FEDERICA: A Virtualization Based Infrastructure for Future and Present Internet Research

Mauro Campanella*

Consortium GARR
Via dei Tizii 6, 00185 Roma, Italy
Mauro.Campanella@garr.it

Abstract. The Europe wide infrastructure managed by the FEDERICA project demonstrates how an infrastructure based on computers and network physical resources, both capable of virtualization, is a flexible, neutral and efficient facility for future and present Internet research. The facility can create virtual resources sets, grouped in virtual infrastructures (slices) according to users' specification. The user has full control on the resources in the assigned slice which can be used for many types of research, from Future Internet clean-slate architectures to security and distributed protocols and applications. The infrastructure has European size and it is capable of federating with other facilities worldwide.

Keywords: Virtualization, NRENs, Future Internet, polymorphic test infrastructure, cloud infrastructures.

1 Introduction

FEDERICA [1] is a European project started in January 2008. It has been engineered to provide support to the research on current and Future Internet technologies and architectures. The project is linked to the European FIRE initiative [4] and the European Future Internet Assembly [5]. Other similar initiatives exists worldwide, e.g. GENI [6] in the United Stated and AKARI [7] in Japan.

Research and experimentation on novel technologies and architectures require new experimental environments that combine flexibility, neutrality, a minimum set of constraints for the researchers, reproducibility and allow full control of the testing environment. The project has distinguishing characters from similar projects, e.g. Onelab [3], in particular, it does not mandate a specific operating system, and the user has full control of the physical networking layers down to the data link layer,

The project is based on the constant developments of National Research and Education Networks (NRENs) [8] in Europe and worldwide, which created a strong multi-domain hybrid network infrastructures with advanced capabilities.

* On behalf of the FEDERICA consortium.

T. Magedanz et al. (Eds.): TridentCom 2010, LNICST 46, pp. 123–132, 2011.

The FEDERICA project uses the NRENs infrastructures and virtualization technologies to create a research facility available to the public and private sector.

In the following, section 2 describes the FEDERICA project framework and architecture. Section 3 details the operational infrastructure. Section 4 lists status challenges and new possibilities. Section 5 concludes the article.

2 The FEDERICA Project

The project is co-funded under the 7th European Community Framework Program. It started 1st January 2008 and lasts 30 months. The project partners include a wide range of stake-holders on network research, NRENs, DANTE, TERENA, academic and industrial research groups and vendors.

2.1 Project Goals and Objectives

The main goal is to support research in present and Future Internet. To achieve this goal, the project set its objectives to:

- Engineer and implement a Europe-wide Infrastructure to be used as a distributed testing facility
- Research in virtualization of e-Infrastructures integrating network and computing resources
- Facilitate technical discussions amongst specialists, in particular arising from experimental results and disseminating knowledge and NREN experience of meeting users requirements
- Contribute with real test cases and results to standardization bodies, e.g. IETF, ITU-T, OGF, TM Forum/IPsphere.

2.2 Requirements

According to its goals, the infrastructure has to support research on the widest range of new technologies and protocols, in particular in networking. To achieve such goals, the infrastructure must:

- Avoid to impose specific technological or architectural constraints to the researchers, as an example avoid mandating the IP protocol. The facility has to create environments that are technology agnostic and neutral (transparent) to new protocols or technologies.
- Ensure reproducibility of the experiments. Given the same initial conditions, the behaviour of a virtual resource should be the same, as a basic principle to obtain the same experimental results. This requirement is considered of particular importance.
- Provide to the user complete control and configuration capabilities within the assigned resources, allowing disruptive testing.
- Open to interconnect or federate with other e-Infrastructures and Internet.

2.3 Infrastructure Framework

The requirements point to two key framework principles for the infrastructure:

- The utilization of a combination of network and computing physical resources.
- The use of virtualization technologies applied both to the computing and network resources.

Virtualization is defined here as the capability to create a virtual version of a physical resource, both in computing and network environments. The virtual resources (e.g. a virtual network circuit, a disk partition, a virtual computer) are typically created by segmenting a physical resource. Virtualization creates unconfigured (clean) virtual resources, e.g. an image of the hardware of a computing element on which (almost) any operating system can be installed, a point-to-point network circuit, a portion of disk space. Those resources can be then tailored and configured to users needs and even moved from a virtualization-aware platform to another.

The framework for such an infrastructure is based on two distinct layers (see Fig. 1 for a pictorial representation):

1. The virtualization substrate. The physical environment which contains all the hardware and software to instantiate the virtual resources.
2. The virtual infrastructures layer, containing all the virtual sets of resources (or slices).

The virtualization substrate is managed and controlled as a single administrative domain. The virtual infrastructures are in principle an unlimited, or very large

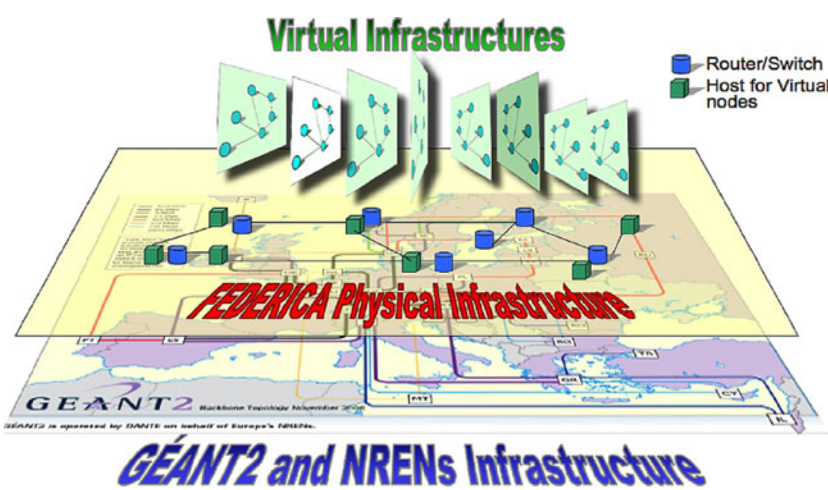

Fig. 1. Pictorial view of the FEDERICA infrastructure

number, restricted by the physical resources available and the requested characteristics. In case of federated facilities, the slices can comprise set of resources provided by different infrastructures.

The architecture defines only two basic resource entities:

1. Data connectivity. In form of a point to point circuit with or without assured capacity guarantees and with or without a data link protocol (a bit pipe).
2. A computing element, offering the equivalent of a computer hardware containing at least RAM, CPU and one network interface and mass storage. The computing element is capable of hosting various operating systems and also perform functionalities (e.g. routing). The computing element is defined as the basic entity. The RAM, NICs, storage are considered characteristics of the entity. This differs from other proposal, in particular for cloud services, where storage is considered a basic entity.

The figure represents the slice in vertical format for sake of clarity and to show that there is no dependency or hierarchy between them. Each slice may contain a virtual resource coming from any part of the substrate.

3 The Infrastructure Implementation

Following the framework outlined above, the FEDERICA infrastructure is implemented in two layers. The substrate an its made of network and computing physical resources. Each physical resource can produce virtual resources, slicing itself. Resource grouped by a topology in sets, or slices, are handled to the the user and compose the other layer.

As a design principle, the infrastructure favours testing of functionalities, protocols and new ideas, rather than providing a laboratory suited to very high performance studies.

The network resource in the wide area are provided by the participating NRENs through the GÉANT [9] infrastructure. Each NREN hosts a point of presence (PoP). A mesh of one Gigabit Ethernet circuits connects the PoPs. The circuits are initially at one Gbps, the capacity has been chosen as a compromise between cost on the wide area network and total capacity. It has been adequate for now to users' requirements. This capacity can be sliced, still creating high-speed links, and although expensive is contributed by the participating NRENs. Most of the circuits are created over the GÉANT SDH equipment using generic framing procedure and virtual concatenation.

Each PoP hosts network and computing elements. The network equipment in the four core PoPs is a programmable high-end router/switch from Juniper Networks. The model is a MX480 with dual CPU and 1 line card with 32 ports at 1Gb Ethernet (8 optical and 24 copper). The MX functionalities include virtual and logical routing, MPLS, VLANs, IPv4, IPv6. Two MX480 are equipped with Ethernet linecards with hardware QoS capabilities to enforce precise QoS at the packet level when requested. Smaller multi-protocol switche/routers (Juniper EX series) are installed in non-core PoPs. The computing equipment (V-Node) is based on off the shelf PC hardware, running virtualization software. Each PC contains 2 x Quad core AMD CPU running at 2 GHz, 32GB RAM, 8 network interfaces, 2x500GB disks. All the interfaces of the V-Nodes are connected to the Juniper routers.

The initial choice of the virtualization software for the V-nodes is the free version of the VMware [10] client software (ESXi) . This choice has been done after a review of other virtualization software (e.g. XEN). In particular it has been evaluated the completeness and structure of the Application Programming Interface, the availability of usage examples, available expertise, free tools to import and convert node images. The capabilities and performance of the free version have been adequate for the current requirements. The major drawback is the more difficult management of the V-Nodes without the commercial software. Such complexity is considered adequate for the initial phase of the project, but will be reviewed as a function of the number of slices and requests for a possible upgrade to the commercial versions of the software.

These building blocks of the substrate pose very few constraints to the user. The virtualization software in the V-nodes can host a large variety of operating systems and tools . It is possible to create an image from a fully configured system, avoiding the need of configuration in the slice. In the current status of the infrastructure the most significant one is that the data link layer is fixed to Ethernet framing. Future development of FEDERICA, according to users' requirements, will implement access to optical equipment to overcome this limitation.

3.1 Topology

The topology is composed of 13 distributed physical sites. Amongst these PoPs, a full mesh of four is equipped with Juniper MX router/switches and it is considered the core. The 9 non-core nodes are equipped by EX switches. The core nodes are equipped by 2 V-Nodes, the non-core PoPs host one node each. The FEDERICA physical topology is depicted in Fig. 2. The design placed particular importance on the resiliency and load balancing of the network, based on GÉANT infrastructure, and resources availability at partners locations. To minimize the load on the physical network resources and hence the interference between virtual resources, the network topology has a high level of meshing.

The FEDERICA substrate is a single administrative domain that contains all the physical resources (point to point circuits, nodes) in all PoPs. The domain does not contain the optical equipment of GÉANT used to transport the circuits between PoPs. It is configured as an IPv4 and IPv6 Autonomous System with both public and private addresses. The infrastructure is permanently connected to Internet using the Border Gateway Protocol and receives full routing tables in the four core PoPs.

The infrastructure is centrally managed and monitored by a Network Operation Centre. The NOC has also the task to create the slices. The monitoring system uses mainly the simple network management protocol to retrieve information for physical and virtual resources. In a PoP a packet flow based monitoring is available

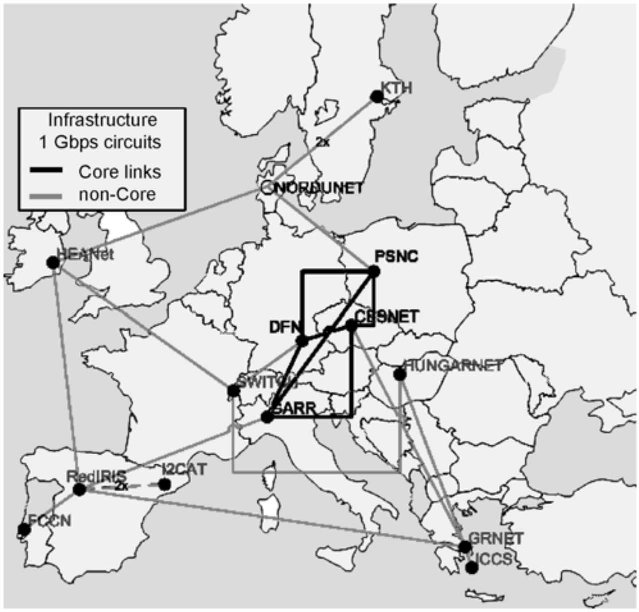

Fig. 2. FEDERICA network topology (November 2009)

3.2 Reproducibiity

The reproducibility and the stability of the behaviour of the virtual resources is a fundamental requirement for quantitative evaluations of new ideas. As an example, a virtual circuit may not be capable of offering a constant, fixed amount of bit per second, and a virtual computer image may not provide a constant CPU usage.

The project has focused the engineering of the infrastructure to provide reproducibility of a single virtual resource. The difference is due to two main independent causes:

- The behaviour of the physical resource supporting the virtual resources, e.g. due to its workload
- The virtualization technology itself, usually a layer placed between the physical resources and the virtual ones

Computing elements in FEDERICA have been chosen to provide specific functionalities in hardware, like virtualization-aware CPUs. The added hardware capabilities ensure a smoother sharing of the CPU by concurrent executing images. As additional measure, a virtual node is usually assigned to a idle core. As network interface cards do not support virtualization of flows in hardware, the V-node has been equipped with the maximum number of interfaces. A running image received an idle interface if available. If a node requires a high level of reproducibility, to such node can be fully dedicated a single core, a fixed amount of physical memory, a fraction of disk and a network interface. Not all the nodes requires such guarantees. In addition to the packet based QoS technologies of the switch/routers, some circuits are connected to packet Quality of Service capable line cards in the Juniper MX. In the rest of the infrastructure, the resources have been adequately increased, to avoid overbooking and minimize contention. It is possible then to create a slice with a set of resources, which exhibits, singularly, a known behaviour in all conditions.

Assuring the reproducibility of a set of connected resources is out of the scope of the infrastructure and it is instead a user responsibility. The complexity increases rapidly with the number of resources involved and it is strongly dipendent on the technologies and architecture. The classic problem of guaranteeing an end-to-end quality of service of an IP flow exemplifies the issue. In case of virtual infrastructures, as in the case of Internet traffic, often the requirements do not mandate strict guarantees, but rather a best effort behaviour. Virtual resource performance measurements are ongoing in FEDERICA.

3.3 Access Policy

The infrastructure is available to public and private researchers on Future Internet technologies. The access to the infrastructure is controlled by a User Policy Board (UPB). The UPB receives all requests for access, analyzes their technical feasibility (not the scientific content) and the current availability of resources and then prioritize them. The motivation for mandating a controlled access is mainly related to security (identification of the user, agreement on an acceptable use policy) and reproducibility (quality of service) guarantees.

3.4 Resource Virtualization and Slice Creation

The process to create a virtual computing system is rather straightforward and can also accept an image provided by the user or on available template of various operating systems. The virtualization capabilities in the network are also evolving, as described in [2]. The article reviews the current research in a Network Virtualization Environment (NVE) and the many challenges associated. The initial choice in FEDERICA is to use Virtual LANs and use QoS techniques for circuit virtualization; multi protocol label switching (MPLS) may be applied when needed.

The slice creation procedure definition is developing to incorporate the users' feedback. The current implementation of the infrastructure is based on manual or semi-automated provisioning of the virtual resources. The manual process is a choice in the step that maps virtual to physical resource. Even if the infrastructure contains a small amount of resources, this step is fundamental to ensure that the performance requirements of the virtual resources is respected and the infrastructure is efficiently used.

The current slice creation process consists of the following steps. First, the researcher that wants to perform an experiment over the FEDERICA infrastructure is required to provide the NOC with the desired topology, including requirements for the nodes and the network (each V-node RAM size, CPU power, mass storage space, topology and bandwidth between the V-Nodes, routing or switching functionalities, protocols). The request may be for un-configured resources, that the user will configure directly, even substituting protocols, or resources with an initial configuration, e.g. IP routing.

Once the NOC receives the slice description and resource requirements, the NOC maps the logical topology requested on the physical topology of the substrate and chooses the sites (PoPs) from which physical resources will be allocated. Besides instantiating all the resources requested by the user, the NOC needs to instantiate an extra virtual machine, that act as a gateway between Internet and the slice: the Slice Management Server. Access control of the Slice Management Server is performed by means of identity credentials managed by a RADIUS server.

The next step for the NOC is to instantiate Ethernet VLANs to connect the slice resources and create the topology required by the researcher. Finally, the NOC needs to setup the Slice Management network for the user that will connect the Slice Management Server to the management interface of each one of the managed resources in the slice (V-Nodes, logical routers, software routers). The connection is performed creating virtual interfaces in all resources and one in the Management Server in the same IP subnet (usually private) and creating an additional VLAN linking them. This subnet is initially the only IP path for the user to connect to the slice resources when accessing from Internet the Management server.

When the NOC has created the slice, it communicates to the researchers the information to access it: the public IP address of the Virtual Slice Management Server, the security credentials, the credentials to access in case of the Juniper logical routers and/or the software routers, and finally the IP addressing scheme of the Virtual Slice Management Network.

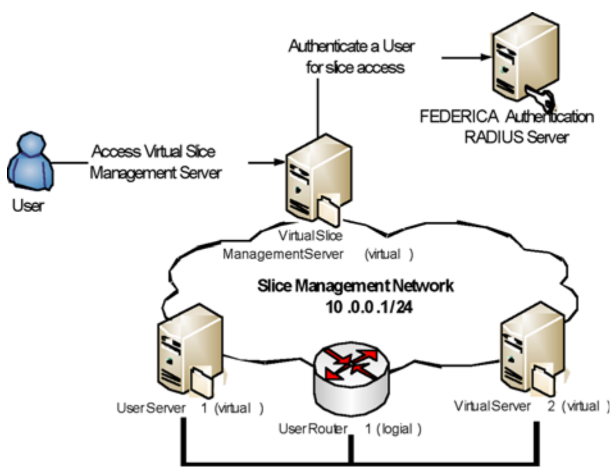

Fig. 3. Slice access example)

In the example in Fig. 3 the user has requested a slice containing two virtual servers connected through a router (created on a Juniper). The NOC created the three resources, connected them through a VLAN (black line at the bottom of the Figure), instantiated the Virtual Slice Management Server and created the Slice Management Network. The slice management network (cloud at the centre of the Figure) is needed to access the resources independently of user's slice topology and the virtual resource configuration. The researcher connects to the Virtual Slice Management Server using the credentials provided by the NOC, and is authenticated by the FEDERICA Authentication RADIUS Server. VMware virtual machines may also be configured to be accessed through remote Virtual Network Console (VNC) connections. By exploiting this mechanism users would have access to the console of their virtual servers, but they would also be able to interact with graphical user interfaces and to even access the BIOS of the server.

All the steps are performed either manually or using a heterogeneous set of tools (web portal for users, VMware Infrastructures application, the remote console of the devices, VNC clients, monitoring tools). A tool bench that provides a unified environment to operate the FEDERICA infrastructure and configure also the slices is being developed, and will be progressively deployed and used by the NOC and the FEDERICA users.

4 Status and Challenges

The FEDERICA infrastructure is now supporting a variety of users, who experiment on monitoring, new routing protocols, advanced control and management of physical and virtual circuit based topologies, energy-aware routing. The current feedback is positive and demonstrates the wide range of applicability of an architecture based on virtualization.

The experience has also suggested a list of developments, which also represent a set of research challenges:

- An increased level of control of each virtual resource behaviour. The performance of the virtual resources is still a function of the hardware and software used.
- Monitoring of the substrate and the resources in the slices and their relationship. Monitoring is considered a fundamental resource for the management of the substrate and for users' analysis of their experiments.
- Automation of the procedures and virtual resource description. An increased level of automation is needed to improve the management of the infrastructure and the efficiency in slice provisioning and management. The resource description and the schema to describe the relation between them and with the physical resource ia a fundamental element to achieve a greater automation and the creation of a FEDERICA service. The standardization of the resource representation is fundamental also for the fast-developing "cloud computing" [11] [12] architecture and services, allowing a possible synergy.
- increase the federation capabilities with other facilities to offer a richer environment to the users.

5 Conclusion

An infrastructure substrate based on virtualization both in computing and network resources is a novel approach to provide an ideal environment for innovative research and services on present and Future Internet. The virtual infrastructures created can be tailored to a very large variety of testing scenarios and present a simple interface to the user. The time needed to experimentally validate an idea can be reduced as well as the debugging and analysis phases.

Such infrastructures demonstrate the capability of current technologies to decouple the functionalities from their physical location, creating cloud infrastructures and granting new possibilities (e.g. the mobility of the routing functionality. The developments of the infrastructure require research on reproducibility behaviour, resource mapping, monitoring and resource standardization.

Acknowledgments. The FP7 project FEDERICA is partially supported by the European Commission under the Grant Agreement No.: RI- 213107. The authors acknowledge the fundamental contribution of all project partners.

References

1. Federated E-infrastructure Dedicated to Research Innovating in Computing Architectures, http://www.fp7-federica.eu
2. Chowdhury, N.M.M.K., Boutaba, R.: Network Virtualization: State of the Art and Research Challenges. IEEE Communications Magazine, 20–26 (July 2009)
3. A European Project based on PlanetLab Europe, http://www.onelab.eu/
4. Future Internet Research and Experimentation, http://cordis.europa.eu/fp7/ict/fire/
5. Future Internet Assembly, http://www.future-internet.eu/
6. Global Environment for Network Innovation, http://www.geni.net
7. AKARI Architecture Design Project, http://akari-project.nict.go.jp/eng/index2.htm
8. For a detailed analysis of NRENs in Europe and their role, see the documents (in particular the TERENA compendium), http://www.terena.org/publications/
9. GÉANT, the European Research and Education backbone, http://www.geant.net
10. VMware, http://www.wmware.com
11. Mell, P., Grance, T.: For a definition of Cloud Computing, see NIST Definition of Cloud Computing. Version 15 (10-7-09), http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc
12. IEEE Computing now: System Virtualization: Cloud Computing: Opportunities and Challenges (October 2009), http://www.computer.org/portal/web/computingnow/articles