRENs provide the networking infrastructure for such services and are expanding into the offering of additional services to the Grid community. In almost all cases, these services are international in geographical scope. There are various types of Grid services such as dedicated optical paths, dedicated point-to-point IP circuits, storage facilities or

[68] Commands, SR. "Deliverables - GN3 - GÉANT." 2013.
<http://GÉANT.archive.GÉANT.net/Media_Centre/Media_Library/Pages/Deliverables.aspx>

[69] Available online at: <www.egi.eu>

computation power (CPUs)[70][71]. The ability to use grid or cloud computing facilities for the processing of stored data is also foreseen as paramount importance in some projects. For example, the environmental sciences research groups are increasingly data rich and, as a data-driven science, will benefit from data services and computation in data analysis. The analytical and modelling platforms need both high-performance computing and distributed grid or cloud computing, implying an underlying requirement for efficient high-performance reliable networking.

2. Computing power: Cloud computing is a paradigm shift in how data centres and service providers are architecting and delivering highly reliable, highly scalable services to their users in a manner that is significantly more agile and cost effective than previous models. This new model offers early adopters the ability to quickly realize the benefits of improved business agility, faster time to market and an overall reduction in capital expenditures. However, enterprises and service providers need to understand what elements their cloud must contain in order to build a truly successful cloud. Cloud services are not yet as common as Grid services. Some NRENs already offer virtualisation services with other NRENs planning to introduce the service. Cloud services are usually managed through some kind of virtual management interface[72].
3. Heterogeneous Systems Support: Service providers and enterprises have requirements around both commodity and proprietary systems when building out their clouds. Not only should cloud management solutions leverage the latest hardware, virtualization and software solutions, but they should also support a data centre's existing infrastructure. Cloud management providers must integrate with traditional IT systems in order to truly meet the requirements of the data centre.
4. Service Management: To fully realise cloud computing functionality in a product offering, it is important that administrators have a simple tool for defining and metering service offerings. A service offering is a quantified set of services and applications that end users can consume through the provider — whether the cloud is private or public. Service offerings should include resource guarantees, metering rules, resource management and billing cycles.
5. Dynamic Workload and Resource Management: In order for a cloud to be truly on-demand and elastic while consistently able to meet consumer service level agreements (SLAs), the cloud must be workload- and resource- aware. Cloud computing raises the level of abstraction to make all components of the data centre virtualized, not just compute and memory. Once abstracted and deployed, it is critical that management solutions have the ability to create policies around workload and data management to ensure that maximum efficiency and performance is delivered to the system running in the cloud. This becomes even more critical as systems hit peak demand. The system must be able to dynamically prioritize systems and resources on-the-fly based on business priorities of the various workloads to ensure that SLAs are met.
6. Reliability, Availability and Security: To be fully reliable and available, the cloud needs to be able to continue to operate while data remains intact in the virtual data centre regardless if a failure occurs in one or more components. Additionally, since most cloud architectures deal

[70] Asadzadeh, Parvin et al. "Global grids and software toolkits: A study of four grid middleware technologies." *arXiv preprint cs/0407001* (2004).

[71] von Laszewski, Gregor, and Kaizar Amin. "Grid middleware." *Middleware for Communications* (2004): 109-130.

[72] "7 Requirements for Building Your Cloud Infrastructure - CIO.com." 23 Feb. 2014 <http://www.cio.com/article/648465/7_Requirements_for_Building_Your_Cloud_Infrastructure>

with shared resource pools across multiple groups both internal and external, security and multi-tenancy must be integrated into every aspect of an operational architecture and process. Services need to be able to provide access to only authorized users and in this shared resource pool model the users need to be able to trust that their data and applications are secure.

7. Integration with Data Centre Management Tools: Within most data centres, a variety of tools are used for provisioning, customer care, billing, systems management, directory, security and much more. Cloud computing management solutions do not replace these tools and it is important that there are open application programming interfaces (APIs) that integrate into existing operation, administration, maintenance and provisioning systems (OAM&P) out of the box. These include both current virtualization tools from VMware and Citrix, but also the larger data centre management tools from companies like IBM and HP.
8. Visibility and Reporting: The need to manage cloud services from a performance, service level, and reporting perspective becomes paramount to the success of the deployment of the service. Without strong visibility and reporting mechanisms the management of customer service levels, system performance, compliance and billing becomes increasingly difficult. Data centre operations have the requirement of having real-time visibility and reporting capabilities within the cloud environment to ensure compliance, security, billing and chargebacks as well as other instruments, which require high levels of granular visibility and reporting.
9. Administrator, Developer and End User Interfaces: One of the primary attributes and successes of existing cloud-based services on the market comes from the fact that self-service portals and deployment models shield the complexity of the cloud service from the end user. This helps by driving adoption and by decreasing operating costs as the majority of the management is offloaded to the end user. Within the self-service portal, the consumer of the service should be able to manage their own virtual data centre, create and launch templates, manage their virtual storage, compute and network resources and access image libraries to get their services up and running quickly. Similarly, administrator interfaces must provide a single pane view into all of the physical resources, virtual machine instances, templates, service offerings, and multiple cloud users. On top of core interfaces, all of these features need to be interchangeable to developers and third parties through common APIs.
10. Storage facilities are required by research groups in the social sciences and humanities areas that have the need for both long-term storage as well as short term usage of their databases and archives; however the detailed requirements may well be different. Typical sizes of the database for the SHARE[73] project are a few hundred Gigabytes. For the CLARIN project[74], 50 to 100 TBytes would be the typical size of the data for language material. With the exception of the linguistic computations of CLARIN, this area has light-weight, computing storage and networking requirements. There is a strong need for good access to the data archives by the academic and research user community, usually via web-based transactions. Data from these archives is made available to a worldwide community of millions via the web, implying the need for a very good routed IP service. Similarly, database information resources are located

[73] "The Survey of Health, Ageing and Retirement in Europe (SHARE), 2002. 25 Mar. 2014.
<http://www.share-project.org/>

[74]"Common Language Resources and Technology (EUDAT), 2011. 25 Mar. 2014.
<http://www.eudat.eu/common-language-resources-and-technology-infrastructure>

in the US and Japan and there is a requirement to exchange data updates on a daily basis. Some notable storage services offered include:

a. Distributed storage specifically for Grid users;
b. Distributed storage for any NREN users;
c. Dedicated/special high-level connectivity to commercial-content servers or commercial content;
d. Hosting of commercial-content servers or appropriate commercial content on the NREN network;
e. Video servers for use by NREN sites;
f. Mirroring of content from outside the NREN network

Among this list mirroring is the most common type of storage facility offered. The most common model is that these storage services are not provided by the NREN but by one or more individual institutions, often in collaboration with the NREN. Radio telescopes are a specific type of e-Science resource for which several NRENs provide connectivity. Such instruments pose special challenges: they are often located in remote areas yet require high-capacity connections due to the huge amounts of data generated[75]. The amount of data collected and stored is increasing exponentially, along with the related need for the bandwidth to transport the data in order to make it available to researchers and users. Because the capacity, throughput, jitter, and delay requirements of the network can be stringent, commercial network providers cannot make these connections available quickly and at an affordable price. NRENs, GÉANT and others involved in providing network connectivity need to collaborate with the user communities to ensure that the networking requirements associated with the deluge of data are well understood. Adequate network services need to be put into place in a timely and economically viable manner. Aspects, such as speed of provision, throughput, privacy, persistence of connection, security and other important parameters need to be addressed.

3.1.4 e-Learning

e-Learning refers to the use of electronic media and information and communications technologies (ICT) in education. e-Learning includes numerous types of media that deliver text, audio, images, animation, and streaming audio, and includes technology applications and processes such as audio and video tape, satellite T.V., CD-ROM, and computer-based learning, as well as local intranet/extranet and web-based learning. Various technologies are used to facilitate e-Learning. Most e-Learning uses combinations of these techniques, including blogs, collaborative software, and virtual learning environments (VLE), also known as learning platforms, utilise virtual classrooms and meetings which often use a mix of communication technologies. A_collaborative software (LMS) is software used for delivering, tracking and managing training and education; for example, tracking attendance, time on task, and student progress.
e-Learning users interact directly through the use of systems such as portals, learning delivery systems, authoring tools, administration interfaces and so on. System functionality required by the end users include retrieving learner information, storing content in a repository, or collaboration services that support the use of collaborative tools and provide services for the creation and management of collaborative sessions, including both synchronous and asynchronous communication as appropriate

[75] Commands, SR. "Deliverables - GN3 - GÉANT." 2013.
<http://GÉANT.archive.GÉANT.net/Media_Centre/Media_Library/Pages/Deliverables.aspx>

(e.g. the use of instant messaging, audio, video, etc) . End users require some common services provided by lower-level functionality which is not education-specific, such as authentication and authorization services. Infrastructure is the underlying network, storage, and processing capability provided for an implementation. Examples of Internet-based learning management systems include Blackboard Inc. and Moodle. These types of LMS allow educators to run a learning system partially or fully online, asynchronously or synchronously. Moodle provides blended learning opportunities as well as platforms for distance learning courses.

3.2 Technical Requirements

Many of the technical requirements imposed to the NREN's infrastructures in order to provide their services are related with characteristics of the traffic produced by the applications. Some others involve network capabilities to be supported by the architecture. The following sections focus on specific technical requirements generated by the provided services and that should be addressed by the NREN's network architecture.

3.2.1 Quality of Service

Distributed resource allocation in networks suffers from a fixed stack and the lack of information from applications. The fixed stack mandates that resource allocation is done once for the full Internet, using the same mechanisms which have to be deployed in all the routers in the path of a packet. Currently, built-in mechanisms that allow the same network to really provide different QoS levels do not exist. Applications have no way of asking for certain QoS parameters, since the sockets API only allows to specify a reliable (TCP) or unreliable (UDP) transport service. This fact does not enable the network to dynamically modify the allocation of the resources accordingly. As a consequence today separate networks that use separate resource allocation policies have to be used in order to support different types of applications with strict quality requirements: networks dedicated to support VoIP, Video on Demand, Data, etc.
The set of applications that provides the different services studied before have very different needs in relation with the QoS to be guaranteed for their traffic. From wiki and blogging, to video-conferencing passing through VoIP there is a huge variation range for the different parameters that shape the QoS (e.g. bandwidth, packet-loss, latency, jitter). Nowadays, the most common technique is to over-allocate physical resources beforehand in order to guarantee certain levels of quality for specific applications, where the whole network must be configured in the same way, generating a larger waste of resources. The repeating structure of layers in RINA, together with the capability of a DIF to be configured to support different kind of QoS for each one of its flows, makes RINA a logic solution to provide multiple QoS services. This means not only that two applications may have different QoS, but that the same application can ask for different QoS (i.e. for a videoconference application, one audio flow and one video flow with different QoS).

3.2.2 Network Virtualisation

Network virtualization is a concept which has attracted interest in the last decade, especially since the rise of computing virtualization (hypervisors, virtual machines, etc.). Network virtualization resembles

computing virtualization concept and such tries to arbitrate the sharing of the network resources in a transparent way for the different applications which use the network. This has been of special interest by ISPs and DCs, but also for the NRENs in order to optimize resource allocation and support services as IaaS, NaaS or Cloud computing.
The reality is that protocols, and in particular the TCP/IP protocols, were designed in the early and mid 70's without having virtualization in mind. The theoretical layered L1/L2/L3/L4 model has proven not to be suitable for virtualization, leading to L2 over L2, L2 over L3, L2 over L4 or L3 over L3 network protocols (i.e. VXLAN[76], NVGRE[77], STT[78]). Management task of even standard protocols has become a complicated task, increased by these non-standard protocols that generate an important overhead for network operators and Data Centre providers, due to its complexity and vendor specific semantics as well as creating vendor lock-in situations.
In RINA, each DIF has all the functionalities needed to provide networking. Two different DIFs mean two different scopes. It is a policy what determines the specific behaviour of a DIF. Virtualization is unneeded in RINA in the sense that a DIF can act as a virtualized network itself by setting the specific policies associated to an expected behaviour. In addition, the recursive architecture formed by an unique type of DIF avoid the cross layer problem previously described.

3.2.3 Mobility and Multi-homing

Increasingly, the Internet is being used as a mechanism for delivering a range of services to specific user-groups. Thus, user access to services is becoming less dependent on the physical location either of the user or of the service. The research and education community is at the forefront of this development. Security is a key issue in this area: it is important to know who wants to access a particular service and who is entitled to do what. This means that authentication and mobility services go hand in hand and that the development of these services can either constrain or stimulate the way other services are developed and delivered to users.
Multi-homing and mobility, which can be seen as dynamic multi-homing, require a complete naming and addressing schema, which is not had today (IPs addresses name the interface but not the node) that differentiates nodes from PoAs (IPs), identifies applications without falling on default names or locations ("well-known ports") and optimizes routing and its routing table's size.
RINA's naming structure is based on Saltzer's OS model[79], where Application names are location-independent to allow an application to move around, node addresses are location-dependent but route-independent. PoAs addresses are by nature route-dependent. Moreover, since RINA is structured in recursive layers, an interesting observation can be made: mapping application names to node addresses is the same mapping than mapping node addresses to PoAs. In other words, for any

[76] K. Duda P. Agarwal L. Kreeger T. Sridhar M. Bursell C. Wright M. Mahalingam, D. Dutt. VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks. Internet-Draft draft-mahalingam-dutt-dcops-vxlan, IETF Secretariat, February 2013.

[77] I. Ganga A. Greenberg G. Lin M. Pearson P. Thaler C. Tumuluri N. Venkataramiah Y. Wang M. Sridharan, K. Duda. NVGRE: Network Virtualization using Generic Routing Encapsulation. Internet-Draft draft-sridharan-virtualization-nvgre-00, IETF Secretariat, September 2011.

[78] J. Gross B. Davie. A Stateless Transport Tunneling Protocol for Network Virtualization (STT). Internet-Draft draft-davie-stt-01, IETF Secretariat, 2012.

[79] J. Saltzer. On the Naming and Binding of Network Destinations. RFC 1498 (Informational), August 1993.

layer N, nodes at the layer N+1 are applications and nodes at the layer N-1 are points of attachment, making this relationship relative.
VMs mobility within a Data Centre in RINA is easy achievable, the only requirement is to make the scope of the DIF to arrive to the new destination server by creating and enrolling a new IPC Process. Networking configuration of the VMs will remain the same since they would continue being part of the same DIF. Another important point highlighted by mobility is the access to the NRENs network by the users using WLANs, which is a common practice (i.e. VPNs). For these reason a strict authorization protocol is required. In addition, following a standard approach would facilitate "roaming" among users of different NRENs when using services on not his own NREN's network.

3.2.4 Scalability

The infrastructure of each NREN varies noticeably depending on various factors, as country, users, services provided, etc. Despite this fact, we look for general solutions, reason why the alternative architectures for NREN's networks must scale and fit the characteristics of each individual infrastructure. Having this in mind, there are several areas where scalability becomes a key requirement:

Routing tables size

Nowadays there is an addressing and routing problem in Internet or bigger infrastructures composed by many networks that inter-communicate. The common point of all these networks is the use of TCP/IP communication protocols, no matter what underlying physical infrastructure is used, causing challenges in scalability, multi homing, and inter-domain traffic engineering.

The issue originates in a fixed stack: there's single network layer scope in the whole Internet based on the inter-domain Border Gateway Protocol (BGP) which is used to exchange routing information between Autonomous Systems (AS). The growth of multi homed hosts and networks along with the growth of internet caused a drastically enlargement in routing table's sizes and convergence times (due to the increase of routing update events and increased address sizes). Since 1990s, the router memory size and forwarding capacity have managed to follow the growth[80]. Although other solutions as Netmap or commodity HW solutions are rising, routers that can handle a million routes and relay packets at very high speed are expensive, and a hardware limit can be reached very soon.

Possible solutions to resolve this issue, that become scalability requirements are: i) to augment the number of routing scopes ii) to switch from the flat Internet routing model to hierarchical routing, separating edge networks from transit networks.

Force approaches

The ever-decreasing cost of networking components (memory, storage, CPUs, bandwidth) has created the habit of solving scalability problems using a brute-force approach: deploying more hardware to mitigate the issues. One of the major sources of inefficiency is the way congestion control is dealt with in current networks. Looking at data transfer, the Internet is one flat network at a planetary scale. Congestion control is just applied once, the control loop being located at TCP, at the endpoints of the network. Control theory says that for a control loop to be effective, the time-to-notify (that is the time

[80] V. Fuller. Scaling issues with routing and multihoming. IRTF RRG meeting, Mar 2007.

between an event is detected and the controller takes a corrective action) should be minimized: the controller should be as close as possible to the controlled entity. TCP does the opposite: it makes the control loop as long as possible, maximizing the time-to-notify and its variance[81]. Moreover, TCP's congestion control doesn't know when the network is congested; therefore it tries to sense congestion (by measuring packet loss and delay). As a result sometimes TCP confuses congestion by delay/loss produced by other factors. Last but not least, applications have no way of telling TCP what is their required sending rate; therefore TCP always tries to use as much bandwidth as it can, until it starts sensing packet loss and delay. All of these properties together results in well-known features: poor isolation of flows, which by design interfere with each other; wasted bandwidth; unpredictable network behaviour at high loads and unsatisfied application expectations (low throughput or high delay). The current approach to solving this issue is to add more bandwidth to networks, so that they keep operating on a predictable region of offered load.

3.2.5 Protocols and programmability

Traditional network architecture is static and requires significant operational investment to manage. Programmability within a network has several benefits that ease to achieve some of the services requirements:

- Reduced long-term costs.
- Ability for applications to maintain information about device capabilities.
- Ability for networks to respond to application status and resource requirements.
- Better allocation of bandwidth and resources.
- Packet prioritization for traffic shaping.
- Improved operational flexibility and enhanced transparency.
- Support for emerging privacy and security technologies.

In RINA, each DIF has all the functionalities needed to provide networking. Policies on these DIFs allow the programmability of their behaviour in all the points formerly described. Moreover, the number of protocols used is bounded. Any protocol can be implemented as a policy in a RINA DIF, which provides the architecture with an intrinsic flexibility as well as security, explained below.

3.2.6 Security

The network architecture should provide security mechanisms inherently. Moreover, such a security framework integrated in the network architecture itself must be usable by the applications. This would avoid changes in network configuration that break security aspects provided to the applications involving manual reconfiguration of the tools providing it (i.e. firewalls when the IP is changed). The current internet architecture was not designed with security purposes. This means that security in TCP/IP have been added as add-ons or separate protocols. This fact, along with the exposed addresses that anyone can see and the use of the well-known ports impose severe security risks. Moreover, this variety of protocols makes them more prone to contain errors or security bugs.
Understanding this situation helps to identify the following security requirements, or desirable capabilities:

[81] J. Day. How in the heck do you lose a layer!? Conference on the Network of the Future (NOF), pages 135–143, 2011.

- Build-in security by design in the network architecture
- A minimum set of robust and well know protocols
- Not well-known ports or public addresses

RINA provides them all. Each DIF can define its own security policies specifically configured to optimize its performance in the scope of the DIF. A DIF will never be able to access and compromise a different DIF since addresses and ports have sense only within a DIF so no information is exposed to others. Finally, RINA defines a small set of protocols used by all the DIFs, so their implementation can be focused and of high quality.

3.2.7 Network Management

In more traditional network management, network reconfigurations are manually performed by administrators based on collected statistics, alarms, or even by changing policies. Administrators have a narrow view of network events. It is difficult to understand large amounts of collected data from the network. This problem worsens as the data-set increases, when more mediation and correlation is required. Moreover, this information is not always public (i.e. between ISP providers or Telco operators).
The proliferation of specialized management solutions many times caused by the equipment manufacturer, tailored to a specific combination of layers in a specific domain is not a scalable approach: network management becomes complex, non-flexible and prone to errors when the network grows. The absence of common, but at the same time extensible, abstractions that capture the behaviour of a layer is key to achieve effective network management.
RINA organises the network in self-contained layers of different scopes. Network managers can have a clean view of one DIF, where traffic from other DIFs is not introducing noise to the analysis. Moreover, the self-contained property means that a DIF can contain all the necessary networking functionalities (e.g. allocation, routing). This property allows the network manager to have a general view of all these network functionalities. For scenarios where many DIFs must be managed, RINA conceives a DIF Management System as a centralized tool to perform management tasks over the systems of the network capable of making complex configuration changes affecting many layers at once and of optimizing the performance of a set of layers working together. The commonality provided by RINA allows multi-layer management to be vastly simplified; thus opening the door to more robust, dynamic, responsive and cheaper network management operations

3.3 NREN Survey

National Research and Education Networks (NRENs) have historically been leaders in applied research in networking. In the mid-1980s, NRENS were early adopters of technological innovations such as packet networking and data networking. Nowadays, packet based networking has become the standard for data networking, and commercial Internet Service Providers (ISPs) have adopted packet based networking.

Now, like in the mid-1980s, a shift in network architecture is proposed by the Recursive InterNetworking Architecture (RINA) community.

In the GN3+ OC IRINA project, a survey was held to gather the necessary information from the NRENs in order to assess the general requirements, the most demanding applications and key service parameters. Based on the survey results an accurate use case tailored to the NREN environment was derived (Section 4). This use case will then incubate two experiments, one reference scenario built on the current state-of-the-art in TCP/IP networking, and one experimental setup based on the IRATI RINA prototype (http://irati.eu). The results of these experiments will be compared to quantitatively assess how RINA stacks up against TCP/IP. By leveraging the benefits that RINA can bring, NRENs can, like in the mid-1980s, be the early adopters of a new network technology.

24 NRENs responded to the survey and results are summarised in three sections. The first section analyses the current topology and applications of a NREN. The second section summarises the future drivers for improvement in a set of future requirements and the last section summarises the expected impacts of these requirements. All gathered data has to be treated confidential, as such no specific organisations are named.

3.3.1 Analysis of current NREN topologies and applications

3.3.2 NREN topologies

In the TERENA Compendium of 2013 published by the GÉANT (GN3+) project it is stated that the typical core capacity of a NREN is now 10Gb/s[82]. Some NRENs have reached 20 or 40 Gb/s and Germany even has capacity of 100Gb/s. 10Gb/s is also the mean capacity up from 2Gb/s in 2008. In the next years it is expected that several NRENs will upgrade their capacity to 100Gb/s. The typical capacity of the links does not tell the whole story as the typical NREN networks form a mesh, with redundant core and access links.

As many NRENs have access to dark fibre, NRENs can increase capacity easily and economically when required. The aggregate length of dark fibre used internally by NRENs in the GÉANT region has increase by more than 10% compared to 2012. Cross-border dark fibre links between NRENs are also continuously developed.

The IRINA survey tries to map the current NREN topologies and applications in a set of questions with regard to the utilization of the available transport infrastructure. An NREN typically connects universities and research institutes to the Internet. But also libraries and museums, primary schools and museums and other institutions can be connected to the Internet via an NREN. According to the TERENA compendium, approximately 88% of all the university-level students in the GÉANT region are connected via a NREN service. The respondents to the survey indicated that most (63%) of the customers are served via a L3 connection, followed by (29%) by a L2 connection. The typical requested bandwidth for a point-to-point L3 service is 1 Gb/s (52%) but 10Gb/s links (22%) are more and more often demanded. Slower connections are nowadays less attractive (<1Mb/s: 9%, 10Mb/s: 4% and 100Mb/s: 10%). The typical requested bandwidth for a L3 VLAN service is similar to the requested

[82] http://www.terena.org/publications/files/TERENA-Compendium-2013.pdf

bandwidth for a L3 point-to-point service although one respondent indicated that the most frequent requested bandwidth is 1Gb/s. For further clarification, a comparison of both services is given in

Figure 4.

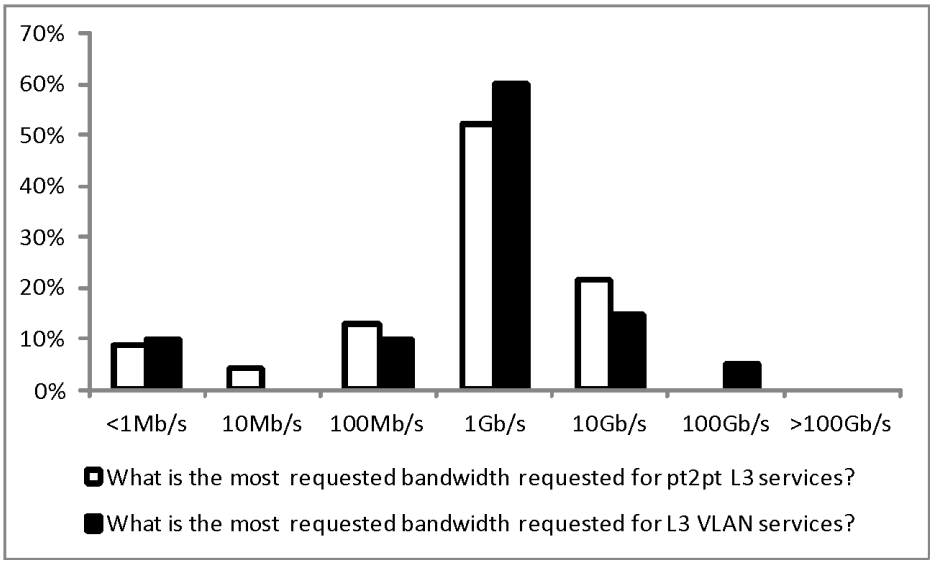


Figure 4: Comparison of most requested bandwidth for L3 connectivity services

When we zoom in at the difference in timescale for a service, the respondents give a similar view for both low-bandwidth services (≤100Mb/s) and high-bandwidth services (≥1Gb/s). Low bandwidth services are however more frequent in the timescale of seconds. An overview is given in

Figure 5.

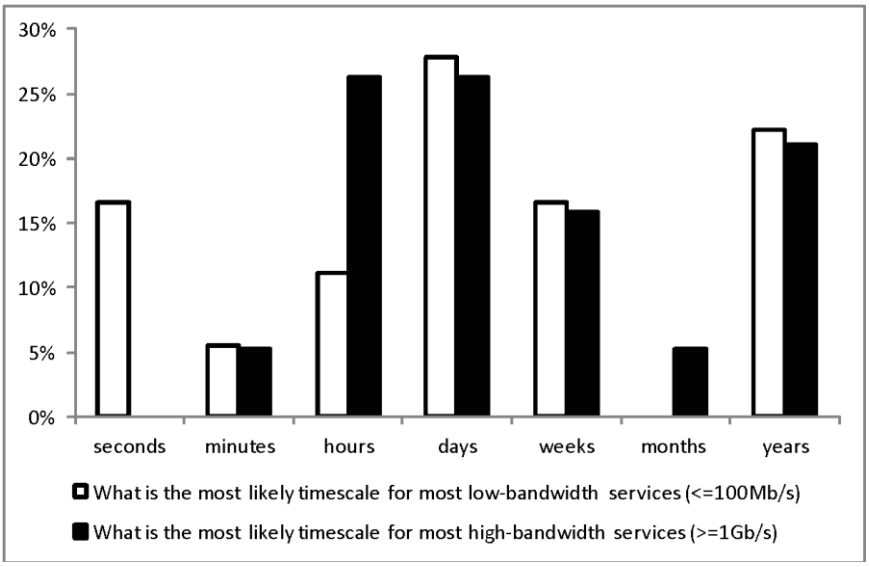


Figure 5: Comparison of most likely timescale for low- and high bandwidth services

Looking in more detail to peering relationships, we note that almost all respondents are interconnected with incumbent ISPs while three quarters are interconnected with the GÉANT backbone and almost three quarters with other NRENs. Over half of the NRENs are interconnected to mobile operators. Five respondents mentioned to be interconnected to Level3, 7 to data centres (e.g. Google) and 9 to local area networks (e.g. CERN).

Peering relations are typically set up for a few days (8 responses) or for a long period (months: 2, years: 8) via a Layer 3 (21 responses) or a Layer 2 (9 responses) interconnection. Over half (13) of the respondents offer multi-homed connections to their peering networks. The main reasons to provide a multi-homed connection is for redundancy (11) and load-balancing (2). NRENs that do not offer multi-homing don't offer the service because there is no-demand (7) or because it is too costly (4).

3.3.3 NREN applications

NREN collaboration infrastructure and related services are helping researchers and providers of higher education to collaborate worldwide. According to the TERENA compendium, four pillars of the NREN are:

1. Numbering schemes and VoIP to connect IP telephony deployments
2. Video- and web-conferencing
3. Group collaboration services
4. Multimedia content repositories.

Other services provided by an NREN are networked e-science resources (e.g. cloud services) and e-learning services. The survey also demanded to the respondents which application services are currently offered and which ones are the most-demanding IP-based services for an NREN (at most 3 services could be selected). An overview is given in Figure 6. Offsite data storage and backup, video conferencing, point-to-point links and VPN or VPLS tunnels are considered as the most demanding services.

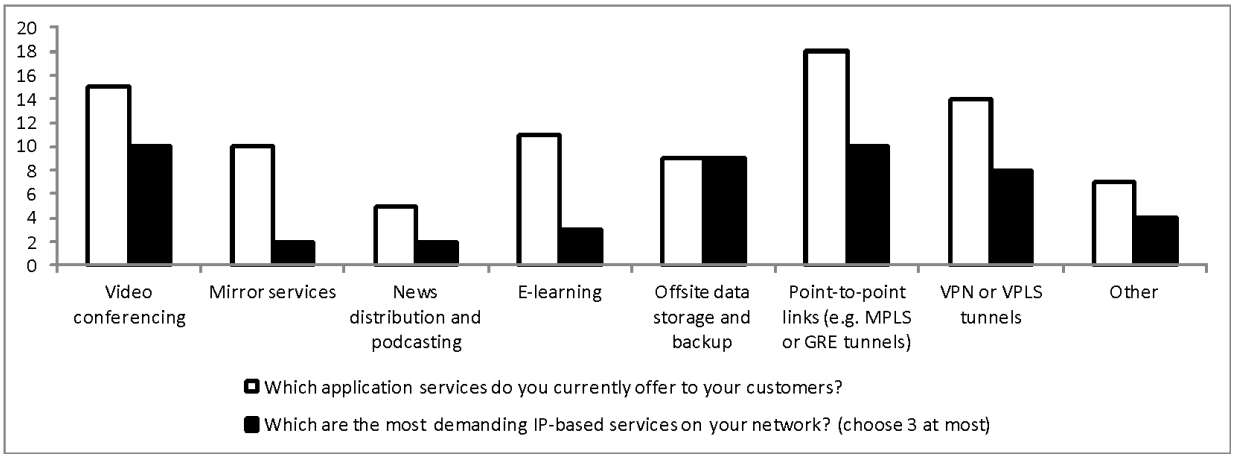


Figure 6: Overview of offered application services and the technological difficulty to offer each IP-based service.

The most stringent requirements for these IP-based services are Bandwidth (33%), Availability (31%), Latency (24%) and Provisioning Time (10%). Mobility is not considered as a stringent requirement (0%).

Ten of the 24 respondents offer cloud services to their customers. 7 offer the cloud service via centralised resources (of which 3 have redundant resources) and 2 via distributed resources.

The NRENs were also asked which operational services they are currently deploying. This is illustrated in

Figure 7.

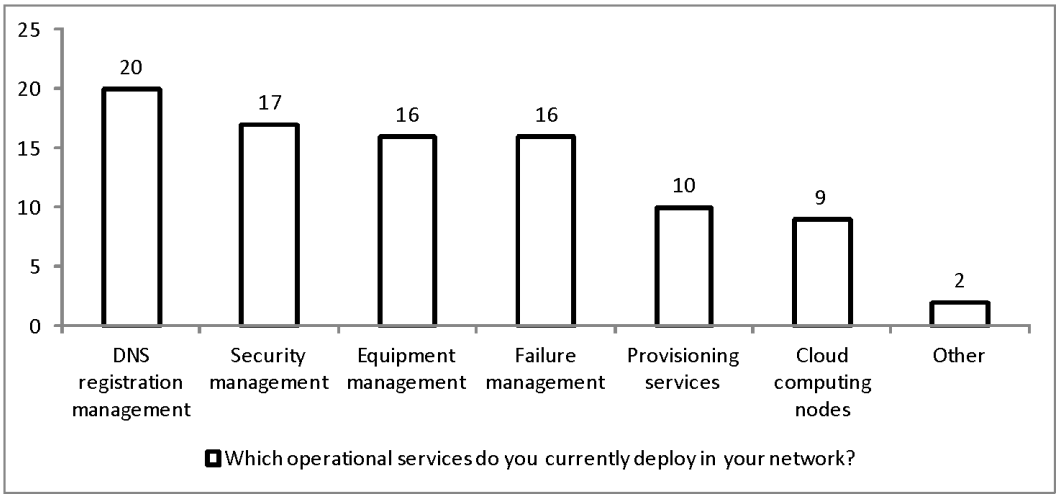


Figure 7: Overview of operational services offered by the NRENs

3.3.4 Future requirements

The NRENs were also asked to give their opinion on which applications will become more important in the future. Offside data storage and backup was identified as the application service that would become most important in the future from the provided list by one in four NRENs, followed by point-to-point links (19%) and VPN or VPLS services and Other (each 14%).

According to the TERENA compendium, sixteen of the GÉANT partner NRENs currently offer cloud services that are not produced via a commercial vendor and eleven more are planning to offer such services. This is also reflected in the survey. Ten of the 24 respondents to the survey currently offer cloud services while all but one is planning to roll out a cloud service in the future. There is also a move from centralised services to distributed cloud services. The survey results are added for further clarification in

Figure 8.

Looking at the time it currently takes to provision a cloud service on the existing infrastructure it is interesting to make a difference between the situation where a cloud service has to be set up within the NRENs network (both institutions are national) and the situation where a cloud service has to be

set up spanning multiple NRENs (institutions are international). In general it takes longer to provision a cloud service that spans multiple NRENs.

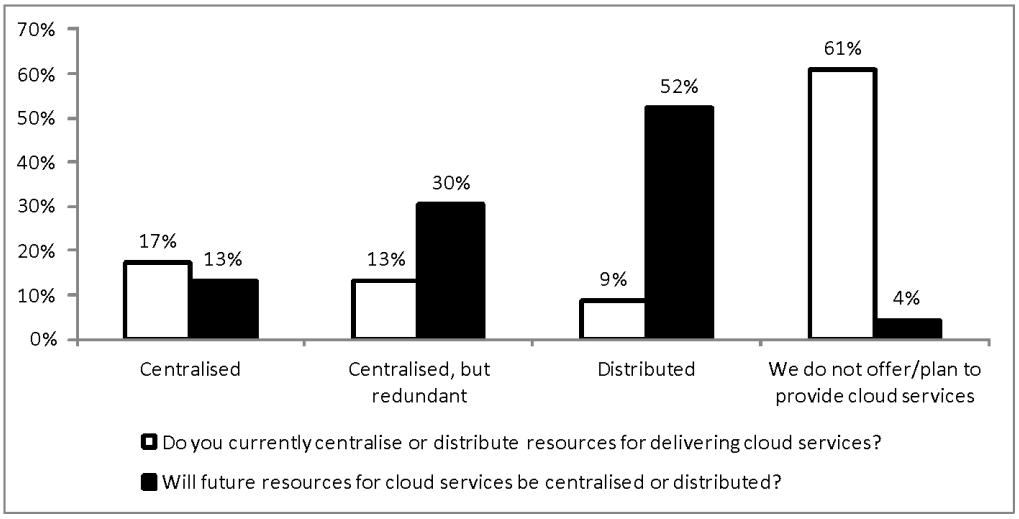


Figure 8: Comparison of current situation and future situation for cloud services

Zooming in on future operational services the respondents were asked which operational services will become the most important in their future network, at most 3 options could be selected. The respondents identified security management as the most important operational service (14 responses), closely followed by cloud computing nodes (12) and equipment management (11). Failure management (7), Provisioning services (7) and mobility management (6) are also considered important services for the future. Two respondents included DNS registration management and 3% include other operational services.

3.3.5 Impacts of future requirements on the NREN networks

In the TERENA compendium (2013), it is stated that until now the digital divide has been determined primarily by considering connectivity. In the future it may be necessary to assess it more in terms of service deployment. In the IRINA survey the NRENs were questioned to identify the most important technical goals for their IP network in the future. At most, 3 technical goals could be selected per NREN: 14 NRENs mentioned the maximization of throughput as one of the most important technical goals, while 11 mentioned improve link utilization and 9 the reduction of latency and the reduction of provisioning time. Improving mobility is mentioned by 5 and a reduction in power consumption by 3.

The shift in usage of the Internet by users will impact the NRENs network. Nowadays users want access to the Internet via a variety of mobile devices. Currently NRENs provide mobility via Wi-Fi but they state in the questionnaire that there is big demand for 3G and 4G services from NRENs. No big impact is expected on the backbone but saturation of the peering points with mobile network operators is a concern. When asked for the most important issues for the IP network of an NREN, bandwidth- and security related issues were most often noted. In terms of bandwidth, several NREN customers have higher and higher bandwidth demands (for specific services). This issue is mainly mitigated with light paths and upgrades to dark fibre to increase link capacity where possible. In terms of security, DDoS

attacks are more frequent which is mitigated by specific countermeasures. Another important issue is the competition and differentiation from commercial ISPs which is mitigated by offering better service quality and higher bandwidth.

4 IRINA use case

The IRINA use case is based on the survey results and the TERENA compendium[83], considering three main aspects: a **topology** consisting of a Pan-European backbone network (GÉANT) interconnecting 33 NREN networks, the **services** deployed over these networks and how **future requirements** impact these services.
The NREN networks are divided into three classes: 8 large NRENs, 16 medium-sized NRENs and 9 smaller NRENs.

Backbone Network		
GÉANT	The backbone network, operated by DANTE on behalf of the European NRENs	
Large-size NRENs		
DFN	Deutsches Forschungsnetz	Germany
RENATER	Réseau national de télécommunications pour la technologie, l'enseignement et la recherche	France
JANET	The UK's research and education network	United Kingdom
GARR	Gruppo per l'Armonizzazione delle Reti della Ricerca	Italy
RedIRIS	Spanish academic and research network	Spain
URAN	Ukrainian Research and Academic Network	Ukraine
PIONIER	Polish Optical Internet	Poland
NORDUnet	Scandinavian NRENs	Denmark, Finland, Iceland, Norway, Sweden

[83] http://www.terena.org/publications/files/TERENA-Compendium-2013.pdf

Medium-size NRENs		
RoEduNet	Romanian Education Network	Romania
SURFnet		Netherlands
BELNET		Belgium
GRNet		Greece
FCCN	Fundação para a Computação Científica Nacional	Portugal
CESNET		Czech Republic
NIIF/HUNGARNET		Hungary
BASNET		Belarus
ULAKBIM	Ulusal Akademik Ağ ve Bilgi Merkezi	Turkey
ACOnet		Austria
SWITCH		Switzerland
BREN	Bulgarian Research and Education Network	Bulgaria
AMRES	Kademska Mreža Srbije	Serbia
SANET	Slovak academic network	Slovakia
HEAnet		Ireland
CARNet	Croatian Academic and Research Network	Croatia

Small-size NRENs		
RENAM	Research and Educational Networking Association of Moldova	Moldova
LITNET		Lithuania
MARNET		Macedonia
ARNES	Academic and Research Network of Slovenia	Slovenia
SigmaNet		Latvia
EENet	Estonian Educational and Research Network	Estonia
MREN	Montenegrin Research and Education Network	Montenegro
RESTENA	Réseau Téléinformatique de l'Education Nationale et de la Recherche	Luxembourg
UoM-CSC		Malta

For each of the NREN types, a representative "reference network" is chosen, based on available information on these networks: based on RENATER for a large NREN, based on RoEduNet for a medium-sized NREN and based on AMRES for a smaller network.

4.1 The backbone network: GÉANT

The physical topology to be used for the backbone network will be the GÉANT topology, as shown in Figure 9. In the use case, GÉANT serves as a transport network for interconnecting the NRENs, not delivering any L3 services. The network consists mostly of 100G and parallel 10G optical lines. For network usage, we use the GÉANT usage Map (
Figure 10), presenting a map of Europe indicating the percentage network traffic to and from each country that is being monitored. The different levels of network traffic are indicated by different colours, which are listed on the map legend. From this map we use the following distribution of aggregated traffic between the NRENs and GÉANT, with peaks towards 80% of the bandwidth:

Aggregated (bidirectional) traffic between the NRENs and GÉANT	
Large-size NRENs	5.0Gb/s, provided by 4x10G links (peak: 32Gb/s)
Medium-size NRENs	750Mb/s, provided by 1x10G links (peak: 8Gb/s)
Smaller NRENs (if directly connected to GÉANT)	400Mb/s, provided by 5x1G links (peak: 4Gb/s)

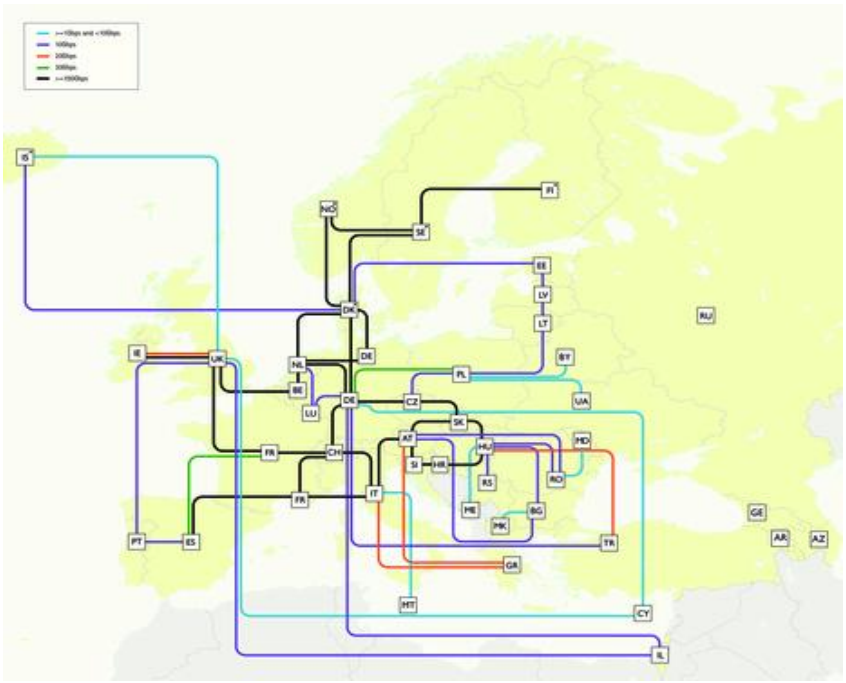


Figure 9: GÉANT topology, including NORDUnet connections (2014)

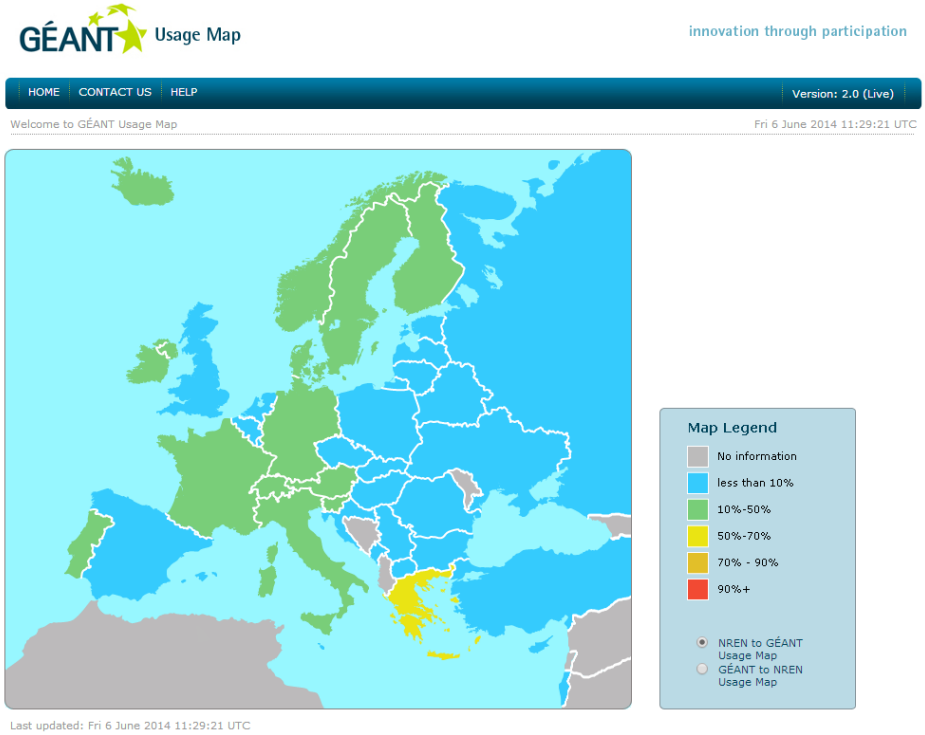


Figure 10: GÉANT usage map

4.2 The NREN networks

In order to keep the data manageable, we select three reference NRENs to be used as a model for the other NRENs in their class.

4.2.1 The Large-size Reference NREN (LSRN)

Large network, based on RENATER[84], it has 12.000 km of dark fibre, 120 links (max 40 WDM channels at 10G), 72 PoPs, 126 10GbE wavelengths deployed, 662 connected institutions over 1346 sites, external connectivity totals around 100Gb/s, directly connected to transit of two service providers (Paris (40G) and Marseille (20G)) and the SFINX IXP (2x 10G).
RENATER is directly interconnected to GÉANT using a 10 Gb/s channel for IP traffic and 10 Gbps for backup (via the Strasbourg-Kehl cross-border fibre with the German NREN, DFN) and has specific interconnections for certain research projects. Apart from the connection to X-WIN/DFN, cross-border fibre exists with BELNET and RESTENA).

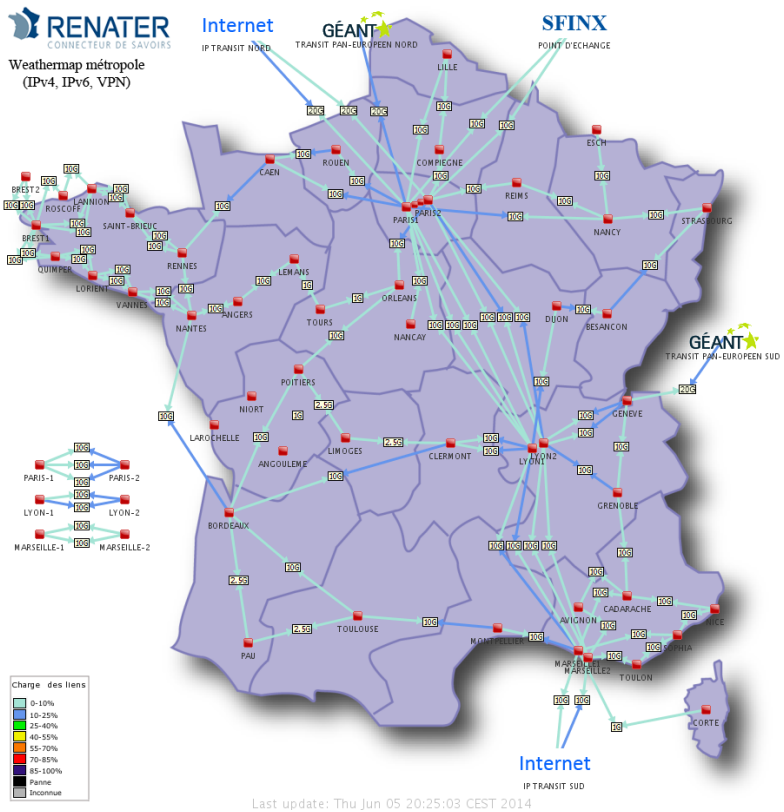


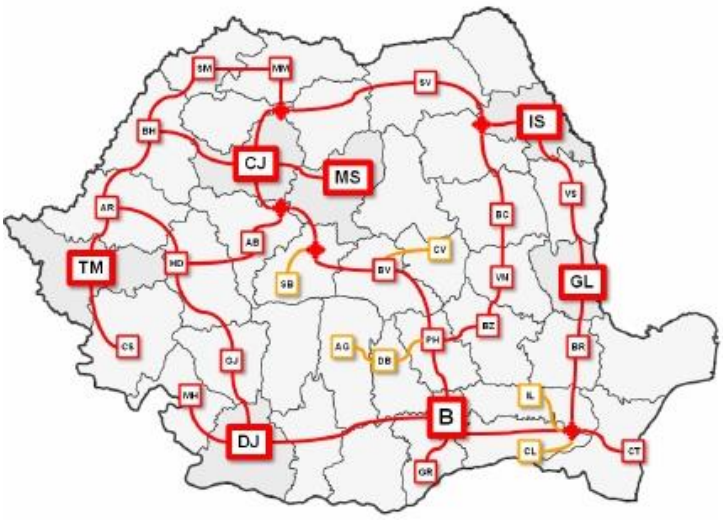
Figure 11: The RENATER network as representation for a Large NREN[85]

[84] http://www.renater.fr/IMG/pdf/RAPPORT_RENATER_2013_Anglais-2.pdf

[85] http://pasillo.renater.fr/weathermap/weathermap_metropole.html

4.2.2 The Medium-size Reference NREN (MSRN)

5636km of Dark Fibre, 41 PoPs (18 ROADM, 23 ADM) + Layer2/Layer3 equipment in all PoPs, 1x 100Gb/s link, 79x 10G links, 60x 8x1Glinks, Connectivity to GÉANT: 10Gb/s through Bucharest PoP. CBF to RENAM (2x10GE + 1x10GE spare).



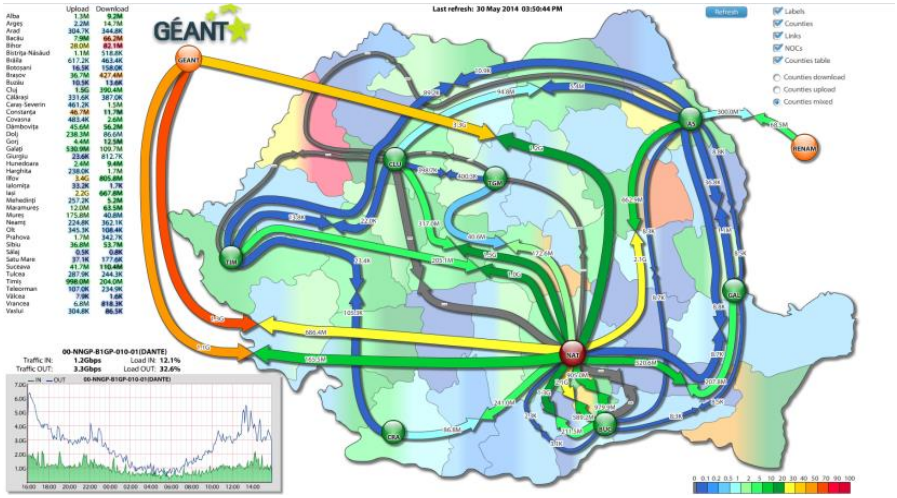


Figure 12: The RoEduNet2 network as representation for a medium-size NREN[86]

[86] https://nmis.roedu.net/weathermap/

4.2.3 The Small-size Reference NREN (SSRN)

Academic Network of Serbia (AMRES) [5] is the national research and education network of Serbia, and has 2150km of Dark Fibre, 54 PoPs, 153 Links, connecting 160 institutes and around 150000 active users. It is Connected to GÉANT through another NREN (HIIF/HUNGARNET).

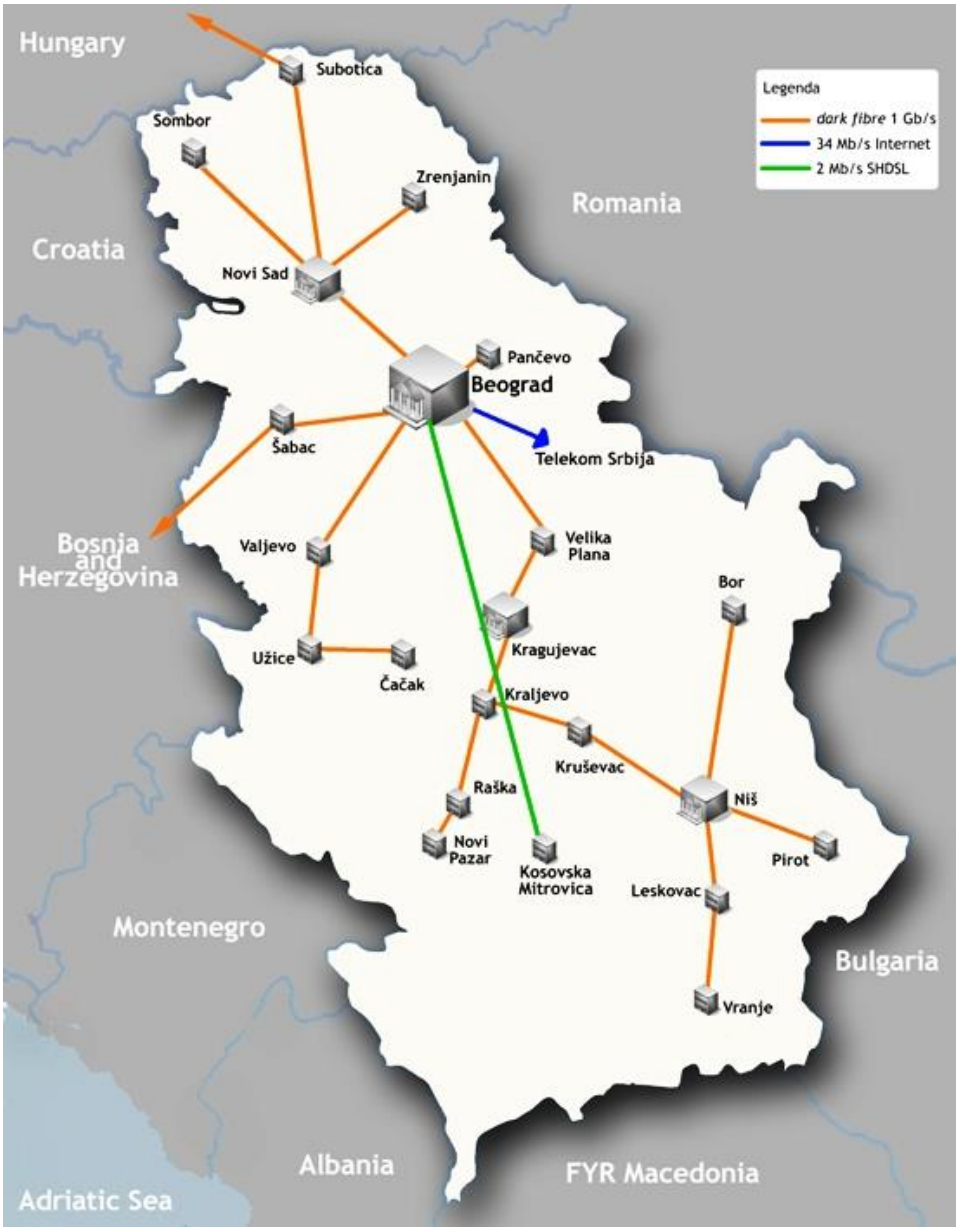


Figure 13: The AMRES network as a representation of a small NREN[87]

[87] http://wiki.ceengine.eu/AMRES

4.2.4 Integrated scenario

A minimal integrated scenario consists of each NREN type interconnected by GÉANT, as shown in

Figure 14. This type of scenario will be used to design the experimental validation of the IRINA software.

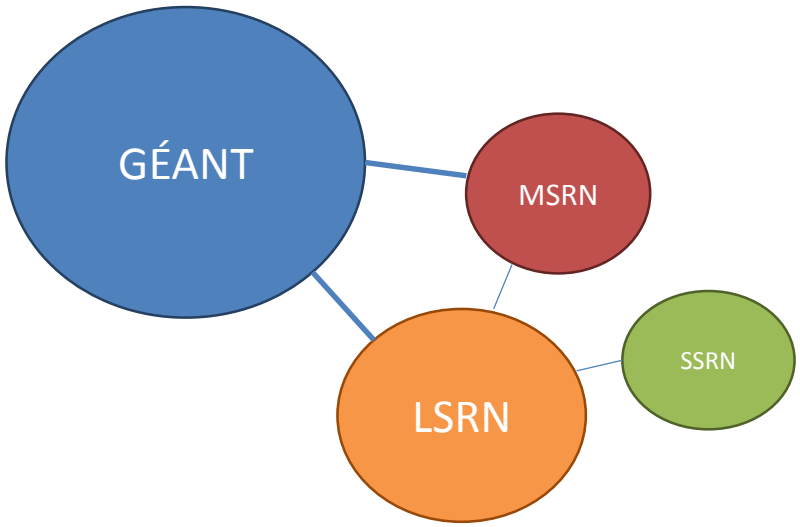


Figure 14: Minimal deployment scenario

4.3 Services

4.3.1 Video Conferencing

For the video conferencing service, we will base on the RENATER deployment, Skype and SCOPIA.
In 2012 RENATER opened RENAvisio+, a high definition video conferencing service offering a reservation guaranteed and optimal encryption, enabling secure meetings. A federated portal was set up in October 2012 to allow access to SeeVogh, a global workstation video conferencing tool. SeeVogh is based on a network of 62 globally interconnected nodes making the real-time collaboration service efficient, robust and stable across the internet. After 1 January 2013, SeeVogh evolved into a commercial cloud-based service.
In 2012, RENATER served 1800 meetings with its new SeeVogh service and 23000 conferences using tits Remote Meeting System, which represented a 50% increase with respect to 2011.[88]

For video conferencing, the following data rates can be observed:

[88] http://www.renater.fr/IMG/pdf/RAPPORT_RENATER_2013_Anglais-2.pdf

Video conferencing Bandwidth[89]		
Skype		
Call type	**Minimum D/U speed**	**Recommended D/U speed**
Calling	30kb/s	100 kb/s
Screen sharing	128 kb/s	300 kb/s
Video calling (HQ)	400 kb/s	500 kb/s
Video calling (HD)	1.2 Mb/s	1.5 Mb/s
Group video (3 people)	D: 512 kb/s U: 128kb/s	D: 2Mb/s U:512 kb/s
Group video (5 people)	D: 2Mb/s U: 128 kb/s	D: 4Mb/s U: 512 kb/s
Group video (7+ people)	D: 4Mb/s U: 128 kb/s	D: 8Mb/s U:512 kb/s
SCOPIA XT Desktop Server		
Number of clients	**SD (kb/s)**	**HD (kb/s)**
1	1536	3584
2	3072	7168
3	4608	10752
n	1536 x n	3584 x n

Table 2: Video conferencing Bandwidth requirements for Skype and Scopia XT

4.3.2 Point-to-point links and VPN services

Ethernet technology is a well known technology that has been deployed for decades in Campus Networks to provide LAN (Local Area Network) services. Due to the lower pricing of Ethernet interfaces compared to the traditional SDH and PDH technologies, Ethernet has become very important in today's MAN (Metropolitan
Area Networks).
Ethernet supports also higher bandwidths with fine granularity which is not available with traditional SDH networks.
Ethernet technology has the following applications and advantages:

- Suitable for transport of IP traffic.
- Provide easy interconnection between LANs.
- Point to point services.

[89] https://support.skype.com/en/faq/FA1417/how-much-bandwidth-does-skype-need

- Aggregation and point to multipoint services.
- Traffic differentiation.
- Network segmentation with VLANs.
- Network security by dedicated links or virtual networks.

4.3.3 Cloud Storage

The cloud storage service for the IRINA project will be based on the SURFDrive[90] service provided by SURFnet: SURFdrive is a personal storage service for the higher education and research community, offering staff, researchers and students an easy way to store, synchronise and share files in the secure and reliable SURF community cloud. Users get at least 100 GB data storage capacity, and can access their files at all times from any location by means of offline synchronisation. Users can also grant guest users access to their personal files. All data transmitted over the networks is encrypted.

Among other functionalities, SURFdrive offers:

- offline synchronisation, ensuring that users have access to their files at all times
- easy and secure file sharing within the higher education and research community
- users can invite guest users to share documents
- real-time insight into document edits by other users
- optimised for smartphone and/or tablet use
- 30-day backup and recovery
- search by metadata and full text search
- 99.5% availability, with 99.9% availability from 2015

4.4 NREN Service deployment

The NREN service deployment for these key services in the IRINA use case was estimated from the information above. The current and future requirements are shown in Table 3 and Table 4.

NREN type	Videoconferencing	Point-to-point and VPN	Cloud storage
Backbone	-	10Gb/s	-
Small NREN	-	100 Mb/s and 1 Gb/s	-
Medium NREN	SD only, 8000 calls/yr	100 Mb/s, 1 Gb/s, 10Gb/s	50% adoption, Centralised

[90] http://www.surf.nl/en/services-and-products/surfdrive/surfdrive.html

Large NREN	50% SD and 50% HD, 25000 calls/yr	100 Mb/s, 1Gb/s, nx10Gb/s	100% adoption, Centralised, redundant

Table 3: Current NREN service deployment

NREN type	Videoconferencing	Point-to-point and VPN	Cloud storage
Backbone	-	10Gb/s and 100Gb/s at L2 and L3	-
Small NREN	SD only	1Gb/s and 10Gb/s	10% adoption, all centralised
Medium NREN	30% SD, 70% HD, 20000 calls/yr	10Gb/s and 40Gb/s	80% adoption, 50% Centralised / 50% Distributed
Large NREN	HD only, 75000 calls/yr	10, 40 and 100Gb/s	100% adoption, Distributed

Table 4: Future NREN service deployment

5 Technologies for the reference NRENs

This section studies GÉANT and the three reference NRENs chosen in the MS1 study in order to analyse: the scope and technologies used by each network; the services provided by these networks and their interconnection with other networks. Once these aspects are understood, we carry out a detailed analysis on how to apply RINA to the scenario in section 6.

5.1 GÉANT

GÉANT is the pan-European research and education network that interconnects Europe's National Research and Education Networks (NRENs). Figure 9 shows the different Points of Presence (PoPs) of the GÉANT Network and how they are interconnected by dark fibre or leased lines.

Figure 15 shows a high-level overview of the GÉANT PoPs design, which is basically composed by layer2/3 routers (Juniper MX) and optical transport equipment (Infinera). This PoP design allows GÉANT to provide the following services to its customer NRENs:

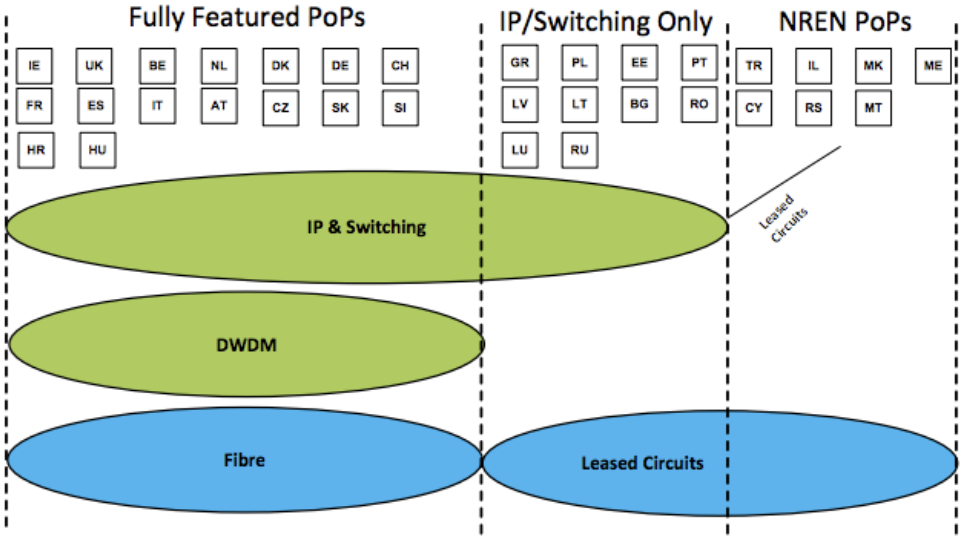


Figure 15: Schema of a GÉANT Point of Presence

- **GÉANT IP**. Best effort IP service (both IPv4 and IPv6 are supported). Provides high-bandwidth international Internet connectivity. It provides general purpose IP transit services to NRENs.
- **GÉANT L3VPN**. Provides a Virtual Private Network service for many-to-many or one-to-many environments.
- **GÉANT Plus**. Point-to-point layer 2 circuits (L2VPN) of assured bandwidth and performance. It is delivered on dedicated VLANs.

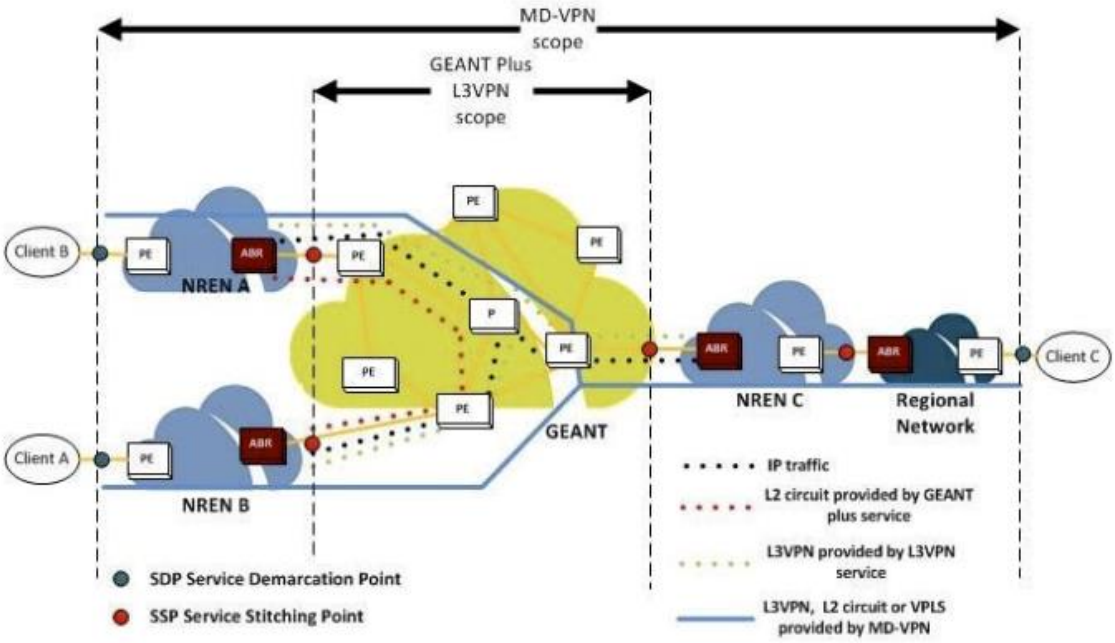


Figure 16: VPN services currently provided by GÉANT

- **GÉANT Lambda**. Provides transparent 10 Gbps or 100 Gbps wavelengths between transmission equipment in GÉANT's PoPs. The customer (NREN) interface type can be Ethernet or SDH.

5.2 Large NREN: RENATER

RENATER has two 10G commodity Internet access points, provided by two operators (Cable&Wireless and Level3). RENATER also operates an Internet eXchange point called SFINX, whose hardware architecture can be seen in

Figure 17. SFINX has two PoPs interconnected via 2x10 GbE links. Internet Service Providers (ISPs) can connect to the IXP nodes via Fast Ethernet, Gigabit Ethernet or 10 Gigabit Ethernet. SFINX provides two BGP route servers to facilitate the establishment of peering sessions within ISPs.

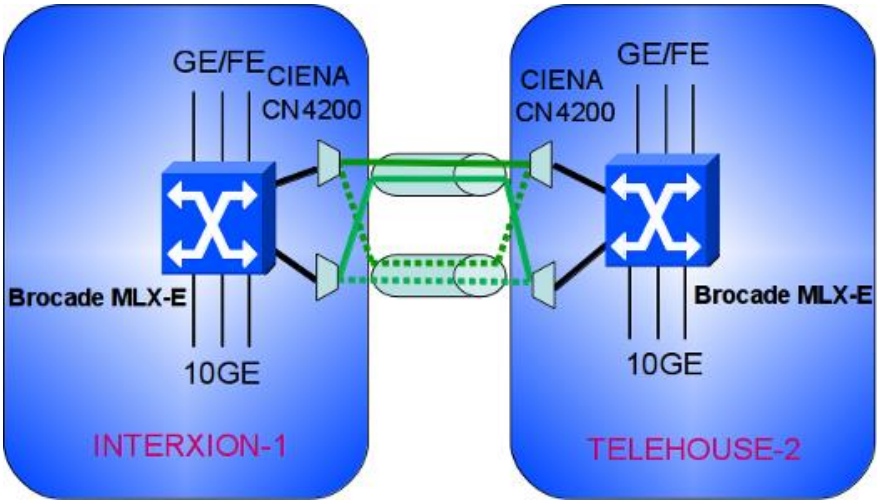


Figure 17: Hardware architecture of the SFINX IXP

User institutions are connected to RENATER i) via a regional or a metropolitan network; ii) directly to the NREN PoP or iii) via other institutions. The first approach is the most usual one. An illustrative example is the one provided by the universities within Paris metropolitan area, which are connected to RENATER via a metropolitan network called RAP (Réseau Académique Parisien).
Regional networks or user institutions connect to RENATER at one of their PoPs. To do so, the connecting institution needs to deploy a piece of equipment and connect to a number of ports of RENATER. Each port can be configured in native mode (no VLANs allowed) or in VLAN mode (which allows running concurrent services in the same port). RENATER provides the following connectivity services:

- **Access to the Internet**, via BGP-4 peering. Three types of announcements can be received from RENATER: full Internet routing table, just a default route or just the prefixes advertised by RENATER.
- **Virtual Private Network** (point-to-point or multipoint connectivity). A number of technologies to implement the VPN service are available, depending on the user requirements: DWDM, L2 EoMPLS, L3VPN (implemented via Virtual Routing and Forwarding), Premium IP, GRE tunnels or IPSec tunnels.

5.3 Medium-sized NREN: SURFnet

Because of information availability, the reference network has been updated to SURFnet. SURFnet7 is connected to locations in the Netherlands and its neighbouring countries via an 11,000-km fibre optic network. Connections between the 21 core locations around the country are illuminated using CPL (Common Photonic Layer) optical equipment from Ciena. This equipment allows for the various fibre optic connections to be illuminated at up to 88 different wavelengths. Each wavelength can transport as much as 100 Gb/s, ensuring ample bandwidth. The SURFnet7 network structure can be decomposed into several layers, as shown in Figure 18.

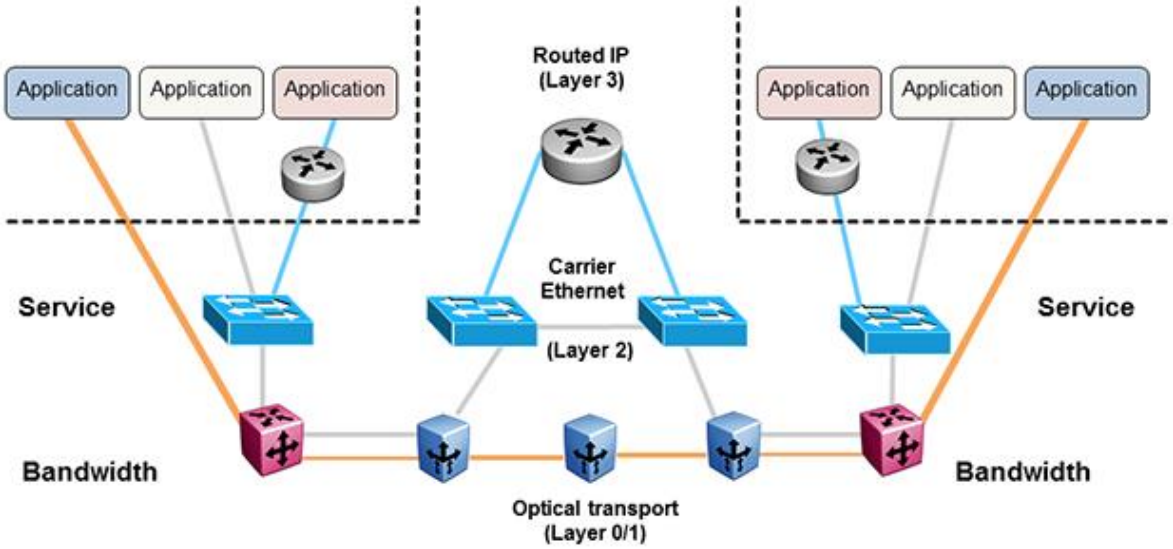


Figure 18: 3-Layer structure of the SURFnet 7 network

The SURFnet network consists of approximately 11,000 km of fibre optic cable. These fibre optic connections serve as the basis for a circuit-based optical network, depicted in

Figure 19. The five largest circuits are colour-designated in the below illustration. The two major SURFnet points-of-presence (PoPs) in Amsterdam make up the heart of the circuit structure. These locations house SURFnet's IP routers and the NetherLight equipment. From here, the SURFnet network establishes connections with external parties such as research networks and commercial providers. The largest participating institutions also form PoPs on the five major circuits. The PoP equipment used to connect all locations is housed at the participating institutions. There are approximately 350 PoPs stationed throughout the Netherlands.

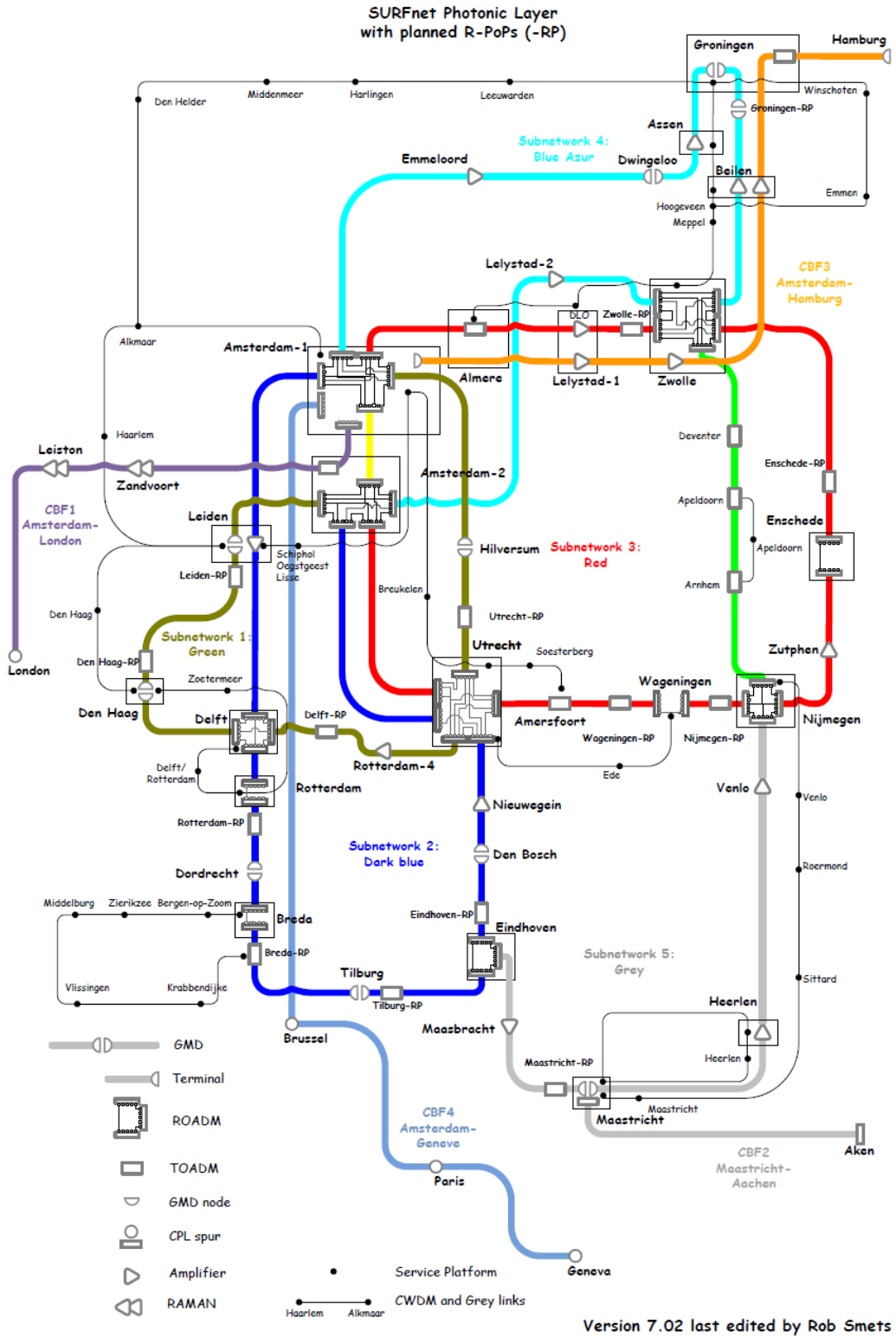


Figure 19: SURFnet7 Common Photonic Layer (layer 0-1)

Layer 2 is used to transport Ethernet services (light paths or IP traffic) over the network. SURFnet7 transports these services using state-of-the-art Carrier Ethernet technology. Amongst other benefits, this innovative technology offers a high degree of flexibility. With Carrier Ethernet, the network can be configured to provide the exact amount of required bandwidth – thus considerably improving the level of efficiency. Participating institutions connect to the SURFnet network via an IP traffic router. Light paths can be connected via a router, switch or definitive host (such as a microscope or super computer).
Layer 3 is used to process the institution's internet services (connections to the rest of the internet). All routers at the institution are linked to the two SURFnet routes at the core locations in Amsterdam via the Ethernet services layer.
In terms of international connections, SURFnet is connected to the Amsterdam Internet Exchange (AMS-IX) with 60 Gb/s, to Netherlight with 10 Gb/s, to BNIX with 2 Gbps, to LINX with 10 Gbps, to GÉANT with 60 Gb/s and to others (such as Google and range of ISPs) with 67 Gbps. For all other traffic, SURFnet maintains connections with upstream providers (2x10 Gb/s with KPN and 2x10 Gb/s with GTT). SURFnet also has cross-border connections in Hamburg, Brussels, Geneva, Paris, London, Hasselt and Aachen.
In order to access the services provided by SURFnet, an institute needs to get physical connectivity to the network. The physical connection to SURFnet7 is the interface on a patch panel (ODF) inside a SURFnet PoP (point of presence) to which an institute connects its hardware. The physical connection also forms the boundary between the infrastructure of the SURFnet network and the institutional network. There are two types of ports an institution can connect to: *a single service port* (SSP) - in which the port can be assigned to a single service - and a *multi-service port* (MSP) - in which the port divides the available bandwidth amongst the number of services configured in the port (up to 10). The different services are distinguished by a VLAN tag. SURFnet provides the following types of connectivity services:
SURFInternet. Reliable Internet connection at 1, 10 or 100 Gbps. Can be configured with static routing or BGP routing. In all cases the institution is connected to the two SURFnet core routers. VRRP (in the case of static routing) and BFD (in the case of BGP routing) are used to provide resiliency.

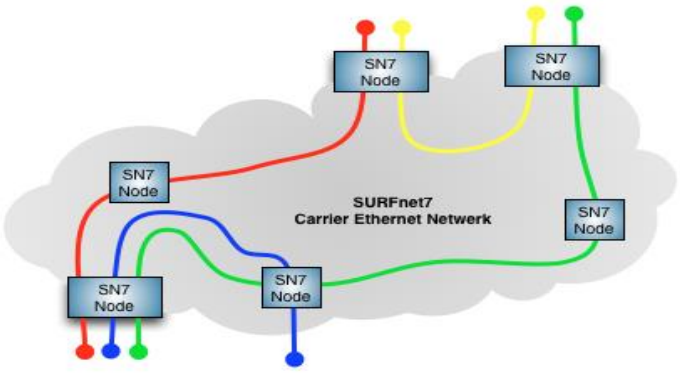


Figure 20: Optical Private Network (OPN)

Light-paths. Lightpaths provide high-capacity private connectivity between two or more nodes of the Carrier Ethernet network of SURFnet7. Lightpaths can be point-to-point with or without protection. A number of lightpaths can be provided together creating an Optical Private Network (OPN), as shown in
Figure 20. Lightpaths can connect to international locations via GÉANT/CBDF or via the NetherLight Open Lightpath Exchange.

5.4 Small NREN: AMRES

The technologies used for AMRES infrastructure are:

- **Optical technology** - is used for data transfer via optical infrastructure (dark fibre) - Gigabit Ethernet technology (1000BASE-X) is used in AMRES, as well as satellite modules and single mode or multimode cables.
- **xDSL VPN** - technology used for accessing the network. It is being accomplished in cooperation with Telekom Srbija.
- **Analogue lines** - used for accessing the network and presents a technology that has been intensively used in the past. HDSL modems are used for data transfer. This technology is problematic due to the stability of functional services.

5.5 Services (video-conferencing)

SeeVogh[91], formerly known as EVO, provides a worldwide videoconference and collaboration services offering: audio, video, instant messaging, chat, captured desktop, whiteboard, shared files and recording/playback. It was designed to provide a global-scale, robust, real-time collaboration service to the Large Hadron Collider (LHC) experiments and other major research and education programs[92].

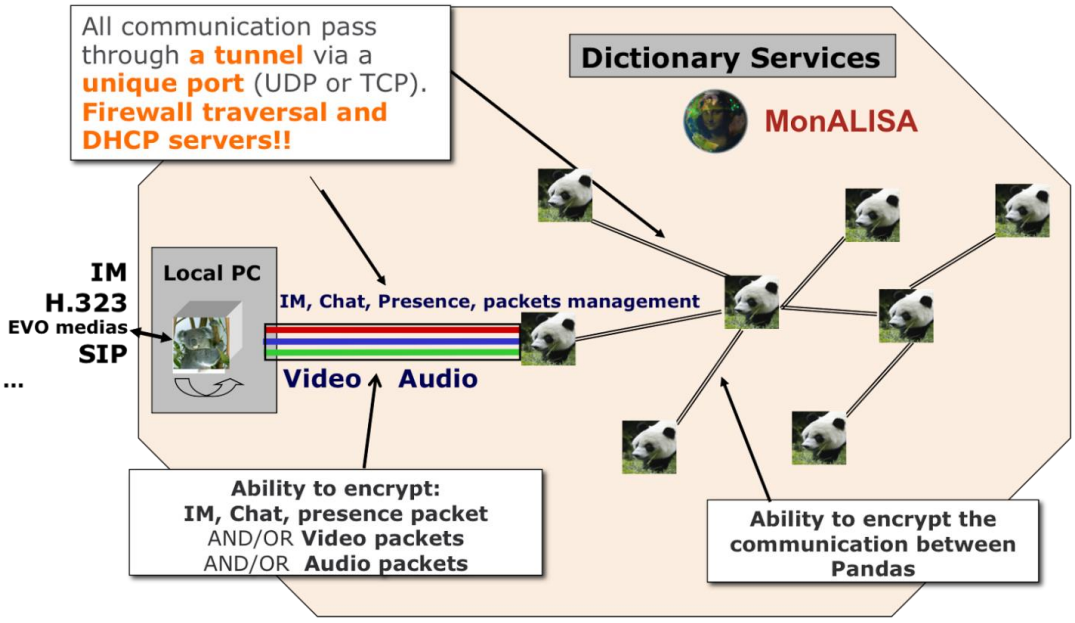


Figure 21: Architecture of the EVO/SeeVogh distributed application

The services provided by SeeVogh are widely used within the research and education community, having an average of 2500 sites connected to the SeeVogh infrastructure per day, running 500 international meetings with up to 200 participants per meeting (this number can be higher since SeeVogh places no limitations on the number of users per meeting.

[91] SeeVogh website, available online at https://seevogh.com

[92] Phillipe Galvez, "From EVO to SeeVogh". TERENA Networking Conference 2011. Available online at https://tnc2011.terena.org/core/presentation/13.

SeeVogh is a distributed application whose core is formed by a worldwide network of peer-to-peer servers called Pandas (currently there are around 70 servers distributed within 25 countries), as shown in

Figure 21. Pandas can encrypt all communication between themselves if the environment they are deployed into is not trusted. The Pandas, also called Unlimited Self-Hosted Instance (USHI), can be thought of a Multi-Conference Unit (MCU), sized by the number of ports. USHIs are cloud instances that can be distributed across geographies on multiple hybrid-clouds and localized to the user community that reduces latency.

Users can access the services provided by SeeVogh via a client installed in their devices, which tunnels all the communication with the network of Pandas (IM, H.323, SIP, etc) via a single TCP or UDP port - in order to facilitate firewall transversal.

6 Applying RINA to the GÉANT and NRENs scenario

6.1 Overview

Figure 22 shows an overview of the different types of networks that have been considered in this use case study, as well as their interconnection points. The figure doesn't display different layers or the architecture of the different networks. We have adopted an NREN-centric approach to describe the NREN and GÉANT ecosystem, analyzing the different types of networks and stakeholders that an NREN connects to. The interactions considered in this study - marked with a labelled white box in the Figure - are briefly introduced in the following lines and later discussed in the next sections of this chapter.

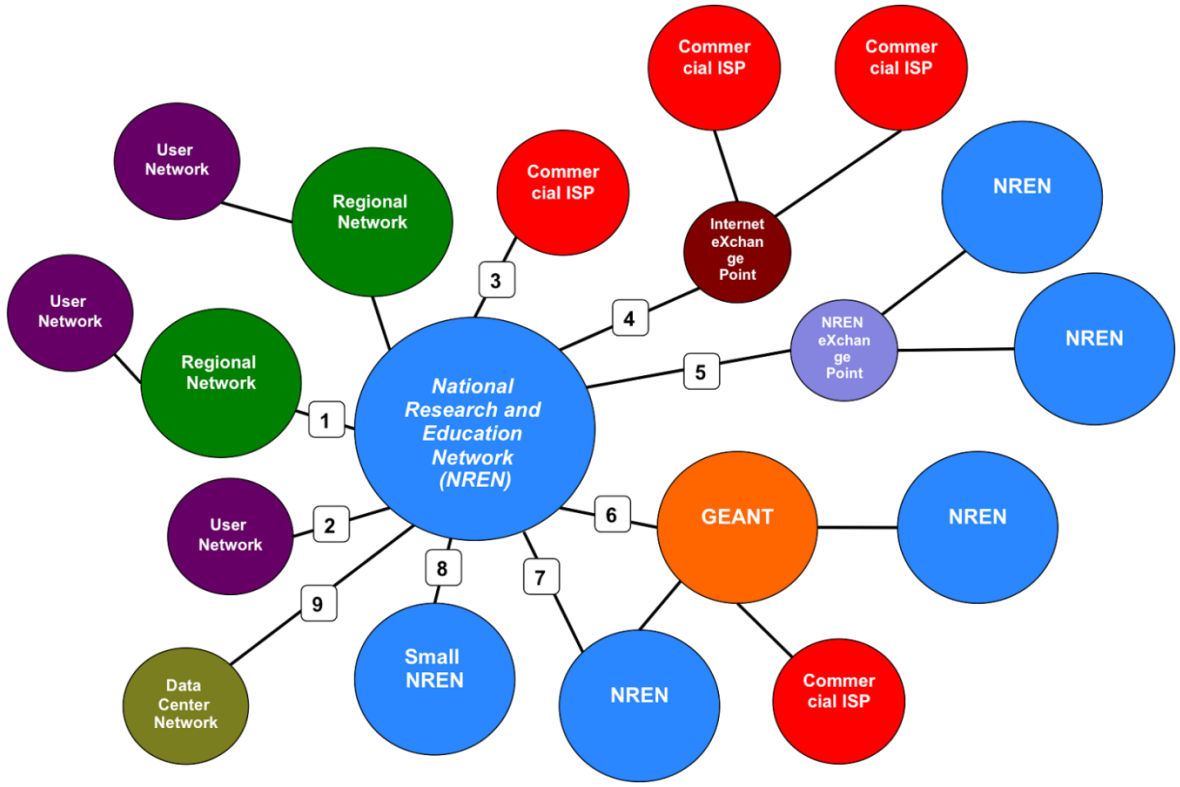


Figure 22: The different types of networks considered in the use case study, and their interconnection points

1. **NREN - Regional Network**. NRENs usually play the role of a national research and education backbone network, interconnecting several smaller networks of regional or metropolitan scope that are in turn connected to the end user networks (universities, research institutions, etc.)
2. **NREN - User Network**. Some large user institutions may directly connect to the NREN and install equipment in one of the NRENs PoPs.
3. **NREN - Commercial ISP**. NRENs establish transit or peering agreements with commercial ISPs on a one by one basis.
4. **NREN - Internet eXchange Point (IXP)**. NRENs can also peer with commercial ISPs using the facilities provided by an IXP network.
5. **NREN - NREN eXchange Point (NXP)**. Although not existing today - as far as the authors of this report are aware - the RINA structure facilitates the creation of NREN eXchange Points in which NRENs can directly connect to each other.
6. **NREN - GÉANT**. The main role of GÉANT is to act as the European backbone for the NRENs, providing the main vehicle for NRENs to connect to each other. GÉANT is also connected to commercial ISPs and provides Internet transit services.
7. **NREN - NREN (peer)**. NRENs connect directly to each other via Cross Border Dark Fibre (CBDF) or other means.
8. **NREN - small NREN ("transit")**. Some small NRENs (such as AMRES) do not connect directly to GÉANT but to another NREN which allows them to reach international destinations.

9. **NREN - Data Centre (DC) network**. Some NRENs own DCs that they use to provide cloud services to their customers.

An example of a side view showing the different layers in the interconnection of a campus network, an NREN, GÉANT and another NREN is provided in Figure 23. More details about the different networks and their interconnection will be given in the following sections; but the Figure should be useful to understand the repeating structure of layers. Sitting on top of the physical layer there are point to point (or multipoint in case of radio technology) DIFs, which provide connectivity between a limited number of systems (two in the case of a cable). Then each network has a number of "internal DIFs" that are used to connect together all the border routers of the network and aggregate the traffic generated by one or more "e-mall DIFs". "E-mall DIFs" are the ones that are designed to support distributed applications, such as the public Internet, VPN DIFs, application-specific DIFs, etc.

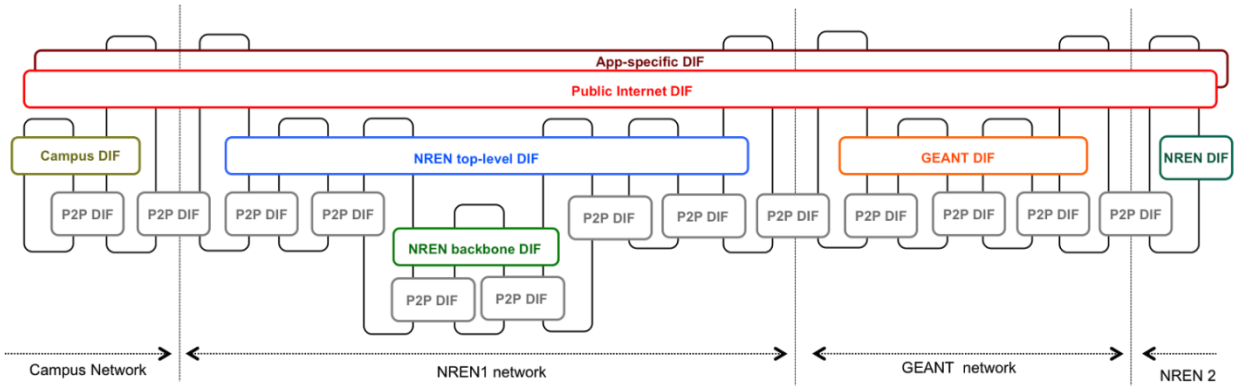


Figure 23: Side view of the different layers in the interconnection of a campus network, and NREN, GÉANT and another NREN. The public Internet or other DIFs are floating on top

As explained in the next section, each network can use an arbitrary number of layers that is most convenient for the network's operational environment (for example, two in the NREN 1 network or one in the case of GÉANT in the example of

Figure 21). The decision will depend on the size of the network, the different types of traffic it needs to transport, and other criteria.

6.2 Internal NREN design

NRENs are usually small/mid-sized networks that play the role of national backbones with a number of PoPs ranging between 20 to 100. They interconnect with customer networks (regional or institutions' networks), GÉANT, other NRENs, ISPs or IXP networks. The services they provide to their customers can be broadly classified between Internet access and VPNs. Taking all these facts into account, we can explore an initial NREN design that leverages the typical RINA structure of DIFs arranged in a necklace, with DIFs with less

scope and carrying traffic that is more and more aggregated as one moves down the layers. This structure is depicted Figure 24, with two levels of DIFs: the '*NREN top-level DIF*' and the '*NREN backbone DIF*'. These DIFs support a number of "e-mall DIFs" that provide IPC services to different applications, such as the general purpose "*Public Internet DIF*" or several application-specific DIFs tailored to the requirements of a different range of applications.

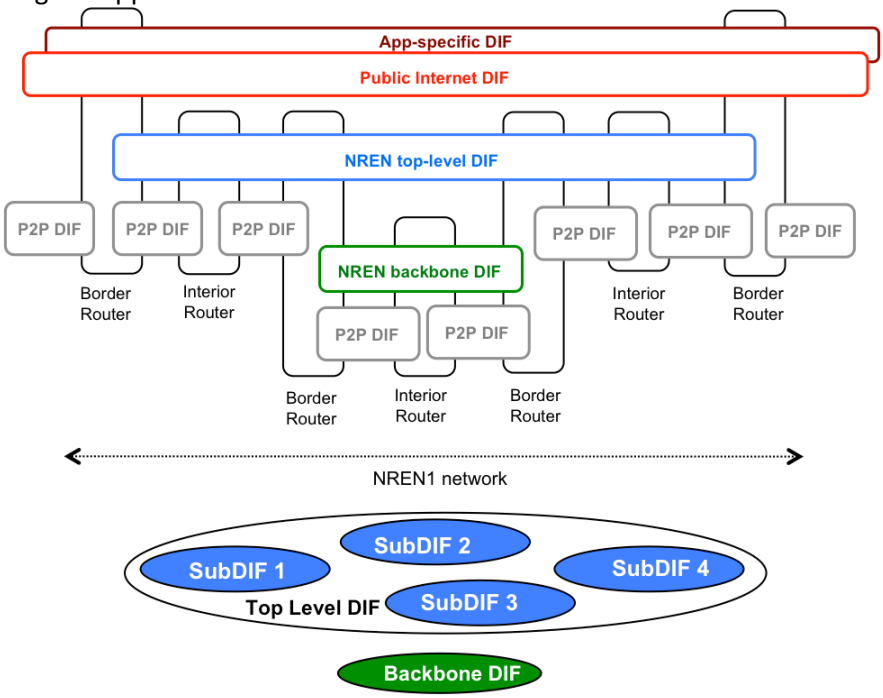


Figure 24: Example of an NREN with two internal layers: top-level DIF and backbone DIF.

Figure 24 shows the side and top views of the "*NREN top-level DIF*" and the "*NREN backbone DIF*". The former is the DIF that provides "end-to-end" connectivity across the NREN, supporting one or more *e-mall DIFs*. It is structured in different sub-DIFs, for example covering different regions of a country, and relies on the backbone DIF to provide connectivity to the routers at the border of each sub-DIF. The "*backbone DIF*" transports the aggregated traffic between the different regions of the "*top-level DIF*". This is just an example of a structure that can be scaled up or down, depending on the needs of the NREN. The structure allows the network designer to bound the sizes of the routing tables in the DIFs [8], to better aggregate different types of traffic or to customize the policies of the different DIFs. Small or even medium sized NRENs will probably have enough with a single DIF for the NREN, large NRENs may directly use the two-level structure depicted in this example. Although it is unlikely in the NREN environment, more levels of DIFs could be inserted; for example one could envision a structure for large ISPs with at least three levels: metropolitan, regional and backbone.

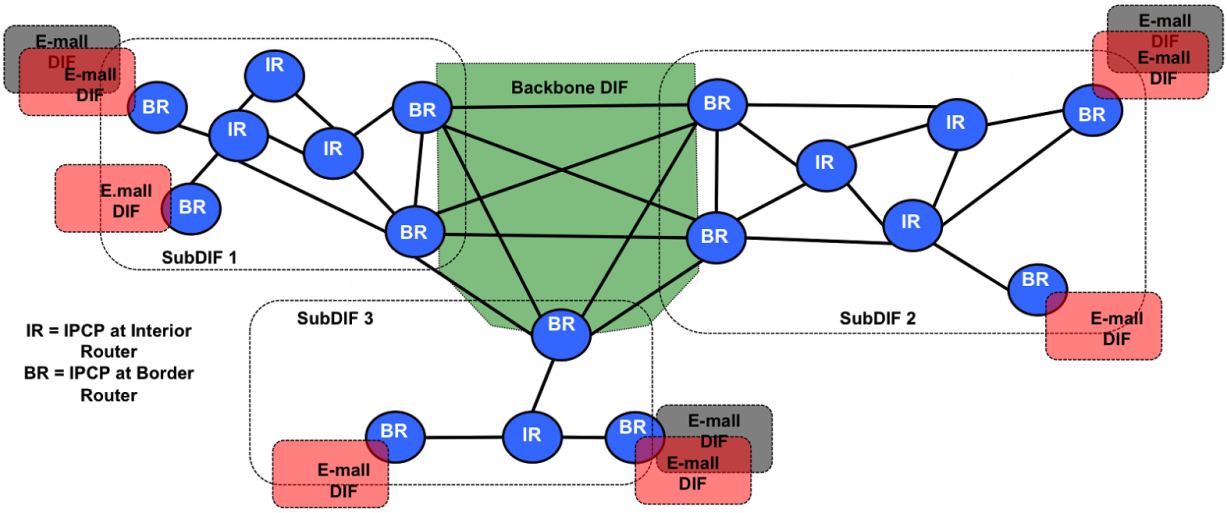


Figure 25 Top-view of the NREN's Top-level DIF.

The connectivity graph of the "*NREN top-level DIF*" is shown in Figure 25, as well as its supported and supporting DIFs. IPC Processes (IPCPs) in this DIF are grouped in different sub-DIFs. There are two types of IPCPs: those belonging to internal routers (within a region) or those belonging to border routers (at the borders of a region). Border routers is where the recursion takes place and there are instances of supported DIFs (*e-mall DIFs*, in red and grey) or supporting DIFs (*the backbone DIF*, in green).

The *backbone DIF* can support a full mesh connectivity between the IPCPs at the border routers of every region, causing all regions to be one hop away from each other. This design facilitates the use of topological addressing [9], employing simple schemas such as <*region, id*> to address the IPCPs in the DIF. Since all inter-region border routers are interconnected by the backbone DIF, IPCPs only need to maintain intra-region routing tables, keeping the size of the routing table bounded. Depending on the requirements of the region, each one can run a link-state or a distance-vector routing protocol inside the region (they don't need to be the same for all regions).

Since this DIF needs to support different *e-mall DIFs* with potentially very different requirements in terms of delay, packet loss, jitter, reliability, etc. the "NREN top-level DIF" needs to provide different QoS cubes in order to efficiently multiplex the traffic offered by the DIFs it supports while still keeping a high utilization. Resource allocation policies inspired in models such as the "*delta-Q*" framework[93] could be particularly useful since they were designed with this goal in mind, and exhibit good properties when the system is under stress due to high offered loads[94].

In terms of the security policies that would be suitable for the *NREN top-level DIF*, it depends on the design options. The design shown in Figure 24 keeps all the IPCPs in the *NREN-top-level DIF* within the NREN systems (in other words, that DIF doesn't span to the customer border routers). In this case authentication policies can be more relaxed since all the systems where the DIFs are instantiated are under the control of the NREN. An alternative design could make the *NREN top-level DIF* span to the first customer router; in this case

[93] D. C. Reeve, "A new blueprint for Network QoS", PhD thesis, 2003. Available online at http://www.cs.kent.ac.uk/pubs/2003/1892/

[94] N. Davies, "Delivering predictable quality in saturated networks", PNSol technical report, 2003. Available online at http://www.pnsol.com

stronger authentication policies (based on symmetric or asymmetric cryptography) would be more appropriate - also depending on the level of trust with the entities participating in the *NREN top-level DIF*. Security-related SDU protection policies such as encryption need not be very strong since the NREN supporting DIFs will either be point to point links or the *NREN backbone DIF* (which is also under full control of the NREN). Probably the highest security risk for the NREN top-level DIF would come from shared supported DIFs such as the *Public Internet DIF*. Therefore the *NREN top-level DIF* and the NREN Management System should keep an eye on the resources used by these DIFs in order to detect and mitigate any potential DoS or DDoS attacks.

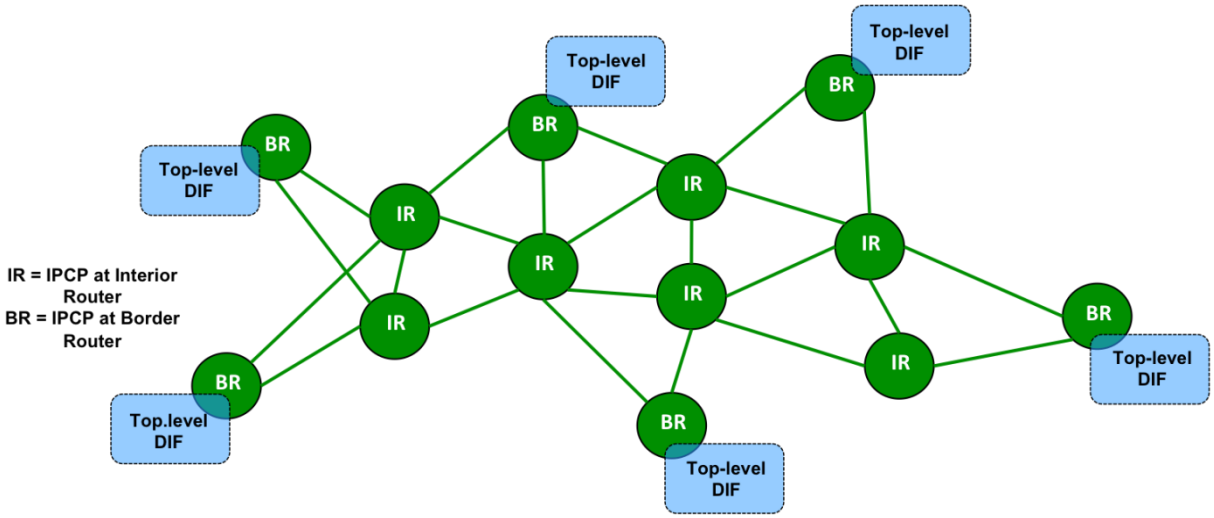


Figure 26 Top-view of the NREN's Backbone DIF.

Figure 26 depicts the connectivity graph of the IPCPs in the backbone DIF. The scope of this DIF should be small enough to allow for link-state routing without any scalability problems. In terms of resource allocation, this DIF would carry highly aggregated traffic, therefore connection-oriented resource allocation policies resembling virtual circuits would probably be the most effective solution. Security should not be a big issue since: i) all the IPCPs of this DIF are instantiated in systems controlled by the NREN; ii) the only user of the DIF is a DIF controlled by the NREN and iii) the backbone DIF relies directly on point to point DIFs.

VPN DIFs can be designed on top of the *NREN top-level DIF* or closer to the hardware, as seen in

Figure 27. The decision will depend on the requirements of the VPN and the capabilities provided by the *NREN top-level DIF*. If these capabilities (modelled as QoS cubes) are not enough to support the VPN, a new DIF that runs in parallel to the *NREN top-level DIF* can be instantiated to implement the VPN service. Note that this approach is more complex to manage and consumes more resources than deploying the VPN DIF on top of the *NREN top-level DIF*, therefore it should only be used if it is really required.

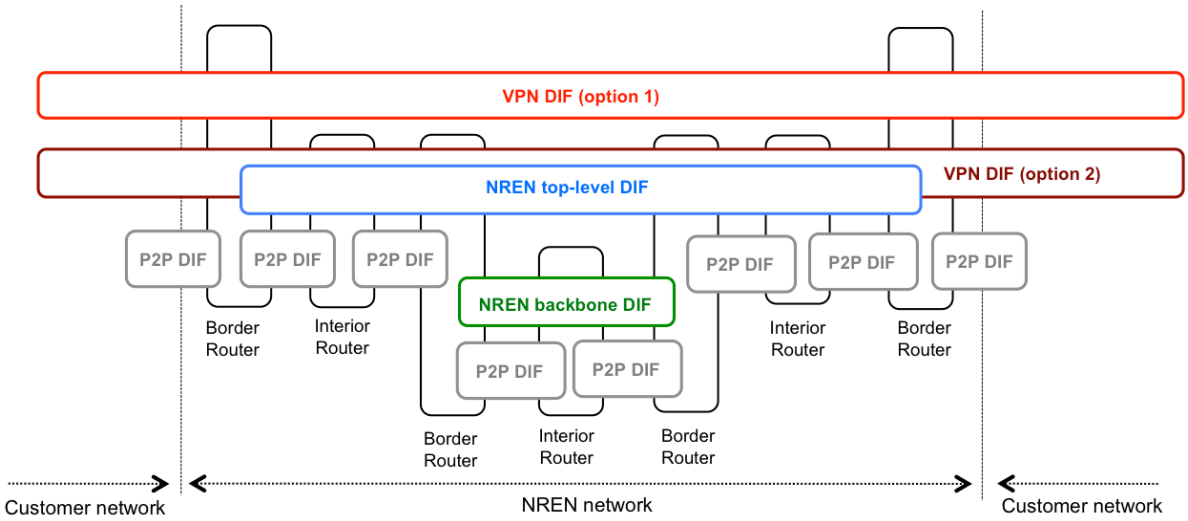


Figure 27: Different options for implementing VPN DIFs.

6.3 Connecting to other NRENs (via GÉANT and CBDF), multi-provider DIFs

The primary option for NRENs to connect to each other is via GÉANT, which is the European backbone network designed to bring NRENs together. GÉANT also provides access to the public Internet as well as VPN services at various levels. Taking all of this into account, the GÉANT network design is more or less equivalent to that of a medium-sized NREN (30-40 PoPs), with the difference that its customers are not regional/campus networks but NRENs and that there is more physical distance between PoPs.

Figure 28 shows an example of the GÉANT network modelled with the RINA architecture. A single DIF (the *GÉANT DIF*) spans all the routers in the GÉANT network and makes it a single resource allocation domain. This DIF can support a number of "e-mall DIFs" on top, such as the *Public Internet DIF*, application-specific DIFs or a number of VPN DIFs. These VPN DIFs will usually have a greater scope than just the GÉANT network, and will typically involve a number of NRENs and even regional and/or campus networks if they span to end users in labs, for example. Therefore the setup of these VPN DIFs will require the collaboration of a number of management domains, as explained in section 6.7. As in the case with NREN networks, VPN DIFs can also be implemented closer to the physical medium if overlaying them over the *GÉANT DIF* is not enough. The *VPN DIF option 2* in

Figure 28 provides a graphical example of this situation.

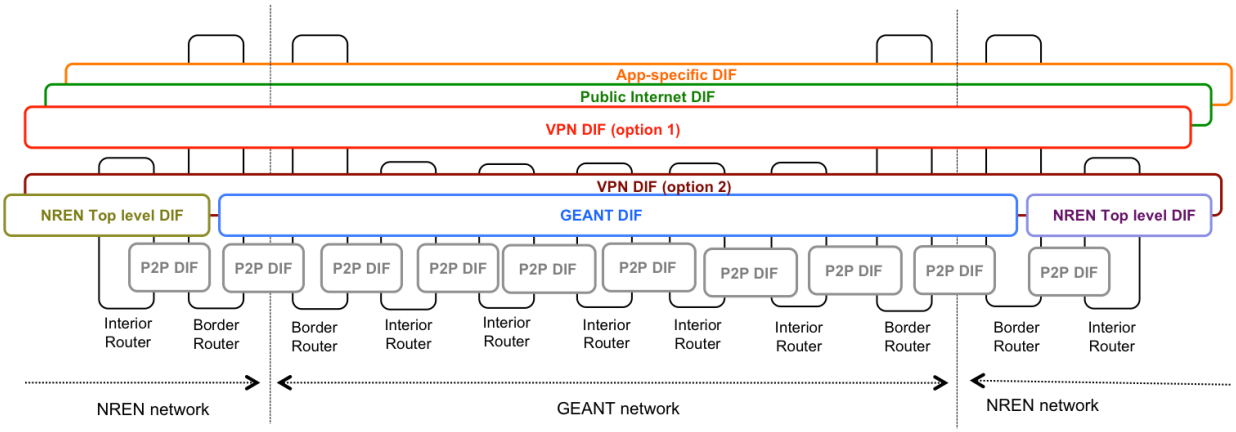


Figure 28: Model of the GÉANT network with the RINA architecture

One of the DIFs available at GÉANT's border routers is the "*public Internet DIF*". NRENs thus can get access to the public Internet via GÉANT, as an alternative or as a complement to the peering with commercial ISPs either directly or via an Internet eXchange Point (IXP). Note that the "public Internet" need not be the only Internet(work) available, therefore other ones could be available via GÉANT. For example, community or application-specific Internets such as "*High-energy physics*", "*HD Videoconferencing*" or "*Radio-astronomers*" to name a few. Each one would have its own policies designed to optimally support the community or application for which the DIF was designed.
In addition to GÉANT, NRENs can also connect directly to each other via Cross Border Dark Fibre (CBDF) or other means. In some cases, where it makes sense for a number of NRENs to reach a common physical location, NREN exchange points could be setup in order to allow different NRENs to directly connect to each other, similar to the way Internet eXchange Points operate.

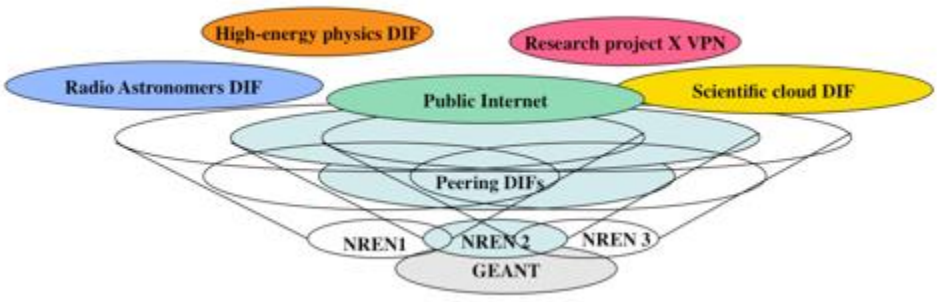


Figure 29: NRENs can operate together federated DIFs specialized to support different applications for the research and education communities

With the collaboration of GÉANT or through direct interconnection NRENs can get together to operate a number of *e-mall* DIFs in common, as illustrated in Figure 29. These NREN federations/alliances can operate international DIFs designed to support a number of specialized applications for the research and education community: acquiring and analysing data from the Large Hadron Collider (LHC), radio-astronomy, e-health, etc. In order to operate one of these federated DIFs, the participating NRENs have to agree on all the policies of the DIF (authentication, access control, data transfer, routing,

addressing, resource allocation, etc) as well as in the definition of the QoS cubes supported by this DIF.

6.4 Connecting to commercial ISPs (and other e-mall providers in general)

NRENs have two main options to connect to commercial ISPs: either a direct peering or using the services provided by an IXP. An IXP provides a common network to which a number of commercial ISPs and their potential customers have access in order to establish peering relationships with each other in a more dynamic fashion. The IXP network can be concentrated in a single physical location or spread over a number of locations (such as the SFINX IXP operated by RENATER, which has two PoPs).

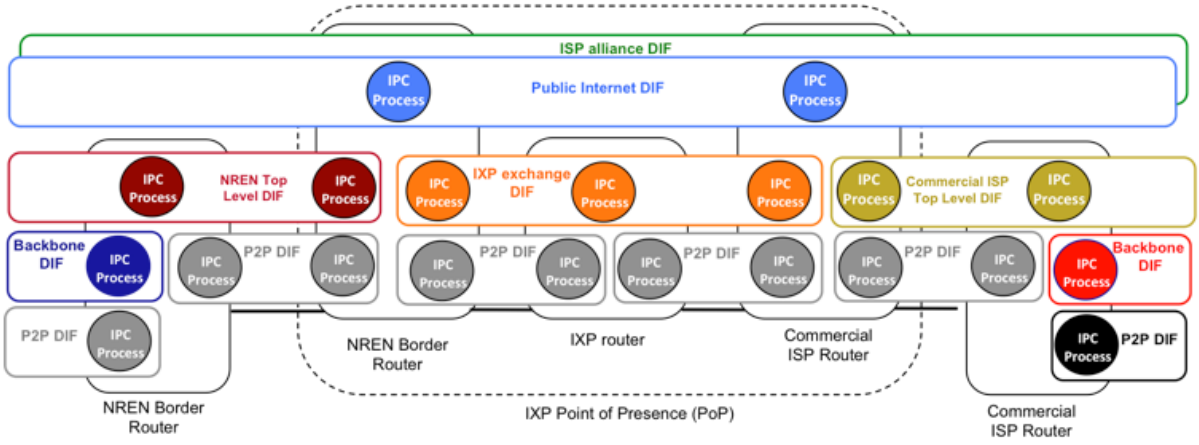


Figure 30: Interconnection of NREN with commercial ISPs via IXP exchange point

Figure 30 illustrates how this structure can be modelled using RINA. The IXP network is made by a series of border routers to which commercial ISPs and their customers can connect, and zero or more internal routers that connect the border routers together. An *IXP exchange DIF* spans the whole IXP network and allows entities attached to the IXP to allocate flows to each other. The obvious benefit for an NREN to connect to commercial ISPs is to get better connectivity to the *public Internet DIF,* but there can be others. ISPs alone or in collaboration with other ISPs can also provide access to other e-mall DIFs specialized for other applications such as voice, HD video, Content Delivery Networks (CDNs) etc. Similarly, commercial ISPs could extend specialized DIFs operated by NRENs to corporate customers wishing to access them or even the general public. For example, some companies could have a collaborative project together with a number of academic institutions researching an innovative application enabled by a DIF with very specific policies.
Figure 31 shows an example of the connectivity graph of an *IXP exchange DIF* in which the IXP has two IPC Processes (IPCPs) in different locations, redundantly connected. Customers of the IXP (NRENs or commercial ISPs) connect to one of the IXP IPCPs.

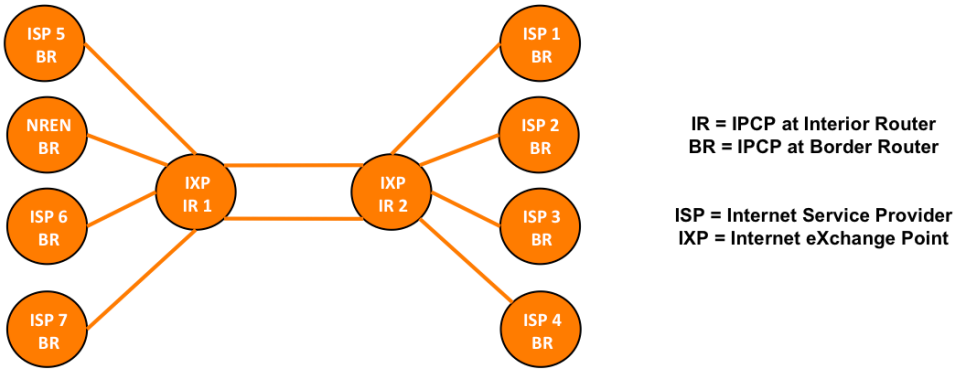


Figure 31: Example connectivity graph of an IXP exchange DIF

6.5 Connecting to customers

NRENs have two main types of customers: regional networks and customer (usually campus) networks. The use case analysis focuses on the second type of customers, although the design would also apply to the case of regional networks with minor variations. There are two main types of services that NRENs provide to their customers: i) access to well-known "e-mall" DIFs of which the public Internet is currently the most well-known example and ii) creation of and/or access to "private" VPN DIFs on demand with certain characteristics.

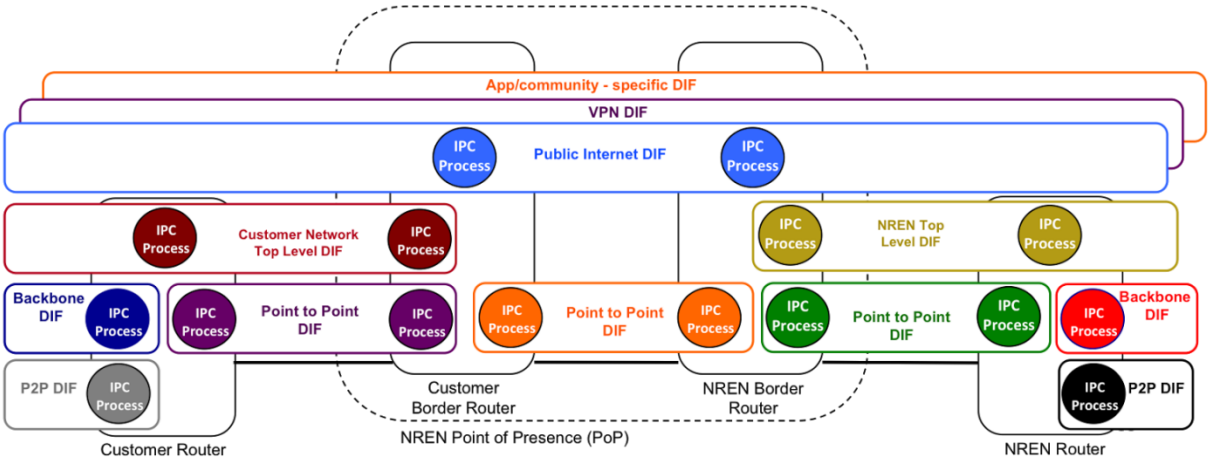


Figure 32: Interconnection between an NREN and a customer network (I)

There are also a number of options that NRENs and their customers can use to connect together in a common PoP.

Figure 32 provides a first example, in which the NREN border router connects only via a Point to Point DIF to the Customer Border Router. The NREN top level DIF finishes at the NREN border router and is not exposed to the customer network. The customer border router has access to several DIFs available via the NREN border router: the *Public Internet DIF*, VPN DIFs and well known application/community

- specific DIFs. In this example all the NREN border routers are members of the *Public Internet DIF*, as is the case of RENATER in the real world.

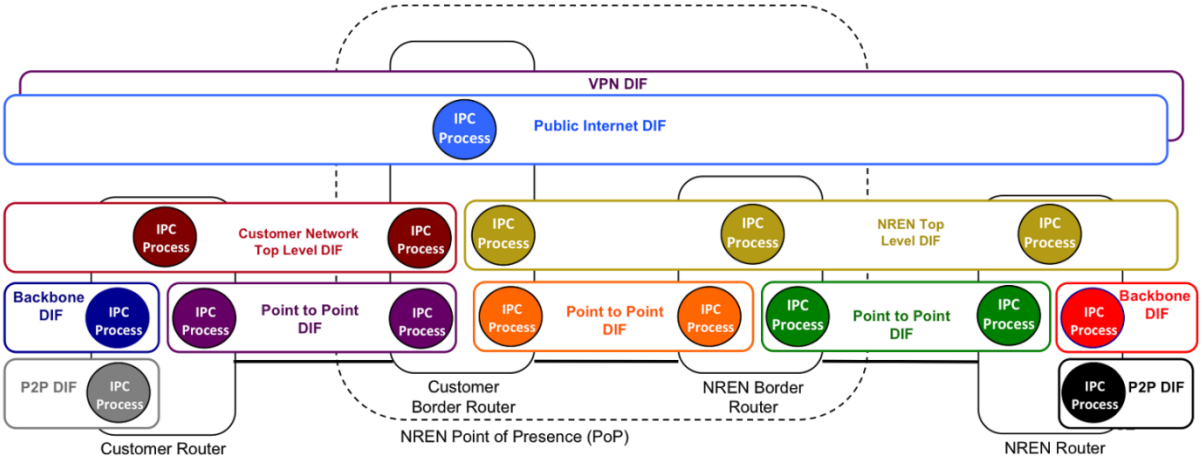


Figure 33: Interconnection between an NREN and a customer network (II)

In another example provided by

Figure 33, the NREN border router does not provide direct access to the *Public Internet DIF*, but *NREN top level DIF*s spans to the customer border router. This allows the NREN to centralize all "public Internet" presence in a few locations (such as in the SURFnet case, in which only two Internet routers are present).

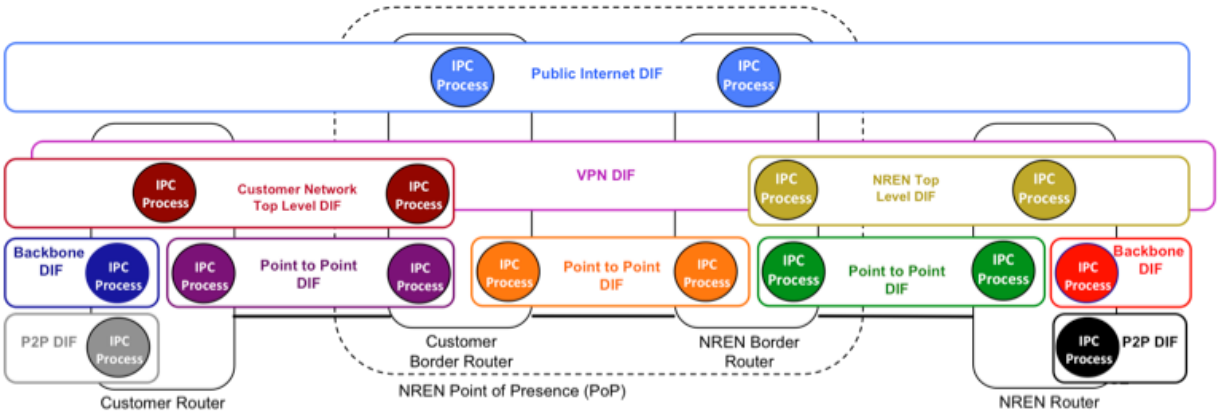


Figure 34: Interconnection between an NREN and a customer network (III)

Figure 34 provides a third example, in which the NREN provides a VPN DIF closer to the lower layers, parallel to the customer and NREN provider top DIFs. The NREN has to publicly expose more internal systems compared to the first example in this section and therefore has higher security risks. In addition to this the VPN has more IPC Processes and is therefore more difficult to manage (instantiate, monitor, destroy). However, there may be cases that require this type of VPN, and its configuration is possible with the RINA architecture.

6.6 Internal Data Centre connectivity

NRENs are in an excellent position to provide high-quality private computation environments for scientific applications to their customers, since - in some cases - they can control all the resources between the data centre (DC) and the customer network.

Figure 35 illustrates an example of this environment, in which the NREN owns a DC directly connected to the NREN top level DIF, which in turn provides connectivity to a customer network. The NREN can provide a computation environment tailored to the customer needs via VPN DIFs.

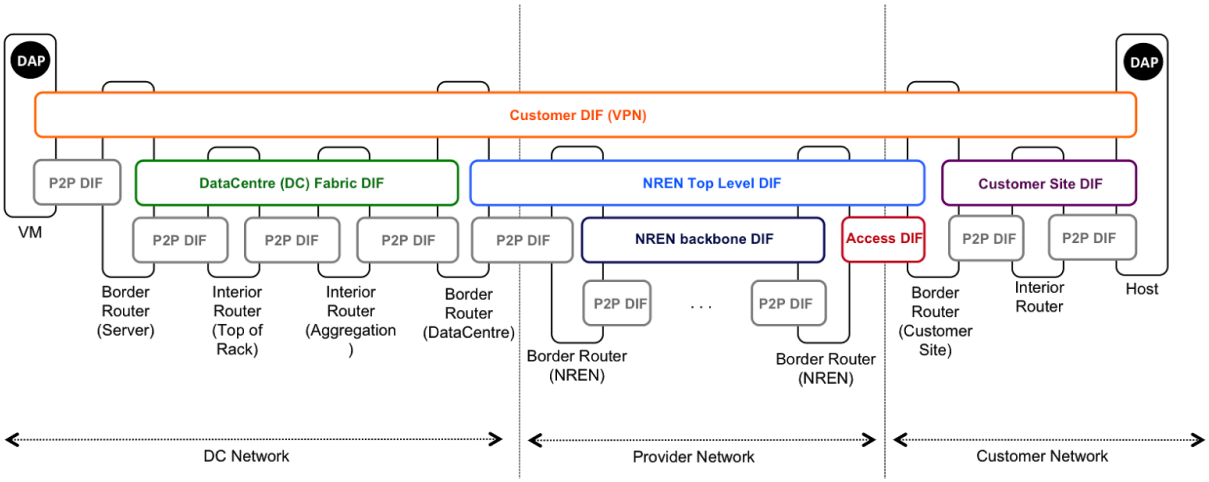


Figure 35: NREN providing customized computing environment to a directly connected customer

The Data Centre Network is modelled using the following systems: i) Virtual Machines (VMs) contain applications; ii) Servers host VMs (and have internal DIFs to communicate to them), and act as border routers for VMs; iii) Top of Rack (ToR) routers interconnect Servers of the same rack; iv) Aggregation (A) routers interconnect ToRs following a multi-stage Close Fabric or leaf-spine connectivity graph, as shown in

Figure 35; v) DC border routers interconnect the DC network to the external world, in this case the NREN network.

Figure 36 shows the connectivity graph of the *DC Fabric DIF*, whose goal is to provide any-to-any connectivity between the DC resources, making the DC a single resource pool. The DC Fabric supports a number of higher-level DIFs (dotted rectangles in the Figure) that partition the DC into different customizable networking domains. The NREN DC has the ability to provide private computing domains that span to the customer's premises, bringing a number of VMs together supported by VPN DIFs (in purple and orange in the Figure). Some VMs may be directly accessible via the public Internet DIF (in red in the Figure).

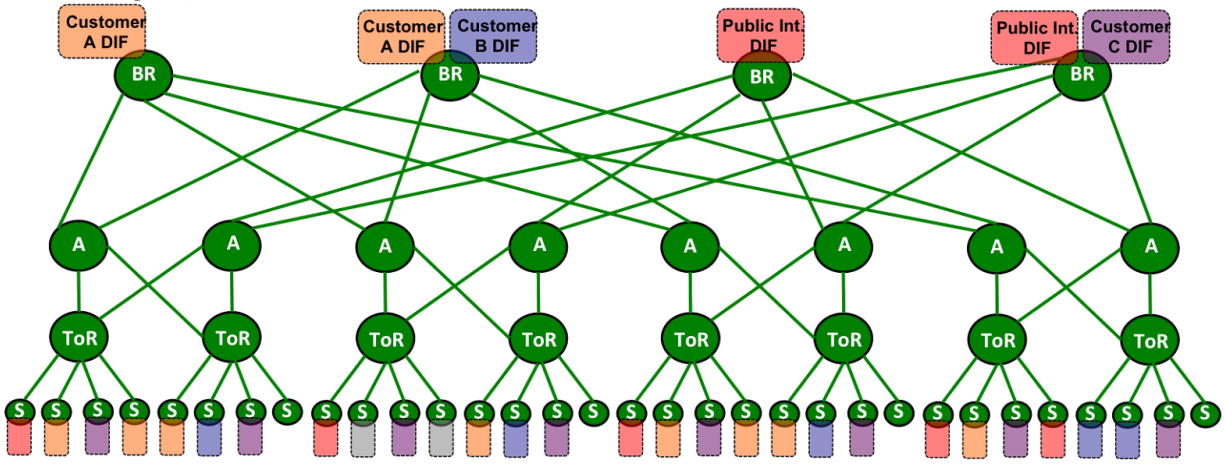


Figure 36: Connectivity graph of the DC Fabric DIF

Policies in the DC Fabric DIF are optimized for the DC environment, and could be variations of the following ones. The connectivity graph could follow a "fat-tree" or "leaf-spine" design[95] in order to guarantee full bisection bandwidth and no oversubscription. Resource allocation policies should be able to effectively multiplex flows of different requirements in terms of capacity, loss and delay since a broad range of applications can be deployed in the DC. The DC designer has full control over the connectivity graph and therefore can assume it is fixed: hierarchical addressing can be used to facilitate forwarding and minimize the size of routing tables. Finally, since this DIF is not exposed outside of the DC and all of its member IPCPs will be instantiated in resources controlled by the DC provider, security policies need not be too strict (encryption may not be required, authentication may be optional).

[95] M. Al Fares, A. Loukissas, A. Vahdat; "A Scalable, Commodity Data Centre Architecture", SIGCOMM 2008.

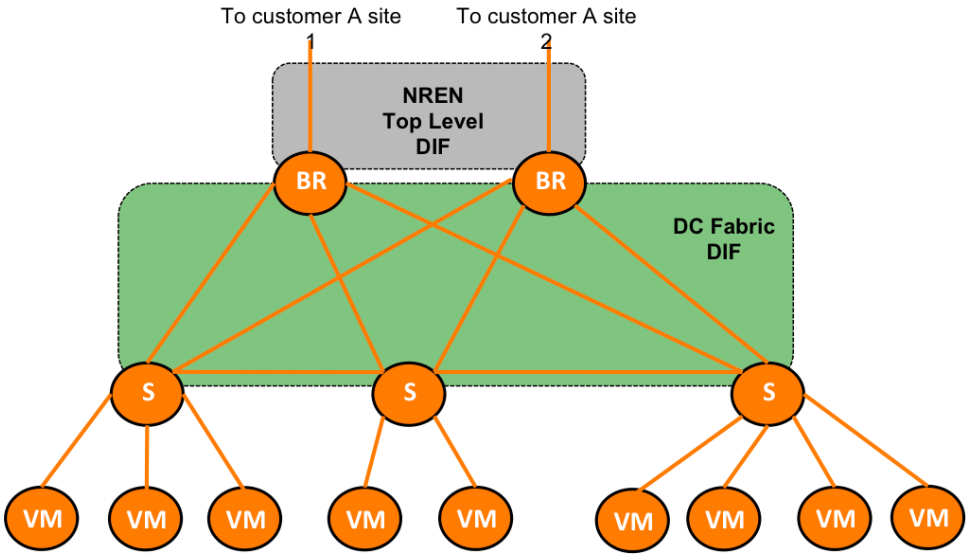


Figure 37: Example partial connectivity graph of a VPN DIF that spans to two sites of a customer

VPN DIFs will typically have an arbitrary connectivity graph and custom policies tailored to the needs of the applications the VPN DIF will support. These VPN DIFs, as shown in

Figure 37, connect customer VMs together and to other computing/storage resources at the customer site(s). DC VPN DIFs are supported by the DC Fabric DIF, the NREN Top Level DIF and one or more DIFs belonging to the customer (not shown in the Figure).

6.7 Network Management

This section discusses a very important aspect of the operation of NRENs and GÉANT: network management. It particularly focuses on the interactions between management domains in order to setup and operate multi-owner DIFs such as VPN DIFs.

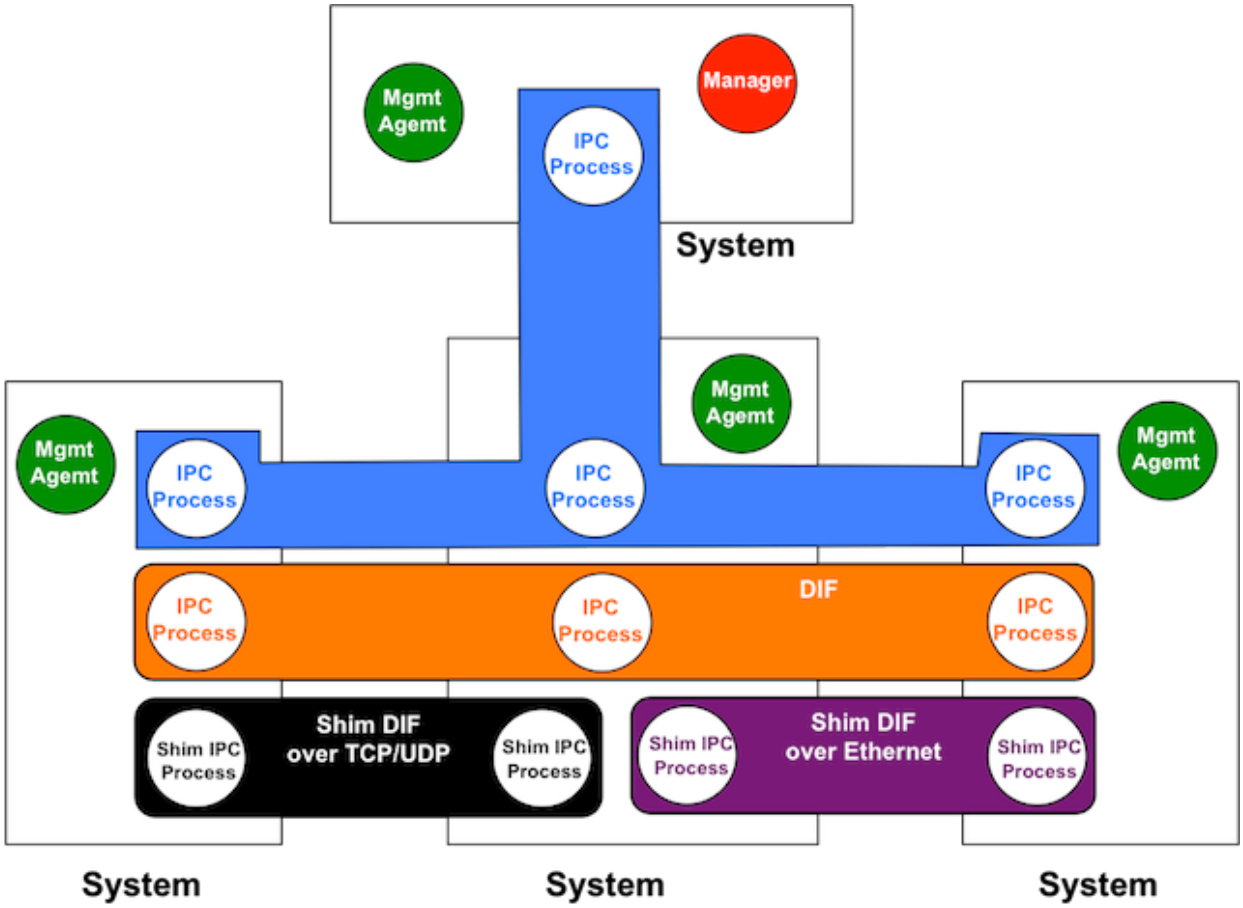


Figure 38: Typical configuration of a centralized management system in RINA

Figure 38 shows an example of the configuration of a management system for a simple RINA network. All the managed computer systems have at least one management agent with full permissions to access the RIBs (Resource Information Bases) of all the IPC Processes in the computer system. This allows the Management Agent (MA) to perform any operation on the IPC Process and to access its internal state; as well as to expose these capabilities remotely. Management Agents will typically be managed by a centralized Manager process (one per management domain), who will operate on its RIB in order to create, destroy and monitor the IPC Processes. The Management Agents and the Manager(s) together form the Network Management DAF (Distributed Application Facility), which is typically supported by a dedicated DIF - although it doesn't need to be. In addition to the traditional centralized configuration, other options for the Network Management System are possible; ranging from fully distributed (no Managers, just Management Agents collaborating on a peer to peer basis) to the centralized option discussed before.

Another interesting property is the ability to instantiate multiple Management Agents in the same computer system. Although one MA needs to be the "main MA" with full permissions, other MAs with less privileges can be instantiated and put under the control of other Management Systems. One possible use case for this feature is the so-called "*Network as a Service*" offering, in which the owner of the computer system may delegate the management of some IPC Processes to its customers. This

would allow companies without infrastructure to become operators of RINA networks, or customers to integrate the management of NREN-provided VPNs with their own Management System.
The biggest advantage that RINA provides to network management is commonality. All layers (DIFs) have a single information model expressed via a common Resource Information Base schema[96]. This schema can be further extended by the different policies that are instantiated in each layer. Commonality between layers reduces complexity, which is the biggest enemy of effective network management. This structure should allow for more predictable behaviour when applying changes to multiple layers, opening the door to a higher level of automation in network operations. The ultimate vision is that network management should just be "monitor and repair" but not "control". The operator should be able to specify different policy configurations for a DIF, for different network conditions. The Management System should just monitor the network conditions and automatically apply the configuration changes when transitioning between known operational regions. Human operators would only need to be involved when the network was operating in unplanned conditions (such as natural disasters).

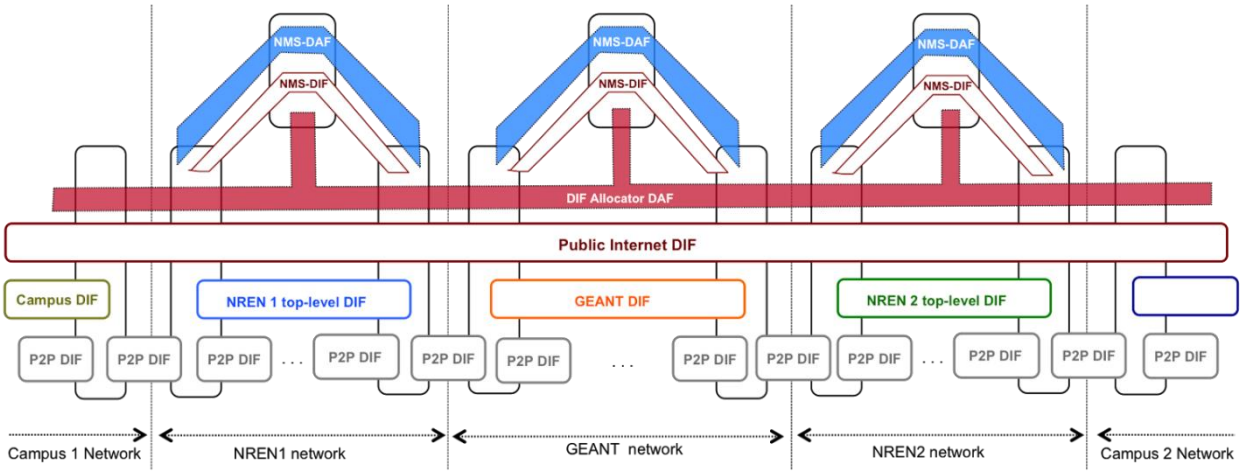


Figure 39: Different Management Domains (NMS DAFs) and the DIF Allocator DAF

Figure 39 shows a simplified example of two campus networks, attached to two NRENs, interconnected via GÉANT. The NMS-DAF of the NRENs and GÉANT is shown in blue, and its respective supporting DIFs in brown. The NMS-DAF of each NREN or GÉANT manages the DIFs that are internal to the NREN or GÉANT network, as well as its portion of DIFs that span multiple management domains, such as the Public Internet DIF or VPN DIFs. The example depicts a centralized NMS-DAF configuration for each domain, in which a single Manager process manages all the computing systems in its management domain via the Management Agents.

NMS-DAFs are restricted to a single management domain, and therefore cannot directly communicate with Managers in other management domains. However, there are some tasks, such as the setup of DIFs spanning multiple management domains - VPNs, for instance - that

[96] The PRISTINE consortium, "Draft specification of the common elements of the management architecture", Deliverable D5.1, July 2014. Available online at http://ict-pristine.eu/?page_id=37

require the collaboration of several management systems, since IPC Processes in various domains have to be created and properly configured to form the new DIF. This is achieved via the "*DIF Allocator DAF*", shown in garnet in

Figure 39. The *DIF Allocator DAF* is a distributed application whose goal is to enable the collaboration between multiple Management Domains in order to setup or modify DIFs spanning multiple networks.

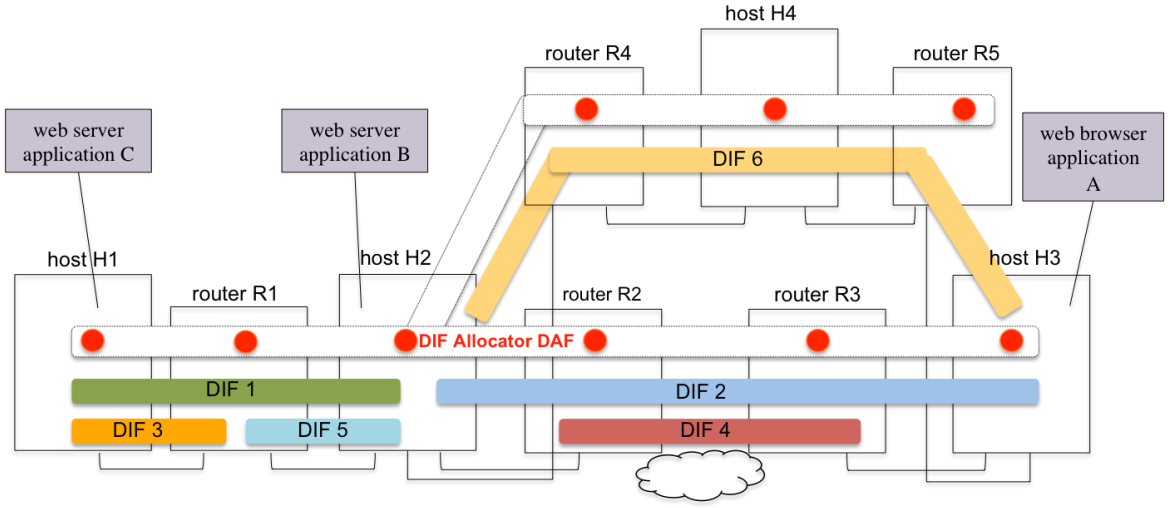


Figure 40: Example of use of the DIF Allocator DAF

The *DIF Allocator DA*F is an example of a Namespace Manager DAF (or NSM-DAF). NSM-DAFs. To manage a name space in a distributed environment requires coordination to ensure that the names remain unambiguous and can be resolved efficiently. The need for the DIF allocator comes from the realisation that all the applications of a certain namespace may not be available in the same DIF (for example, the public Internet), but they need to be discoverable anyway. An example of the DIF Allocator DAF is shown in

Figure 40. A web browser application named "A" in host "H3" wants to establish a flow to a web server application "C" in host "H1", therefore it invokes the flow allocation primitive of the RINA API. The RINA software in "H3" checks that application "C" is not available from any of the DIFs locally available in "H3", and therefore submits a request to the DIF Allocator DAF. The DIF Allocator DAF is structured to maintain a distributed mapping of application process names to DIF name, so the query is forwarded through the DIF Allocator DAF until it reaches the DIF Allocator DAF process of the system in which the instance of application "C" is executing (in host "H1"). Hosts "H1" and "H3" don't share a DIF in common, therefore the DIF allocator DAF has to either enlarge an existing DIF so that it has enough scope for "A" and "C" to communicate, or create a new DIF on top of other existing ones. The creation of the new DIF will involve the interaction of the DIF Allocator DAF with multiple Management Systems, since new IPC Processes will have to be instantiated and configured. More details about how the DIF Allocator DAF works are provided in "Layer discovery in RINA networks"[97] (just take into

[97] E. Trouva, E. Grasa, J. Day, S. Bunch, "Layer discovery in RINA networks". 17th International Workshop on Computer Aided Design and Modeling, CAMAD 2012 (Barcelona, September 2012).

account that the *DIF Allocator DAF* is called *Inter DIF Directory or IDD* in this document, which was its old name).
NRENs, GÉANT and their customers could jointly setup and operate one or more DIF Allocator DAFs dedicated to manage one or more scientific/research application namespaces and to create VPN DIFs supporting these applications. These DIF Allocator DAFs would locate applications through a set of DIFs and collaborate with the Network Management Systems of NRENs, GÉANT and customer networks to dynamically setup, grow or shrink VPNs that connect together instances of that application across several management domains.

6.8 Application-specific DIFs

This section analyses an example of an application-specific DIF setup to support the operation of a distributed collaboration application such as the one described in section 1.5. The SeeVogh distributed application infrastructure is currently setup as a peer to peer network of USHI instances overlaid over the Internet. The USHI instances themselves are responsible for routing the traffic in the USHI peer to peer network. This is a scenario that can be used with the RINA architecture as well, but another interesting design is also possible: create an application-specific DIF that supports the operation of the distributed collaboration application.

The "*SeeVogh DIF*" allows all USHI instances to be one hop away from each other - releasing the application from having to perform IPC-specific tasks. It also allows the collaboration application to rely on the IPC capabilities provided by the network, such as multi-homing, mobility or multicast; as well as to leverage dedicated hardware capable of processing large amounts of traffic if required.

Figure 41 shows an example of a possible connectivity graph of the "*SeeVogh DIF*", as well as all the N-1 DIFs it relies on. USHIs are running on VMs hosted in NRENs data centres, which have access to the NREN top-level DIF. Campus Networks can also run USHI instances or just SeeVogh clients that allow researchers to use the collaboration services provided by the SeeVogh infrastructure. The "SeeVogh" DIF may add an extra layer of security by requiring the authentication of any IPC Process joining it.

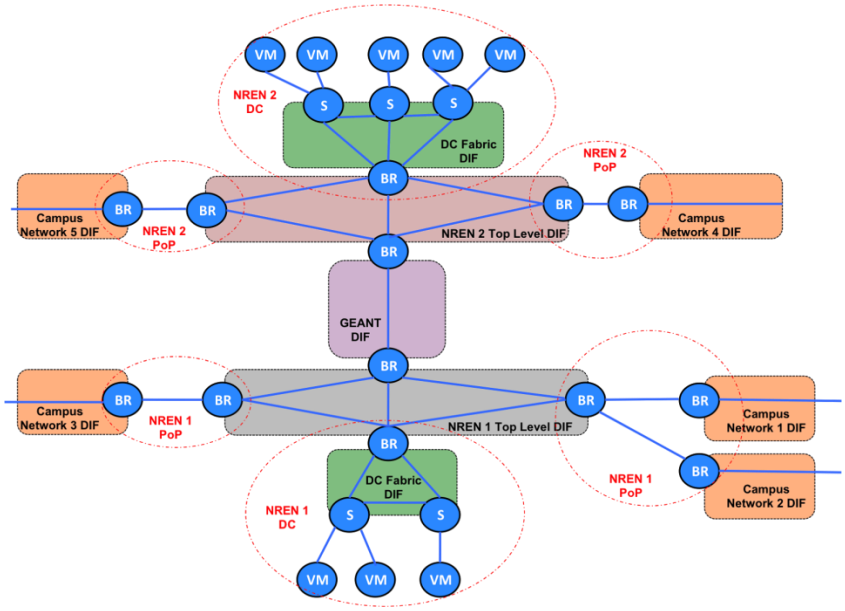


Figure 41: Example connectivity graph of a SeeVogh DIF

7 RINA Traffic Generator for simulation of video traffic

T3.2 identified that the IRATI stack lacks of a traffic generation tool complying with the requirements of the IRINA experimentation objectives since it only provides a RINA-based tool - the `rina-echo-time` application - which only performs basic ping functionality and provides rudimentary bandwidth testing capabilities. In order to comply with IRINA's requirements two main options were available: modify an existing traffic generation tool or write a new one from scratch. Modifying an existing tool would require either rewriting part of its existing code - mainly adapt its networking related functionalities to the RINA API - or divert its socket API calls to functions wrapping the RINA API that would be mimicking the sockets POSIX signatures.
The following sections present the analysis of the existing traffic generation tools that were considered for the porting activities.

7.1 Existing traffic generation tools

IRINA identified three possible tools for porting to RINA: *netperf*[98], *D-ITG*[99] and *Ostinato*[100].

7.1.1 netperf

`netperf` is a benchmark that can be used to measure the performance of many different types of networking. It provides tests for both unidirectional throughput and end-to-end latency. The environments currently measurable by `netperf` include:

- TCP and UDP via BSD Sockets for both IPv4 and IPv6
- DLPI
- Unix Domain Sockets
- SCTP for both IPv4 and IPv6

[98] http://www.netperf.org/netperf/

[99] http://traffic.comics.unina.it/software/ITG/

[100] https://code.google.com/p/ostinato/

`netperf` does not support advanced traffic models for simulating applications, only simple packet generation for bandwidth and latency testing. This functionality of `netperf` is already mostly present in the `rina-echo-time` application developed by IRATI.

7.1.2 D-ITG

D-ITG (Distributed Internet Traffic Generator) is a platform capable to produce traffic at packet level accurately replicating appropriate stochastic processes for both IDT (Inter Departure Time) and PS (Packet Size) random variables (exponential, uniform, cauchy, normal, pareto).
D-ITG supports both IPv4 and IPv6 traffic generation and it is capable to generate traffic at network, transport, and application layer.
D-ITG currently supports the following operating systems:

- Linux (Ubuntu, Debian, Fedora, CentOS, OpenWRT)
- Windows (XP, Vista, 7)
- OSX (Leopard)
- FreeBSD

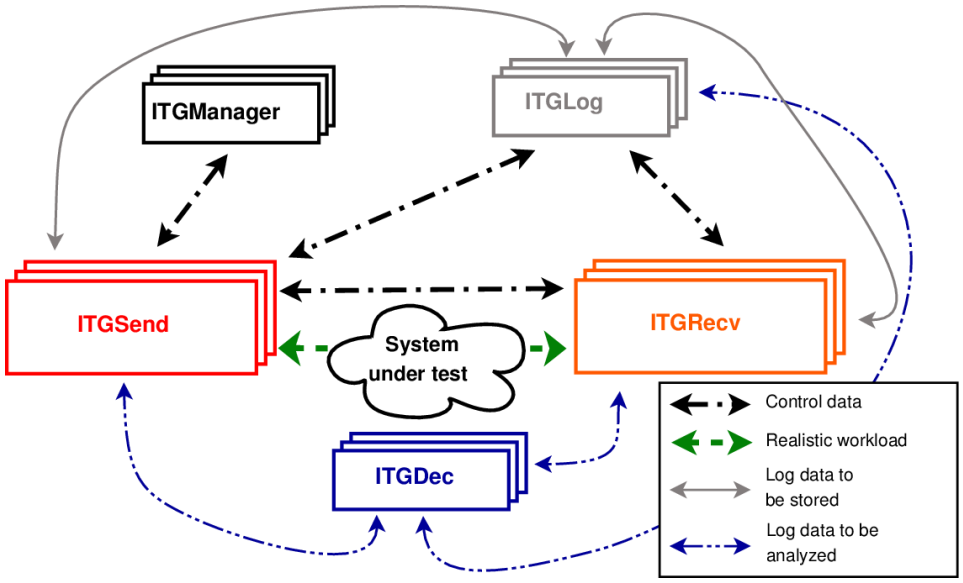


Figure 42: D-ITG

Analysis of the D-ITG tool led to the following conclusions:

- It has support for multi-user scenarios
- Powerful traffic generation and logging system
- Easy management from a central location
- Albeit with a modular architecture in general, the low-level parts of the traffic generator components (i.e. ITGSend and ITGRecv) are tightly bound and intertwined, preventing an affordable approach for the necessary modifications that would allow it to run in a RINA environment.

7.1.3 Ostinato

Ostinato is an open-source, cross-platform network packet crafter/traffic generator and analyser with a friendly GUI. It can craft and send packets on several streams with different protocols at different rates. It supports Windows, Linux, BSD and Mac OS X and the most common standard protocols such as Ethernet/802.3/LLC SNAP/VLAN (with QinQ), ARP, IPv4, IPv6, IP-in-IP a.k.a IP Tunnelling (6over4, 4over6, 4over4, 6over6), TCP, UDP, ICMPv4, ICMPv6, IGMP, MLD and any based protocol (HTTP, SIP, RTSP, NNTP etc.)
Analysis of the Ostinato tool led to the following conclusions:

- It supports a lot of different protocols, and allows detailed analysis, but this is not really that useful for IRINA
- it allows configuring stream rates, bursts, no. of packets but does not seem to support advanced traffic models.

7.2 IRINA test tool: rina-tgen

Following the analysis of existing software, we concluded that it would require less effort to implement a basic traffic generation algorithm into a new application to emulate the video service identified in the use case. These algorithms will be taken from the IEEE 802.16 study group, which analysed different services such as VoIP and Video traffic, and validated models based on Interrupted Poisson Processes (IPP), Interrupted Renewal Processes (IRP) and Interrupted Discrete Processes (IDP). By having multiple independent such processes, HTTP and FTP can be simulated as 4IPP; VoIP can be simulated as IDP, 2IDP, 4IDP; Video can be simulated as 2IRP[101].
The `rina-tgen` tool is currently supporting constant bit rate and random traffic with poisson distributed interarrival times. It is available as Free Open Source Software under the GÉANT license[102]. It links with the BOOST C++ libraries[103] and uses TCLAP[104] for parsing the command line arguments. The IRATI stack must be installed for `rina-tgen` to compile.

```
USAGE:

./rina-tgen   [-l] [--interval <unsigned integer>] [-c <unsigned integer>] [--
duration <unsigned integer>] [--rate <unsigned integer>] [-s <unsigned integer>]
[--distribution  <string>]  [--poissonmean  <double>]  [--qoscube  <string>]  [-d
<string>]  [--client-api <string>] [--client-apn <string>] [--server-api <string>]
[--server-apn <string>] [-r] [--sleep] [--] [--version] [-h]

WHERE:
```

option		value	effect
-l	--listen		Run in server (consumer) mode
	--interval	<uint>	Report statistics every x SDUs (server)
-c	--count	<uint>	Number of SDUs to send, 0 = unlimited

[101] http://ieee802.org/16/tg3/contrib/802163c-01_30r1.pdf, accessed March 2015

[102] http://github.com/IRATI/traffic-generator

[103] http://www.boost.org/

[104] http://tclap.sourceforge.net/

	--rate	<uint>	Bitrate to send the SDUs, in kb/s, 0 = no limit
-s	--size	<uint>	Size of the SDUs to send (bytes)
	--distribution	<string>	Distribution type: CBR, poisson
	--poissonmean	<double>	The mean value for the poisson distribution used to generate inter-arrival times, default is 1.
	--qoscube	<string>	Specify the qos cube to use for flow allocation.
-d	--dif	<string>	The name of the DIF to use (empty means 'any DIF')
	--client-api	<string>	Application process instance for the client
	--client-apn	<string>	Application process name for the client
	--server-api	<string>	Application process instance for the server
	--server-apn	<string>	Application process name for the server
-r	--register		Register the application to any DIF
	--sleep		Sleep instead of busywait between sending SDUs
--	--ignore-rest		Ignores the rest of the labelled arguments following this flag.
	--version		Displays version information and exits.
-h	--help		Displays usage information and exits.

Table 5: Command line options for the rina-tgen tool (v1.0.1)

8 Deployment in the iLab.t test bed

8.1 Test bed scenario

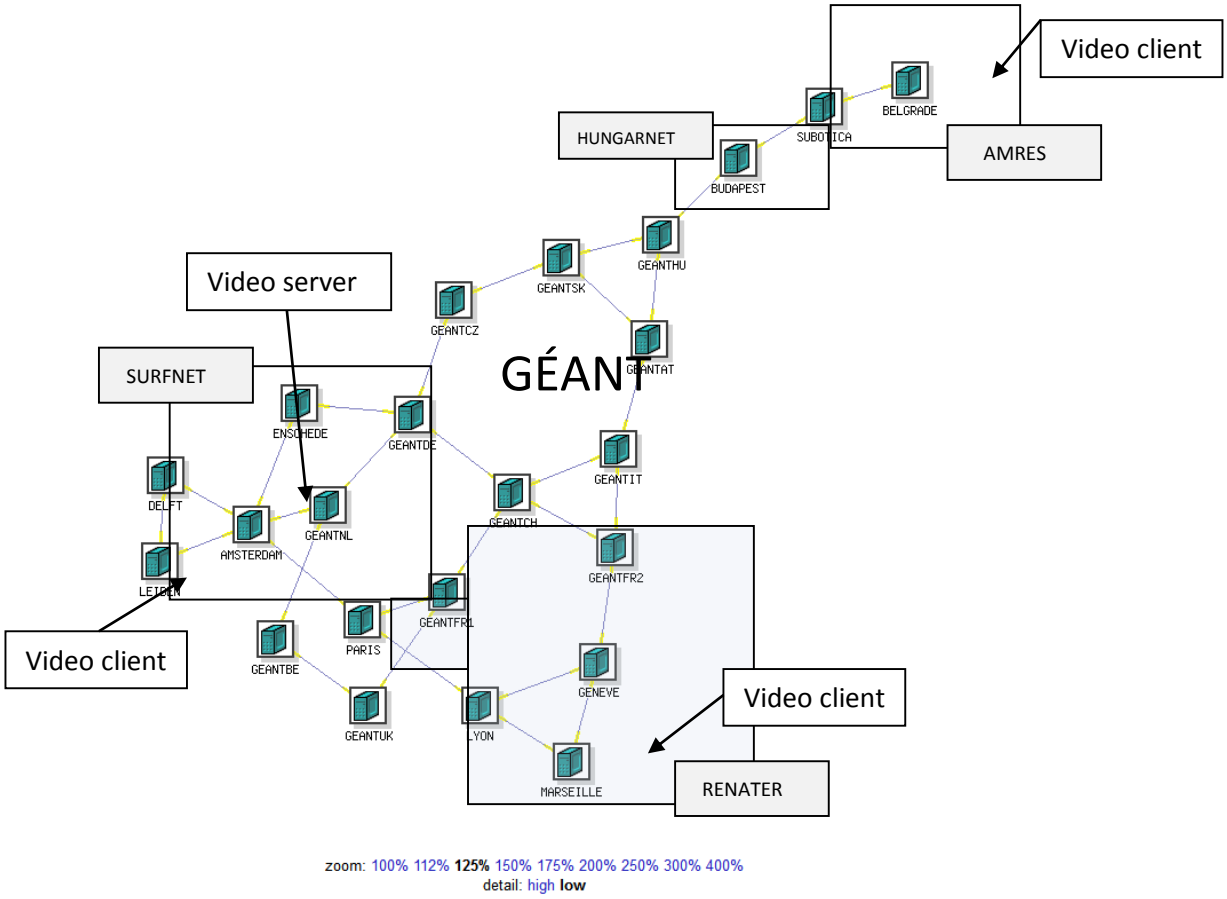


Figure 43: Test bed deployment

Figure 43 shows the testbed implementation of the minimal scenario (

Figure 14). Based on this deployment, we identified two aspects to be important to illustrate for IRINA: the functioning of the traffic generator, developed within the project (Section 8.2) as an early proof-of-concept. To show a clearer benefit of RINA for videoconferencing, we look at a deployment where

the server (SeeVogh) is running in a virtualised environment inside a datacentre. We demonstrate the progress beyond the state-of-the-art by analysing the efficiency of VM-to-host communication (Section 8.3).

8.2 `rina-tgen` between client and server

Here, we demonstrate the behaviour of the rina-tgen application developed by IRINA; which mimics Video traffic using the poisson distribution.

The server will report bandwidth statistics after every 100 SDU's it receives.

```
./rina-tgen --listen --interval 100
```

For video applications, a data rate of up to 8Mb/s is recommended (See Table 2). We set the duration to 5 seconds, and run the test with a large SDU size. The test is performed three times using a poisson distribution with different mean values, and once with constant bitrate.

```
#/usr/local/irati/bin$ ./rina-tgen --size 1470 --rate 8000 --duration 5 --distribution
poisson --poissonmean 0.1

32018(1426778619)#concurrency (DBG): Lockable created successfully
32018(1426778619)#concurrency (DBG): Lockable created successfully
32018(1426778619)#logs (DBG): New log level: INFO
32018(1426778619)#netlink-manager (INFO): Netlink socket connected to local port 32018
32018(1426778619)#traffic-generator (INFO): starting test
32018(1426778624)#traffic-generator (INFO): sent statistics: 3306 SDUs, 4859820 bytes in
5012700 us
32018(1426778624)#traffic-generator (INFO):     => 7.7560 Mb/s

#/usr/local/irati/bin$ ./rina-tgen --size 1470 --rate 8000 --duration 5 --distribution
poisson --poissonmean 1
32020(1426778633)#concurrency (DBG): Lockable created successfully
32020(1426778633)#concurrency (DBG): Lockable created successfully
32020(1426778633)#logs (DBG): New log level: INFO
32020(1426778633)#netlink-manager (INFO): Netlink socket connected to local port 32020
32020(1426778633)#traffic-generator (INFO): starting test
32020(1426778638)#traffic-generator (INFO): sent statistics: 3397 SDUs, 4993590 bytes in
5002410 us
32020(1426778638)#traffic-generator (INFO):     => 7.9859 Mb/s

#/usr/local/irati/bin$ ./rina-tgen --size 1470 --rate 8000 --duration 5 --distribution
poisson --poissonmean 10
32022(1426778645)#concurrency (DBG): Lockable created successfully
32022(1426778645)#concurrency (DBG): Lockable created successfully
32022(1426778645)#logs (DBG): New log level: INFO
32022(1426778645)#netlink-manager (INFO): Netlink socket connected to local port 32022
32022(1426778645)#traffic-generator (INFO): starting test
32022(1426778650)#traffic-generator (INFO): sent statistics: 3387 SDUs, 4978890 bytes in
5001234 us
32022(1426778650)#traffic-generator (INFO):     => 7.9643 Mb/s

#/usr/local/irati/bin$ ./rina-tgen --size 1470 --rate 8000 --duration 5 --distribution CBR
32024(1426778658)#concurrency (DBG): Lockable created successfully
32024(1426778658)#concurrency (DBG): Lockable created successfully
32024(1426778658)#logs (DBG): New log level: INFO
32024(1426778658)#netlink-manager (INFO): Netlink socket connected to local port 32024
32024(1426778658)#traffic-generator (INFO): starting test
32024(1426778663)#traffic-generator (INFO): sent statistics: 3403 SDUs, 5002410 bytes in
5000940 us
32024(1426778663)#traffic-generator (INFO):     => 8.0024 Mb/s
```

Table 6: Client side statistics running rina-tgen

From the client output (Table 6), we see that the tool outputs traffic at a rate close to the requested rate. The CBR is of course the most accurate. Using higher values for the mean of the poisson distribution smooths out the traffic faster.

```
#/usr/local/irati/bin$ ./rina-tgen --listen --interval 100
22827(1426778598)#concurrency (DBG): Lockable created successfully
22827(1426778598)#concurrency (DBG): Lockable created successfully
22827(1426778598)#logs (DBG): New log level: INFO
22827(1426778598)#netlink-manager (INFO): Netlink socket connected to local
port 22827
#RESULTS FOR POISSON. MEAN 0.1
22827(1426778619)#traffic-generator (INFO): New flow allocated [port-id =
1]
22827(1426778619)#traffic-generator (INFO): Starting test:
        duration: 5
        count: 0
        sduSize: 1470
22827(1426778619)#traffic-generator (INFO): 100 SDUs in 132493 us => 8.8759
Mb/s
22827(1426778619)#traffic-generator (INFO): 100 SDUs in 132293 us => 8.8894
Mb/s
22827(1426778619)#traffic-generator (INFO): 100 SDUs in 146611 us => 8.0212
Mb/s
22827(1426778619)#traffic-generator (INFO): 100 SDUs in 147210 us => 7.9886
Mb/s
22827(1426778620)#traffic-generator (INFO): 100 SDUs in 220222 us => 5.3401
Mb/s
22827(1426778620)#traffic-generator (INFO): 100 SDUs in 88175 us => 13.3371
Mb/s
22827(1426778620)#traffic-generator (INFO): 100 SDUs in 161870 us => 7.2651
Mb/s
22827(1426778620)#traffic-generator (INFO): 100 SDUs in 73623 us => 15.9733
Mb/s
22827(1426778620)#traffic-generator (INFO): 100 SDUs in 220104 us => 5.3429
Mb/s
22827(1426778620)#traffic-generator (INFO): 100 SDUs in 103195 us =>
11.3959 Mb/s
22827(1426778620)#traffic-generator (INFO): 100 SDUs in 117467 us =>
10.0113 Mb/s
22827(1426778621)#traffic-generator (INFO): 100 SDUs in 191050 us => 6.1555
Mb/s
22827(1426778621)#traffic-generator (INFO): 100 SDUs in 146951 us => 8.0027
Mb/s
22827(1426778621)#traffic-generator (INFO): 100 SDUs in 205581 us => 5.7204
Mb/s
22827(1426778621)#traffic-generator (INFO): 100 SDUs in 146959 us => 8.0022
Mb/s
22827(1426778621)#traffic-generator (INFO): 100 SDUs in 59318 us => 19.8253
Mb/s
22827(1426778621)#traffic-generator (INFO): 100 SDUs in 234591 us => 5.0130
Mb/s
22827(1426778622)#traffic-generator (INFO): 100 SDUs in 146968 us => 8.0017
Mb/s
22827(1426778622)#traffic-generator (INFO): 100 SDUs in 117569 us =>
10.0026 Mb/s
22827(1426778622)#traffic-generator (INFO): 100 SDUs in 220451 us => 5.3345
Mb/s
22827(1426778622)#traffic-generator (INFO): 100 SDUs in 161857 us => 7.2657
Mb/s
22827(1426778622)#traffic-generator (INFO): 100 SDUs in 117402 us =>
10.0169 Mb/s
22827(1426778622)#traffic-generator (INFO): 100 SDUs in 161657 us => 7.2747
Mb/s
22827(1426778622)#traffic-generator (INFO): 100 SDUs in 117913 us => 9.9735
Mb/s
22827(1426778623)#traffic-generator (INFO): 100 SDUs in 220372 us => 5.3364
Mb/s
```

```
22827(1426778623)#traffic-generator (INFO): 100 SDUs in 117345 us =>
10.0217 Mb/s
22827(1426778623)#traffic-generator (INFO): 100 SDUs in 102869 us =>
11.4320 Mb/s
22827(1426778623)#traffic-generator (INFO): 100 SDUs in 161668 us => 7.2742
Mb/s
22827(1426778623)#traffic-generator (INFO): 100 SDUs in 161867 us => 7.2652
Mb/s
22827(1426778623)#traffic-generator (INFO): 100 SDUs in 190865 us => 6.1614
Mb/s
22827(1426778624)#traffic-generator (INFO): 100 SDUs in 205780 us => 5.7148
Mb/s
22827(1426778624)#traffic-generator (INFO): 100 SDUs in 117565 us =>
10.0030 Mb/s
22827(1426778624)#traffic-generator (INFO): 100 SDUs in 132254 us => 8.8920
Mb/s
#RESULTS FOR POISSON. MEAN 1
22827(1426778633)#traffic-generator (INFO): New flow allocated [port-id =
2]
22827(1426778633)#traffic-generator (INFO): Starting test:
        duration: 5
        count: 0
        sduSize: 1470
22827(1426778633)#traffic-generator (INFO): 100 SDUs in 158640 us => 7.4130
Mb/s
22827(1426778634)#traffic-generator (INFO): 100 SDUs in 149937 us => 7.8433
Mb/s
22827(1426778634)#traffic-generator (INFO): 100 SDUs in 136679 us => 8.6041
Mb/s
22827(1426778634)#traffic-generator (INFO): 100 SDUs in 164599 us => 7.1446
Mb/s
22827(1426778634)#traffic-generator (INFO): 100 SDUs in 123429 us => 9.5277
Mb/s
22827(1426778634)#traffic-generator (INFO): 100 SDUs in 126399 us => 9.3039
Mb/s
22827(1426778634)#traffic-generator (INFO): 100 SDUs in 146916 us => 8.0046
Mb/s
22827(1426778634)#traffic-generator (INFO): 100 SDUs in 147068 us => 7.9963
Mb/s
22827(1426778635)#traffic-generator (INFO): 100 SDUs in 149916 us => 7.8444
Mb/s
22827(1426778635)#traffic-generator (INFO): 100 SDUs in 154301 us => 7.6215
Mb/s
22827(1426778635)#traffic-generator (INFO): 100 SDUs in 135216 us => 8.6972
Mb/s
22827(1426778635)#traffic-generator (INFO): 100 SDUs in 173429 us => 6.7809
Mb/s
22827(1426778635)#traffic-generator (INFO): 100 SDUs in 180771 us => 6.5055
Mb/s
22827(1426778635)#traffic-generator (INFO): 100 SDUs in 144025 us => 8.1652
Mb/s
22827(1426778635)#traffic-generator (INFO): 100 SDUs in 152831 us => 7.6948
Mb/s
22827(1426778636)#traffic-generator (INFO): 100 SDUs in 139603 us => 8.4239
Mb/s
22827(1426778636)#traffic-generator (INFO): 100 SDUs in 148448 us => 7.9220
Mb/s
22827(1426778636)#traffic-generator (INFO): 100 SDUs in 136681 us => 8.6040
Mb/s
22827(1426778636)#traffic-generator (INFO): 100 SDUs in 158723 us => 7.4091
Mb/s
22827(1426778636)#traffic-generator (INFO): 100 SDUs in 135201 us => 8.6982
Mb/s
22827(1426778636)#traffic-generator (INFO): 100 SDUs in 138131 us => 8.5137
Mb/s
22827(1426778636)#traffic-generator (INFO): 100 SDUs in 120528 us => 9.7571
Mb/s
22827(1426778637)#traffic-generator (INFO): 100 SDUs in 154306 us => 7.6212
Mb/s
22827(1426778637)#traffic-generator (INFO): 100 SDUs in 176361 us => 6.6681
Mb/s
```

```
22827(1426778637)#traffic-generator (INFO): 100 SDUs in 154309 us => 7.6211
Mb/s
22827(1426778637)#traffic-generator (INFO): 100 SDUs in 138149 us => 8.5125
Mb/s
22827(1426778637)#traffic-generator (INFO): 100 SDUs in 163138 us => 7.2086
Mb/s
22827(1426778637)#traffic-generator (INFO): 100 SDUs in 149922 us => 7.8441
Mb/s
22827(1426778638)#traffic-generator (INFO): 100 SDUs in 136658 us => 8.6054
Mb/s
22827(1426778638)#traffic-generator (INFO): 100 SDUs in 132285 us => 8.8899
Mb/s
22827(1426778638)#traffic-generator (INFO): 100 SDUs in 157212 us => 7.4803
Mb/s
22827(1426778638)#traffic-generator (INFO): 100 SDUs in 135244 us => 8.6954
Mb/s
22827(1426778638)#traffic-generator (INFO): 100 SDUs in 139613 us => 8.4233
Mb/s
#RESULTS FOR POISSON. MEAN 10
22827(1426778645)#traffic-generator (INFO): New flow allocated [port-id =
3]
22827(1426778645)#traffic-generator (INFO): Starting test:
        duration: 5
        count: 0
        sduSize: 1470
22827(1426778645)#traffic-generator (INFO): 100 SDUs in 150696 us => 7.8038
Mb/s
22827(1426778645)#traffic-generator (INFO): 100 SDUs in 153482 us => 7.6621
Mb/s
22827(1426778646)#traffic-generator (INFO): 100 SDUs in 145343 us => 8.0912
Mb/s
22827(1426778646)#traffic-generator (INFO): 100 SDUs in 145172 us => 8.1007
Mb/s
22827(1426778646)#traffic-generator (INFO): 100 SDUs in 141960 us => 8.2840
Mb/s
22827(1426778646)#traffic-generator (INFO): 100 SDUs in 152293 us => 7.7220
Mb/s
22827(1426778646)#traffic-generator (INFO): 100 SDUs in 141937 us => 8.2854
Mb/s
22827(1426778646)#traffic-generator (INFO): 100 SDUs in 148867 us => 7.8997
Mb/s
22827(1426778646)#traffic-generator (INFO): 100 SDUs in 141265 us => 8.3248
Mb/s
22827(1426778647)#traffic-generator (INFO): 100 SDUs in 144774 us => 8.1230
Mb/s
22827(1426778647)#traffic-generator (INFO): 100 SDUs in 143477 us => 8.1964
Mb/s
22827(1426778647)#traffic-generator (INFO): 100 SDUs in 142637 us => 8.2447
Mb/s
22827(1426778647)#traffic-generator (INFO): 100 SDUs in 140033 us => 8.3980
Mb/s
22827(1426778647)#traffic-generator (INFO): 100 SDUs in 153757 us => 7.6484
Mb/s
22827(1426778647)#traffic-generator (INFO): 100 SDUs in 152454 us => 7.7138
Mb/s
22827(1426778647)#traffic-generator (INFO): 100 SDUs in 149347 us => 7.8743
Mb/s
22827(1426778648)#traffic-generator (INFO): 100 SDUs in 144610 us => 8.1322
Mb/s
22827(1426778648)#traffic-generator (INFO): 100 SDUs in 148841 us => 7.9010
Mb/s
22827(1426778648)#traffic-generator (INFO): 100 SDUs in 145378 us => 8.0893
Mb/s
22827(1426778648)#traffic-generator (INFO): 100 SDUs in 147706 us => 7.9618
Mb/s
22827(1426778648)#traffic-generator (INFO): 100 SDUs in 150931 us => 7.7916
Mb/s
22827(1426778648)#traffic-generator (INFO): 100 SDUs in 150635 us => 7.8070
Mb/s
22827(1426778648)#traffic-generator (INFO): 100 SDUs in 138290 us => 8.5039
Mb/s
```

```
22827(1426778649)#traffic-generator (INFO): 100 SDUs in 149754 us => 7.8529
Mb/s
22827(1426778649)#traffic-generator (INFO): 100 SDUs in 152846 us => 7.6940
Mb/s
22827(1426778649)#traffic-generator (INFO): 100 SDUs in 142102 us => 8.2757
Mb/s
22827(1426778649)#traffic-generator (INFO): 100 SDUs in 153010 us => 7.6858
Mb/s
22827(1426778649)#traffic-generator (INFO): 100 SDUs in 154317 us => 7.6207
Mb/s
22827(1426778649)#traffic-generator (INFO): 100 SDUs in 150802 us => 7.7983
Mb/s
22827(1426778649)#traffic-generator (INFO): 100 SDUs in 148286 us => 7.9306
Mb/s
22827(1426778650)#traffic-generator (INFO): 100 SDUs in 151963 us => 7.7387
Mb/s
22827(1426778650)#traffic-generator (INFO): 100 SDUs in 151965 us => 7.7386
Mb/s
22827(1426778650)#traffic-generator (INFO): 100 SDUs in 142996 us => 8.2240
Mb/s
#RESULTS FOR CONSTANT BIT RATE
22827(1426778658)#traffic-generator (INFO): New flow allocated [port-id =
4]
22827(1426778658)#traffic-generator (INFO): Starting test:
        duration: 5
        count: 0
        sduSize: 1470
22827(1426778658)#traffic-generator (INFO): 100 SDUs in 143991 us => 8.1672
Mb/s
22827(1426778658)#traffic-generator (INFO): 100 SDUs in 146947 us => 8.0029
Mb/s
22827(1426778658)#traffic-generator (INFO): 100 SDUs in 146989 us => 8.0006
Mb/s
22827(1426778658)#traffic-generator (INFO): 100 SDUs in 146975 us => 8.0014
Mb/s
22827(1426778658)#traffic-generator (INFO): 100 SDUs in 146985 us => 8.0008
Mb/s
22827(1426778658)#traffic-generator (INFO): 100 SDUs in 147013 us => 7.9993
Mb/s
22827(1426778659)#traffic-generator (INFO): 100 SDUs in 146982 us => 8.0010
Mb/s
22827(1426778659)#traffic-generator (INFO): 100 SDUs in 146980 us => 8.0011
Mb/s
22827(1426778659)#traffic-generator (INFO): 100 SDUs in 146946 us => 8.0029
Mb/s
22827(1426778659)#traffic-generator (INFO): 100 SDUs in 147008 us => 7.9996
Mb/s
22827(1426778659)#traffic-generator (INFO): 100 SDUs in 146934 us => 8.0036
Mb/s
22827(1426778659)#traffic-generator (INFO): 100 SDUs in 147003 us => 7.9998
Mb/s
22827(1426778659)#traffic-generator (INFO): 100 SDUs in 146967 us => 8.0018
Mb/s
22827(1426778660)#traffic-generator (INFO): 100 SDUs in 146964 us => 8.0020
Mb/s
22827(1426778660)#traffic-generator (INFO): 100 SDUs in 146971 us => 8.0016
Mb/s
22827(1426778660)#traffic-generator (INFO): 100 SDUs in 146977 us => 8.0013
Mb/s
22827(1426778660)#traffic-generator (INFO): 100 SDUs in 146952 us => 8.0026
Mb/s
22827(1426778660)#traffic-generator (INFO): 100 SDUs in 146973 us => 8.0015
Mb/s
22827(1426778660)#traffic-generator (INFO): 100 SDUs in 146945 us => 8.0030
Mb/s
22827(1426778660)#traffic-generator (INFO): 100 SDUs in 146959 us => 8.0022
Mb/s
22827(1426778661)#traffic-generator (INFO): 100 SDUs in 146992 us => 8.0004
Mb/s
22827(1426778661)#traffic-generator (INFO): 100 SDUs in 146950 us => 8.0027
Mb/s
```

```
22827(1426778661)#traffic-generator (INFO): 100 SDUs in 146973 us => 8.0015
Mb/s
22827(1426778661)#traffic-generator (INFO): 100 SDUs in 146962 us => 8.0021
Mb/s
22827(1426778661)#traffic-generator (INFO): 100 SDUs in 147002 us => 7.9999
Mb/s
22827(1426778661)#traffic-generator (INFO): 100 SDUs in 146957 us => 8.0023
Mb/s
22827(1426778661)#traffic-generator (INFO): 100 SDUs in 146977 us => 8.0013
Mb/s
22827(1426778662)#traffic-generator (INFO): 100 SDUs in 146961 us => 8.0021
Mb/s
22827(1426778662)#traffic-generator (INFO): 100 SDUs in 146969 us => 8.0017
Mb/s
22827(1426778662)#traffic-generator (INFO): 100 SDUs in 146971 us => 8.0016
Mb/s
22827(1426778662)#traffic-generator (INFO): 100 SDUs in 146971 us => 8.0016
Mb/s
22827(1426778662)#traffic-generator (INFO): 100 SDUs in 146937 us => 8.0034
Mb/s
22827(1426778662)#traffic-generator (INFO): 100 SDUs in 146964 us => 8.0020
Mb/s
22827(1426778663)#traffic-generator (INFO): 100 SDUs in 146970 us => 8.0016
Mb/s
```

Table 7: Server-side statistics running rina-tgen

The output from the server illustrates this point further: with a mean for the poisson distribution of 0.1, measured bandwidth rates go between 5.3 and 19.8 Mb/s, while with a mean value of 10, the values range between 7.6 and 8.5 Mb/s (note that the intervals change, the rates are measured per 100 packets).

We also analysed the traffic using `tcpdump`[105] and `Wireshark`[106]. The traffic trace taken during the experiment is shown in

Figure 44. The ordinate shows the time (seconds), the abscis the received badwidth during each 0.1 second interval (in bytes). The total traffic average (as measured using CBR) is slightly above 100 kiloByte every 100 ms, or slightly above the rate of 8 Mb/s. This is because Wireshark displays the total throughput (including the header overhead); not only the goodput (useful payload).

This graph clearly shows the effect of the mean value for the poisson distribution, and the stability of the CBR generator.

This test shows that RINA in general and the IRATI prototype implementation specifically can support traffic patterns associated with video services as a very early proof-of-concept.

[105] http://www.tcpdump.org/

[106] https://www.wireshark.org/

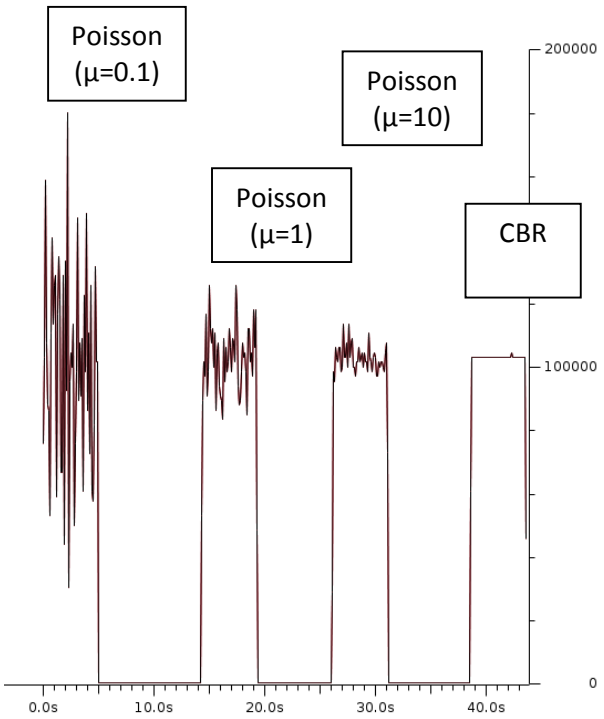


Figure 44: Wireshark I/O Graph of rina-tgen; measured at the server node

8.3 Inter-VM communication performance

Virtual machine networking is commonly implemented by providing VMs with (virtual) Network Interface Cards (NICs) emulated in the Hypervisor. The emulated NIC forwards VM packets to/from the Hypervisor's TCP/IP stack. The Hypervisor usually connects to the emulated NIC through a special (software) network interface. In order to connect the emulated NIC with other VMs hosted by the same Hypervisor or with the external network, the Hypervisor's software interfaces are bridged to other host interfaces (physical interfaces and/or software interfaces associated to other emulated NICs) using software switches - e.g. OpenVSwitch or the standard Linux in-kernel bridge. Each Hypervisor may host many bridges in order to build arbitrary network topologies for VMs.
In RINA, networking is IPC between application processes. As a consequence, there is no need to emulate a NIC to connect the VM stack to Hypervisor stack. It is enough to design an ad-hoc shim DIF that provides VM-to-Hypervisor point-to-point connectivity, directly using shared memory or message passing mechanisms provided by the Hypervisor itself. This is possible because the DIF abstraction is at a higher level than the NIC abstraction. Therefore, while a physical machine will typically have one or more shim DIF over Ethernet or WiFi as lowest level network access, a VM will have one or more shim DIF for Hypervisors.
Exploring the possibilities of using hypervisor internal mechanisms other than the traditional networking subsystem for VM-to-Hypervisor communication allows for better performance and provides an easier manageable solution. Unlike traditional VM networking, the shim DIF for Hypervisor is not restricted by the limitations of the Ethernet technology.
Among the advantages of shim DIF for Hypervisors over traditional NIC emulation:

- No need to implement complex and expensive NIC emulation.

- No need to generate and assign MAC addresses, which can become an issue at scale, especially for large DCs.
- No need to create and configure software L2 bridges to connect VMs and Hypervisor physical NICs together.
- Users of the shim DIF are not restricted to the Ethernet MTU (maximum payload of 1500 bytes or 9000 bytes if Jumbo frames are used). Actually, this restriction is commonly bypassed in traditional VM networking by using the TCP Segmentation Offloading (TSO) features offered by emulated NICs. However, this is basically a workaround that adds complexity and it is not needed by the shim DIF over Hypervisor.
- No need to perform TCP/UDP checksumming in the emulated NIC (checksum offloading), since shared memory communication is protected from corruption by other means (Error Correcting Code memory). Checksumming is not actually performed by modern paravirtualized NIC (e.g. virtio-net, xen-netfront), but again this is a workaround that is not needed by the shim DIF over Hypervisors.

The IRATI shim DIF for Hypervisors is built on top of the Virtual Message Passing Interface (VMPI) VM-to-Hypervisor shared-memory communication mechanism. A VMPI device is used to implement the point-to-point link and it is seen as a special device on both VM and Hypervisor. Each VMPI device is assigned two identifiers, one on the VM OS and the other on the Hypervisor OS. The first identifier is necessary to distinguish multiple VMPI devices in the same VM, while the second one is required to distinguish between the multiple VMPI devices (assigned to possibly different VMs) on the same Hypervisor. Nevertheless, the scope of those identifiers is confined to a single OS, so that the management is far easier than MAC management. The scope of MACs needs to be unique on the L2 domain in which the NICs exist - that may be a large segment of the DC infrastructure, involving multiple hypervisors.
In order to assess the performance gain that can result from deploying the shim DIF for Hypervisors, some comparative tests have been executed to measure VM-to-Hypervisor throughput performance. Three sessions of tests have been carried out, as explained below. QEMU/KVM has been chosen for the testbed, since it is supported by the shim DIF for Hypervisors provided by the IRATI prototype. The testbed involves a single QEMU VM (the guest), running on a physical machine (the host).
The first two tests sessions assess UDP throughput performance at variable packet size, therefore testing the performance of traditional VM networking. The guest is endowed with an emulated NIC, whose corresponding tap device is bridged to the host stack through a Linux in-kernel software bridge. The Linux bridge is accessible in the host stack through a bridge interface (e.g. *br0*). Once the bridge interface and the guest NIC have been given an IP address on the same IP subnet – they are on the same L2 domain – the netperf benchmarking tool is used to measure UDP performance between the host and the guest. In particular, the netperf server (netserver) listens on the bridge interface, while the netperf client runs in the guest. The only difference between the first and the second test session is the model of the emulated NIC. In the first session, an Intel e1000 NIC is used, which is implemented by QEMU by emulating the hardware behaviour – e.g. NIC PCI registers, DMA, packet rings, offloadings, etc. The second test session makes use of the a *paravirtualized* NIC model, the virtio-net device. Paravirtualized devices don't correspond to real hardware, instead they are explicitly designed to be used by virtual machines, in order to save the Hypervisor from the burden of emulating real hardware. Paravirtualized devices allows for better performances and code reusability – the virtio standard also provides paravirtualized disk, serial console, number generator, etc. Despite being more virtualization-friendly that e1000 (or other emulated NICs like r8169 or pcnet2000), the guest OS still sees the virtio-net adapter like a normal ethernet interfaces, with all the complexities involved – MAC, MTU, TSO, checksum offloading, etc.

The third test session shows instead the performance of the shim DIF for Hypervisors. The test scenario involves a shim IPC process on the host and the corresponding one on the guest. The two shim IPC processes forms an instance of shm DIF for Hypervisors. No normal IPC processes are used in this scenario, the applications can run directly over the shim DIF. This is a consequence of the flexibility of the RINA architecture, since the application can use the lowest level DIF whose scope is sufficient for the communication (guest-to-host in this case) and that provides the required QoS. In these tests the host runs the rina-echo-time application in server mode, while the guest runs the rina-echo-time application in client mode. Each test run consists in the client sending an unidirectional stream of PDUs of fixed size. Measurement have been taken varying the PDU size.
As shown in figure 44, the shim DIF for Hypervisors outperforms both e1000 and virtio-net NIC cases, validating the analysis done in this section – that is simpler and cleaner architecture allows for better performance. Note that while Ethernet MTU is set to 1500 in the first two sessions, it is possible to go beyond the limit because of the UDP Segmentation Offloading (UFO) feature – supported by both e1000 and virtio-net models. This features allows the NIC (real hardware or emulated) to accept UDP packets that does not fit into a single 1500 bytes Ethernet frame, where the NIC arranges for doing the necessary segmentation in hardware. It's interesting to note that in the virtio-net case, this segmentation is not really carried out, since there is no real Ethernet cable to deal with, but the oversized packet is directly forwarded to the host stack, which is able to process and deliver to the receiving application (netserver) without further segmentations. This is clearly an optimization, but can also be seen as a workaround that is not necessary when RINA comes into the picture.

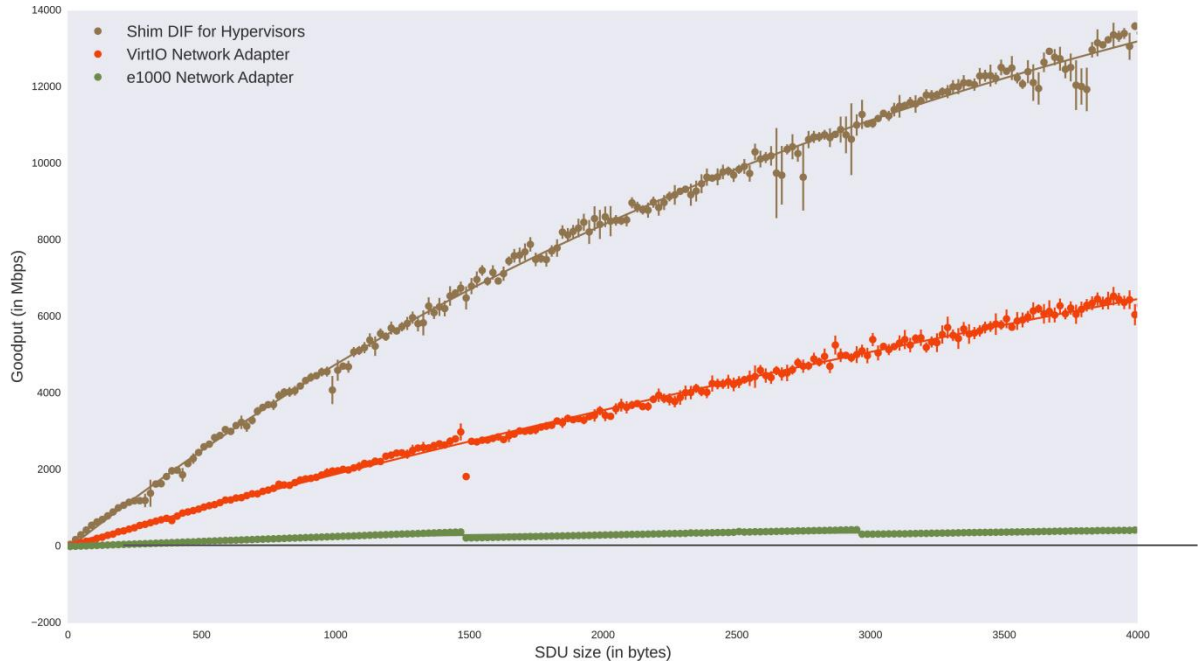


Figure 45: Inter-VM communication performance (single host)

We evaluate the communication performance by measuring the round-trip time between a VM and its host for various SDU sizes.

Figure 46 shows the average RTT measured over 100 "pings" for SDU sizes up to 4070 bytes (the current limitation of the Shim for HV). The linear regression line shows that adding a normal DIF

(stacking another layer) incurs a small penalty (roughly 20-30) µs each way that is dependent on SDU size (as may be expected by the processing of larger packets).

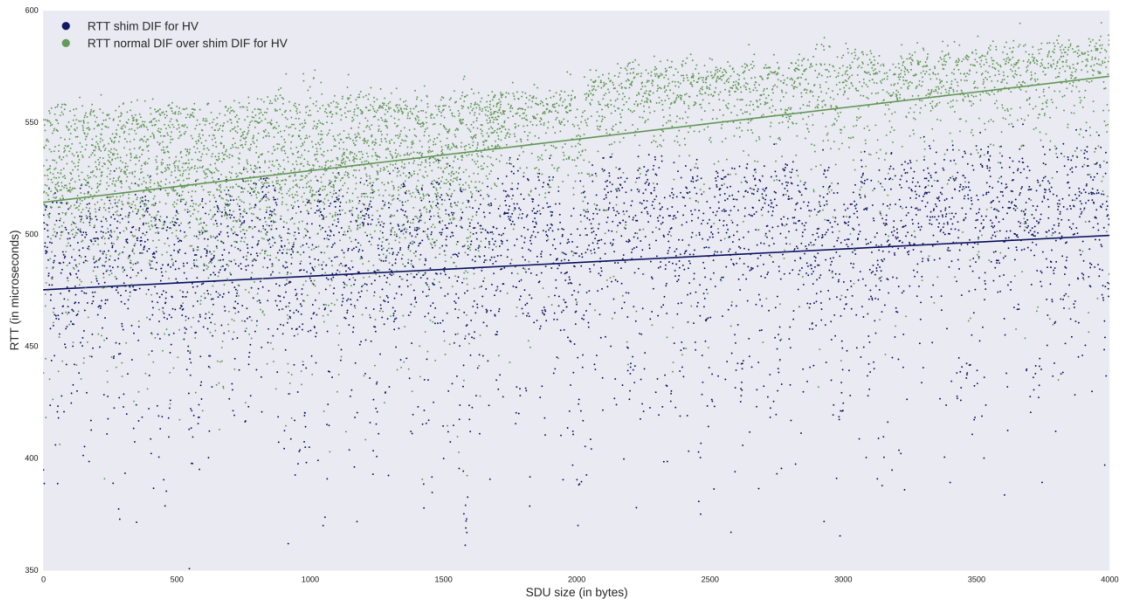


Figure 46: RTT between VM and Host for SDU sizes between 0 and 4070 bytes

9 Conclusions

There has been much debate regarding the design and implementation of next generation network architectures. Two very different approaches have been proposed in the literature, namely a clean slate approach and an evolutionary approach. Rexford and Dovrolis[107] highlight that evolutionary research attempts to understand the behaviour of the current Internet and its inherent problems and try to resolve them according to two major constraints, firstly backward compatibility with existing technologies and secondly incremental deployment of new technologies. On the other hand clean-slate research aims to design a completely new "Future Internet" architecture that is a major improvement on the current Internet without being constrained by existing Internet technologies. Day[108] contends that the idea of the Internet "evolving" to accommodate solutions to its inherent problems is not a sufficient solution and that no amount of patching is going to solve its fundamental flaws rather a clean slate approach to future Internet design should be undertaken.
The IRINA project set out to reach four clear Objectives, which were all achieved:

1. Perform a comparative study of RINA against the current state of the art (Section 2)
2. Build a use case how RINA would be used in an NREN scenario (Sections 3-6)
3. Improve the current prototype and showcase a lab trial proof-of-concept (Sections 7-8)
4. Disseminate the project approach and organize a workshop

IRINA has investigated clean-slate network architectures taking a holistic approach, not only focusing on particular features or point solutions. The goal was to identify not only different architectures that perform better than the current one for certain situations and certain features; but to identify an architecture that tries to improve on the current Internet as a whole. The procedures and cost for deployment and adoption of an architecture are the most important factors that have been taken into account by IRINA in the SWOT analysis.
Building the use case brought us in closer contact with the NREN community, where we received generous support and enthusiastic responses to our survey; greatly strengthening our reference scenario.
As a result of the project, the RINA community now has a more potent tool available for experimentation in the `rina-tgen`, which will live on and be further extended as part of the IRATI software suite. The lab tests showing that the shim DIF for Hypervisors outperforms both e1000 and virtio-net NIC cases show that a simpler and cleaner architecture does allow for better performance for virtual networking solutions when compared to TCP/IP.
IRINA participated in various national and international events. The project kickoff for the project took place at a special event during the GN3plus symposium in Vienna, October 8-10, 2013. IRINA then participated in the GN3plus/TERENA JRA1 Architecture workshop, organised by NORDUnet in Kastrup on November 20th-22nd 2013, and presented the project concept and objectives. The IRINA project was presented to the RINA community at the second RINA workshop, held in Dublin 28th and 29th of January 2014. A paper, entitled "RINA: An Opportunity for NRENs to Lead Internet Research" provided an opportunity to present IRINA to the wider TERENA community at TNC2014 Dublin on May 22nd

[107] Rexford, Jennifer, and Constantine Dovrolis. "Future Internet architecture: clean-slate versus evolutionary research." *Communications of the ACM* 53.9 (2010): 36-40.

[108] Day, John. "How in the Heck do you lose a layer!?." *Network of the Future (NOF), 2011 International Conference on the* 28 Nov. 2011: 135-143.

2014. At the second GN3plus/TERENA JRA1 Architecture workshop in Kastrup, November 11th -13th 2014, the project presented and demonstrated the IRATI prototype on which it is building, while further engaging in discussions with the JRA1 participants. IRINA was presented together with the other ongoing RINA projects at IETF91 November 9-14 in Honolulu, HI, during the sdnrg session. On December 8th 2014, IRINA participated in a half-day tutorial on RINA organised by IRATI and held at the IEEE GLOBECOM conference in Austin, TX. IRINA was the main organiser together with the PRISTINE project, of the third RINA workshop, held in Ghent on January 28th 2015. IRINA organised another RINA tutorial at DRCN 2015 on March 24th 2015 in Kansas City, MO. Finally, IRINA results will be submitted to IEEE Communications Magazine as part of a joint IRATI-IRINA publication, and will be presented during a session on RINA at the TERENA Networking Conference, June 15th-18th in Porto.
Through the Open Calls, the GÉANT association and GN3+ project gave us a valuable opportunity to work with the NREN community and showcase RINA technology to a potential stakeholder. The IRATI software is under continuing development, mainly through the FP7 ICT PRISTINE project. IRINA took a small step in providing a lightweight proof-of-concept of RINA in an NREN environment. The next step is developing proof-of-concept applications (or porting existing applications) that can be deployed as a showcase in a Future Internet test bed provided by FIRE or GÉANT.

Glossary

ANA	Autonomic Network Architecture
API	Application Programming Interface
AS	Autonomous System
BFD	Bidirectional Forwarding Detection
BGP	Border Gateway Protocol
CBDF	Cross Border Dark Fibre
CDN	Content Distribution Network
CE	Control Element
CloNe	Cloud Networking
CPL	Common Photonic Layer
DAF	Distributed Application Facility
DC	Data Centre
DCN	Data Centre Network
DDoS	Distributed Denial of Service
DIF	Distributed IPC Facility
DNS	Domain Name System
DNSSEC	Domain Name System Security Protocol
DMS	DIF Management System
DoS	Denial of Service
DWDM	Dense Wavelength Division Multiplexing
EID	Endpoint Identifier
EoMPLS	Ethernet over MPLS
FD	Forwarding Directive
FIA	Future Internet Assembly
FE	Forwarding Element
FEM	Forwarding Element Manager
FIRE	Future Internet Research and Experimentation
IETF	Internet Engineering Task Force
IPC	Inter-Process Communication
IPCP	IPC Process
ISP	Internet Service Provider
IXP	Internet eXchange Point
LFB	Logical Function Block
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
ICN	Information Centric Networking

IP	Internet Protocol
ISP	Internet Service Provider
LISP	Locator/ID Separation Protocol
MCU	Multi Conference Unit
MPLS	Multi Protocol Label Switching
MSP	Multi Service Port
NaaS	Network as a Service
NAT	Network Address Translation
NF	Network Function
Ncore	NEBULA Core
NDM	Named Data Network
NDP	NEBULA Data Plane
NetInf	Network of Information
NfaaS	Network Functions-as-a-Service
NFV	Networks Function Virtualization
NMS	Network Management System
NREN	National Research and Educational Network
NVENT	NEBULA Virtual and Extensible Networking Techniques
NREN	National Research and Educational Network
OconS	Open Connectivity Service
OPN	Optical Private Network
OSPF	Open Shortest Path First
PEST	Political, Economic, Social & Technological
PoA	Point of Attachment
PoP	Point of Presence
PSOC	The Pouzin Society
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RINA	Recursive InterNetwork Architecture
RLOC	Routing Locator
SAIL	Scalable and Adaptive Internet Solutions
SCTP	Stream Control Transmission Protocol
SDH	Synchronous Digital Hierarchy
SDK	Software Development Kit
SDN	Software Defined Network
SDU	Service Data Unit
SLA	Service Level Agreement
SOTA	State Of The Art
SSP	Single Service Port
SWOT	Strengths, Weaknesses, Opportunities, Threats
ToR	Top of Rack
VDC	Virtual DC
VLAN	Virtual Local Area Network
VNF	Virtualized Network Function
VoIP	Voice over IP
VPN	Virtual Private Network
VRF	Virtual Routing and Forwarding
VRRP	Virtual Router Redundancy Protocol

WDM	Wavelength Division Multiplexing
XIA	eXpressive Internet Architecture