27-01-2012

Deliverable DJ2.5.1: Network Factory Footprint

Deliverable DJ2.5.1

Contractual Date:	30-11-2011
Actual Date:	27-01-2012
Grant Agreement No.:	238875
Activity:	JRA2
Task Item:	Task 5
Nature of Deliverable:	R (Report)
Dissemination Level:	PU (Public)
Lead Partner:	SWITCH
Document Code:	GN3-11-203
Authors:	V. Ajanovski (MARNET), K. Baumann (SWITCH), A. Chuang (i2CAT/RedIRIS), F. Farina (GARR), P. Gasner (RoEduNet), B. Jakimovski (MARNET), J. Jofre, (i2CAT/RedIRIS), R. Krzywania (PSNC), A. Sevasti (GRNET), V. Smotlacha (CESNET), M. A. Sotos (RedIRIS), S. Ubik (CESNET), A. Weinert (PSNC)

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7 2007–2013) under Grant Agreement No. 238875 (GÉANT).

Abstract

This deliverable documents the proposed scope of the GÉANT Network Factory, the technologies currently available to realise it, and business aspects of indicative implementation scenarios as a basis for further study, in anticipation of the outcomes of the ongoing procurement activities for the next-generation GÉANT.

Table of Contents

Executive Summary 1

1 Introduction 4
1.1 A GÉANT Network Factory 4
1.2 About this Document 6
1.2.1 Terminology 7

2 Network Factories: Existing Experience 8
2.1 GÉANT 8
2.1.1 GÉANT Capabilities and Hardware 9
2.2 FEDERICA 13
2.2.1 Objectives 13
2.2.2 Infrastructure Overview 14
2.2.3 Features 15
2.2.4 Issues and Limitations 18
2.2.5 FEDERICA Business Model, Implementation Costs and Maintenance 19
2.3 Other Initiatives 24
2.3.1 Evaluation 25
2.4 Conclusions 38

3 Network Factory-Enabling Technologies for GÉANT 39
3.1 Network Factory-Enabling Technologies at Layer 1 39
3.1.1 Current GÉANT Layer 1 Technology 39
3.1.2 GÉANT Slicing Options at Layer 1 40
3.2 Network Factory-Enabling Technologies at Layer 2 44
3.2.1 Current GÉANT Layer 2 Technology 44
3.2.2 GÉANT Slicing Options at Layer 2 45
3.2.3 Emerging Technologies 46
3.2.4 Summary 51
3.3 Network Factory-Enabling Technologies at Layer 3 53
3.3.1 Current GÉANT Layer 3 Technology 53
3.3.2 GÉANT Slicing Options at Layer 3 54
3.3.3 Emerging Technologies 55
3.3.4 Summary 57
3.4 Monitoring Functionality 58

4 Foundations of a Network Factory Business Case 59
4.2 Strategic Fit 60
4.3 Selected Network Factory Scenarios 60
4.3.1 OpenFlow-Based Network Factory 61
4.3.2 Network Factory Delivering Layer 3 Slices 63
4.4 Risk Assessment and Analysis 64

5 Conclusions 72

Appendix A FEDERICA 73
A.1 Physical Resources 73
A.2 User Support and Portal 75
A.3 User Access and Slice Management 77
A.4 Slice Creation and Management Process in Detail 78
A.5 Details of FEDERICA PoP setup 79

References 80

Glossary 83

Table of Figures

Figure 2.1: Virtual network creation over GÉANT 9
Figure 2.2: Graphical representation of FEDERICA [FED_DJRA2.1] 14
Figure 3.1: GÉANT and the NRENs 10 G fibre map – at the end of 2009 40
Figure 3.2: Bus topology with ROADM 42
Figure 3.3: Reconfigurable Optical Add-Drop Multiplexer (4 lambdas) 43
Figure 3.4: EoMPLS schematic 46
Figure 3.5: Example of an OpenFlow-based network 49
Figure 3.6: Flow tables virtualisation through FlowVisor 49
Figure 3.7: Example integration of OpenFlow switches in Network Factory PoPs 51
Figure 3.8: GRE tunnelling for the Layer 3 Network Factory 54
Figure 4.1: OpenFlow-enabled Network Factory scenario architecture 61
Figure 4.2: Layer 3 services Network Factory scenario architecture 63
Figure A.1: FEDERICA logical topology 73
Figure A.2: The User Portal architecture 76

Figure A.3: Service LAN and SSH access for a user's slice 77
Figure A.4: FEDERICA virtual circuit creation with VLANs 79

Table of Tables

Table 0.1: Summary of slicing options and technologies at Layer 1, Layer 2 and Layer 3 2
Table 1.1: Terminology 7
Table 2.1: Network Factory-related aspects of GÉANT network services 11
Table 2.2: FEDERICA customer types, incentives and benefits 20
Table 2.3: Capital costs 23
Table 2.4: Operational costs 24
Table 2.5: Evaluation table 1: technical philosophy 25
Table 2.6: Evaluation table 2: available hardware 29
Table 2.7: Evaluation table 3: physical substrate sharing 29
Table 2.8: Evaluation table 4: virtualisation 30
Table 2.9: Evaluation table 5: slicing 31
Table 2.10: Evaluation table 6: project focus 32
Table 2.11: Evaluation table 7: slicing and virtualisation network layers 33
Table 2.12: Evaluation table 8: application usage 35
Table 2.13: Evaluation table 9: GÉANT applicability 37
Table 3.1: Summary of Layer 1 slicing options 44
Table 3.2: Summary of Layer 2 slicing technologies 53
Table 3.3: Summary of Layer 3 link slicing technologies 58
Table 4.1 Hardware-based OpenFlow v1.0 specification support 62
Table 4.2 Overall Network Factory business case risk analysis 67
Table 4.3 OpenFlow-based Network Factory business case risk analysis 69
Table 4.4 Network Factory delivering Layer 3 slices business case risk analysis 71

Executive Summary

This deliverable documents the proposed scope of the GÉANT Network Factory and describes the current and emerging technologies that can be used to realise it, as well as business aspects of indicative implementation scenarios as a basis for further study.

A Network Factory is a facility and its associated services for delivering to researchers logical and physical networks on top of the GÉANT production environment. Delivering such facilities and/or services has been a common direction for research and education infrastructures globally, to address user requirements for shared, partitioned physical network and IT infrastructures that provide secure and isolated application-specific infrastructures. The aim of the GÉANT Network Factory is to provide a framework for delivering dedicated segments or "slices" of physical resources to researchers whose projects or applications require specific network parameters that do not fit within the standard GÉANT services. The underlying concept is to provide the Network Factory as a Service (NFaaS), ideally giving researchers complete control of the resources in their slice and allowing them to experiment within their slices on top of the GÉANT backbone.

While focusing on the challenges presented by a multi-domain environment, the initial implementation of the GÉANT Network Factory will use mainly resources from the GÉANT backbone, allowing end users to utilise the single-domain GÉANT Network Factory facilities through remote connections from their local environment. It is envisaged that in later stages, the GÉANT Network Factory will be enriched by resources from participating National Research and Education Networks (NRENs). Detailed requirements for NRENs to participate in a multi-domain Network Factory are a subject of future work as an incremental step, following that of the Network Factory infrastructure and service definition/deployment over the GÉANT backbone.

To define requirements, best practice and a model for the GÉANT NFaaS, which may be subject to revision once the results of the ongoing procurement process for the next-generation GÉANT backbone is complete, JRA2 T5 has considered a range of projects and initiatives in which infrastructures equivalent to a Network Factory have been implemented. These include the current GÉANT network, whose hybrid infrastructure is already capable of creating slices using existing connectivity services at Layer 1 (GÉANT Lambda), Layer 2 (GÉANT Plus) and Layer 3 (GÉANT IP), depending on the user requirements and the GÉANT Points of Presence involved. Other projects evaluated were ANI, VINI, OFELIA, AKARI, GENI, PASITO and, in particular, FEDERICA. The FEDERICA case is examined in greater detail as it was an initiative driven by the GÉANT-NREN community for which more detailed information exists and of which most business case aspects are expected to be applicable to the GÉANT Network Factory as well. The overall analysis has shown the desirability of supporting research at 100 Gbps, and has indicated that it is common practice to deploy testbeds for network research that allow overlay networks to be established over a production environment, often using OpenFlow technology.

In reviewing the technologies applicable to GÉANT for realising a Network Factory infrastructure and services at Layer 1, Layer 2 and Layer 3, JRA2 T5 has considered not only the Network Factory-enabling technologies currently deployed in GÉANT, but also technologies that are or soon will be commercially available. The slicing options and technologies at each layer that are relevant to GÉANT[1], are summarised in Table 0.1 below.

Layer 1 Slicing Options	Layer 2 Slicing Options	Layer 3 Slicing Options
Static Lambda without OEO	802.1q VLAN, IEEE 802.1ad	IP and GRE tunnels
ROADMs without OEO	EoMPLS	Differentiated IP services
ROADMs and OEO	EoSDH	Implementation of Layer 3 links with GÉANT Layer 1 and 2 services
	MPLS-TP	Hardware-based logical routers
	PBB/PBT	Servers hosting software-based routers
	OpenFlow	

Table 0.1: Summary of slicing options and technologies at Layer 1, Layer 2 and Layer 3

The preliminary business case for a GÉANT Network Factory indicates that there is a good strategic fit from the perspective of network capabilities, user requirements, current trends among R&E infrastructures, and an enriched GÉANT services portfolio. General critical success factors include compatibility with the installed infrastructure and services, quality of service offerings, and take-up. Due to an evolving technology matrix, the document sets the foundations for a business case using two indicative scenarios as different implementation options rather than an exhaustive list: an OpenFlow-based Network Factory and a Network Factory delivering Layer 3 slices. Each has been assessed from the point of view of technology, financial factors, and risk; general Network Factory risks have also been considered. The results of these assessments should not be regarded as the final outcome; rather, they are a basis for moving on to the production of specialised business case outputs and to undertaking the study and design phases of the Network Factory solution for GÉANT, incorporating the outcomes of the GÉANT backbone evolution process.

While both business and technical aspects of implementing a GÉANT Network Factory as presented in this document need to be further analysed and assessed, for a short-term solution it has been decided to pursue an OpenFlow-based Network Factory, deployed on top of the current GÉANT backbone in a way that ensures its viability over the future GÉANT backbone. The decision takes into account the potential presented globally by OpenFlow testbeds, the low cost implications and the minimum requirements imposed on the GÉANT production environment. A detailed business case and technical implementation planning for the short-term solution are already underway. At the same time, JRA2 Task 5 will elaborate a business case, technical specification and implementation of a long-term Network Factory solution, in parallel to the next-generation GÉANT procurement and deployment works. The planned long-term business case analysis is expected to provide more insight into all aspects of a full-scale Network Factory design and deployment over GÉANT, upon which decisions driving the design and implementation choices will have to be made. In the meantime, the

[1] The deliverable does not attempt to cover all possible technologies for slicing at different layers, only those relevant to GÉANT.

OpenFlow-enabled Network Factory will serve as a starting point for validating NFaaS concepts and making preliminary offerings available to the user community.

1 Introduction

One of the objectives of the GN3 project is to provide a next-generation pan-European network and related services that meet the communications needs of research communities in all fields. Such needs include both a transport facility for production data and a network environment where experiments can be conducted.

To prevent the production traffic of commodity services from being disrupted by high-bandwidth applications and experiments, it makes sense to separate them. This also enables researchers to modify the behaviour of infrastructure elements, such as traffic routing, which could not be realised on the production infrastructure.

The GÉANT network could meet these requirements by including the technology and policy for a Network Factory, that is, a facility and its associated services for delivering to researchers logical and physical networks on top of the GÉANT production environment.

Delivering such facilities and/or services has been a common direction for Research and Education (R&E) infrastructures globally (more details are provided in "Network Factory Footprint: Third-Party Initiatives" [NFEval]):

- Internet2, Indiana University and Stanford University have committed to delivering a Software-Defined Network (SDN) infrastructure, a common facility for delivering virtual networks to researchers [NDDI].
- The Global Environment for Network Innovations (GENI) is an infrastructure designed to support experimental research in networks [GENI].
- OpenFlow in Europe Linking Infrastructure and Applications (OFELIA), within the EU's FP7 ICT programme, is working on an experimental facility for researchers in the form of a test network [OFELIA].
- ESnet's Advanced Networking Initiative (ANI) program delivers a testbed emulating a real network with capabilities to support a wide range of communications research [ANIBEDWEB].

1.1 A GÉANT Network Factory

The hybrid infrastructure from which the GÉANT network is built is capable of creating logical and physical networks that can be considered independent of the production infrastructure, but that share its physical elements. Such networks over GÉANT can therefore be considered as comprising a Network (or Infrastructure)

Factory. Apart from network resources (circuits and network equipment), these infrastructures can also contain other resources such as computing equipment[2].

Work by JRA1 Task 4 (Future Network, Current and Potential Uses of Virtualisation) (GN3 Deliverable DJ1.4.1: "Virtualisation Services and Framework Study" [GN3_DJ1.4.1]) has shown that the emergence of new applications (for example, scientific applications that require 10 G or even 100 G connectivity and applications with strict computing and network resource requirements) requires physical network and IT infrastructures to be shared and partitioned in order to provide secure and isolated application-specific infrastructures.

The objective of a GÉANT Network Factory is to provide a framework for delivering slices of physical resources to researchers, ideally giving them complete control of the resources in their slice and allowing them to experiment within their slices on top of the GÉANT backbone. The underlying concept is to provide the Network Factory as a Service (NFaaS).

The Network Factory will support applications and user communities that require specific network parameters that do not fit within the standard GÉANT services. These include, for instance, dynamic user infrastructures that require frequent reconfiguration, and testbeds with some of the production network properties (such as physical topology) but isolated from the production network traffic and operations. Advanced applications additionally require from a Network Factory service guaranteed high capacity and particular Quality of Service (QoS) parameters, such as low loss and jitter[3]. This is particularly important for non-traditional user groups, for example, from the arts, humanities and health science sectors. A detailed user requirements survey and analysis specific to the Network Factory service is proposed as a matter of high priority for the next phase.

User groups and communities with potential use cases for the GÉANT Network Factory have not been systematically addressed, as this is not within the scope of JRA2 Task 5 (Multi-Domain Network Service Research, Network Factory). They are expected to emerge out of the European research community on network technologies (Future Internet Research & Experimentation – FIRE) but also from other fields, ranging from arts and humanities to health science and seismology.

Network Factory users should meet the following criteria:

- Their specific demands cannot be satisfied by the production network and existing services.
- The infrastructure they require is multi-domain with respect to functional and/or geographical scope.
- Their proposed application is innovative or represents a new type of network use.
- Their proposed application and/or research could disrupt the production network and existing services.

JRA2 Task 5 will deliver a GÉANT Network Factory solution by:

- Providing a basic framework that can evolve into a permanent infrastructure for performing tests that is independent of the production environment.

[2] The virtualisation of computing elements is outside the scope of the GÉANT Network Factory and should be obtained by users through other computing-oriented projects.
[3] While guaranteed high capacity and QoS parameters are offered by existing GÉANT services, the Network Factory users have further requirements that mean those services cannot meet all their needs. Indeed this is one of the criteria for Network Factory users.

- Focusing on the challenges presented by a multi-domain environment, to understand how the Network Factory service can be implemented between different domains.

The implementation of the Network Factory infrastructure will be based on the partitioning of network resources. It will feature a set of "slices" as defined in *Terminology* on page 7. In this deliverable, slicing technologies applicable to the GÉANT Network Factory core will be grouped according to the networking layer in which they operate (Layer 1, Layer 2 and Layer 3) and will be analysed separately.

It is important to note that in a general case within the GÉANT context, slices may involve more than one administrative domain and thus the Network Factory may evolve to a multi-domain entity. However, this is not part of the GÉANT Network Factory work so far. The Network Factory will be created using mainly resources from the GÉANT backbone, allowing end users to utilise the single-domain GÉANT Network Factory facilities through remote connections from their local environment. It is envisaged that in later stages, the GÉANT Network Factory will be enriched by resources from participating National Research and Education Networks (NRENs). Detailed requirements for NRENs to participate in a multi-domain Network Factory are a subject of future work as an incremental step, following that of the Network Factory infrastructure and service definition/deployment over the GÉANT backbone.

It is important to note also that the GÉANT Network Factory work is not intended to provide prescriptive input to the current transmission and switching equipment procurement process. Rather, the outcomes of the Network Factory work are dependent on it and on the ongoing migration from the current GÉANT architecture to the next-generation backbone that it will accomplish.

1.2 About this Document

In order to document the proposed scope of the Network Factory and the technologies suitable for realising it, this deliverable:

- Reviews the current GÉANT network, the services it provides, its capabilities and hardware for supporting a Network Factory. See Chapter 2 *Network Factories: Existing Experience* on page 8.

- Evaluates existing research/experimentation facility projects and initiatives, particularly FEDERICA, in order to identify technical and business aspects that could be applied or replicated within the GÉANT Network Factory environment. See Chapter 2 *Network Factories: Existing Experience* on page 8.

- Discusses possible deployment options for the Network Factory on top of the GÉANT network, considering technologies that are currently available and emerging standards. See Chapter 3 *Network Factory-Enabling Technologies for* GÉANT on page 39.

- Provides the foundations of a business case for delivering a future NFaaS, including indicative options from a technical point of view. See Chapter 4 *Foundations of a Network Factory Business Case* on page 59.

The deliverable is not intended as a detailed description of or architecture for a Network Factory, since the work is still in the service strategy phase. It is a preliminary study that provides the foundation for a business case. Specifications will follow once the business case has been approved and the design phase begins.

1.2.1 Terminology

Table 1.1 below provides the definitions of key terms as used in this document.

Term	Definition
Network Factory	A facility and its associated services for delivering to researchers logical and physical networks on top of the GÉANT production environment.
Slice	In the context of network infrastructure, a dedicated partition or segment of the physical network infrastructure. A slice ideally exposes the same set of functionalities as the physical object whose behaviour it reproduces. However it is allocated with a subset of the physical object's resources. Slicing isolates a logical resource from the activities of the other logical instances and can be obtained by exclusive partitioning of the physical substrate (for example, reserving a number of ports in a router).
Substrate	The set of physical resources upon which a Network Factory is implemented, including backbone links, network elements and servers.
Virtualisation	In the context of network and computing infrastructure, virtualisation is the creation of a virtual version of a physical resource (e.g. network, router, switch, optical device or computing server), based on an abstract model of that resource with somewhat limited capabilities and often achieved by partitioning (slicing) and/or aggregation. A virtual infrastructure is a set of virtual resources interconnected together and managed by a single administrative entity. [GN3_DJ1.4.1]

Table 1.1: Terminology

2 Network Factories: Existing Experience

This chapter considers a range of projects and initiatives in which infrastructures comparable to a Network Factory, as defined in the *Introduction* on page 4, have been implemented, in order to identify best practices and aspects that could be applied or replicated in the GÉANT Network Factory. The findings serve as a basis for defining the requirements of a Network Factory (NF) over GÉANT, albeit subject to revision on the basis of the results of the ongoing procurement process for the next-generation GÉANT backbone implementation. Still they provide a list of aspects to consider in choosing a model for the Network Factory as a Service. The chapter is divided into three sections:

- A description of the GÉANT backbone network, its services and slicing capabilities.
- A detailed description of the FEDERICA project, which explains the slicing of network resources on a separate infrastructure deployed on top of the GÉANT substrate. This section also briefly considers business model approaches and cost estimates for FEDERICA.
- A summary of other Network Factory initiatives, in the form of a comparison table.

2.1 GÉANT

GÉANT is a pan-European backbone network that interconnects National Research and Education Networks (NRENs) across Europe and also provides worldwide connectivity through links with other regional networks. GÉANT offers connectivity services at Layer 1 (L1), Layer 2 (L2) and Layer 3 (L3) of the ISO/OSI reference model, i.e.:

- GÉANT Lambda

  A 10 Gbps (10 GE or OC-192/STM-64) point-to-point transparent wavelength service. A 40 Gbps wavelength service has been successfully tested and is also available; a 100 Gbps wavelength service is planned. GÉANT Lambda is a Layer 1 service. As this service is provided directly over the Dense Wavelength-Division Multiplexing (DWDM) transmission platform, it is only available to NRENs that are part of the GÉANT fibre cloud.

- GÉANT Plus

  The GÉANT Plus service offers point-to-point sub-lambda circuits of between 155 Mbps and 10 Gbps (typically at 1 Gbps) across an existing pre-provisioned network. (GÉANT Plus circuits are also known as light paths or Ethernet Private Lines (EPLs)). GÉANT Plus is a Layer 2 service.

- GÉANT IP

  The GÉANT IP service offers IP / Multi-Protocol Label Switching (MPLS) connectivity at speeds of up to 40 Gbps (100 Gbps is planned). While GÉANT IP is a Layer 3 service, it also offers Layer 2 Virtual Private Network (VPN) capabilities built on the common IP infrastructure yet delivered to the users as dedicated protected circuits.

The backbone network is built over leased dark fibre, lit with DWDM technology using a system owned and operated by DANTE, combined with managed services (wavelength and based on Synchronous Digital Hierarchy (SDH)) provided by commercial telecommunications operators. GÉANT services terminate in Points of Presence (PoPs), which are attached to NREN infrastructure, in order to ensure services are reachable by end users and campus networks.

Further information about the current GÉANT architecture and service portfolio is available in GN3 Deliverable DS1.1.1,2 “Final GÉANT Architecture” [GN3_DS1.1.1,2]. DS1.1.1,2 also presents architecture options and recommendations for the next-generation GÉANT network, taking into account current and future requirements and opportunities for improvement such as those afforded by technology developments.

2.1.1 GÉANT Capabilities and Hardware

2.1.1.1 *Services*

One of the areas of focus for GÉANT’s capabilities with regard to supporting a Network Factory implementation is the creation of virtual networks between various GÉANT PoPs by delivering a set of logical links that will be seen by end users as a dedicated infrastructure.

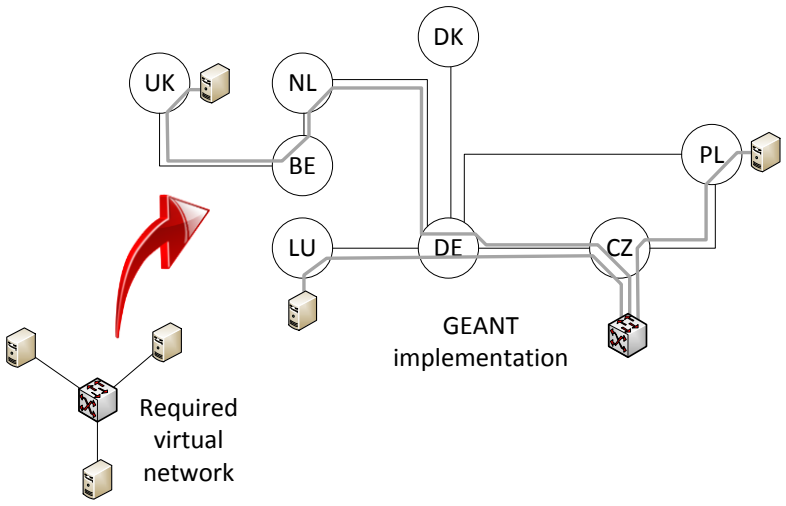

Figure 2.1: Virtual network creation over GÉANT

Figure 2.1 shows an example of virtual network implementation over the GÉANT infrastructure, with a simple star topology comprising of a central switch and three remote servers connected to it. Such a topology can be implemented over the GÉANT infrastructure by delivering logical links between PoPs. It is important to point out

that the current GÉANT infrastructure can deliver only network connectivity, i.e. logical links, between PoPs, while provisioning of servers and additional switching capabilities are left with the NRENs and end users. In theory, any of the GÉANT services mentioned in Section 2.1 can be used to implement such a virtual topology. In practice, however, not all services are available at all GÉANT PoPs at the time of writing this document, and each of the services has advantages and disadvantages. These are explained below based on the example of Figure 2.1.

The logical link from PL to CZ cannot be implemented with the GÉANT Lambda service because PL is not in the GÉANT fibre cloud. Both the IP/MPLS and GÉANT Plus services are, however, available. In comparison, a logical link between LU and CZ cannot be implemented via IP/MPLS or the GÉANT Lambda services, as only the GÉANT Plus service is available at LU. UK and CZ are fully featured PoPs and can be linked using any of the services, depending on user requirements. A full list of the technologies available at each PoP is given in Figure 2.2 in Section 2.1 of [GN3_DS1.1.1,2].

Each GÉANT service has different implications for the network and end users. These are compared in Table 2.1 and discussed in the sections that follow.

	GÉANT Lambda	GÉANT Plus	GÉANT IP
Capacity:	10 Gbps or 40 Gbps (no other scale).	155 Mbps up to 10 Gbps (at 155 Mbps increments).	Up to 40 Gbps, limited capacity control.
Resiliency/backup:	Optional.	No (technically possible, but due to lack of interest not supported).	Provided by standard IP/MPLS mechanisms.
Additional equipment required at NREN site:	Yes.	No.	No.
Network layers used (transparency[4]):	Layer 1.	Layer 2 (users can still use some Layer 2 features).	Layer 3.
Connectivity to non-GÉANT sites:	No.	Yes.	Yes.
Point-to-multipoint:	No.	No.	Yes.
Not available in:	Bulgaria (BG) Cyprus (CY)	Bulgaria (BG) Cyprus (CY)	None[5]

[4] Transparency here refers to network layers that are under GÉANT control that cannot be overridden by end users. The more transparent the service, the lower is the network layer available to end users to configure themselves.

[5] For the following routerless PoPs, IP/MPLS logical link functionality is limited: Belgium (BE), Croatia (HR), Cyprus (CY), Ireland (IE) , Israel (IL), Luxembourg (LU), Macedonia (MK), Malta (MT), Montenegro (ME), Portugal (PT), Russian Federation (RU), Serbia (RS), Slovakia (SK), Slovenia (SI), Turkey (TR).

	GÉANT Lambda	GÉANT Plus	GÉANT IP
	Estonia (EE) Greece (GR) Israel (IL) Latvia (LV) Lithuania (LT) Macedonia (MK) Malta (MT) Montenegro (ME) Poland (PL) Romania (RO) Serbia (RS) Turkey (TR)	Estonia (EE) Israel (IL) Latvia (LV) Lithuania (LT) Macedonia (MK) Malta (MT) Montenegro (ME) Romania (RO) Serbia (RS) Turkey (TR)	
Time to set up:	10 weeks (due to hardware procurement requirements).	5 working days.	Dependent on specific service implementation (e.g. Layer 2 Virtual Private Network (VPN)).

Table 2.1: Network Factory-related aspects of GÉANT network services

Capacity Control

All services offer some degree of capacity control. However, the GÉANT IP service is the most difficult to manage in this respect. Although it is over-provisioned by design, so as to allow small-to-medium-sized traffic flows over an uncongested path, it is nonetheless a “best effort” service, with no capacity or deterministic performance guarantees for pure IP connectivity. The GÉANT Lambda and Plus services are much better suited for controlling capacity attributes. However, the Lambda service does not scale: only 10 Gbps or 40 Gbps wavelengths are available. In contrast, the GÉANT Plus service scales from 155 Mbps up to 10 Gbps, providing flexibility with 155 Mbps increments. Given that 1 Gbps circuits are the most commonly used, the GÉANT Plus service would seem to be the best option for the example in Figure 2.1. However, at the time of writing, it is highly likely that the attributes of the GÉANT Plus service will be revised in the new GÉANT backbone.

Resilience and Lead Time

The level of resiliency offered by the GÉANT connectivity services ranges from none to full. The GÉANT Plus service does not support resiliency. Although it is technically possible to implement and manage backup circuits, the NRENs showed limited interest in this option and the service is therefore implemented without protection. The GÉANT IP service is resilient in the case of hardware failure or fibre cuts, providing backup protection against circuit failure at up to the full subscribed capacity on an appropriate interface and using advanced routing equipment to ensure fast recovery from unexpected events. For the GÉANT Lambda service, it is possible to request a backup path that travels through different physical network elements and is fully diverse to

the primary path. The Lambda service is, therefore, the most reliable. However, it requires the installation of additional optical equipment at the NREN site, which takes time (an estimated 10 weeks for the circuit to be implemented) and makes this service the most expensive option. Neither GÉANT Plus nor GÉANT IP requires any additional hardware, and the implementation of the circuits can be quite quick (around 5 days for GÉANT Plus).

Transparency

Each service is implemented at a different layer of the ISO/OSI network reference model, and therefore imposes different restrictions on using the features of underlying layers. For example, if a logical link is implemented on a 1 GE circuit using 802.1q VLAN technology, users are not normally allowed to define their own VLANs within the logical link (unless 802.1ah is supported along the circuit) or to configure Ethernet parameters along the circuits, since those are under provider control. The GÉANT Lambda service is implemented through DWDM technology, providing a pure Layer 1 wavelength which gives users an opportunity to use any features of Layer 2 and higher. GÉANT Plus services are implemented over Layer 2, and some of the Layer 2 features (e.g. double VLAN tagging) may not be available in logical links. If the GÉANT IP service is used, but without "* over IP" encapsulation, only features of Layer 3 and higher can be managed by the end users. In addition, users are not allowed to modify Layer 3 attributes of the logical link itself. While in some respects the GÉANT IP service provides the most restrictive features for the end users, its infrastructure also supports Layer 2 emulation (L2 VPN[6]) which provides users with control of Layer 2 features and higher.

Apart from the above considerations, service availability at particular sites must also be taken into account in choosing the most appropriate service for delivering logical links.

2.1.1.2 *Hardware*

The physical GÉANT infrastructure uses Alcatel-Lucent and Juniper equipment.

At Layer 1, the DWDM equipment used to light the leased and owned fibre is the ALU 1626 Light Manager (LM). Two trials with 40 Gbps links have been successfully tested on part of the GÉANT fibre footprint, and the ALU 1626 LM is in theory capable of supporting 100 Gbps transmission, provided significant enhancements are implemented. The cost-effectiveness of using DWDM technology and upgrading the currently deployed system to deliver 40 Gbps and 100 Gbps capacities is being investigated and compared with other options by GN3 SA1 (Network Build and Operations) as part of the transmission and switching equipment procurement for the new GÉANT backbone.

Layer 2 services are implemented with the ALU 1678 Metro Cross Connect (MCC), which offers both SDH and Ethernet implementations of the logical links but does not support 40 Gbps or 100 Gbps. SA1 is reviewing the optoelectrical switching technology deployed on GÉANT as part of the transmission and switching equipment procurement for the new GÉANT backbone. In most PoPs, 10 x 1 Gbps interface cards are available. However, in some PoPs only 1 x 10 Gbps interfaces are installed. Such PoPs therefore require an additional hardware switch, which needs to tag 802.1q VLANs in order to separate circuits from each other within the same physical

[6] As mentioned in the GÉANT IP service summary in Section 2.1 on page 5, although the L2 VPN is a Layer 2 service, it is classified as an Layer 3 service in the GÉANT services portfolio since it is implemented with the infrastructure delivering Layer 3 services.

port. As mentioned in *Transparency* above, this may influence the Layer 2 features available to end users, if the GÉANT Plus service is chosen for slicing.

At L3, Juniper M160 and T640 routers are used to implement GÉANT IP services. The equipment was re-procured two years ago and is future-proof in terms of its ability to support 100 Gbps in a scalable manner.

Further information about the hardware implications of supporting 40 Gbps and 100 Gbps services can be found in [GN3_DS1.1.1,2].

2.2 FEDERICA[7]

2.2.1 Objectives

FEDERICA (Federated E-infrastructure Dedicated to European Researchers Innovating in Computing network Architectures [FEDERICA]) was a European Commission co-funded project under the 7th Framework Programme. It started in January 2008 and officially ended in October 2010.

The main project objective was to create an e-Infrastructure for research into Future Internet, giving researchers complete control over a set of resources in a slice composed of network and computational resources, allowing even disruptive experiments and placing the fewest possible constraints on researchers [VInfra]. Researching the use of virtualisation in e-Infrastructures and facilitating collaboration between experts were also key goals.

FEDERICA is the first initiative to deliver an infrastructure and services compliant with the Network Factory concept involving key players from the European NREN community. As such, it is presented in greater detail here not only because more detailed information has been available to JRA2 T5 but also because most business case aspects are expected to be applicable to the GÉANT Network Factory as well.

[7] This section covers the objectives of FEDERICA and provides an overview of the infrastructure, service features, issues and limitations at the time of writing, as well as business aspects of the initiative. Further details – on implementation (physical elements, PoP setup), functionality (access and management), operations (slice creation) and user support – are provided in Appendix A on page 75.

2.2.2 Infrastructure Overview

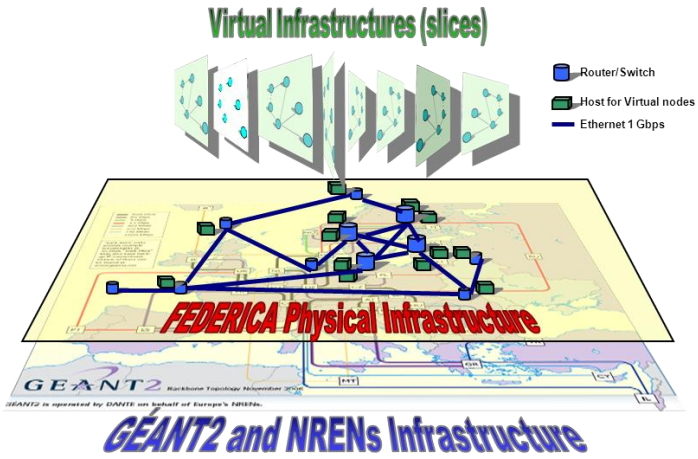

Figure 2.2: Graphical representation of FEDERICA [FED_DJRA2.1]

FEDERICA's setup relies on two key framework choices:

- The presence of physical network and computing resources (servers). These resources form the substrate of the infrastructure.
- The use of virtualisation/slicing technologies applied both to computing and network resources. Virtualisation allows virtual, non-configured resources to be created, e.g. an image of the hardware of a computing element on which (almost) any operating system can be installed, an emulated point-to-point network circuit, a portion of disk space. Those resources are then tailored to various needs.

This framework leads to a design in which the infrastructure comprises of two distinct layers:

- The FEDERICA substrate: The physical infrastructure which contains all the hardware and software required to create the virtual resources.
- The layer containing the user slices: Each slice contains the virtual resources and the initial network topology that connects them as per user requirements.

The FEDERICA substrate is a single administrative domain. The user slices (also known as Virtual Infrastructures (VIs) in FEDERICA's terminology) may, in principle, be unlimited; in practice, the number of instances is high, but is restricted by the physical resources available and the characteristics requested for each slice. FEDERICA slices are completely isolated from each other and behave exactly as if they were physical infrastructures from the point of view of each user experiment.

Two basic resource entities are defined as follows:

- Connectivity in the form of a point-to-point circuit with or without assured capacity and with or without a data link protocol (a "bit pipe").

- A computing element, offering the equivalent of computer hardware containing at least RAM, CPU and one network interface. Mass storage is optional, although usually available. The computing element is capable of hosting various operating systems and also of performing functions within the slice (e.g. routing, monitoring). Routing devices can be realised by computing elements, as they can be implemented using virtual machines or virtualisation capabilities offered by the FEDERICA substrate devices.

To minimise the load on the physical resources and the interference between virtual resources, the network topology has a high level of meshing. As most of the network interfaces for servers do not support virtualisation in hardware, additional physical interfaces are installed on the servers.

A detailed description of the FEDERICA PoPs and the infrastructure's network topology can be found in the FEDERICA deliverable "Update on the FEDERICA Infrastructure" [FED_DSA1.3]. A summary of the physical devices that make up the FEDERICA substrate are presented in A.1 *Physical Resources* on page 73.

The service architecture of FEDERICA follows the Infrastructure as a Service (IaaS) paradigm, exposing the resources of the substrate through virtualisation, both for computing elements and the network devices of a slice.

FEDERICA has two major services:

- Provisioning of Layer 3 Best Effort or Quality of Service (QoS) IP slices with the option of both pre-configured and unconfigured resources (i.e. routing protocols, operation systems, QoS parameters).
- Provisioning of Layer 2 slices in the form of Layer 2 data pipes (i.e. emulated Ethernet links) and unconfigured resources (i.e. empty virtual machines, clear routing tables, no network protocols).

2.2.3 Features

FEDERICA's features can be summarised as follows:

- FEDERICA's substrate is fully controlled and managed by the FEDERICA Network Operations Centre (NOC). This allows the user to control the resources at Layer 2 and above within their slices (depending on the layer at which the FEDERICA service is delivered). At the moment, raw resources in terms of data pipes that are available to the users are Layer 2 Ethernet links.

- Full control over the connectivity resources of the FEDERICA substrate (a mesh of GÉANT Plus circuits) allows specific QoS parameters to be assured for the logical links within the user slices and realistic transmission delays to be experienced. Currently the links run up to 1 Gbps but this may be upgraded.

- FEDERICA users have full flexibility and freedom to choose any networking protocol or operating system to be installed on their slices. Of course only certain options make sense depending on the FEDERICA service used (Layer 2 or Layer 3).

- FEDERICA can ensure the reproducibility of the testing environment and of the conditions of the user experiments at a different location or time. Repeatability of the experiments, in the sense of obtaining the same results given the same initial conditions at any time, can also be ensured.

- The overall architecture is federation-ready in line with the Slice-Based Federation Architecture [SFA] concept, a framework of controls and definitions that permits a federation of slice-based network substrates to interoperate and originates from the GENI initiative. Moreover, FEDERICA provides non-web-based, federated Secure Shell (SSH) access to all its resources, supported by the Single Sign-On (SSO) infrastructures already in place within the R&E community [FED_DJRA2.1, FED_DJRA2.3].

A detailed description of FEDERICA infrastructure use cases is available online [FED_DNA2.3, FedCase].

2.2.3.1 *Slice Creation and Management Process*

This section summarises the information provided in the FEDERICA deliverables [FED_DSA2.1] and [FED_DSA2.3].

The slice creation process relies on the information provided by the users and on the current state of the substrate. The goal is not simply to create the slice the user has requested, but also to optimise the use of the substrate, ensure a level of reproducibility of the computing resources, even if the user did not explicitly request it, and to avoid any adverse influence on other slices.

Depending on the user requirements, the technology used to create the network topology of each slice may differ from VLANs to MPLS, with routing or without. There is therefore no single predefined procedure and each slice has to be engineered separately. By default, the topology is created using VLANs, with routing provided by a routing-instance configured on a physical substrate router (see also *Router Virtualisation* on page 18).

The step-by-step procedure for creating a slice using VLAN technology without guaranteed capacity is summarised in [FEDdocs]. An overview is provided in A.4 *Slice Creation and Management Process in Detail* on page 78.

The NOC continues to support the user throughout the experiment, for example, if technical issues arise or if a blocked resource with no user connectivity needs to be reloaded. All slice configurations have a backup on a storage server in case of failures. User configuration backup and checkpoints for the virtual hosts are not yet implemented and are the responsibility of the user.

The NOC decommissions the slice when it is no longer needed.

2.2.3.2 *Virtualisation in FEDERICA*

As stated in Section 2.2.2, virtualisation is a key feature in FEDERICA and the FEDERICA infrastructure. It is used to provide slices: virtual resources over the physical substrate allocated on network devices, routers, switches and servers (the latter in the form of virtual machines). The slices:

- Have a loose or no dependency on the specific physical location of the exploited substrate devices or on a specific physical entity.
- Permit on-the-fly reconfiguration, cancellation and creation of resources in the slice (e.g. a routing element).
- Are often created by off-the-shelf components that offer embedded virtualisation functionalities.

Virtualisation in FEDERICA offers the following benefits:

- Testing new protocols on a network slice upon a production physical infrastructure guarantees more realistic, reliable results than canonical test activities performed in standalone environments such as a laboratory, a geographically limited testbed or a testbed not able to provide reproducibility of resource behaviour.

- Care is taken so that the provision of new slices to end users does not impact the configuration of slices that are already in place.

- In addition, the FEDERICA services allow for the migration of virtual routers between different physical locations in the FEDERICA infrastructure, that is, between the substrate devices that provide the resources in a slice, which simplifies existing network maintenance tasks.

Network Virtualisation

FEDERICA network virtualisation can be explained by the example of a slice containing only two hosts connected by a single circuit. Creating a virtual link between the two virtual systems requires the following steps:

1. Connecting the network interface(s) in the virtual hosts to one of the physical interface(s) in the hosting platform.
2. Creating a virtual circuit from one host to the other, with a specified assured capacity or with a best effort QoS.

The first aspect that the NOC considers in these steps is how to avoid congestion. The NOC follows the rule of assigning one physical Network Interface Card (NIC) for every virtual network card of the virtual machines. This can be done by creating a software bridge for every logical interface. This practice limits the number of virtual machines that can be hosted on the same server, but allows 1 Gbps capacity to be granted to a virtual NIC. If the user's slice does not require capacity guarantees, more virtual NICs can be mapped on the same physical link out of the server.

The technologies used to slice 1 Gbps physical capacity of substrate links are:

- MPLS. If capacity guarantees are needed, hardware features in the Juniper routers can be exploited.
- Ethernet VLANs: as for MPLS, but with the number of VLANs available for all the slices in FEDERICA limited to 4096.
- Physical circuits: limited to one circuit of 1 Gbps.
- IP-packet-based policies: limited to IP-based slices.

QoS guarantees for slice circuits are provided by enabling associated features of the substrate network elements.

Router Virtualisation

Router virtualisation refers to several routers hosted on the same physical device, sharing low-level resources like router CPUs, memory and forwarding engine.

FEDERICA core routers provide two kinds of virtual routers:

- Virtual router. This approach allows a new routing instance to be allocated inside the Juniper router. With virtual routers only one general routing table is available.
- Logical system. This feature segments a physical router so that its slices can be configured to operate as multiple independent routers. This approach provides flexible routing segmentation by cloning the processes of the routing protocol daemons. Multiple logical devices act as completely independent routers.

Host Virtualisation

The virtualisation of servers is based on hypervisor software. The hypervisor selected to manage the virtual machines in FEDERICA is VMware ESXi version 3.5. It has been selected because of its performance and minimal hardware limitations compared to competitors. In addition, it provides effective remote management APIs.

FEDERICA software routers are implemented as virtual machines using Ubuntu Server as the main OS and XORP as the routing tool. XORP [XORP] has been selected because it is an open-source tool and supports a larger number of protocols than similar software tools. Users can manage software routers through command line interfaces.

Monitoring the Slices

In FEDERICA the physical substrate is monitored, extending the monitoring capabilities to the virtual slices. As virtual slices are created, the physical connectivity among the devices participating in the slice is validated. At the same time, the virtual connectivity within and between slices also needs to be validated. Monitoring features are available in the infrastructure, both for hosts, the physical servers, and links. Statistics about the substrate and the slices are collected by the HADES (Hades Active Delay Evaluation System) network measurement suite and are exposed both to the NOC and to the users who own a slice.

See [FED_DJRA1.2] for details about the FEDERICA monitoring system.

2.2.4 Issues and Limitations

FEDERICA's main issues and limitations relate to scalability, automation of slice creation, and multi-hop links.

Compared to PlanetLab or commercial clouds, the scalability of the FEDERICA virtual infrastructure is limited by the underlying physical substrate. As a consequence of this, user access to the FEDERICA slices has to be governed with care by the User Policy Board.

In addition, FEDERICA's automation for instantiating virtual resources is limited. As noted in Section 2.2.3.1, the process of slice creation is not trivial, as it needs to optimise the use of the substrate, whilst complying with user requirements. Link usage also has to be considered. It is desirable to limit the use of core links in order to ensure that the slice behaves as predictably as possible, even if reproducibility has not been explicitly requested. Provisioning takes place on a slice-by-slice basis, with extra care for the assignment of VLAN and IP plane ranges. If a slice requires logical routers, it constrains the virtual network topology, since the logical router must be hosted in one of the core PoPs. Overbooking of network resources also has to be avoided.

The manual setup of the resources means that FEDERICA provides neither the Platform as a Service (where various pre-configured images and network scenarios can be provided from a repository), nor the Application as a Service (where a set of applications can be pre-selected by the users).

In addition, although the FEDERICA enabling technologies are suitable for point-to-point virtual links, in the case of virtual links with QoS requirements implemented over multi-hop physical links additional planning is needed to avoid resource congestion. Only through careful manual engineering by the NOC can each virtual network topology be deployed suitably. These are further reasons why it was not possible to expose Platform as a Service (PaaS) functionality to the FEDERICA users.

2.2.5 FEDERICA Business Model, Implementation Costs and Maintenance

This section summarises the information provided in FEDERICA Deliverable DJRA2.1 "Architectures for virtual infrastructures, new Internet paradigms and business models" [FED_DJRA2.1]. It is provided here because the business case aspects of FEDERICA, as an initiative driven by the GÉANT-NREN community, are expected to be applicable to / reusable by the GÉANT Network Factory as well.

2.2.5.1 *Approaches to the FEDERICA Business Model and Economic Issues*

The FEDERICA team produced a general overview of the latest business model taxonomies and conducted a survey of the business-related concepts considered by related projects and frameworks, in order to make recommendations for a FEDERICA-specific potential business model.

FEDERICA can be viewed as a provider that offers a platform for running experiments over a virtual network, with a virtually private infrastructure. FEDERICA customers using this service are expected to be:

- Industrial or academic institutions or project consortia wishing to test their innovative approaches.
- Industrial partners offering new software and/or hardware, possibly in collaboration with other industrial partners or projects. The benefits for such customers (who are also providers) may be the ability to test and fine-tune their equipment or software and its interoperability features, reach a set of customers, introduce or help the adoption of a new technology, etc.

The FEDERICA business model should therefore focus on business requirements for customers aiming to run experiments over the platform, such as a specification of charging schemas and SLAs offering the right incentives in terms of resource usage, quality levels selected, etc.

In fact, such customers might also have their own business models. For instance, a customer could define a complex network topology needed for experimental purposes, reserve a FEDERICA slice to implement it, and then offer this slice as a platform to other projects.

The three main types of FEDERICA customers, and their incentives and benefits, are summarised in Table 2.2.

Customer Type	Incentives	Benefits
Project or industrial / academic institutions	Test innovative research approaches.	Access to the right experimental infrastructure.
Provider of new hardware or software	Test and promote new products.	Introduction of new technology to potential customers. Testing and fine-tuning of the product in a real and challenging environment.
Provider of basic infrastructure	Attain more customers.	Revenue.

Table 2.2: FEDERICA customer types, incentives and benefits

The requirements of a customer wishing to run experiments over the FEDERICA platform can be met by combining the resources offered by FEDERICA with the customer/user's own infrastructure, i.e. with a local testbed. Increasingly, large business and science projects need dedicated networks for specific applications and high data volume grids. They want to be able to manipulate the network in the same way they can manipulate the application. FEDERICA can add value by providing slices, resources and tools.

The sharing of resources is a much-debated issue. For example, if a specific number of slices employ a physical link with a specific capacity, there are two ways to share this link:

- By statically allocating capacity between them. This may lead to under-utilisation.
- By not offering any capacity guarantee at all. This may influence the experiments' results, depending on the conditions.

In order to deal with the resolution of resource conflicts, a practical incentive mechanism based on charging is proposed. Indeed, charging is the most effective incentive mechanism in any system in which resources are limited, and can lead each customer/consumer to demand only the resources actually needed, while operating in a simple and distributed way.

Charging can employ either actual or virtual money (i.e. tokens), and should be complemented by a mechanism for determining the prices of the various resources per time unit. In particular, prices (either in money or in tokens) can be:

- Fixed over long time periods, thus giving rise to the risk of exhaustion of the supply of certain resources.
- Fluctuating according to the relation between demand and supply, either of which can vary.

FEDERICA wishes to offer the resources for free, rather than charging actual money. Thus renewable tokens can be used as an incentive mechanism in order to avoid exhaustion of resources. Essentially these tokens amount to a renewable right to use a certain amount of each of the resources for free during a specified period of time. This approach has similar requirements to an approach where real money is used, in the sense that demand may fluctuate and thus the amount of resources each token "buys" may vary over time. As a simple example, assume that tokens for capacity are awarded on an hourly basis, but demand varies with the time of day. A token should "buy" less capacity during the busy hour than off peak. The tokens should not be storable, as this would allow a consumer who has been inactive for a long time to reserve a large portion of the resources at once. It should finally be noted that even if real money is charged, prices should be low enough to ensure that resources are affordable and utilised. Charging should function primarily as a control mechanism rather than as a means to generate revenue.

Prioritisation of the various users for accessing resources can be attained by means of charging, even if no actual money is employed. In such a case, users are price takers, that is, prices are non-negotiable, although they may vary over time in order to attain market clearing.

An alternative mechanism is to run auctions to determine resource prices and allocation, thus also avoiding the possibility of exhausting the supply of a resource. There are several auction mechanisms that could be used in the context of FEDERICA. The main source of complexity here is that each user has to reserve a multitude of resources, both communication and computational. Thus, there is a large range of possible combinations. If sealed-bid combinatorial auctions are employed, then the complexity of determining the winner will be prohibitive. However, if simultaneous ascending auctions are employed, for example, for each slot, then it is likely that it will last for many rounds, and it will be hard for bidders to employ a meaningful strategy due to the multitude of available options. One way of simplifying the problem of auction-mechanism design is to run auctions only for the scarce resources attracting high demand, while selling the less scarce ones at fixed prices since it is unlikely that they will be exhausted.

The following is a summary of the key steps required to implement the ideas discussed above:

- Define an abstract resource model. This includes defining the various resource types (exclusive, shared, offering minimum guarantees, etc.) and the resource units (slices) that will be offered to customers. Once this model is stable, the economic mechanisms will be defined independently of resource implementation details.

- Define the time model. This complements the resource model by specifying the time slots over which resources are granted.

- Define a demand model. This includes specifying how customers express their experiment requirements in terms of resources and time. It should be consistent with the specific needs of FEDERICA customers.

- Define an inventory/resource allocation model. This specifies the economic mechanisms to be deployed in the context of the resource and demand models defined above for allocating resources over time (present and future).

- Define an adaptation strategy. Since the demand in such a system will be shaped once the above mechanisms are in place, room for the mechanism to adapt and improve performance must be factored in. Hence simple strategies need to be implemented that will learn from the actual behaviour of the system and improve its success in reserving and allocating resources over time.

Of course, all the above mechanisms will be valuable in the case of competition, i.e., when demand exceeds supply.

2.2.5.2 *Estimates of Capital and Operational Costs of a FEDERICA-like Infrastructure*

FEDERICA has been built using off-the-shelf hardware and software. The objective was to maximise virtualisation capabilities and functionalities, rather than to maximise the performance of the servers and network equipment. (Performance is still a consideration, however, and is optimised by the use of core routers/switches that are capable of line-rate switching and have the full set of functionalities and software capabilities.)

The capital and operational costs of building and operating a FEDERICA-like infrastructure are summarised in the tables below, as a result of a JRA2 T5 enquiry. The purpose of this assessment is to obtain a real-life estimation of the CAPEX and OPEX implications for an infrastructure comparable to the GÉANT Network Factory and its associated services.

The costs are given as an indicative range, since there are many configuration options for each type of asset (for example, to provision servers of higher capacity or extended circuits). Operation has a fixed minimum cost, then it grows approximately in line with the number of personnel involved.

Type of asset	Unit cost (indicative range)	Notes	Units in FEDERICA
V-Nodes	2K – 8K euro	The total cost is not just the sum of the number of units, but also of their configuration. A working solution's cost can be expected to be closer to the minimum value of the range provided.	25
Routers/switches	3K – 20K euro (small)	Small routers/switches are suitable for non-core PoPs with a maximum of 2 servers. The related software includes at least the VLANs, OSPF, MPLS protocols and virtual routing instances. The cost estimates relate to an Ethernet-based switch with routing functionality and none or minimal expandability.	12
	50K – 200K euro (medium)	Medium routers/switches have additional functionality, e.g. BGP, HW-based QoS, line-rate switching, multiple line cards.	4

Type of asset	Unit cost (indicative range)	Notes	Units in FEDERICA
Software licences	0.2K – 1K euro	FEDERICA decided to use the free version of the virtualisation software. If the infrastructure is large, management may be improved by installing commercial licences and using the management platform for the specific virtualisation software (additional cost of 30-50K euro).	1
Installation of circuits	0.6K – 1.5K euro	Includes personnel expenses, cabling.	
System installation	12 months – 24 months effort	Mainly person effort and cannot be evaluated without a detailed configuration. For FEDERICA it was quite time consuming and can be estimated between 12 months and 24 months of expert effort, mainly due to the long delivery time of the GÉANT Plus circuits. The remainder of the deployment (hypervisors, User Portal, Web Site, Monitoring) took about 6 PM.	n/a

Table 2.3: Capital costs

Type of Asset	Unit Cost (Indicative Range)	Notes	Units in FEDERICA
V-Nodes maintenance licences	7% – 15% of non-discounted initial cost	Only needed when the guarantee expires.	-
Routers/switches (hardware)	10% – 15% of non-discounted initial cost	Hardware maintenance costs vary according to the requested response time. In FEDERICA this was next business day.	
Routers/switches (software)	-	The cost depends on the hardware type and on the enabled options.	
NOC	50K – 100K euro	A minimum of two dedicated personnel is needed. This includes overall system maintenance, software and hardware updates, and service operation (creation of new user slices).	
User support (non NOC)	50K – 100K euro	The task is to support the user in the initial phase of using the facility. A minimum of one person is needed.	

Type of Asset	Unit Cost (Indicative Range)	Notes	Units in FEDERICA
Housing equipment (including power, cooling)	50 – 200 euro	This cost might be embedded in an existing contract if the number of units is small.	
Wide area network connectivity at 1 Gbps	40K – 200K euro / year	The range given is very wide due to the wide variation of cost. The exact value depends on different factors (e.g. the country, circuit length, and reliability). In FEDERICA all the circuits have been provided through GÉANT and a precise estimate is difficult.	

Table 2.4: Operational costs

The tables show that the operational costs are the larger and also dependent upon the specific types of infrastructure components.

The cost to maintain and evolve hardware and software should also be evaluated. The software development effort to produce a workable environment for the user and NOC operations is particularly significant and can be estimated in the order of 500K to 1000K euro in the first two years.

2.3 Other Initiatives

This section presents a summary of relevant third-party projects and initiatives that JRA2 Task 5 evaluated in order to identify technical aspects and best-practices that could be applied or replicated within the GÉANT Network Factory production environment.

The complete evaluation data gathered by JRA2 Task 5 is available online [NFEval].

2.3.1 Evaluation

	What is the technical philosophy of the project / initiative?
ANI	ANI provides a testbed where specialised networking research with 100 G requirements can be performed. The testbed is initially an isolated network for experimenting. Later it will be connected to the production-level network, but it will still function as a separate entity with the philosophy of being configurable, breakable, reservable, able to be reset, and dedicated to research projects.
VINI	VINI [VINI] is a PlanetLab-like testbed where users can configure virtual topologies within their slices. A VINI virtual topology consists of virtual machines (aka “slivers” on PlanetLab [PlanetLab]) connected by point-to-point virtual links.
OFELIA	The philosophy of the OFELIA project is to create an experimental facility that allows researchers not only to experiment on a test network but also to control the network itself, precisely and dynamically.
AKARI	AKARI [AKARI] has developed a pool of testbeds for the Japan NGN. Two technologies are provided: • A software-only testbed (CoreLab), derived from PlanetLab Japan and sharing the server nodes. • A testbed built on special purpose devices (VNodes) connected through dedicated links.
GENI	GENI is an infrastructure designed to support experimental research in network technologies. It allows research on multiple layers of abstraction, from the physical substrates through the architecture and protocols to networks of people, organisations, and societies. The core concepts for the suite of GENI infrastructures are programmability, virtualisation and other forms of resource sharing, federation and slice-based experimentation.
PASITO	PASITO is a fixed infrastructure deployed over the production network with connected routers, switches and servers, all dedicated to this infrastructure. The infrastructure is not open so only the groups that are part of the PASITO network can use it. The infrastructure is designed to allow researchers to perform different kinds of tests that cannot be performed in local testbeds due to resource limitations.

Table 2.5: Evaluation table 1: technical philosophy

	What is the available hardware? (Layer 1, Layer 2, Layer 3)?

ANI	• GMPLS-enabled DWDM devices (Layer 0-1)
	• Layer 2 switches supporting OpenFlow
	• Layer 3 Routers (Juniper M7i)
	  ○ OSCARS compatible, MPLS-enabled
	• Test and measurement hosts
	  ○ Virtual Machine-based test environment
	  ○ 4 x 10 G access for test hosts initially (eventually 40 G and 100 G)

VINI	VINI consists of 42 nodes at 27 sites.
	VINI Node Specifications:
	• National LambdaRail:
	  ○ Acme 8X15LT2-DC 1U rackmount server with DC48V power
	  ○ Dual Intel Xeon 2.8 GHz, 800 MHz system bus
	  ○ 4 GB RAM
	  ○ 2 Seagate 250GB SATAII drives
	  ○ Supermicro IPMI 2.0 remote management card
	• Internet2:
	  ○ HP DL320g5 1U rackmount server
	  ○ 2.4 GHz dual-core Intel Xeon 3060
	  ○ 4 GB RAM
	  ○ 2 160 GB SATA drives
	  ○ iLO2 remote management

	What is the available hardware? (Layer 1, Layer 2, Layer 3)?
OFELIA	**[OFELIAIsland]** **Island i2CAT:** • 5 x NEC IP8800/S3640-24T2XW switches • 3 x SuperMicro server SYS-6016T-T • 2 x SuperMicro server SYS-6016T-T **Island TUB:** • 4 x NEC IP8800/S3640-48TW • 1 x NEC IP8800/ S3640-48TW(reserve), • 1 x HP 5406 zl Pro Curve with a 24 Port SFP+ Module • 1 x IBM X3650 M3 with 2x Xeon 4x **Island IBBT:** • iLab.t Virtual Wall: ○ 100 nodes (dual CPU core processors) at 2.0 GHz ○ 4 GB RAM ○ 4 disks of 80 GB ○ 60 6 x 1 GbE and 40 4 x 1 GbE experimental interfaces connected to Force10 E1200 switch (576 x 1 GbE + 8 x 10 GbE ports) • w-iLab.t wireless testbed: ○ Across three 15 m x 91 m floors, 200 PCEngine Alix3c3 (500 connected to two 5 dBi dual band antenna **Island UEssex:** • 4 x NEC IP8800/S3640-24T2XW switches • 3 x Extreme Black Diamond 12804 Switches

	What is the available hardware? (Layer 1, Layer 2, Layer 3)?
	• 3 x ADVA ROADMs • 1 x Calient Diamond Wave fibre switch • 4 nodes (2 x Intel Xeon E3110) • JPEG 2000 4K video coder and decoder • Anritsu Traffic analyser **Island ETHZ:** • 3 NEC IP8800/S3640-24T2XW • 2 optical 10 GB interfaces • Machine for the FlowVisor
AKARI	• Layer 0-1, no information provided. • CoreLab: PlanetLab tools with GRE-tap tunnels and virtual OpenFlow switch. • VNode: GRE encapsulation, support for MPLS, VLAN, and OpticalPath foreseen (not yet implemented).

GENI	• OpenWave:
	  ○ HP ProCurve 5400 Switch
	  ○ NEC IP8800 Switch
	  ○ Toroki Lightswitch 4810
	  ○ Quanta L4BG Switch
	• ORBIT:
	  ○ NEC WiMax Base Station
	• ShadowNet (ProtoGENI):
	  ○ Juniper M7i Router

PASITO	Mainly Layer 2 and Layer 3 hardware, with very limited access to Layer 1 equipment. Layer 1 is only deployed between two nodes and not across

	What is the available hardware? (Layer 1, Layer 2, Layer 3)?
	the whole network. Layer 2 Cisco and Juniper switches and also a CRS-1 and different Juniper M-series routers are available.

Table 2.6: Evaluation table 2: available hardware

	Does the project / initiative share the physical substrate with the production environment?
ANI	No.
VINI	No
OFELIA	No, OFELIA is a dedicated testbed.
AKARI	CoreLab: Yes. VNode: No.
GENI	Yes.
PASITO	Yes.

Table 2.7: Evaluation table 3: physical substrate sharing

	Does the project / initiative use virtualisation? If yes, to what extent?
ANI	The application hosts have a virtualisation environment and all test applications will run under virtual machines.
VINI	VINI currently uses Trellis [Trellis], a set of extensions to the PlanetLab kernel. More precisely, NetNS [NetNS] supports virtualisation of the network stack. Currently, VINI users are only able to request virtual topologies (i.e. nodes connected via virtual links hosted on sites connected via physical links) that mirror the physical network connectivity.
OFELIA	Yes. Network virtualisation by OpenFlow.
AKARI	CoreLab supports OpenFlow-enabled network virtualisation. In addition, specific topologies can be created through OpenFlow virtual machines. VNode: A custom hardware device that enables network virtualisation through slicing of its cards and servers. The virtual machines are used exclusively to build the routing logic (e.g. using Quagga) that drives the device forwarding engine.
GENI	GENI uses virtualisation technology, mainly using OpenFlow on switches and routers. Also, for the WiMax ORBIT solution, every GENI slice runs as a separate virtual machine, emulating its own router and performing IP routing, or alternatively implementing novel non-Internet protocols.
PASITO	The project uses virtualisation mainly to provide virtual machines to the researchers. Virtualisation is also available in routers but is limited by hardware capabilities.

Table 2.8: Evaluation table 4: virtualisation

	Does it use slicing? Slicing of physical resources (e.g. the backbone itself) or slicing of virtualised network resources.
ANI	N/A. The researchers have direct access to the networking equipment.
VINI	VINI is a PlanetLab-like testbed where users can configure virtual topologies within their slices. VINI supports simultaneous experiments with arbitrary network topologies on a shared physical infrastructure.
OFELIA	Network nodes that are OpenFlow 1.0 capable, in most cases slices of NEC IP8800/S3640 switches. A (virtual) machine that will run the OpenFlow controller controls the network slice. (Virtual) machines that act as network endpoints. Researchers will be allowed to SSH into these virtual machines to conduct their experiments (e.g. traffic generation and analysis).
AKARI	CoreLab: no VNode: some slicing must be implicitly supported. As the number of available FastPath cards is limited, the forwarding engine must be sliced among a fixed number of virtual machines. However, no explicit indication about this issue is provided in the AKARI documentation.
GENI	GENI uses slicing on virtualised network resources.
PASITO	The PASITO network infrastructure can be considered as a slice of the RedIRIS10 production network as it has been built deploying VLANs and L2 VPNs over RedIRIS10. Users can request network and computer resources and deploy their own topology over PASITO infrastructure.

Table 2.9: Evaluation table 5: slicing

	Does the project focus on the network or does it also take computing into account?
ANI	The project is intended as a testbed for research into networking itself, but also as a testbed for projects related to other services (mainly data transfer protocol and services) in a 100 G environment.
VINI	Both computing and network resources. A VINI virtual topology consists of virtual machines connected by point-to-point virtual links. Applications running inside the VINI slice can send and receive traffic over the virtual topology, and also control how packets are forwarded within the topology.
OFELIA	The project focuses on the network: OpenFlow controllers, which control the flows, and OpenFlow software 1.0 installed on NEC routers/switches.
AKARI	The project focuses on the network.
GENI	The project focuses on the network.
PASITO	There are mainly networking experiments but computing resources are also available.

Table 2.10: Evaluation table 6: project focus

	What network layers does the system use for slicing or virtualisation?
ANI	N/A
VINI	Layer 2 and Layer 3
OFELIA	Layer 2 and Layer 3
AKARI	Layer 2 and Layer 3 according to the available documentation. There are some references to Layer 1 slicing, but no details are provided.
GENI	Layer 2 and Layer 3
PASITO	Layer 2 and Layer 3

Table 2.11: Evaluation table 7: slicing and virtualisation network layers

	Where and for what kinds of application has the system been used (if already implemented)?
ANI	The following projects currently use the testbed: • Advance Scheduling of Multi-Domain Dynamic Circuits • Usability Investigations for High-Energy Physics Analysis • Securing Network Services using DASH • Testing high-speed protocol PERT over a real 10 Gbps network • Scalable Optical Networking with OpenFlow • Measuring Energy Efficiency In Networks
VINI	VINI allows users to test their applications under virtual topologies (i.e. nodes connected via virtual links hosted on sites connected via physical links) that mirror the physical network connectivity. They must access the VINI RSpec (Resource Specification). The VINI RSpec lists the VINI sites, describing the nodes they host, the unallocated capacity for each node and all capacity limits of each node's network interface card.
OFELIA	Using OpenFlow software 1.0 and FlowVisor tools.
AKARI	N/A
GENI	• PlanetLab • ProtoGENI: Emulab-based network and distributed computing testbeds • CMULab wireless networking testbed • ORBIT wireless networking testbed • Diverse Outdoor Mobile Environment (DOME): programmable optical network experiment environment • Breakable Experimental Network (BEN): programmable optical network experiment environment • KanseiSensorNet: Extreme Scale Motes (XSM) based sensor network testbed • ViSE: outdoor wide-area sensor/actuator network testbed

	Where and for what kinds of application has the system been used (if already implemented)?
PASITO	The following are current projects in PASITO: • Large information transfers in IP networks • Network virtualisation • IPv6 services • Multiservice networks • Multimedia flows • Multiprotocol collaboration and experiences • Monitoring, provisioning and management tools • Next Generation Optical networks • AAA new services

Table 2.12: Evaluation table 8: application usage

	What part of the project / initiative is applicable to or could be replicated in the GÉANT Network Factory?
ANI	The project is ongoing and much of it (even some basic concepts) has been redefined during the past year. At the moment it seems that various proposals for using the testbed for research are dealt with in a case-by-case manner. Each project has to go through a project proposal submission process and has to be accepted. After acceptance the access to resources is dedicated and regulated via a Google calendar reservation system. A systematic approach of the degree of a factory is not in place. At its current level of development, therefore, ANI can be relevant to the GÉANT NF for obtaining use-case ideas, and possible backbone infrastructure that will allow networking research projects in 100 G speeds.
VINI	VINI does not really comply with the GÉANT NF model as it is built on top of the PlanetLab infrastructure, a research testbed consisting of nodes interconnected through Internet2. Thus, it does not have its own network resources and operational environment. Furthermore, it is based on simulated resources.
OFELIA	OFELIA is using OpenFlow as a technology for separating experimental traffic from productive traffic. In this way, researchers can try new routing protocols, security models, addressing schemes and even alternatives to IP, leaving production traffic unaffected. Using OpenFlow to establish an overlay network over the production environment is a principle to be considered by the GÉANT NF.
AKARI	AKARI as a whole is still being defined but the relevant testbeds have already been deployed. As both the software and the hardware solutions rely on dedicated code and special-purpose devices with dedicated infrastructure, there are no infrastructure elements that can be adopted “as is” by the GÉANT NF. Virtual infrastructures are configured by users using XML configuration files describing the requested slices. Once the slices are available (no details about the instantiation process are available), they are accessed and managed through web portals. Such features should be taken into account for the GÉANT NF.
GENI	The scope of GENI as a multi-disciplinary initiative makes the identification of specific best practices relevant to the GÉANT NF quite challenging. For network slicing, GENI mainly uses OpenFlow, thus providing additional evidence of the potential of this technology. It is expected that a more detailed analysis of the GENI developments (such as the Slice Facility Architecture (SFA) model) will be required in subsequent steps of the GÉANT NF design and specification work.

	What part of the project / initiative is applicable to or could be replicated in the GÉANT Network Factory?
PASITO	The project provides a fixed infrastructure for researchers without interfering with the production network. It uses resources both from the RedIRIS10 backbone and institutions connected to RedIRIS. By agreement, this hardware can be used by all PASITO participants and it is maintained by each one of the PoPs participating in PASITO. This model could be extended to GÉANT and the NRENs.

Table 2.13: Evaluation table 9: GÉANT applicability

2.4 Conclusions

The analysis of the projects and initiatves relevant to the GÉANT Network Factory concept has identified various practices for providing slices of computing and network resources to the end-user researchers. It appears to be desirable for the GÉANT Network Factory supporting infrastructure to allow for research at 100 Gbps. It is also a common practice to deploy testbeds for network research that allow overlay networks to be established over a production environment, with OpenFlow being a widely deployed technology.

A detailed study of FEDERICA has provided significant insight into several aspects of a solution for the delivery of Network Factory functionality to end users by the GÉANT-NREN community. These include the combination of network and computing resources as well as virtualisation technologies, a model for defining and delivering user slices, and service offerings at different layers with differing degrees of freedom for experimentation by end users. However, replication of the technology choices made by FEDERICA in the GÉANT Network Factory is not recommended as the landscape is evolving quickly and there is room for improvement. Furthermore, the operational as well as provisioning models in FEDERICA are far from fully functional, while many of the associated business aspects are not fully addressed.

Best practices with applicability to the GÉANT Network Factory case, as summarised in Table 2.13, in combination with the following chapter *Network Factory-Enabling Technologies for GÉANT*, and the outcomes of the ongoing procurement for the next-generation GÉANT, will all have to be incorporated into the future GÉANT Network Factory work to elaborate a business case and implement a long-term solution.

3 Network Factory-Enabling Technologies for GÉANT

This chapter presents the technologies applicable to GÉANT for realising a Network Factory infrastructure and services. The scope is not limited to GÉANT as it stands today, but takes into account technologies that are or soon will be commercially available. The analysis also takes into account the possible evolution of GÉANT as foreseen based on the recommendations made in the architecture study carried out by SA1 ([GN3_DS1.1.1,2]). The recommendations include features such as resiliency and robustness to failures, ease and speed of reconfiguration, and, most pertinently for the Network Factory, upgrades to the current optical layer, enhancing the physical topology, reviewing the switching layer technologies, and upgrading the IP layer equipment. According to this study, any technology choice for GÉANT should not preclude virtualisation. (Further details are given in [GN3_DS1.1.1,2].)

The chapter is divided into three sections, each of them focusing on technologies at a different network layer:

- The Layer 1 section describes how to implement virtual network facilities over the lowest possible layer in the GÉANT network, with the emphasis on optical links.
- The Layer 2 section describes the creation of virtual infrastructures with Layer 2 capabilities using features of Ethernet and MPLS, and emerging technologies such as Provider Backbone Bridge (PBB) and Provider Backbone Transport (PBT).
- The Layer 3 section describes how to implement virtual network facilities with Layer 3 functionality for users.

3.1 Network Factory-Enabling Technologies at Layer 1

3.1.1 Current GÉANT Layer 1 Technology

The DWDM technology currently used for the GÉANT backbone complies with ITU-T standards regarding wavelengths and spacing. The finest granularity that permits 100 Gbps per lambda circuit is a 50 GHz spacing grid. Alien lambdas (as defined in the GN3 Deliverable DJ1.2.1: “State-of-the-Art Photonic Switching Technologies” [GN3_DJ1.2.1]), could also be transported over the GÉANT backbone, as some experiments have already demonstrated (see Section 2.5.8 of [GN3_DJ1.2.1]).

Many NRENs manage their own dark fibre backbone, so the transport of lambdas across Europe is now an established fact (see Figure 3.1[8]). Some GÉANT PoPs are still only interconnected by leased lambdas, but the process of extending lit fibre is ongoing [GN3_DS1.1.1,2].

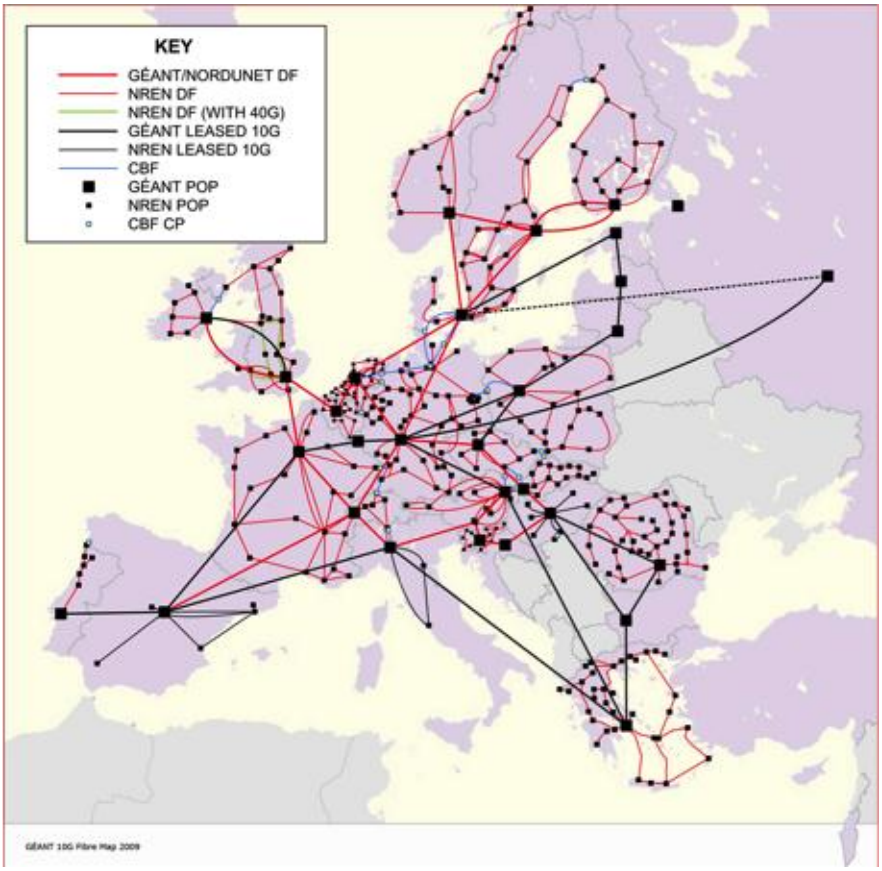

Figure 3.1: GÉANT and the NRENs 10 G fibre map – at the end of 2009

3.1.2 GÉANT Slicing Options at Layer 1

3.1.2.1 *Technology and Service Definitions*

The light beam whose wavelength complies with ITU-T grid standards and is transported across the dark fibre is called a “lambda beam”. The lambda beam is the carrier of the modulation signal that includes data. Using modulation, a coding convention is established by which the bits are recognised by the interfaces. The interfaces receive and transmit the modulated lambda beam (line cards/interfaces) to the fibre and deliver the data to the end user (client cards/interfaces) at Layer 2 of the OSI model (Ethernet or SDH). The end users have a “lambda circuit” or “lambda” established between them, but they are constrained to use the same vendor interfaces since the coding is vendor-specific [GN3_DJ1.2.1].

[8] The figure shows the NREN 10 G fibre footprint at the end of 2009. An updated map is being produced; however, it will not be available in time for the publication of this deliverable.

Thus a lambda circuit is delivered at Layer 2, a modulated lambda beam is transmitted on Layer 1, and the optical carrier signal (defined by such physical characteristics as wavelength and frequency spectrum) is called Layer 0. It is very difficult to delimit the border between Layer 1 and 2 strictly, since all optical network equipment is built by manufacturers according to their own coding/modulation methods, for which there are no standards.

Layer 0 is managed by the optical control plane and allows the fibre network manager to choose the light path of the lambda beams, to insert/extract lambda beams into/from fibre, and to react in the case of fibre failure. Each vendor has their own optical control plane implementation, so a unifying management system is required.

A typical service implemented at Layer 0 is the Photonic service; at Layer 1 it is the Lambda service.

> The **Photonic service** provides an end-to-end connection between two or more points in the network and is characterised by its photonic path and allocated capacity. The photonic path is a physical route on which light travels from one end point to one or more other end points. Allocated capacity is a part of the system spectrum that is reserved all along the photonic path for users of the Photonic service. It is important to carry signals over the network with minimal – if any – interventions, so that the processing at the end point will depend only on the application.
>
> The **GÉANT Lambda service** provides private, transparent 10 Gbps wavelengths between any two GÉANT NREN PoPs connected to the GÉANT fibre cloud so that NRENs can then develop their own higher-level network layers.

3.1.2.2 *Optical Slice Using Static Lambda without OEO Regeneration*

For point-to-point and bus topologies of dark fibre, the simplest method for providing static lambda circuits between end sites is by an Optical Add Drop Multiplexer (OADM), as shown in Figure 3.2.

An OADM allows new lambdas to be added to or existing ones to be decoupled from the multiplexed beam transmitted through the fibre. An OADM device has two directions, assigned to the two fibre pairs connected to the device.

Usually, for this type of scenario, most lambdas pass unchanged through the OADM, but some of them are dropped from the multiplexed beam and, as a consequence, new ones can be added at the same OADM towards another direction.

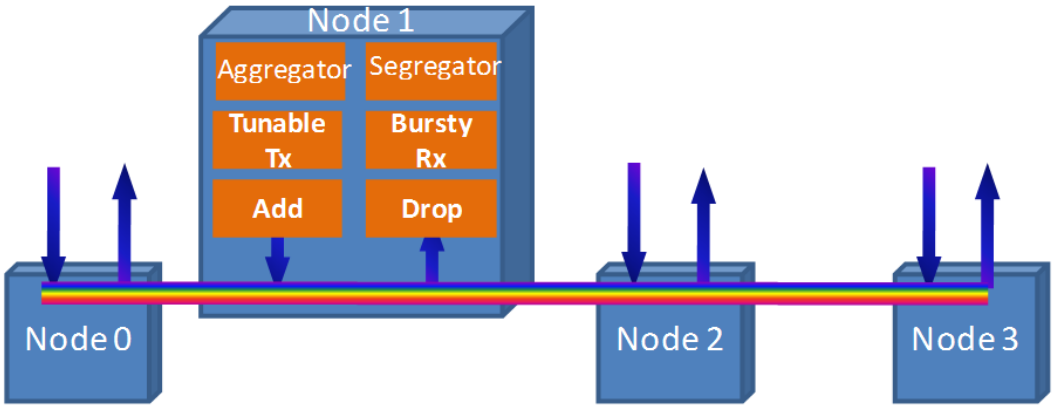

Figure 3.2: Bus topology with ROADM

A link slice using a lambda without Optical-Electrical-Optical (OEO) conversion is the basic building block of the Photonic service. This slice has specific features:

- The signal modulation output is identical to the modulation input even at analogue level.
- The propagation delay is the shortest possible.
- The carrier frequency ("colour") does not change – the slice is even able to transport a coherent optical signal. (The light spectrum can be divided into colours, according to frequency [LightCol].)
- The optical path is fixed, therefore the path parameters do not change.
- Both ends of a created lambda circuit have to be equipped by the same optics manufacturer.

This type of slice is required for specific network services, e.g. transfer of highly accurate time signal transfer and stable frequency [OptTime, OptStable]. It can also be used by applications that use non-standard modulation or a coherent end-to-end optical signal.

3.1.2.3 *Optical Slice Using ROADMs without OEO Regeneration*

Reconfigurable Optical Add-Drop Multiplexers (ROADMs) enable remote configuration of wavelengths at any node. The network operator is able to choose (using configuration utilities) which lambda is added, dropped or passed through the site. Typically, a ROADM node is based on Wavelength Selective Switch (WSS) technology, with additional components. Figure 3.3 shows the principal schema of a 4-channel ROADM.

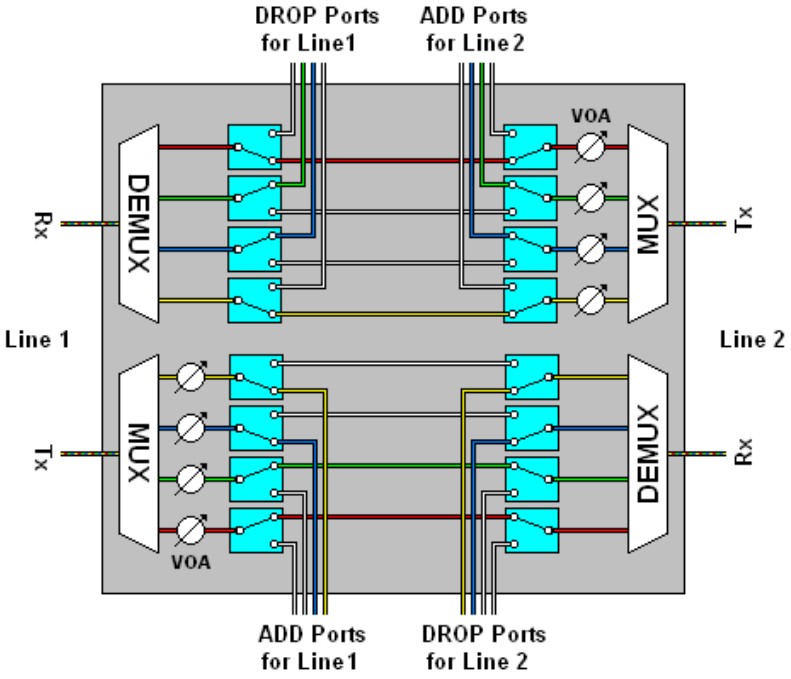

Figure 3.3: Reconfigurable Optical Add-Drop Multiplexer (4 lambdas)

ROADMs behave like an Optical Cross-Connect (OCX) device, but the switching is done between the lambdas rather than between the light beams at the fibre end. The number of switching directions (e.g. optical link interfaces) is often referred to as degrees of switching, generally associated with the transmission fibre pairs. Each degree requires an additional WSS element.

The features and application area of this type of link slice are similar to slices using static lambda (here again, both ends of a created lambda circuit have to be equipped by the same optics manufacturer). The only difference is that the optical path is not fixed, as the control plane can create alternative connections.

3.1.2.4 *Optical Slice Using ROADMs and OEO*

This scenario deals with the most common optical network design, where data channels are regenerated at an OEO node. A regenerator converts the optical signal to an electrical signal which is then transmitted to the next optical link, possibly at a different lambda. One advantage of this is the unlimited range of the optical network.

Unlike the case of a slice without OEO, in this case standard modulation and coding that are compatible with the regenerator must be used. This scenario is typically used for lambda services at Layer 1. It can also be utilised by alien wavelengths.

3.1.2.5 *Summary*

Table 3.1 summarises the differences and features of the three options for GÉANT slicing at Layer 1.

Layer 1 Slicing Option	OEO Regenerator	Range	Service	Layer
Static Lambda without OEO	No	1500 – 2000 km	Photonic + Lambda	0
ROADMs without OEO	No	1500 – 2000 km	Photonic + Lambda	0
ROADMs and OEO	Yes	Unlimited	Lambda	1

Table 3.1: Summary of Layer 1 slicing options

3.2 Network Factory-Enabling Technologies at Layer 2

Layer 2 link slicing can be implemented in GÉANT with the use of the following technologies: IEEE 802.1q VLAN, L2 MPLS VPN and GÉANT Plus circuits.

In addition to the Layer 2 link slicing technologies currently supported by GÉANT, emerging technologies can also be considered: MPLS – Transport Profile (MPLS-TP) and PBB/PBT.

OpenFlow [OF] is a promising network element slicing technology at Layer 2.

3.2.1 Current GÉANT Layer 2 Technology

GÉANT currently supports flexible and efficient provisioning of services at Layer 2 [GN3_DS1.1.1,2], such as point-to-point sub-lambda circuits (referred to as Ethernet Private Lines (EPLs) but also commonly known as lightpaths) with a granularity in the order of 1 Gbps, delivered on GE client interfaces. At the time of writing, GÉANT Plus is the dominant GÉANT Layer 2 service.

GÉANT Plus service instances are usually extended across NREN infrastructures to customer sites within the relevant national territories. They are used where sub-wavelength requirements (currently less than 10 Gbps) exist, with an order lead time of a few working days. The mapping of the client signals to SDH Time Division Multiplexed (TDM) trails is performed using the Generic Framing Protocol – Framed (GFP-F) encapsulation of Ethernet into virtually concatenated groups of VC-4 trails (a “next-generation SDH” concept). There is no facility for packet-oriented statistical multiplexing and limited tolerance to bursty traffic profiles.

In addition to SDH-based Ethernet services, the GÉANT fibre footprint could also support other Layer 2 services using alternative technologies. These technologies include Ethernet over MPLS (EoMPLS), implemented with QoS and TE over high-capacity trunks, Carrier Ethernet based on MPLS-TP or Provider Backbone Bridge Traffic Engineering (PBB-TE) profiles as well as TDM point-to-point links implemented over NG-OTN (using equipment capable of ODU-x switching). For more details about the possible enabling of such technologies on the current GÉANT backbone, see [GN3_DS1.1.1,2].

3.2.2 GÉANT Slicing Options at Layer 2

User requirements studies (such as those influencing [GN3_DS1.1.1,2]) indicate that the majority of users expect 1/10 GE connectivity, with termination ports at the NREN equipment connected to the GÉANT PoPs. Connectivity can be point-to-point or point-to-multipoint. Some use cases require dedicated capacity as a result of slicing; others require a dedicated virtualised network infrastructure composed of both links and nodes.

GE link slicing technologies, as presented below, impose no special hardware or connectivity requirements for the current GÉANT.

3.2.2.1 *802.1q VLAN*

IEEE 802.1q or VLAN tagging is a networking standard developed by the IEEE 802.1 Work Group for creating independent logical networks within a physical network. The standard also allows multipoint topologies to be created. With 802.1q, VLAN-based slicing offers neither traffic isolation nor QoS by default. However, some technologies can be combined with VLANs in order to provide such capabilities. For example, 802.1p can be activated upon request [8021P].

Some problems can arise with VLAN tagging in multi-domain environments. The 802.1ad standard or Q-in-Q solves this by allowing double VLAN tags to be used in an Ethernet frame [QinQ]. In the Network Factory context, Q-in-Q should be activated/supported by default in order to avoid multi-domain conflicts.

The use of VLANs for link slicing within GÉANT would require configuration at the edge Ethernet interfaces and activation of Q-in-Q for carrying multiple slice VLANs over the backbone in a scalable manner.

3.2.2.2 *MPLS-Based Slicing: EoMPLS*

Multi-Protocol Label Switching (MPLS) is a packet-switching technology that allows integration of link layer (Layer 2) switching with network layer (Layer 3) routing [MPLS-EoMPLS]. With MPLS, data is transferred over any combination of Layer 2 technologies, using any Layer 3 protocol. MPLS is a highly scalable, protocol-agnostic data-carrying mechanism.

The main benefits of this technology are independence from any particular data link layer technology and no need for multiple Layer 2 networks to satisfy different types of traffic.

For the Layer 2 slicing scenario, MPLS technology would be used to implement virtual links on top of the GÉANT physical Layer 2/Layer 3 equipment. There are different solutions to implementing Layer 2 virtual links (or L2 MPLS VPNs). One of them is Ethernet over MPLS (EoMPLS).

EoMPLS is a tunnelling mechanism that transports Layer 2 Ethernet frames over an MPLS network. EoMPLS encapsulates Ethernet frames in MPLS packets and forwards them across the MPLS network. Each frame is transported as a single packet and the provider edge routers connected to the backbone add and remove labels as appropriate for packet encapsulation [MPLS-EoMPLS].

A Network Factory exploiting EoMPLS would create an abstraction layer on top of the physical topology where each virtual circuit would correspond to a virtual link. Therefore it would allow the creation of virtual links for different users and the allocation of different slices on top of the same physical equipment. An example of this is shown in Figure 3.4.

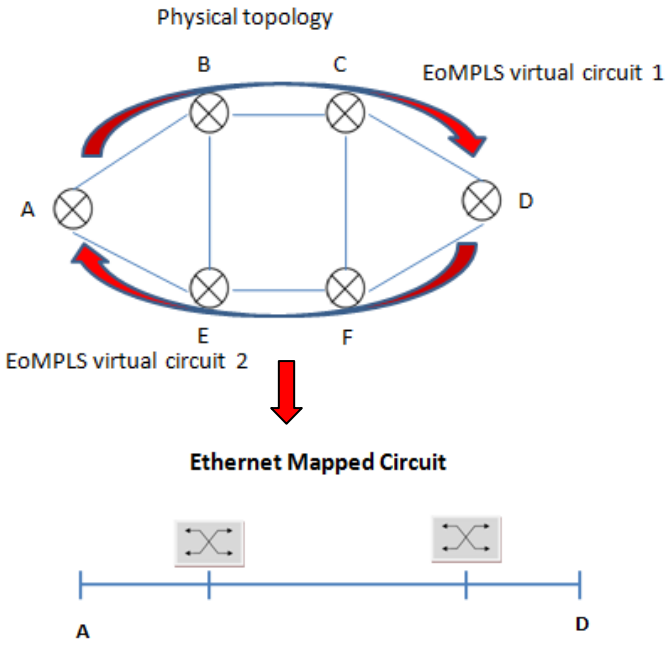

Figure 3.4: EoMPLS schematic

QoS by using three experimental bits in a label to determine the priority of packets and Traffic Engineering (TE) using RSVP-TE, are possible extensions of EoMPLS (see also [GN3_DS1.1.1,2]) as currently supported by GÉANT, so as to provide Layer 2 link slicing with traffic isolation and QoS capabilities.

3.2.2.3 *GÉANT Plus Circuits (EoSDH)*

GÉANT Plus exploits Generic Framing Procedure to encapsulate different Layer 2 technologies (mostly Ethernet over SDH (EoSDH)). Time-Division Multiplexing (TDM) is used to divide the link capacity and therefore link slicing is offered. This technology is currently valid for link slicing at Layer 2. However, it could become unavailable in future because of the decommissioning of the current SDH devices, and would probably be replaced by services based on MPLS or carrier-grade profiles.

3.2.3 Emerging Technologies

The technologies described in this section are not currently available in GÉANT. However, they are emerging standards with unique features that make them attractive for evolution of the network substrate and for the functionalities the Network Factory will provide. In particular, OpenFlow, as an emerging network element

slicing technology, imposes the use of special-purpose switches and computing elements running the protocols that allow the slices to be controlled.

3.2.3.1 *MPLS-TP*

MPLS-TP emerged from IETF and ITU-T to give service providers an environment similar to that of carrier technologies (SONET/SDH or optical transport networks (OTN)) only implemented with MPLS. The result of these efforts is an extension of GMPLS known as MPLS – Transport Profile (MPLS-TP) [MPLS-TPJ, MPLS-TPC]. MPLS-TP builds on the MPLS/GMPLS protocol suite, adding a few extensions to address transport network requirements; the devices supporting MPLS will also be MPLS-TP-enabled after an update of their software. In principle, the common characteristics of the supporting technologies allow the Network Factory to be implemented utilising both IP MPLS and MPLS-TP on the same physical infrastructure, offering both an IP- and MPLS-based VPN substrate and a circuit-based substrate.

The most relevant differences between MPLS and MPLS-TP [MPLS-TP] are:

- Explicit support for bi-directional paths: transport networks commonly use bi-directional circuits, and MPLS-TP also mandates the support of bi-directional LSPs. In addition, MPLS-TP must support point-to-multipoint paths.

- Support for MPLS-only data plane: MPLS nodes usually run IP on all interfaces because they have to support the in-band exchange of control-plane messages. MPLS-TP network elements must be able to run without IP in the forwarding plane. In addition, MPLS-TP network elements have to support out-of-band management over a dedicated management network.

- Data forwarding within an MPLS-TP network element must continue even if its management or control plane fails, similarly to the way high-end routers allow non-stop forwarding. Going a step further, MPLS-TP nodes should be able to work with no control plane, with paths across the network computed solely by the network management system and downloaded into the network elements.

3.2.3.2 *PBB/PBT Carrier Grade Profiles*

A competitor technology to MPLS-TP in connection-oriented Ethernet provisioning is the family of IEEE 802 extensions denoted as Provider Backbone features.

Provider Backbone Bridge (PBB) is an Ethernet data-plane technology invented in 2004 by Nortel Networks. It is sometimes known as MAC-in-MAC because it involves encapsulating an Ethernet datagram inside another one with new source and destination addresses. In PBB, MAC addresses are stacked in a similar way to VLAN ids. Backbone MAC addresses are added to every frame (customer standard MAC addresses are moved into the payload and become transparent), larger VLAN ids are used for service tags (24 bits instead of 12 as used in 802.1Q) and a new id for backbone VLANs is introduced.

Based on PBB, a derived technology has been defined: Provider Backbone Transport (PBT). PBT is a connection-oriented technology and network management architecture. It defines methods to emulate connection-oriented networks by providing connection segments over a packet-switched network. Key data-

plane differences from PBB are related to how a switch's forwarding table is configured and how broadcasting is managed (in order to avoid flooding of frames). PBT has been presented to IEEE802 and a new project has been approved to standardise it under the name of Provider Backbone Bridge Traffic Engineering (PBB-TE) (IEEE 802.1Qay), a modification to PBB. A description of the evolution of these standards with their characteristics is reported in [PBBTE].

Unlike MPLS-TP, PBB-TE only supports Ethernet. The PBB-TE standard will be developed focusing on the following targets:

- MAC learning, spanning tree and similar Ethernet inefficiencies will be replaced with new, more efficient solutions (inspired by IS-IS) to calculate optimal and redundant paths. These features will allow the creation of end-to-end circuits with backbone solutions similar to SDH VCs.

- ITU and IETF standards for resiliency will be adopted in order to define new carrier OAM standards.

Both MPLS-TP and PBB/PBT technologies could be valid Layer 2 link slicing options for a future GÉANT Network Factory implementation.

3.2.3.3 *OpenFlow: Links and Network Element Slicing*

OpenFlow is a technology that allows virtual switching devices to be created. Thus, slicing at the network element level is supported by defining non-overlapping flow tables for every user accessing a network element.

On the top of OpenFlow devices, a controller computing element is expected to control uniformly and collectively the network elements of an OpenFlow infrastructure. In particular, a controller adds and removes flow entries from the flow table on behalf of slice users. For example, a static controller might be a simple application running on a PC to statically establish flows to interconnect a set of test computers for the duration of an experiment. An example of an OpenFlow-based network is shown in Figure 3.5.

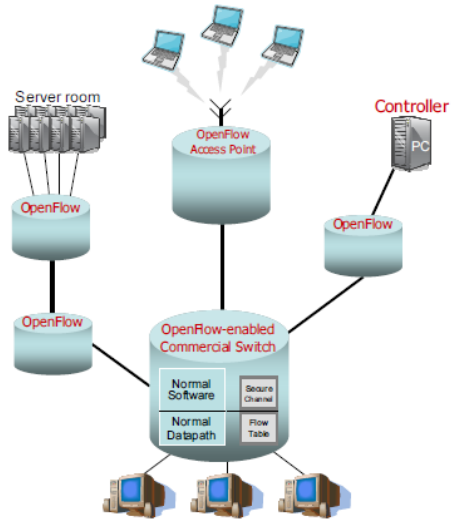

Figure 3.5: Example of an OpenFlow-based network

Slicing of the flow tables can be implemented through the virtualisation of the controller systems. This can be done by deploying a special-purpose hypervisor called FlowVisor [FV]. FlowVisor hosts multiple guest OpenFlow controllers, one controller per slice, making sure that a controller can observe and control its own slice, while isolating one slice from another (both the data path traffic belonging to the slice, and the control of the slice). An example of controller slicing is shown in Figure 3.6.

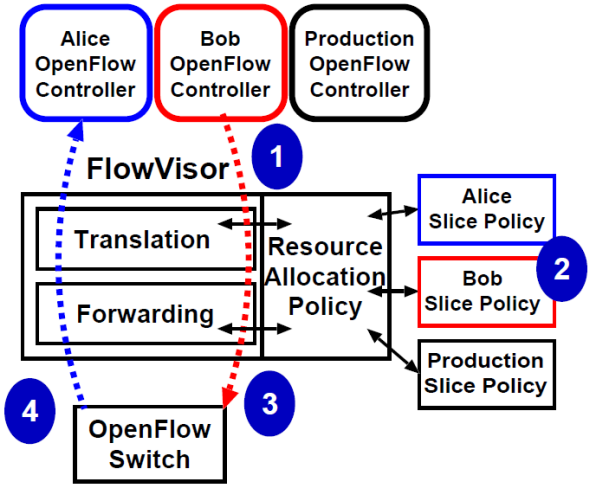

Figure 3.6: Flow tables virtualisation through FlowVisor

In this example, every virtual controller (1) transparently messages each user's slicing policy (2) to the devices and rewrites (3) the flow tables. Messages from the switches (4) are only forwarded to guests if they match their slice policy.

An OpenFlow-enabled Network Factory would require a substrate of network elements enabled to support the specification. Currently the device families that support the OpenFlow v1.0 specification are HP Procurve 5400zl, NEC IP8800, and Toroki Lightswitch. For a list of the available implementations see [Components]. As an alternative, software solutions like the minimal OpenFlow Linux deployment or the more complete Open vSwitch layer can be adopted.

OpenFlow can be mixed with the Layer 2 connectivity slicing technologies already discussed (see [OFMPLS, OFSwitch]): since OpenFlow switches are equipped with 1 Gbps and 10 Gbps ports, any Layer 2 connectivity slicing technology delivering Ethernet links could be adopted by a Network Factory solution to extend its functionalities with OpenFlow switch virtualisation. For testing of protocols over a virtual network, any link slicing technology can be used. For advance applications, which need high capacity or other QoS characteristics, a deterministic link slicing technology must be used, such as EoSDH or lambdas.

In detail, the way to interconnect OpenFlow-enabled network elements through Layer 2 link slices depends on the underlying network topology: every Layer 2/Layer 3 network element could be paired with an OpenFlow switch.

For example, by exploiting Q-in-Q features, OpenFlow-handled traffic can be isolated from production traffic over a backbone, with OpenFlow switches communicating through point-to-point Layer 2 logical links. Such a configuration creates an overlay (a logical topology) over the physical topology, with the OpenFlow switches connecting to Layer 2/Layer 3 devices with a hard loop in the form of a VLAN trunk. A similar configuration has been deployed by JGN2+ in Japan [OFJap].

Similarly, the OFELIA project [OFELIA] interconnects OpenFlow switches with 1 G Ethernet links, separating the testbed islands from the production traffic across partners using L2 VPNs.

An example of how OpenFlow switches could be integrated in Network Factory PoPs is shown in Figure 3.7.

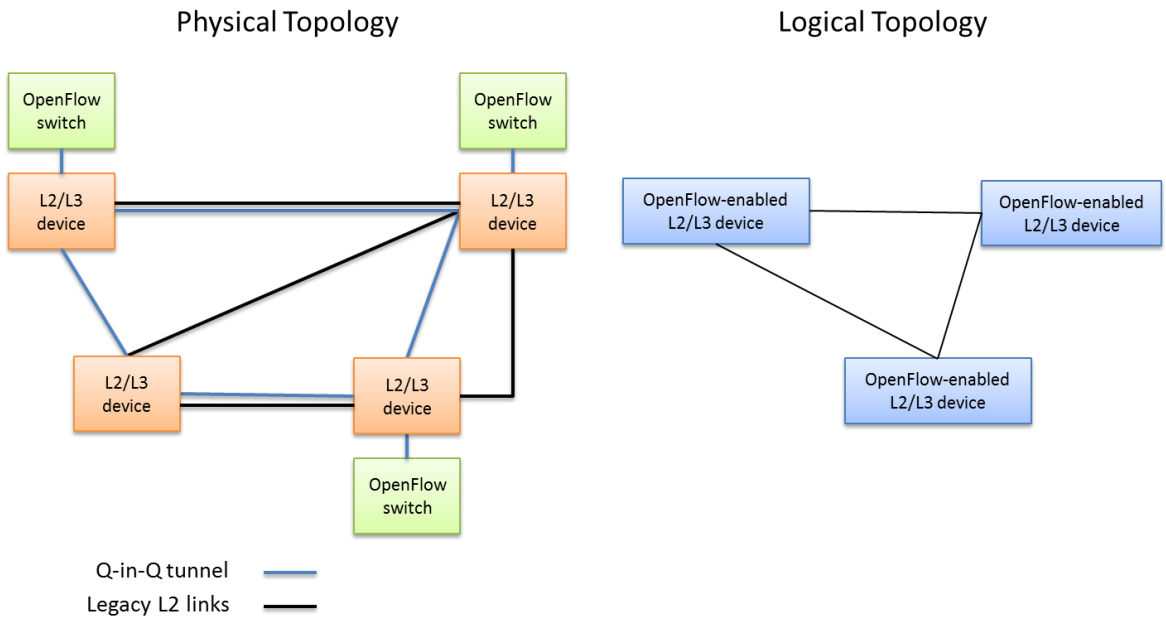

Figure 3.7: Example integration of OpenFlow switches in Network Factory PoPs

Software implementations and FPGA kernels seem to be mature enough to be effectively used in virtual infrastructures. In particular, users could instantiate OpenFlow virtual machines in their slices, controlling them from their home institutions through controller tools like NOX [NOX]. For the users requesting an efficient data plane based on OpenFlow, NetFPGA cards could be reserved by the slices [OFNetFPGA].

A different approach to software-based OpenFlow would be the instantiation of a permanently distributed virtual switch using Open vSwitch [OpenvSwitch]. Again a FlowVisor server could help to share the Open vSwitch service among different users. Being purely software-based, the instantiation and the management of the virtual switch partitions could be automated and exposed to the users along IaaS principles. Performance issues in flow forwarding due to the software implementation should be carefully investigated for this solution.

3.2.4 Summary

Table 3.2 below summarises the attributes of the Layer 2 slicing technologies described in the previous sections.

Technology	Topology	Implementation/End-User Interfaces	Capacity	Traffic Isolation/QoS
802.1q VLAN, IEEE 802.1ad	Multipoint	Ethernet interface on the user end-point, Q-in-Q should be activated by default, no auto-negotiation to	Any value up to the capacity of the physical 1–10 GE ports involved. Parameters along	No isolation, no QoS by default. 802.1p can be activated upon request.

Technology	Topology	Implementation/End-User Interfaces	Capacity	Traffic Isolation/QoS
		avoid issues. Large MTU can be configured. Backup only through spanning tree.	the circuits are under provider control.	
EoMPLS	Multipoint	Resiliency/backup may be implemented. Unknown time for implementation: should be short for L2 VPNs.	No capacity slicing/guarantees. However, MPLS-TE can be used to divert traffic to less loaded links.	MPLS VPNs provide traffic isolation. EoMPLS with QoS and TE optimise isolation and quality.
EoSDH	Point-to-point circuits between GÉANT PoPs	Resiliency is technically available (through SDH), but not supported because of lack of interest. It does not require any additional hardware. The implementation time for circuits is short (around 5 working days).	Scales from 155 Mbps (SDH) to 10 Gbps (SDH/Ethernet), flexibly in increments of 155 Mbps. 1 Gbps circuits are the ones mostly used.	Isolation and guaranteed capacity.
MPLS-TP	Multipoint	MPLS-TP in principle provides the same resiliency level as IP/MPLS. In addition, MPLS-TP does not require IP for the control plane.	Should scale according to the capacity of the substrate.	There are no specific details on QoS support. See MPLS and EoMPLS capabilities.
PBB/PBT	Multipoint	PBT/PBB is an Ethernet-only technology; it is expected to reuse the same backup solutions.	Should scale according to the capacity of the substrate.	PBB enhances VLAN; the same level of isolation & QoS is expected.
OpenFlow	Layer 2 switch slicing; point-to point; multipoint	OpenFlow is an open protocol for programming the flow tables on different switches, routers with	No real capacity slicing abilities as this is a network element slicing technology.	Traffic isolation is achieved by network element slicing.

Technology	Topology	Implementation/End-User Interfaces	Capacity	Traffic Isolation/QoS
		Ethernet interfaces.		

Table 3.2: Summary of Layer 2 slicing technologies

3.3 Network Factory-Enabling Technologies at Layer 3

In a Layer 3 slice, users expect interconnected (through Layer 2 and/or Layer 3 technologies) routers and client interfaces to connect to. At present, GÉANT does not offer routing resources that could be dedicated to users, as security imposes a number of challenges to the separation of production and user-managed routing resources on the same physical infrastructure.

User requirements for a Layer 3 Network Factory are expected to include guaranteed link capacity (unaffected by production traffic on shared physical resources) and a static topology (not relying on current operational routing tables). Thus a GÉANT Network Factory at Layer 3 should include logical systems/routers with independent routing tables and configurations dedicated to each slice. A user should have privileges to alter a slice's routing configuration, queuing and forwarding features as well as to install his own processes to analyse packets, modify routing procedures, modify routing engines, and process packets.

End users may be interested in performing network-related experiments within the Layer 3 slice, or in having a virtual Layer 3 transport network between distant locations so that they can conduct application experiments. In the latter case, it should be possible to fully control user access to network elements and links, as slice configuration management is totally administered by the Network Factory services provider.

3.3.1 Current GÉANT Layer 3 Technology

The GÉANT IP service is designed to provide a robust high-capacity solution to the international connectivity requirements of the majority of academic users. It provides a resilient service in the case of hardware failure or fibre cuts, and uses advanced routing equipment to ensure fast recovery from unexpected events. The features of GÉANT IP include [GÉANTIP]:

- IPv4: The GÉANT network provides transit to all IPv4 unicast packets to and from connected NRENs and towards international partners.
- IPv6: GÉANT forms part of the world's first next-generation Internet network and many of GÉANT's NRENs already provide IPv6.
- Virtual private networks: Multi-Protocol Label Switching (MPLS) is the technology framework used by GÉANT to set up VPNs over the IP network, providing a pre-determined protected route over multiple network domains.
- Multicast: In GÉANT, data traffic with the appropriate multicast group destination will be recognised and treated accordingly.

3.3.2 GÉANT Slicing Options at Layer 3

3.3.2.1 *IP and GRE Tunnels*

Implementing a Network Factory with Layer 3 connectivity introduces some constraints that are typical for IP networks. Connectivity between routers and end points can be implemented by setting a public IP address and relying on GÉANT and Internet routing capabilities. However, this provides no isolation and users will be aware of existing nodes/routers which are not under their control (GÉANT or third-party routers which forward the traffic).

A small upgrade to this is to define Generic Routing Encapsulation (GRE) tunnels between routers and between routers and end points. These emulated one-hop links will provide Layer 3 connectivity with two additional features:

- User traffic is logically isolated from other traffic, and nodes are hidden from user control.
- Both public and private IP addressing schemas can be used, which are isolated from GÉANT and Internet traffic.

The GRE concept is depicted in Figure 3.8.

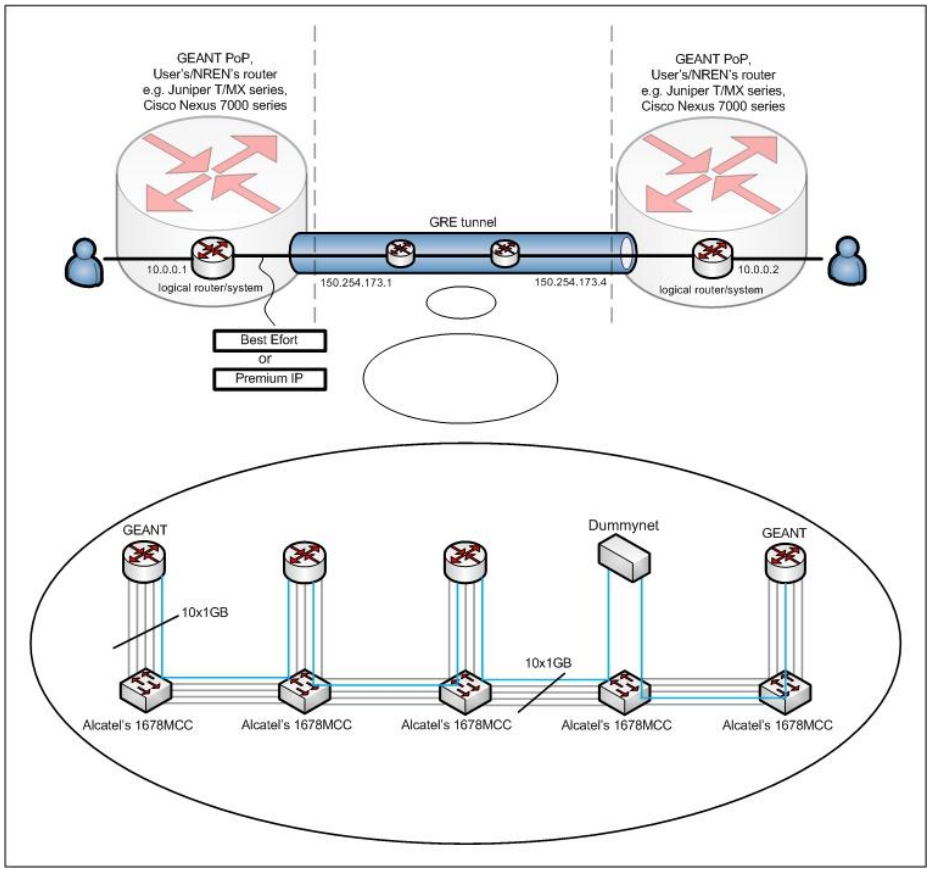

Figure 3.8: GRE tunnelling for the Layer 3 Network Factory

The GRE tunnel hides the real Layer 3 topology from the user, simplifying the network and delivering only what the user requested.

Both approaches have limited functionality as they are subject to IP network constraints, including unpredictable capacity, jitter, reordering, no control over traffic paths, traffic congestion and race conditions. However, this environment is close to the current GÉANT offerings and may be a useful test ground for some end users and their applications.

3.3.2.2 *Link Slicing with Layer 1 and 2 Technologies*

Users may request protected guaranteed-capacity links interconnecting Layer 3 equipment within their slice, isolated from other traffic, and also with constant delay characteristics and lack of reordering. In this case, Network Factory links would have to be delivered using Layer 1 or Layer 2 link slicing technologies as previously presented. Circuits would be terminated at the Network Factory Layer 3 slice-supporting routers with a properly configured internal stitching to the specific logical router of each slice.

3.3.3 Emerging Technologies

The technologies described in this section – hardware-based logical routers and servers hosting software-based routers – are not currently available in GÉANT. However, they are emerging and particularly relevant to Network Factory solutions at Layer 3.

3.3.3.1 *Hardware-Based Logical Routers*

For Layer 3 Network Factory service delivery, hardware-based logical routers are a network element slicing solution. A logical router may behave as an independent routing daemon within a physical router, with a separate, unique copy of the routing protocol process (RPD) and forwarding tables. Crashes of a single user's routing process would have no impact on other operational router instances. Depending on the logical implementation, isolation and management capabilities vary. A further advantage of such a solution is that users can access and configure a particular logical system that is assigned to the user slice. Logical routers may have hardware separation at the data plane but may share the same forwarding information base (FIB) resources, so that scalability of prefixes or next hops is still an issue [LogicSys].

Vendor solutions with logical router capabilities, with different properties and limitations, include:

- Juniper Networks' Inc. M and MX series.
- Cisco Systems' Inc. CRS-1 and ASR 9000 routers.

Protected System Domains (PSDs) [PSD], compatible with Juniper T-series routers, allow both data plane and control plane separation between logical routers and dedicate a separate routing engine (CPU complex, route processor) to each logical router. One of the biggest advantages of this solution is that each logical router has its own scalability and is also a discrete managed entity. For instance, a user can upgrade the OS version in one logical router, while the others remain unaffected.

Virtualisation may not be limited to the data or control plane, but may also involve the management plane and software/hardware component partitioning. It can be implemented in such a way that groups of dedicated software processes with dedicated hardware are delivered to the end user in the form of a slice to provide virtualised control and data planes within an independent management context. This provides true segmentation of network traffic, context-level fault isolation, and management through the creation of independent hardware and software partitions. Each configured logical router presents itself as a unique device to connected users within a physical switch. The logical router runs as a separate logical entity within the chassis, maintaining its own unique set of running software processes, having its own configuration, and being managed by a separate administrator [VArch]. The applicability of such solutions for a backbone-based Network Factory should be further investigated, subject to their availability as an outcome of the ongoing procurement activities for the next-generation GÉANT.

3.3.3.2 *Servers Hosting Software-Based Routers*

Software-based network element slicing at Layer 3 can be supported by servers hosting software-based routers. Such servers would be equipped with a high number of cores (dual 8x or more) and large RAM availability (32 GB or more) as well as open source hypervisor OSs, such as Xen [Xen] or KVM, to run software routers like Quagga or XORP. In addition, to simplify management, only predefined software router templates should be provided (e.g. Quagga routers with one core and a fixed amount of memory (2 GB, 4 GB or 8 GB) according to the size of the slice).

The number and type of physical network interfaces in servers hosting software-based routers would be an important parameter for a Network Factory infrastructure. Virtual routers with no quality of service support could be sharing the same physical NIC, leaving the management of the available capacity to the hypervisor. Exclusive usage of physical NICs could be offered for virtual routers with performance requirements. For slices that need both high performance and flexibility at Layer 2, the servers should permit the integration of reconfigurable NICs (like NetFPGAs [NetFPGA]) in order to validate new routing protocols on efficient devices.

An indicative allocation of physical interfaces on servers offering Network Factory software-based routers would be:

- 8 – 12 Ethernet cards for experimental purposes (1 G or 10 G depending on the PoPs).
- 1 x NetFPGA card with 4 x 1 Gbps (or 4 x 10 Gbps) ports.

In addition to routing network elements, it is possible to deploy a dummynet node [Dummy], which can alter the characteristics of real traffic by introducing delay, packet drop according to a predefined pattern, reordering, etc. A dummynet is usually deployed as an OS which forwards traffic somewhere in the slice, i.e. a software router. Users may want to have access to this box to be able to configure their own network-disturbing schemas, or can request a pre-configured box, hidden somewhere in their topology. The visibility of the box to the end users depends on the users' requirements. The only issue is forcing traffic to pass through such a node, which in practice means that the user end points should be connected directly to the dummynet boxes, or that Layer 2 MPLS circuits should be configured to use pre-defined routing.

3.3.4 Summary

Table 3.3 summarises the attributes of the Layer 3 link slicing technologies described in the previous sections.

Layer 3 Slicing Option	Topology	Implementation/ End-User Interfaces	Capacity	Traffic Isolation/QoS
IP and GRE tunnels	Multipoint	Ethernet interface at the user end point.	Unpredictable capacity, jitter, reordering, no control over traffic paths, traffic congestion and race conditions.	No isolation and users will be aware of existing nodes/routers. GRE tunnels logically isolate user traffic from other traffic, and hide nodes from user control.
Differentiated IP services	Multipoint	Ethernet interface on the user end point.	Packets transferred with priority over the allocated capacity.	Guaranteed level of network performance, no guaranteed service attributes, claims only to provide lower loss, delay and jitter.
Implementation of Layer 3 links with GÉANT Layer 1 and 2 services	Point-to-point circuits between GÉANT PoPs	Ethernet interface on the user end point.	Guaranteed capacity can be requested depending on the available interfaces. It is also possible to deliver higher capacities with 40GE/100GE interfaces deployed in the GÉANT network.	Protected guaranteed quality links offering constant delay and lack of reordering.
Hardware-based logical routers	Layer 3 network element slicing; multipoint	Delivering dedicated physical interfaces to the client side and shared physical interfaces at the line side.	Guaranteed capacity is possible, depending on the implementation.	Isolation is possible, depending on the implementation.

Layer 3 Slicing Option	Topology	Implementation/ End-User Interfaces	Capacity	Traffic Isolation/QoS
Servers hosting software-based routers	Layer 3 network element slicing; multipoint	Both sharing and exclusive usage of physical NICs by users.	Possible through the exclusive usage of physical NICs by virtual routers with performance requirements.	Possible through the exclusive usage of physical NICs by virtual routers with performance requirements.

Table 3.3: Summary of Layer 3 link slicing technologies

3.4 Monitoring Functionality

A framework like perfSONAR should be adopted and engineered to monitor Network Factory slices, regardless of the software/hardware solution or Layer 2/Layer 3 technologies used.

Monitoring should include current and historical utilisation and quality of service metrics for Network Factory links and network elements. The exact engineering of a monitoring solution depends heavily on the technical choices made for the implementation of the Network Factory infrastructure as analysed throughout this chapter and is a subject of future work.

4 Foundations of a Network Factory Business Case

4.1 Introduction

This chapter establishes the foundations of a business case for a GÉANT Network Factory infrastructure and services. Due to an evolving technology matrix, the result of the ongoing migration from the current GÉANT architecture to the next-generation backbone, indicative scenarios are given as different implementation options for the business case rather than an exhaustive list. Technical aspects of the different options under evaluation are based on the relevant technologies as presented in Chapter 3 Network Factory-Enabling Technologies for GÉANT.

Financial assessments refer to the deployment and operational costs of each Network Factory realisation scenario and are limited to the identification of the relevant assessment factors. The risk assessment attempts to identify, classify and control the risks associated with the Network Factory initiative overall but also the indicative scenarios in particular. The results of these assessments should not be regarded as the final outcome; rather, they are a basis for moving on to the production of specialised business case outputs and to undertaking the study and design phases of the Network Factory solution for GÉANT, incorporating the outcomes of the GÉANT backbone evolution process.

It is important to stress that the indicative scenarios presented here refer to a single-domain Network Factory solution over GÉANT, in which end users are expected to utilise the Network Factory facilities through remote connections from their local environment. This approach has been followed not only to ensure the brevity of the scenarios presented but also as a first step in an incremental process to progress from a single-domain to a multi-domain Network Factory involving resources from NRENs, for which technical (such as features supported) and business aspects (such as operational requirements) will need to be further investigated.

The chapter is divided into four sections:

- Overall strategic fit of a GÉANT Network Factory infrastructure and services.
- Selected Network Factory scenarios – an overview of indicative Network Factory implementation scenarios identifying factors for financial assessment (CAPEX, OPEX).
- Risk assessment and analysis – identification and classification of risks for a GÉANT Network Factory realisation in general but also of the specific indicative scenarios put forward.

4.2 Strategic Fit

The hybrid infrastructure from which the GÉANT network is built is capable of creating logical and physical networks that can be considered independent of the production infrastructure, but that share its physical elements. At the same time, researchers from various scientific disciplines around Europe require network resources and functionality that do not fit within the standard production-level GÉANT service portfolio.

The goal of the GÉANT Network Factory infrastructure and services design and specification is the delivery of a network environment where experiments can be conducted without affecting production services, giving researchers complete control of the resources allocated, allowing them to experiment on top of the GÉANT backbone and utilise its resources. Not only current needs, but also the trend among R&E infrastructures globally of delivering similar facilities and/or services, make the present time appropriate for the Network Factory work in GÉANT.

By establishing the foundations of a business case for a GÉANT Network Factory and services, it is expected that the GÉANT Service Area will be enriched with offerings focused on and tailored to the needs of the European R&E community, by contributing visibly and significantly to research in different disciplines of science and to emerging applications.

At the time of writing, only general critical success factors for a Network Factory business case can be defined, including (as a non-exclusive list):

- Compatibility with the installed infrastructure and services.
- SLAs and quality of service offerings.
- Timely specification and deployment.
- Number of research/experimentation initiatives supported when in production.

4.3 Selected Network Factory Scenarios

In the absence of concrete outcomes as to the link and network element slicing technologies that will be provided by the next-generation GÉANT, two indicative scenarios for a GÉANT Network Factory implementation are presented in this section, so that the foundations of a business case can be established. For each of the two scenarios, the GÉANT backbone is expected to provide physical resources and functionality as needed, namely Network Factory traffic transport as well as control and management. The Network Factory service itself is expected to support:

- Description of virtual infrastructure (slice) requirements from the user.
- Announcement of Network Factory infrastructure and service access points to potential users.
- Provisioning of a set of management tools for users to control the behaviour and the characteristics of their allocated slices.
- A financial model to compensate for the use of physical resources supporting the Network Factory.

4.3.1 OpenFlow-Based Network Factory

The OpenFlow-based Network Factory is envisaged to provide logical Layer 2 circuits and virtual switching elements to users for implementation of their slices, as well as hosting of user controllers when requested.

It is envisaged that a substrate hosting the OpenFlow-based Network Factory will be implemented over a subset of the GÉANT PoPs. OpenFlow switches will be deployed within or next to GÉANT PoPs and interconnected utilising Layer 2 link slicing over GÉANT. An example of such a configuration is provided in Figure 4.1.

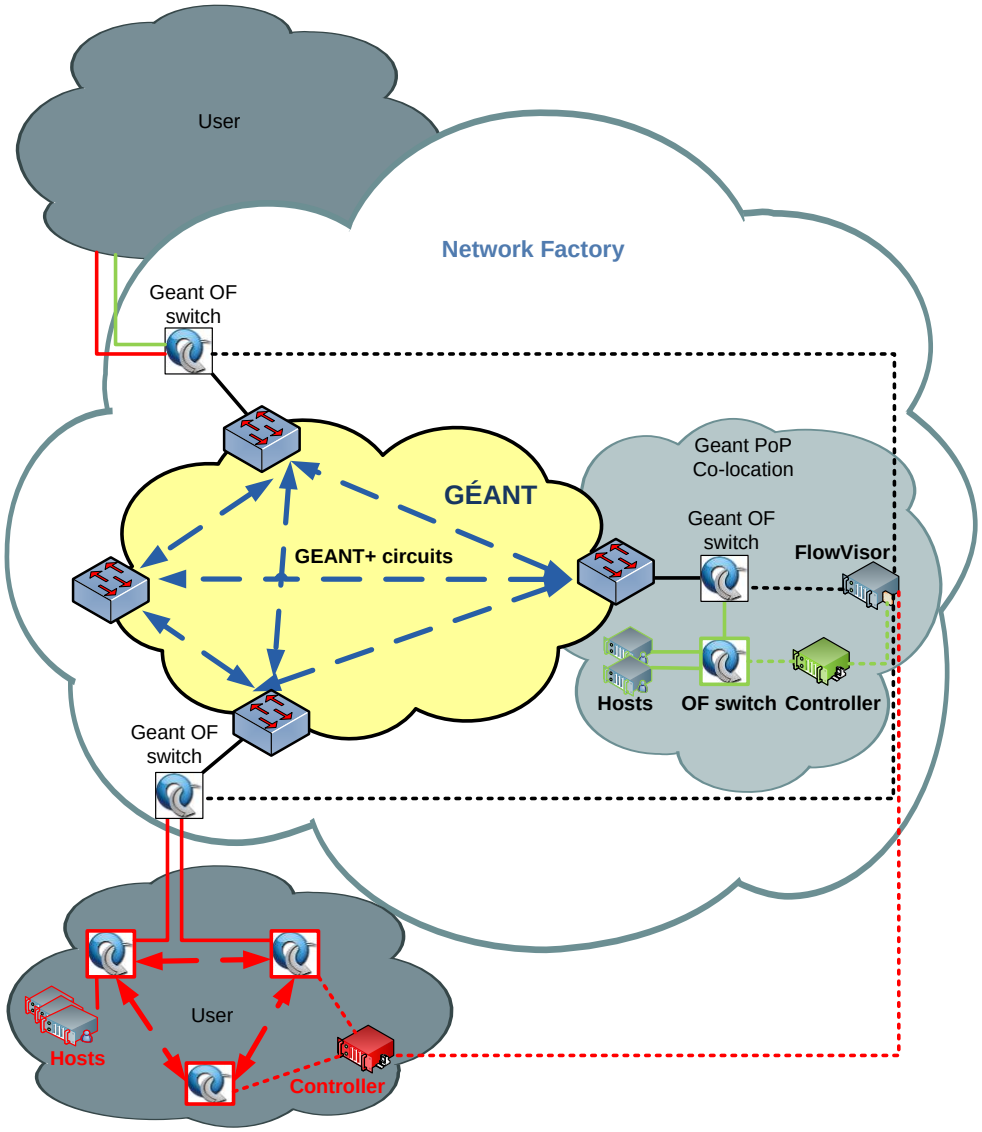

Figure 4.1: OpenFlow-enabled Network Factory scenario architecture

At the time of writing, Layer 2 link slicing for supporting the OpenFlow deployment on top of the GÉANT backbone can be implemented with GÉANT Plus circuits. However, this is almost certain to change in the next-

generation GÉANT. The availability of 10 Gbps circuits would be particularly appealing for Future Internet and other communities that rely on high-bandwidth data transfers (for example, user groups from the arts, including live HD music performances). EoMPLS can be seen as a viable future solution. It is also important to plan so that forthcoming releases of OpenFlow with new features (v1.1 with MPLS tag handling and v1.2 with support for IPv6 and PBB) can be supported by the selected Layer 2 link slicing technology in GÉANT. Depending on the technologies selected for implementing the OpenFlow substrate on top of GÉANT, the number of physical ports reserved at GÉANT PoPs will differ. These are all open issues related to different implementation options for the business case.

With regard to the OpenFlow switching infrastructure, at the time of writing both hardware (see Table 4.1) and software implementations available support OpenFlow v1.0 specification only.

Model	Characteristics
NEC Univerge PF5240	L2/L3 edge switches with 48 10/100/1000 ports + 4 1000/10000 ports in compact 1U form factor.
Pronto 3290	48 x 1 Gbps ports, 4 SPF+ ports, bootable CF card slot, OpenFlow support provided by either of the following 2 firmware: Pica8 firmware or Stanford Reference Software.
Pronto 3780	Pure OpenFlow switch, with no legacy features, 48 x 10 Gbps SFP+ ports.
HP Procurve 5400zl series	Extendable up to 288 x 1 GE ports, and up to 48 x 10 GE ports (12 x 4-port modules).
HP Procurve 6600 series	Fixed 24 (or 48) x 1 Gbps ports and 4 GBIC, or 24 x 10 GE ports (24XG box).

Table 4.1 Hardware-based OpenFlow v1.0 specification support

The availability of implementations of OpenFlow v1.0 specification and subsequent versions at the time of designing and deploying an OpenFlow-based GÉANT Network Factory is an open issue, affecting also the OpenFlow substrate implementation details, slicing options and functionality delivered to users. A major design consideration is whether the switching architecture should be hardware-based or software-based (deployed on servers). Despite the implementation of OpenFlow switches, controllers and virtualisation of controllers through a FlowVisor for the isolation of user slices in the control plane will have to be server-based.

Capital expenditure (CAPEX) for a Network Factory implementation scenario based on OpenFlow is heavily dependent upon the available specification implementations at the time of deployment and the choices made. In the case of hardware-based OpenFlow switches, the cost of purchasing and maintenance has to be anticipated. In the case of software-based OpenFlow switches, equipment costs relate primarily to servers with virtualisation capabilities for hosting software switches. Operating expenditure (OPEX) is also dependent upon the implementation choices made and cannot be estimated at the time of writing.

Finally, deployment planning of an OpenFlow-based Network Factory infrastructure is again dependent upon the solution selected. A hardware-based deployment must anticipate ordering and delivery periods of OpenFlow-compliant network equipment, while a software-based deployment is anticipated to have shorter delivery periods for servers but longer setup and configuration periods.

4.3.2 Network Factory Delivering Layer 3 Slices

A Network Factory offering Layer 3 slices to users can be implemented based on vendor-supported logical/virtual router functionality. For the scenario presented here, logical routers and link slices at Layer 2 (e.g. utilising EoMPLS) are delivered to users for implementing their slices.

In this scenario, the Network Factory substrate will be implemented over a subset of the GÉANT PoPs. At the time of writing, deploying dedicated vendor solutions (routers) with logical router capabilities next to GÉANT PoP production equipment is considered a realistic approach (see Figure 4.2). Depending on the outcomes of the evolution of the GÉANT backbone, and also of a thorough business, technical and operational evaluation following that, logical router functionality for the Network Factory could, in the future, be delivered by production network elements (with Layer 2/Layer 3 functionality) of the GÉANT backbone.

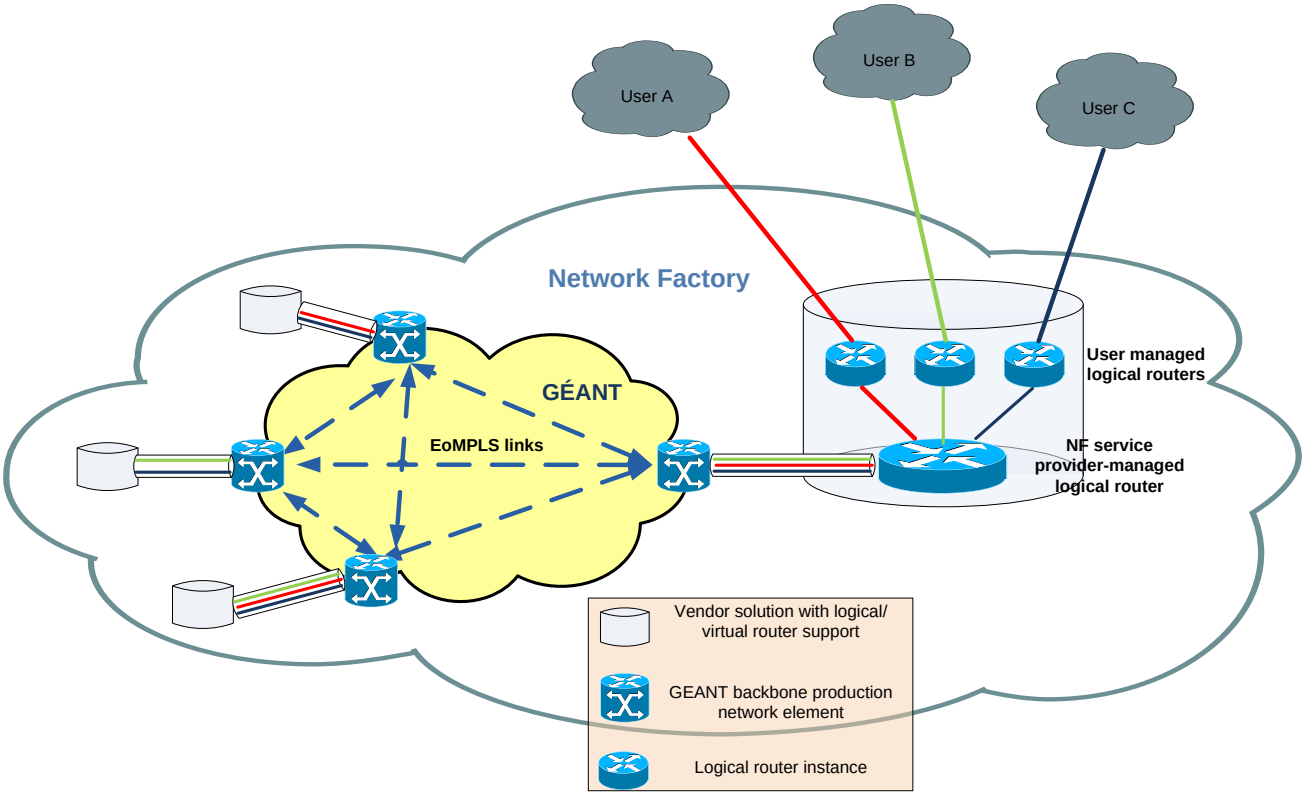

Figure 4.2: Layer 3 services Network Factory scenario architecture

The generic architecture presented in Figure 4.2 is only indicative. Different configurations and capabilities for service offerings and management depend upon the vendor-specific support for slicing of physical routers into logical/virtual instances. Differences concern both the functionality and management of logical routers within a physical router platform (e.g. configuration, allocation, traffic separation) and also the management of L2 VPNs interconnecting logical routers within the Layer 3 slice of a single user.

Capital expenditure (CAPEX) for this Network Factory implementation scenario is based heavily on the specific vendor platform selected. At the time of writing, logical/virtual router solutions are constantly evolving; a proper

CAPEX estimate therefore depends on the offerings available at the time of implementation as well as the scale of the Network Factory infrastructure (e.g. number of physical PoPs, uplink capacity to the GÉANT backbone). The solution selected to segregate traffic among slices within the substrate – i.e. physical versus logical (the case of Figure 4.2) – affects costs in terms of the physical interfaces that will have to be deployed.

Different substrate management and service provisioning implications are expected to influence the CAPEX only minimally; however they will affect OPEX to a variable extent, depending on how much of the substrate will be user-controlled. For example, subject to the platform features offered, it could be possible for users to expand or reduce the physical distribution of logical router instances within their slice themselves, thus minimising the substrate management overhead.

Deployment planning of a Network Factory able to provision Layer 3 slices is quite predictable (assuming an operational GÉANT backbone), as it involves procurement, ordering and deployment of state-of-the-art vendor-supported hardware in the periphery of GÉANT and incremental extensions for interconnections at the GÉANT PoPs.

4.4 Risk Assessment and Analysis

Table 4.2 summarises the overall, generic risks associated with the Network Factory business case.

Risk Identification			Risk Assessment		Risk Control	
Id	**Name**	**Description**	**Probability**	**Impact**	**Avoidance**	**Reaction**
G-R1	GÉANT backbone procurement will not deliver equipment that supports Network Factory functionality.	The backbone equipment under procurement will not support logical/virtual link/element provisioning, such as: • Isolated access to virtual hardware instances, without affecting production services. • Management capabilities for virtual hardware configuration, e.g. configurable resource sharing among virtual instances.	Low	High	• Keep track of procurement process and hardware specifications. • Influence procurement requirements /specifications with Network Factory requirements.	• Redesign/adapt the Network Factory architecture. • Deploy a Network Factory infrastructure that is parallel to the production environment.
G-R2	Network Factory operational procedures will not be implemented within GN3 lifetime.	Operational procedures for the Network Factory infrastructure and services impose an overhead on existing GÉANT operations. Lack of implementation of such operational procedures will result in service unavailability.	Medium	Medium	• Design the Network Factory operational procedures in consultation with Operations. • Keep procedures and service maintenance as simple as possible. • Raise awareness/ stress the importance within Operations of the service, its	• Secure resources necessary for the implementation and support of the operational procedures for the Network Factory. • Support Operations in the implementation of operational procedures.

Risk Identification			Risk Assessment		Risk Control	
Id	**Name**	**Description**	**Probability**	**Impact**	**Avoidance**	**Reaction**
					requirements and procedures.	
G-R3	Service operations and maintenance will be manual to a large extent.	Depending on the implementation scenario adopted, it is expected that the Network Factory service operations will involve manual interventions to a significant extent, as none of the solutions is off-the-shelf. It will be unrealistic to expect full automation of the service and short service delivery times in the early stages of service delivery.	High	Medium	Evaluation of implementation scenarios should consider implications on service provisioning as a high priority.	At the options evaluation and design phases, identify automation optimisation opportunities.
G-R4	OPEX estimates (see sections 4.3.1 and 4.3.2) are insufficient.	The variety of alternative solutions in combination with a rapidly evolving environment in terms of technologies and capabilities offered are challenging accurate OPEX estimates, especially due to the highly differing operational requirements of different solutions.	Low	Medium	Automation of procedures will decrease OPEX.	Perform a further analysis of implementation scenarios and OPEX implications.
G-R5	CAPEX estimates (see sections 4.3.1 and 4.3.2) are insufficient.	The variety of alternative solutions in combination with a rapidly evolving environment in terms of technologies and capabilities offered in general but also in particular to the next-generation GÉANT are challenging accurate CAPEX	Low	Medium		Perform a further analysis of implementation scenarios and CAPEX implications

Risk Identification			Risk Assessment		Risk Control	
Id	**Name**	**Description**	**Probability**	**Impact**	**Avoidance**	**Reaction**
		estimates.				
G-R6	The Network Factory service will not satisfy user needs.	User requirements survey and analysis have not been addressed at all as part of the Network Factory service strategy and design phases. This endangers the extent to which users will value the service, as well as the actual use of it, and means it may not be competitive compared with other offerings.	Medium	High	Begin a user requirements survey and analysis work item within GN3 specific to the Network Factory service as high priority.	Examine existing use cases globally and put forward suggested use cases relevant to the GÉANT user community.
G-R7	Lack of uptake of the Network Factory business case.	Lack of engagement with user communities in early stages and of awareness-raising activities will result in low uptake and thus in low return on investment.	Medium	Medium	Raise awareness of the GÉANT Network Factory initiative among user communities.	

Table 4.2 Overall Network Factory business case risk analysis

Table 4.3 provides a risk assessment of the OpenFlow-based Network Factory scenario.

Risk Identification			Risk Assessment		Risk Control	
Id	**Name**	**Description**	**Probability**	**Impact**	**Avoidance**	**Reaction**
OF-R1	OpenFlow hardware	Vendor roadmaps for solutions with	Low	High	Software	Where delivery of

Risk Identification			Risk Assessment		Risk Control	
Id	**Name**	**Description**	**Probability**	**Impact**	**Avoidance**	**Reaction**
	solutions with enhanced features will not be available.	support for emerging versions of the OpenFlow specifications are not guaranteed.			implementation options are also under consideration. Software implementation will provide the necessary functionality, but is likely to increase the impact of risk OF-R4.	OpenFlow-compliant hardware switches with required functionality is uncertain, software implementations can be adopted.
OF-R2	Layer 2 slicing options at the substrate are limiting.	Depending on the OpenFlow specification available and adopted, the Layer 2 slicing at the GÉANT backbone can impose limitations on the features offered by the Network Factory substrate.	Low	Medium	Current version 1.0 of the OpenFlow specification supports only VLANs. New versions of the specification (v1.1 is under test and v1.2 is under definition) will allow a wider range of slicing technologies like MPLS to be used.	OpenFlow-based Network Factory design will include a migration plan for forthcoming versions of the OpenFlow specification.
OF-R3	OpenFlow protocol support does not provide the required functionality.	Hardware switches with OpenFlow support do not provide fully the necessary functionality, as specified in OpenFlow specifications.	Medium	Low	Specific roadmaps for hardware switches with OpenFlow support should be requested from vendors.	OpenFlow functionality can be supported by extending software implementations.

Risk Identification			Risk Assessment		Risk Control	
Id	**Name**	**Description**	**Probability**	**Impact**	**Avoidance**	**Reaction**
OF-R4	The implementation does not demonstrate high performance.	Performance issues associated with hardware and software implementations of the OpenFlow specification are not extensively documented.	Medium	Medium	Performance goals for the Network Factory OpenFlow-based scenario need to be precisely defined and investigated.	Initial scenario analysis includes preliminary performance evaluation of the different implementation options (hardware-based, software based on physical servers, software based on virtualised servers).
OF-R5	Generic Network Factory use cases with high impact are not supported by an OpenFlow-based implementation.	Lack of user requirements survey and analysis puts the impact of an OpenFlow solution at risk.	Medium	Low	See risk G-R8.	Based on the results of the user requirements survey and analysis proposed in G-R8, a thorough evaluation of the OpenFlow solution will be performed.

Table 4.3 OpenFlow-based Network Factory business case risk analysis

Table 4.4 provides a risk assessment of the scenario of a Network Factory delivering Layer 3 slices.

Risk Identification	Risk Assessment	Risk Control

Id	Name	Description	Probability	Impact	Avoidance	Reaction
L3-R1	Implications of L3 slice creation for isolation and security.	L3 slice delivery presents a number of challenges in terms of isolation and security when coexisting with a L3 production environment.	Low	High	• The scenario presented here, where logical router functionality is delivered by different physical network elements to those offering L3 production services, is expected to minimise implications. • Care should be taken to assure isolated and secured access from non-management networks (i.e. by use of proxies, firewalls, etc.).	Elaboration of the technical aspects of the relevant business case should focus on a detailed specification of isolation, security and user access solutions.
L3-R2	Lack of resource utilisation management in virtual router solutions offered.	Virtual router technology solutions examined have no strict resource utilisation policy, so that the exact performance of services delivered is not guaranteed.	High	Medium	Research how to monitor and manage virtual routing resources and slice isolation.	• Work with vendor(s) to identify solutions. • Investigate advanced resource utilisation monitoring solutions.

Risk Identification			Risk Assessment		Risk Control	
Id	**Name**	**Description**	**Probability**	**Impact**	**Avoidance**	**Reaction**
L3-R3	The management model of the L3-based Network Factory infrastructure and services is inefficient.	The implications of an infrastructure at Layer 3 for which management is split between users and GÉANT Operations are unknown at the time of writing.	Medium	High	Classify vendor solutions in terms of management capabilities of logical routers.	Elaborate the management model of the infrastructure as part of further elaboration of the business case.
					•	

Table 4.4 Network Factory delivering Layer 3 slices business case risk analysis

5 Conclusions

Based on findings within the GN3 project and specifically JRA2 Task 5, and also taking into account equivalent initiatives worldwide, a GÉANT Network Factory infrastructure and associated services are needed and recommended.

The challenges are both business- and technology-oriented, in some cases interdependent, in other cases not. A detailed study of the outcomes of the FEDERICA project, as the first relevant initiative with key players from the NREN community, has revealed a number of achievements but also several functional and business aspects for consideration. At the same time, the maturity of appropriate technologies is, in some cases (such as OpenFlow), inversely proportional to the potential offered in the context of a Network Factory. Embedding Network Factory operations in production environment operations also deserves particular attention.

Both technical and business aspects as presented in this document need to be further analysed and assessed. Emerging state-of-the-art technologies offered by equipment vendors have been identified as relevant to the Network Factory concept, with varying feature-support and operational characteristics. For a long-term Network Factory solution tailored to GÉANT, based on the resources and functionality of the next-generation GÉANT backbone, full analysis is not possible at the time of writing. However, it should be possible soon based on the outcomes of the ongoing procurement process.

For a short-term solution, it has been decided to pursue an OpenFlow-based Network Factory, deployed on top of the current GÉANT backbone in a way that ensures its viability over the future GÉANT backbone. The decision has been made taking into account the potential presented globally by OpenFlow testbeds, the low cost implications and the minimum requirements imposed on the GÉANT production environment. Hence, a detailed business case and technical implementation planning for the short-term solution are already underway.

To conclude, the next steps of JRA2 Task 5 following the study outcomes presented in this document are the elaboration of a business case, technical specification and implementation of a long-term Network Factory solution, in parallel to the next-generation GÉANT procurement and deployment works, as well as the elaboration of a shorter-term business case, followed by delivery and implementation of an OpenFlow-enabled Network Factory over the current backbone.

The planned long-term business case analysis is expected to provide more insight into all aspects of a full scale Network Factory design and deployment over GÉANT, upon which decisions driving the design and implementation choices will have to be made. In the meantime, the OpenFlow-enabled Network Factory will serve as a starting point for validating NFaaS concepts and making preliminary offerings available to the user community.

Appendix A FEDERICA

A.1 Physical Resources

FEDERICA servers are homogeneous across the entire infrastructure, with SUN X2200M2 servers selected as the physical platform. The server includes 2 x AMD Quad Core Opteron 1.9 GHz CPUs with virtualisation support, 2 x 500 GB HD drives, and up to 64 GB of RAM memory. The machines are operated from a VMware virtualisation platform, which allows any chosen operating system upon each virtual machine.

The network infrastructure is built out of 14 Points of Presence (PoPs) located across Europe and interconnected with dedicated 1 Gbps Layer 2 (Ethernet over SDH) circuits. The logical topology of FEDERICA is shown in Figure A.1 below.

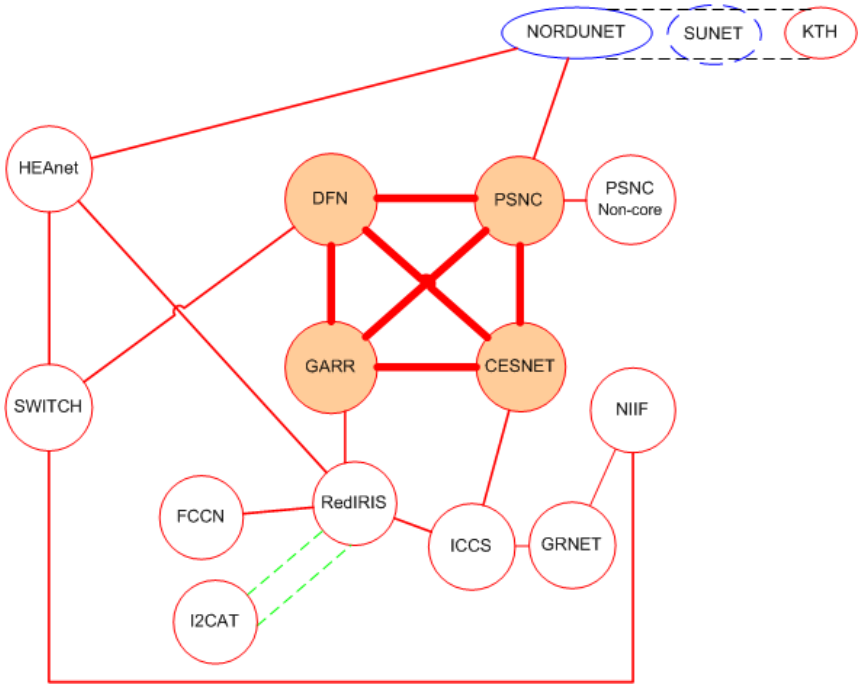

Figure A.1: FEDERICA logical topology

There are 4 core PoPs (DFN – Germany, PSNC – Poland, GARR – Italy, and CESNET – Czech Republic, shown in orange in Figure A.1), which are interconnected in a full mesh topology. The GÉANT Plus circuits

that terminate on GÉANT PoP equipment are delivered with fibre to the FEDERICA core PoP equipment, without crossing any local NREN hardware or intermediate Layer 2/Layer 3 hardware. The purpose of this is to implement those links at the lowest possible level, in order to maximise the Layer 2 functionality delivered over the FEDERICA substrate links to the end users. The core PoPs have no such restriction and some of them use NREN or MAN resources (within the city where GÉANT PoPs are located) to forward the inter-PoP circuits to their FEDERICA end points.

The non-core PoPs are:

- KTH – Sweden
- HEAnet – Ireland
- SWITCH – Switzerland
- FCCN – Portugal
- I2CAT – Spain
- RedIRIS – Spain
- ICCS – Greece
- GRNET – Greece
- NIIF – Hungary
- PSNC – Poland

Each PoP is built according to the same architecture, where all virtualisation servers are attached to a central Ethernet switch, which provides access to the inter-PoP circuits (details are provided in A.5 *Details of FEDERICA PoP setup* on page 79. Juniper MX480 routers are installed in the core PoPs, while Juniper EX3200 switches are installed in non-core PoPs.

The MX480 routers provide crucial functionality for FEDERICA operation and virtualisation capabilities, in addition to their regular switching and routing features. First, core PoPs provide Internet peering points through the MX480 routers (BGP is used, and FEDERICA is a separate AS). Second, they are used to implement virtual routers. This feature is commonly used to deliver Layer 3 FEDERICA services to end users. Such virtual routers are far more efficient than software routers (e.g. Quagga [Quagga] or XORP [XORP]), as the physical hardware resources of MX480 are used, which are optimised for routing efficiency. More details about the implementation of virtual routers in FEDERICA are given in *Router Virtualisation* on page 18.

In the case of FEDERICA services delivered to the user at Layer 3, the FEDERICA NOC cannot access end-user server-based logical router internals, and therefore cannot replace, or assist users with the Layer 3 configuration of their slices. The only exception to this rule is the slice management VLAN, where the virtual hosts' management interfaces have a predefined IP configuration. As described in A.3 *User Access and Slice Management* on page 77, users can access any host using its management interface, while accessing the dedicated slice management VLAN. Those interfaces are optional (as an alternative, users can access the machines using the VMware VPN console) and are not part of the user-requested topology, but are a parallel additional feature dedicated to the control and configuration of the user slice.

If a user needs to access entities beyond the FEDERICA infrastructure, core PoP routers could provide IP routing functionality on peering interfaces, forwarding the traffic to the local NREN and then to the Internet

directly. Such a configuration, however, requires additional security constraints inside both the FEDERICA physical substrate and the peering NREN, as the slice can use any IP addressing schema, either public or private, and those addresses cannot be advertised with FEDERICA BGP to the Internet. The slice must be completely isolated and unreachable from the public Internet.

More detailed information about FEDERICA's physical infrastructure is available in [FED_DSA1.3].

A.2 User Support and Portal

FEDERICA user support comprises of two main areas [FED_DSA2.3, FED_DSA2.1]:

- Consultancy and advice, which start before the user receives a slice and continue throughout the experiment. The initial phase is led by the User Policy Board (UPB), with the NOC as technical consultant.
- Slice creation and management. This task is the more complex. Each slice needs to be configured over the substrate and may need additional configuration in the virtual system images. (Further details are provided in Section 2.2.3.1 on page 16 and in A.4 below.)

The initial user consultancy and advice are based on the documents submitted to the UPB, and are usually provided via the NOC mailing list and the User Portal, which is the main information access tool. The main components of the User Portal are shown in Figure A.2.

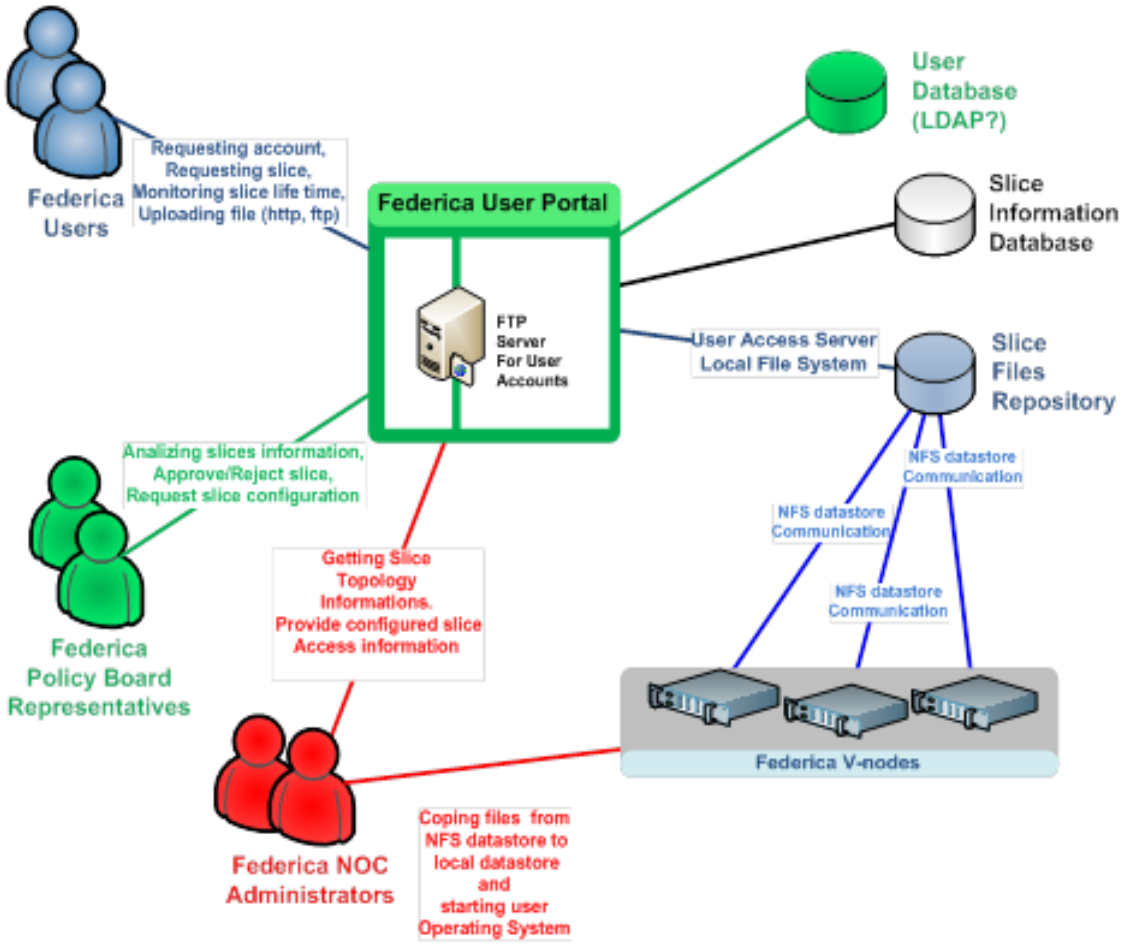

Figure A.2: The User Portal architecture

Access to the FEDERICA User Portal is restricted to authorised users. Each user is identified in the system by their own account. This account provides information about personal data and specifies system privileges. After login, the list of possible actions a user can perform depends on the user type. In the first version of the FEDERICA User Portal, three types of users were envisaged:

- FEDERICA User

  The end user of the FEDERICA infrastructure. To gain access to the FEDERICA User Portal, the user needs to send a registration request. The user then requests a slice and provides all the information required for the slice to be created.

- FEDERICA User Policy Board Representative

  A member of the UPB who has been allocated to a particular user. The UPB as a whole decides whether to approve or reject the requested slices and initiates the slice configuration at the NOC or (in the future) using the automated slice management system.

- FEDERICA Network Operations Centre Administrator

The NOC Administrator is responsible for creating slices within the FEDERICA substrate. (Further details are provided in Section 2.2.3.1 on page 16 and in A.4 below.)

A.3 User Access and Slice Management

End users need to access the virtual hosts and virtual network elements residing in their allocated slice. These virtual resources are located in a logically separate private network, which is not accessible via the public Internet. A gateway mechanism is needed to provide controlled entry to the slice.

To manage the virtual devices, all virtual hosts and virtual network elements have a management interface. The management interfaces are connected through an Ethernet VLAN, called the Service LAN. To make the Service LAN accessible from the public Internet, a virtual host is created for each Service LAN on the User Access Server (UAS). The UAS has two interfaces: one that connects to the public Internet and one to the Service LAN. The only functions provided by the UAS are SSH Server and SSH Client. End users log in to their respective UAS using SSH. Then, through the Service LAN, they open a new SSH session from the UAS to the virtual hosts and virtual network elements in their slice. The architecture is shown in Figure A.3 below. The UAS runs a Linux Operating System with a fixed configuration installed by the NOC.

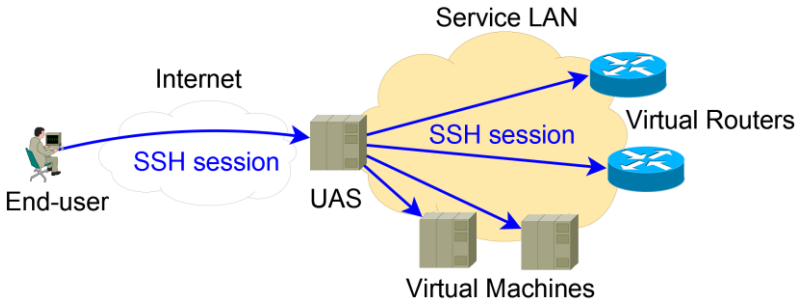

Figure A.3: Service LAN and SSH access for a user's slice

To access the virtual hosts and network elements in their slice, the user therefore needs an SSH client on their own PC and the following information:

- Username/password – created and provided by the NOC Administrator, through the FEDERICA User Portal.
- Internet-reachable IP address of the UAS – configured and provided by the NOC Administrator, through the FEDERICA User Portal.
- Service LAN IP addresses of the virtual hosts. During slice design/creation, the NOC Administrator allocates a private IP subnet for the Service LAN and gives the details to the user. The user creates the detailed addressing plan for the Service LAN and gives this to the NOC. The NOC configures the virtual resources based on this plan. The NOC also configures a local DNS service (dnsmasq) on the UAS, which maps IP addresses to the names of the slice's resources. In this way, name-based identification is possible (and preferred).

- Usernames/passwords of the virtual hosts. The FEDERICA NOC Administrator installs and configures the Operating Systems on the virtual hosts based on the installation instructions provided by the user.

The fact that the user can access the virtual hosts only through the SSH service implies that a working Operating System must be present on the virtual host when the user first logs on to the virtual host. This also means that the user cannot select and install the Operating System on the virtual hosts of his slice (at least not at the first stage of FEDERICA).

A.4 Slice Creation and Management Process in Detail

The procedure is carried out by the NOC Administrator.

1. Having received a slice description from the end user, provided in a structured manner through the User Portal, assign a name to the slice. Document D5 – "Resource description" is the starting point for the user's network topology. (Document D5 is summarised in [FED_DSA2.1].)

2. Create a Resource Tracker (RT) case for the slice. The RT-case in the trouble ticket system will serve as the tracker for the slice during implementation and operation.

3. Create a Wiki page for the slice. The Wiki page provides structured detailed information about the slice for the NOC.

4. Assign the VLAN range, Service LAN IP range, and slice IP range (an IP range for the slice itself).

5. Select substrate links for use by the slice. The layout of links can be done rather freely, though there are some limitations for slices that include logical router instances and physical substrate routers. The link layout is tightly coupled with the virtual host layout.

6. Assign virtual hosts to V-Nodes, optimising the use of the substrate while complying with the user's requirements.

7. Create the virtual links/switches/routers. The creation of the links and the creation of instances in the switches/routers are closely inter-related; they therefore need to be configured at the same time.

8. On the V-Nodes the right VLAN needs to be tied to the right virtual NIC (vm-nic) interface.

9. Create the slice management network. In a similar way to creating links for a slice, the Service LAN for the slice needs to be created. The difference is that the Service LAN is a single broadcast domain, and thus consists only of switched routing-instances. All the links share the same VLAN tag, the first VLAN in the assigned range for the slice. The Service LAN originates from the User Access Server in FEDERICA, uas.poz.pl.net.fp7-federica.eu.

A.5 Details of FEDERICA PoP setup

Using a VLAN Q-in-Q [QinQ], which is available in most of the FEDERICA PoPs, two levels of VLAN stacking can be used for configuration. This has the benefit that the substrate links are sliced using VLAN technology and each slice's virtual link can still carry one level of VLAN identifier defined by the end users.

Each slice's virtual links are implemented by defining a tagged VLAN on the Juniper switching equipment, which connects a distant PoP, as shown in Figure A.4.

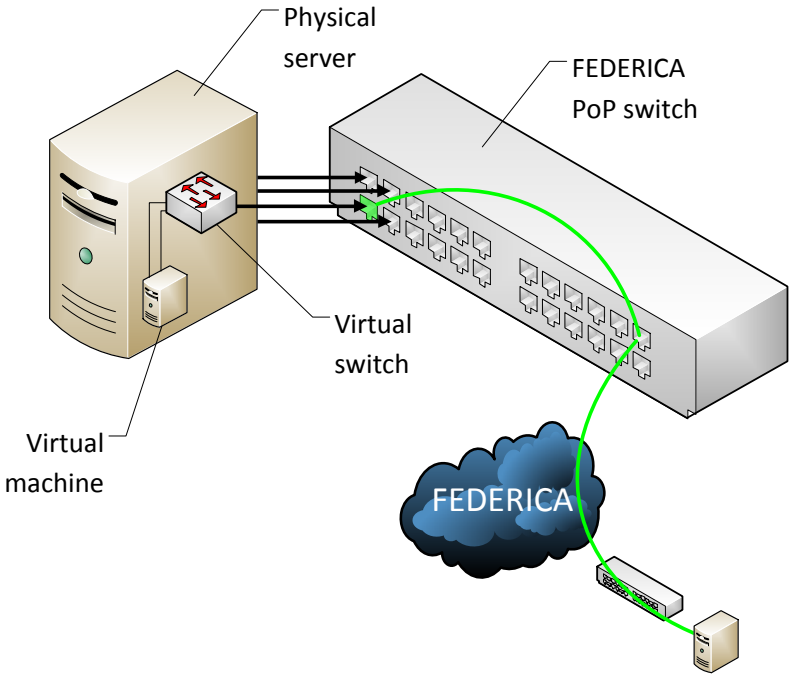

Figure A.4: FEDERICA virtual circuit creation with VLANs

A virtual machine configured on a physical server connects to a virtual switch, which then forwards slice traffic to one of the physical interfaces of the server and then to the switch interface. The traffic on the switch port is then tagged with a VLAN identifier, and the VLAN can be used to connect to the distant FEDERICA PoP through the FEDERICA substrate. As an alternative to the VLAN-based virtual links, the MPLS Circuit Cross-Connect (CCC) feature can be used (EX3200 supports only CCC, while MX480 also provides Translational Cross-Connect (TCC) features for configuration). An MPLS LSP can be set up between the interfaces of the EX3200 switches, providing transparent channels for the end users (behaving as virtual links). Such a configuration delivers more Layer 2 functionality to the end users, i.e. freedom to configure any VLAN, or VLAN stack on traffic within their slice. This approach is implemented, for example, between the RedIRIS and I2CAT PoPs.

References

[8021P]	http://www.javvin.com/protocol8021P.html
[AKARI]	http://akari-project.nict.go.jp/eng/concept-design/AKARI_fulltext_e_preliminary_ver2.pdf
[ANIBEDWEB]	https://sites.google.com/a/lbl.gov/ani-testbed
[Components]	http://yuba.stanford.edu/foswiki/bin/view/OpenFlow/Deployment/Components
[Dummy]	http://info.iet.unipi.it/~luigi/dummynet/
[FED_DJRA1.2]	FEDERICA Deliverable DJRA1.2: “Solutions and protocols proposal for the network control, management and monitoring in a virtualized network context” http://www.fp7-federica.eu/documents/FEDERICA-DJRA1.2_v2.1.pdf
[FED_DJRA2.1]	FEDERICA Deliverable DJRA2.1: “Architectures for virtual infrastructures, new Internet paradigms and business models” http://www.fp7-federica.eu/documents/FEDERICA-DJRA2.1.pdf
[FED_DJRA2.3]	FEDERICA Deliverable DJRA2.3: “Virtual network architectures” http://www.fp7-federica.eu/documents/FEDERICA-DJRA2.3-final.pdf
[FED_DNA2.3]	FEDERICA Deliverable DNA2.3: “FEDERICA Usage Report” http://www.fp7-federica.eu/documents/FEDERICA-DNA2.3.pdf
[FED_DSA1.3]	FEDERICA Deliverable DSA1.3: “Update on the FEDERICA Infrastructure” http://www.fp7-federica.eu/documents/FEDERICA-DSA1.3-reduced.pdf
[FED_DSA2.1]	FEDERICA Deliverable DSA2.1: “FEDERICA SA2 User Support” http://www.fp7-federica.eu/documents/FEDERICA_DSA2.1-final.pdf
[FED_DSA2.3]	FEDERICA Deliverable DSA2.3: “FEDERICA SA2 Final Report” http://www.fp7-federica.eu/documents/FEDERICA-DSA2.3-v3-r.pdf
[FEDCase]	P. Szegedi, J. Ferrer Riera, J. A. García-Espín, M. Hidell, P. Sjödin, P. Söderman, M. Ruffini, D. O’Mahony, A. Bianco, L. Giraudo, M. Ponce de Leon, G. Power, C. Cervelló-Pastor, V. López, S. Naegele-Jackson, “Enabling Future Internet Research: The FEDERICA Case”, IEEE Communications Magazine, submitted for publication http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5936155
[FEDdocs]	https://ihttps://intranet.geant.net/sites/Research/JRA2/T5/Documents/Documents_FEDERICA/FEDERICA-DSA2.3-v3.doc.zip [access restricted to GN3 participants]
[FEDERICA]	http://www.fp7-federica.eu/
[FV]	http://www.openflow.org/downloads/technicalreports/openflow-tr-2009-1-flowvisor.pdf
[GÉANTIP]	http://www.geant.net/Services/ConnectivityServices/Pages/GEANTIP.aspx
[GENI]	http://www.geni.net/
[GN3_DJ1.2.1]	L. Lange Bjørn (NORDUnet), K. Bozorgebrahimi (UNINETT), E. Camisard (RENATER), P. Gasner (RoEduNet), M. Hůla (CESNET), M. Karásek (CESNET), R. Lund (NORDUnet), R. Nuijts (SURFNET), R. Octavian (RoEduNet), P. Škoda (CESNET), S. Šíma (CESNET), P. Turowicz (PSNC), K. Turza (PSNC), S. Tyley (DANTE), J. Vojtěch (CESNET), V. Vraciu

	(RoEduNet), G. Zervas (University of Essex), GN3 Deliverable DJ1.2.1: "State-of-the-Art Photonic Switching Technologies" http://www.geant.net/Media_Centre/Media_Library/Media%20Library/GN3-10-122%20DJ1%202%201v1%200%20State%20of%20the%20Art%20Photonic%20Switching%20Technologies_Read%20Only.doc
[GN3_DJ1.4.1]	M. Campanella (GARR), P. Kaufman (DFN), F. Loui (RENATER), R. Nejabati (University of Essex), C. Tziouvaras (GRNET), D. Wilson (HEANET), S. Tyley (DANTE), GN3 Deliverable DJ1.4.1 "Virtualisation Services and Framework Study" http://www.geant.net/Media_Centre/Media_Library/Media%20Library/GN3-09-225%20DJ1.4.1v1.0%20Virtualisation%20Services%20and%20Framework%20Study.pdf
[GN3_DS1.1.1,2]	L. Altmanova (CESNET), T. Breach (NORDUnet), M. Campanella (GARR), M. Carboni (GARR), M. Enrico (DANTE), L. Fischer (NORDUnet), R. Pekal (PSNC), J. Radil (CESNET), R. Sabatino (DANTE), M. Scarpa (GARR), S. Sima (CESNET), T. Szewczyk (PSNC), R. Tuminauskas (LITNET), C. Tziouvaras (GRNET), J. Vojtech (CESNET), GN3 Deliverable DS1.1.1,2: "Final GÉANT Architecture" http://www.geant.net/Media_Centre/Media_Library/Media%20Library/GN3-10-279_DS1.1.1,2%20Final%20GEANT%20Architecture_v1.pdf
[LightCol]	http://www.efg2.com/Lab/ScienceAndEngineering/Spectra.htm
[LogicSys]	http://www.juniper.net/techpubs/en_US/junos10.0/information-products/topic-collections/feature-guide/logical-systems-overview-solutions.html
[MPLS-EoMPLS]	http://www.cisco.com/en/US/docs/switches/metro/catalyst3750m/software/release/12.1_14_ax/configuration/guide/swmpls.pdf
[MPLS-TPC]	Cisco Systems, "Understanding MPLS-TP and Its Benefits" http://www.cisco.com/en/US/technologies/tk436/tk428/white_paper_c11-562013.pdf
[MPLS-TPJ]	Juniper Networks, "MPLS Transport Profile (MPLS-TP): A Set of Enhancements to the Rich MPLS Toolkit" http://www.juniper.net/us/en/local/pdf/whitepapers/2000406-en.pdf
[MPLS-TP]	http://searchtelecom.techtarget.com/tip/MPLS-and-MPLS-Transport-Profile-MPLS-TP-The-technology-differences?ShortReg=1&mboxConv=searchTelecom_RegActivate_Submit&
[NDDI]	http://www.internet2.edu/network/ose/
[NetFPGA]	http://netfpga.org/
[OFNetFPGA]	Activating OpenFlow on NetFPGA http://www.openflow.org/wk/index.php/NetFPGA
[NetNS]	http://lxc.sourceforge.net/index.php/about/kernel-namespaces/network
[NFEval]	GN3-11-387 "Network Factory Footprint: Third-Party Initiatives" https://www.geant.net/Media_Centre/Media_Library/Media%20Library/GN3-11-387_Network-Factory-Footprint_Third-Party-Initiatives_v1.0.doc
[NOX]	http://noxrepo.org/wp/
[OF]	www.openflow.org/documents/openflow-wp-latest.pdf
[OFELIA]	http://www.fp7-ofelia.eu/
[OFELIAIsland]	http://www.fp7-ofelia.eu/assets/IslandsinventoryPhaseIOpenCall.pdf
[OFJap]	http://www.apan.net/meetings/HongKong2011/Session/Slides/fit/7.pdf
[OFMPLS]	http://www.openflow.org/wk/index.php/OpenFlowMPLS
[OFSwitch]	http://www.openflow.org/documents/openflow-spec-v1.1.0.pdf
[OpenvSwitch]	http://openvswitch.org/

[OptStable] O. Lopez, A. Haboucha, F. Kéfélian, H. Jiang, B. Chanteau, V. Roncin, Ch. Chardonnet, A. Amy-Klein and G. Santarelli, “Cascaded multiplexed optical link on a telecommunication network for frequency dissemination”, Optics Express, Vol. 18, Issue 16, pp.16849-16857(2010)

[OptTime] CESNET Press Release, “A new method of accurate time signal transfer demonstrates the capabilities of all-optical networks”, Prague, April 1, 2010 http://www.ces.net/doc/press/2010/pr100401.html

[PBBTE] http://www.fujitsu.com/downloads/TEL/fnc/whitepapers/UnderstandingPBBTE.pdf

[PlanetLab] http://www.planet-lab.org

[PSD] http://www.juniper.net/techpubs/en_US/junos9.2/topics/concept/psd-virtual-hardware-router.html

[Quagga] http://www.quagga.net/

[QinQ] http://h20430.www2.hp.com/other/procurve/cn/zh/network/pdfs/QinQWhitePaper_Nov_07_WW_Eng_Ltr.pdf

[SFA] http://groups.geni.net/geni/wiki/SliceFedArch

[Trellis] S. Bhatia, M. Motiwala, W. Muhlbauer, Y. Mundada, V. Valancius, A. Bavier, N. Feamster, L. Peterson, and J. Rexford, “Trellis: a platform for building flexible, fast virtual networks on commodity hardware”, In Proceedings of the 2008 ACM CoNEXT Conference, Madrid, Spain, December 2008

[VArch] Cisco Systems, “Cisco Nexus 7000 Series Virtualization Architecture: Multi-Degree Virtualization Enabling Resource Consolidation” http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9402/ps9512/brochure_cisco_nexus_7000_series_virtualization_arch.pdf

[VInfra] Mauro Campanella, Vasilis Maglaris, Martin Potts, “Virtual Infrastructures in Future Internet in Towards the Future Internet - Emerging Trends from European Research”, pp 63-73, IOS Press, 2010 http://www.booksonline.iospress.nl/Content/View.aspx?piid=16465

[VINI] A. Bavier, N. Feamster, M. Huang, L. Peterson, and J. Rexford, “In VINI veritas: realistic and controlled network experimentation”, In Proceedings of the 2006 conference on applications, technologies, architectures and protocols for computer communications, Volume 36, Issue 4, October 2006 http://www.vini-veritas.net/

[Xen] http://xen.org/

[XORP] http://www.xorp.org/

Glossary

ALU 1626 LM	Alcatel-Lucent 1626 Light Manager (DWDM equipment)
ALU 1678 MCC	Alcatel-Lucent 1678 Metro Core Connect (optoelectrical switching equipment)
ANI	Advanced Networking Initiative
API	Application Programming Interface
AS	Autonomous System
ASR	Aggregation Services Routers
BE	Best Effort
BEN	Breakable Experimental Network
BGP	Border Gateway Protocol
CCC	Circuit Cross-Connect
CPU	Central Processing Unit
CRS	Carrier Routing System
DASH	Angel Secure Content Delivery and Host Authentication
DOME	Diverse Outdoor Mobile Environment
DWDM	Dense Wavelength Division Multiplexing
EoMPLS	Ethernet over MPLS
EoSDH	Ethernet over SDH
EPL	Ethernet Private Line
EU	European Union
FEDERICA	Federated E-infrastructure Dedicated to European Researchers Innovating in Computing network Architectures
FIB	Forwarding Information Base
FPGA	Field-programmable Gate Array
FP7	EU's Seventh Framework Programme for Research and Technological Development
G	Generation
GB	Gigabyte
Gbps	Gigabit per Second
GE	Gigabit Ethernet
GENI	Global Environment for Network Innovations
GFP-F	Generic Framing Protocol – Framed
GHz	Gigahertz
GMPLS	Generalised Multi-Protocol Label Switching
GRE	Generic Routing Encapsulation
HADES	Hades Active Delay Evaluation System
HD	High Definition
IaaS	Infrastructure as a Service

ICT	Information and Communications Technology
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IP	Internet Protocol
IS-IS	Intermediate System to Intermediate System
ISO	International Standards Organisation
ITU	International Telecommunication Union
ITU-T	International Telecommunication Union, Telecommunication Standardization Sector
JRA1	GN3 Joint Research Activity 1 Future Network
JRA2	GN3 Joint Research Activity 2 Multi-Domain Network Service Research
JRA1 Task 4	JRA1 Task 4 Current and Potential Uses of Virtualisation
JRA2 Task 5	JRA2 Task 5 Network Factory
km	Kilometre
KVM	Kernel-based Virtual Machine
L	Layer
LBE	Less Than Best Effort
LSP	Label Switched Path
M Series	Juniper Networks M-Series Multiservice Edge Routers
MAC	Media Access Control
MAN	Metropolitan Area Network
MPLS	Multi-Protocol Label Switching
MPLS-TP	Multi-Protocol Label Switching – Transport Profile
MTU	Maximum Transmission Unit
MX Series	Juniper Networks MX-Series 3D Universal Edge Routers
NDDI	Network Development and Deployment Initiative
NF	Network Factory
NFaaS	Network Factory as a Service
NGN	Next Generation Network
NG-OTN	Next-Generation Optical Transport Network
NIC	Network Interface Card
NOC	Network Operations Centre
NREN	National Research and Education Network
OADM	Optical Add Drop Multiplexer
OAM	Operation, Administration and Maintenance
OCX	Optical Cross-Connect
ODU	Optical Channel Data Unit
OEO	Optical-Electrical-Optical
OFELIA	OpenFlow in Europe: Linking Infrastructure and Applications
OS	Operating System
OSI	Open Systems Interconnection
OSPF	Open Shortest Path First
OTN	Optical Transport Networks
PaaS	Platform as a Service
PBB	Provider Backbone Bridge
PBB-TE	Provider Backbone Bridge Traffic Engineering
PBT	Provider Backbone Transport

PC	Personal Computer
perfSONAR	Performance Focused Service Oriented Network Monitoring Architecture
PERT	Performance Enhancement Response Team
PM	Person Month
PoP	Point of Presence
PSD	Protected System Domain
Q-in-Q	802.1q Tunneling
QoS	Quality of Service
R&E	Research and Education
RAM	Random-Access Memory
ROADM	Reconfigurable Optical Add-Drop Multiplexer
RPD	Routing Protocol Process
RSVP-TE	Resource Reservation Protocol - Traffic Engineering
RT	Resource Tracker
SA1	GN3 Service Activity 1 Network Build and Operations
SDH	Synchronous Digital Hierarchy
SDN	Software-Defined Network
SONET	Synchronous Optical Network or Synchronous Optical Networking
SSH	Secure Shell
SSO	Single Sign-On
STM	Synchronous Transfer Module
TCC	Translational Cross-Connect
TDM	Time Division Multiplexed
TE	Traffic Engineering
UAS	User Access Server
VC	Virtual Container
VI	Virtual Infrastructure
VLAN	Virtual Local Area Network
VM	Virtual Machine
VoIP	Voice over IP
VPN	Virtual Private Network
WSS	Wavelength Selective Switch
XSM	Extreme Scale Motes