Towards the Future Internet
G. Tselentis et al. (Eds.)
IOS Press, 2010

doi:10.3233/978-1-60750-539-6-63

Virtual Infrastructures in Future Internet

Mauro CAMPANELLA[a1], Vasilis MAGLARIS[b] and Martin POTTS[c]
[a] *Consortium GARR, Via dei Tizii 6, 00185 Roma, Italy*
[a] *National Technical University of Athens, Zografou, 157 80, Athens, Greece*
[a] *Martel Consulting, Dorfstrasse 97, 3073 Gümligen, Bern, Switzerland*

Abstract. Virtualization in both computer systems and network elements is becoming increasingly part of the Internet. Apart from posing new challenges, virtualization enables new functionalities, and hints at a future Internet architecture, which contains a physical, but polymorphic, substrate on which various parallel infrastructures can be created on demand. This article explores such a new "architecture of virtual infrastructures". The FEDERICA project and the recent evolution trends in the National Research and Education networks are taken as examples of this new type of network infrastructure, which is an evolution of the classic network.

Keywords. Virtualization, networking, NREN, GÉANT, FEDERICA, infrastructure, Future Internet architecture, IaaS, cloud computing

Introduction

Recent advances in virtualization technologies, enabled by powerful hardware in ASICs and CPUs, have made it extremely easy to decouple the operating system from the physical components in computers. On a single hardware platform, more than one operating system can be active at the same time and the fairness of the sharing is enhanced by dedicated hardware. The enabler is usually a thin software layer (e.g. XEN[1], VMware[2], KVM[3]), which abstract the physical resources to a standard (virtual) system. Such advances create more degrees of freedom for the end users and, at the same time, more capabilities. The "cloud computing" [4] [5] paradigm is a clear example of the new type of services available.

In data networks, virtualization has also been present since the beginning as a way to decouple the physical infrastructure from the service offered and it is still subject to research and development [6]. A single network infrastructure, like e.g. the GÉANT[7] backbone in Europe is currently hosting many virtual networks. Network virtualization technologies like MPLS (Multiprotocol Label Switching) and VLANS are common in commercial networks.

At the same, the use of the Internet is evolving towards a very content-rich platform, in which computing power and data easily accessible from everywhere are the key elements. When virtualization technologies are added to all the basic elements, the new Internet infrastructure becomes even more useful and efficient. It enables a more efficient use of the physical resources through resource "slicing", and makes the logical topology of important functionalities (routing, monitoring, computing and storage) less

[1] Corresponding Author.

dependent on the physical topology, by introducing dynamic capabilities. Such an environment can be considered to be a polymorphic substrate capable of hosting user equipment and able to adapt rapidly to users' needs and events. Such an architecture exhibits a fabric (a physical substrate) composed of computing and network elements, each capable of virtualization.

The FEDERICA project [8][9][10] is operating an infrastructure fully based on such an architecture as a service for Future Internet researchers and a platform for evaluating, understanding and validating this new architecture.

In the following, section 1 summarizes the developments in the NREN networks relevant for this article. Section 2 provides an overview of the FEDERICA architecture and section 3 its implementation. Section 4 consider a possible architecture for virtual infrastructures and section 5 lists challenges. Section 6 briefly compares with the cloud computing environment and section 7 concludes the article.

1. Experience in the NRENs

The purpose of the NRENs[11] is to serve and support researchers in networking. The users' requirements are increasingly for services that allow communities to work together ubiquitously, dynamically and rapidly. As a consequence, the development of NRENs in Europe and worldwide has created a strong multidomain hybrid network infrastructure with advanced capabilities. These networks no longer provide to the users just IPv4 and IPv6 connectivity, but are also capable of creating many "virtual", parallel, infrastructure for the users on the same physical footprint. Such infrastructures are enabling better or novel uses of dedicated facilities, such as radio telescopes, large particle colliders and fusion reactors. Virtualization in various flavours is used to slice capacity both at the packet (VLANs, MPLS) and the circuit level (using Generic Framing Procedure as an example). Virtualization is also used to decouple the IP topology from the physical topology and to perform traffic engineering.

In recent years the developments related to virtualization are in two main areas:

- Multidomain services, such as Bandwidth-on-Demand, security and authentication. The services are either in the trial or production phase. In particular the dynamic circuit network (DCN) [12] in ESnet and Internet2 and the AutoBAHN[13] service in GÉANT deal with the creation of virtual circuits on large scale. The challenge is to provide a real-time service, reducing the set-up time of a point-to-point virtual circuit. These multidomain services require the development of new standards for protocols, resource representation and functionalities in equipment, as the NRENs have to create networks in an environment, which is intrinsically multi-vendor and multi-technology.
- Monitoring, management and control of the infrastructure. These ancillary services become very complex when the same physical infrastructure hosts many virtual networks. The approach has been to develop a modular, multidomain, distributed monitoring system (PerfSONAR[14]). The capability to monitor the virtual elements of the network is strongly determined by the functionalities provided by the network equipment.

The NRENs are continuing to reduce the number of transmission protocols used. At the same time they are increasing the number of Interdomain protocols and the use of virtualization to provide advanced services. The NREN infrastructures are already supporting environments based on virtualization, which are evolving in size and complexity. The introduction of computing elements in the network topology is mandated by the need to control and manage all the virtual infrastructures created on the same multidomain physical environment and it will permit a step further in the virtualization of the infrastructure functionalities like routing, monitoring, capacity offering.

2. FEDERICA

Research on Future Internet technologies and architectures has become a hot topic in computer science. There are many initiatives worldwide, e.g. GENI [15] in the United Stated and FIRE [16] in Europe. Such research requires new environments that combine flexibility and a minimum set of constraint for the researchers. FEDERICA (Federated E-infrastructure DEdicated to Researchers Innovating in Computer Architectures) is a European Commission co-funded project started in January 2008 and operational until June 2010, that created a scalable, Europe-wide, clean slate, infrastructure to support experiments on Future Internet. Its architecture is a working example of an infrastructure completely based on virtualization and it is described in the following.

2.1. Project Goals and Objectives

The main project objective is to create an e-Infrastructure for researchers on Future Internet allowing researchers a complete control of set of resources in a “slice”, enabling disruptive experiments and placing the lowest possible set of constraints to researchers. Research on the use of virtualization in e-Infrastructure and facilitating the collaboration between experts are also key goals.

2.2. Architecture

2.2.1. Requirements

As the scope is focused on a research environment on new technologies, the following set of requirements for the infrastructure have been assumed:

- Be technology agnostic and neutral (transparent) to allow disruptive and novel testing, as to not impose constraints to researchers. The requirement is valid for all networking layers, not just the application layer and extends to the operating system used.
- Ensure reproducibility of the experiments, i.e. given the same initial conditions, the results of an experiment are the same. This requirement is considered of particular importance.
- Provide to the user complete control and configuration capabilities within the assigned resources.
- Allow more than one user group to run experiments at the same time, without interference.
- Open to interconnect / federate with other e-Infrastructures and the Internet. This last requirement plans for access to slices, access to the control of slices, management of experiments, interoperability and migration testing.

2.2.2. Framework and Design.

The requirements suggest two key framework choices for the infrastructure, which are at the core of the design:

- The simultaneous presence of computing and network physical resources. These resources form the substrate of the infrastructure.
- The use of virtualization technologies applied both to computing and network resources. Virtualization will allow creating virtual, un-configured resources.

Virtualization is defined as the capability to create a virtual instance of a physical resource, both in the computing and network environment. The virtual resources (e.g. a virtual circuit, a disk partition, a virtual computer) are usually created by segmenting a physical resource. Virtualization may create non-configured (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 to various needs and even moved from one virtualization-aware platform to another.

Such framework leads to a design in which the infrastructure is considered to be built in two distinct layers:

1. The virtualization substrate. The physical infrastructure which contains all the hardware and software capable to create the virtual resources;
2. The layer containing all the virtual infrastructures. Each containing the virtual resources and the initial network topology connecting them.

The virtualization substrate is a single administrative domain. The virtual infrastructures (VI or "slices") may be in principle unlimited, in practice a large number, restricted by the physical resources available and the requested characteristics for the slice.

Two basic resource entities are defined:

1. Connectivity. In the 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, mass storage is optional, although usually available. The computing element is capable of hosting various operating systems and also to perform functionalities (e.g. routing).

To minimize 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 computers do not have virtualization assistance in hardware, more physical interfaces are installed. The infrastructure has been designed to favour testing of functionalities, protocols and new ideas, rather than providing a laboratory for very high performance studies.

Following the framework outlined above, FEDERICA is designed in two layers. The lower layer is the substrate, which is made of physical resources, both network and computing elements, each capable of creating "virtual" resources of their kind. The resource sets, or slices, managed by the user, compose the upper layer.

Given the sophisticated NREN network architecture, a distributed infrastructure can be engineered, with various Points of Presence on the top of the GÉANT [7] backbone, interconnecting several NRENs in Europe. Figure 1 shows pictorially the design of the infrastructure built on top of the existing NREN and GÉANT production environment. The virtual infrastructures (slices) are shown on the top of the picture. More than one slice is active at the same time.

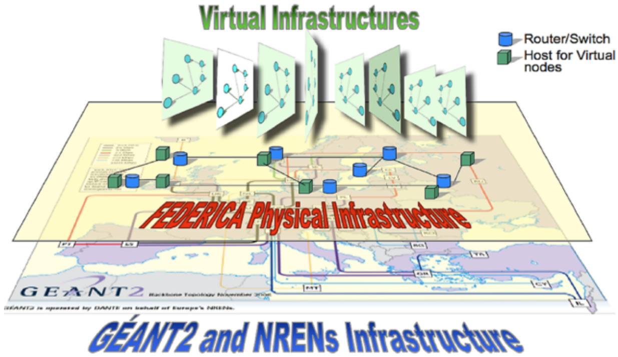

Figure 1: Pictorial representation of FEDERICA.

Figure 1 represents the slice in vertical format for the 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. The same physical node, as an example, can provide virtual systems to more than one slice. A virtual router can be created in a Juniper node (ensuring complete independence between the virtual routers) or by a virtual system running the routing suite.

3. The FEDERICA Infrastructure Implementation

The infrastructure is built using:

- A mesh of Ethernet circuits at 1 Gigabit per second provided by the GÉANT backbone. The circuits are initially at 1 Gbps as this capacity allows slicing to relatively high -speed links and yet is still affordable as contribution by the participating NRENs. Most of the circuits are created over SDH using generic framing procedure and virtual concatenation. Fig. 2 represents the current topology.
- Network equipment. Programmable high -end routers/switches: Juniper Networks MX480 with dual CPU and one line card with 32 ports at 1Gbps Ethernet. The MX functionalities include virtual and logical routing, MPLS, VLANs, IPv4, IPv6. The MX480 are installed in 4 core Points of Presence and 2 MX480 are equipped with Ethernet line cards with hardware QoS capabilities. Smaller multi-protocol switches (Juniper EX series) are installed in non core PoPs.
- Computing equipment. PC based nodes (V-Nodes) running virtualization software, capable of implementing e.g., open source software routers and emulating end user nodes. Each PC contains 2 x Quad core AMD running at 2 GHz, 32GB RAM, 8 network interfaces, 2x 500GB disks. The V-Nodes are connected to the Juniper routers.

The initial choice of the virtualization software for the V-Nodes is VMware [2], the free version of ESXi. This choice has been made after a review of other virtualization software (e.g. XEN). In particular the Application Programming Interface, the availability of usage examples and expertise and an upgrade path to better management using a commercial version of the software were important aspects of the evaluation process. The capabilities and performance of the free version have been adequate for the current requirements.

These building blocks of the substrate pose very few constraints to the user. In the current status of the infrastructure the most significant one is that the data link layer is fixed to Ethernet framing. Future development may permit access to optical equipment to overcome this limitation.

3.1. Topology

The topology is composed of 13 physical sites. Of these Points of Presence (PoP) a full mesh of 4 is equipped with MX router/switches and is considered as the core. The 9 non-core nodes are equipped by EX switches.

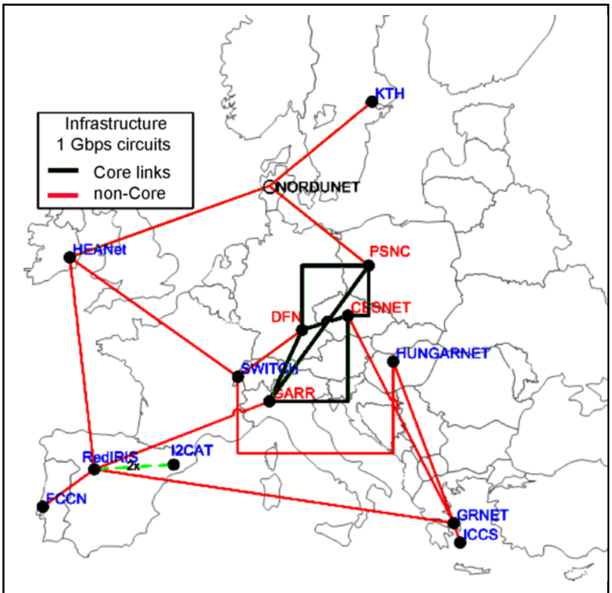

Fig. 2: FEDERICA topology on a map of Europe

The core nodes are equipped with 2 V-Nodes the non-core PoPs host 1 node each. The physical topology is depicted in Fig. 2. The design placed particular importance on the resiliency and load balancing of the network, based on GÉANT's infrastructure, and resources availability at partners' locations.

The substrate is configured as an IPv4 and IPv6 Autonomous System with both public and private addresses. The infrastructure is 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 infrastructure (substrate) is a single 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.

3.2. Resource Virtualization and Slice Creation

The process to create a virtual system is rather straightforward and can be based on an image provided by the user or on template of various operating systems. The virtualization capabilities in the network are also evolving, as described in [6]. The article reviews the current research in a Network Virtualization Environment (NVE) and the many challenges associated. The initial choice is to use VLANs and use QoS techniques for circuit virtualization; MPLS may be applied when needed.

The slice creation procedure definition is constantly developed and may change slightly to incorporate the feedback received after the first user feedback. The slice creation includes a manual step to map the virtual resources to the physical substrate. The step is manual to ensure that the mapping ensures the best reproducibility of the behaviour of the virtual resources.

The current slice creation process consists of the following steps. First, the researcher that wants to perform an experiment 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).

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. 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.

3.3. User Access And Support

When the NOC has finished the slice creation process, they inform the researchers that the slice is ready to use. In the example in Figure 3 the user has requested a simple slice consisting of two virtual servers connected through a Juniper logical router. The NOC has already setup these three resources, connected them through a VLAN (black line at the bottom of the Figure), instantiated the Virtual Slice Management Server and

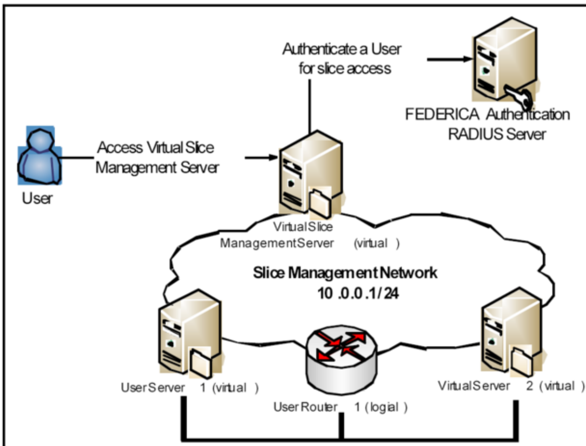

Fig. 3: Researcher accessing a FEDERICA slice

created the Slice Management Network (the cloud at the centre of the figure). The researcher connects to the Virtual Slice Management Server using the credentials provided by the NOC, and is authenticated against the FEDERICA Authentication RADIUS Server. If the authentication is successful, the user can access all his/her nodes via the management IP interfaces.

Besides remote access to the resources, another complimentary mechanism is under investigation. Access to VMware virtual machines can be granted through remote VNC connections (the virtual machine clients would connect to a special port of the physical machine where VMware is installed). 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; i.e. they would have full control of the virtual machine.

During the initial operations, all the steps explained 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).

However, a tool bench that provides a unified environment to operate the infrastructure and to use the slices is being developed, and will be progressively deployed and used by the NOC and the users.

4. An architecture for virtual infrastructures

The project's architecture described in section 2.2.2 represents an implementation of a domain based on virtualization and capable of a large set of services, similar to the ones provided by cloud computing and augmented to include wide area networks communication. The substrate creates the possibility to host many infrastructures, each configured differently and simultaneously active. Domains containing such infrastructures can connect between them at the substrate level to offer common services. Such infrastructure can also connect with domains not based on virtualization, like the current Internet. The interconnection can be done by the substrate and by each slice in parallel, in this case a careful planning of the traffic routing and addressing is needed. The flexibility in configuration of the virtual environments, usage optimization, scalability and resiliency of the infrastructure make it appealing for a large variety of uses.

5. Challenges

The initial list of challenges for virtual infrastructures which has been identified are:

- *The reproducibility and the stability of the behaviour of a single virtual resource*. In the case of a researcher experimenting on a virtual infrastructure, the behaviour of each virtual resource is key to allow reproducible experiments. Two independent causes may generate instability:
  - The sharing of the physical resource with other virtual resources
  - The virtualization technology itself, usually a layer placed between the physical resources and the virtual ones.

  Hardware assistance for virtualization has been introduced recently in computers to reduce these differences. Since 2005, the main CPU manufacturers have added

virtualization friendly extensions, in particular related to protection rings. Hardware enhancements have been introduced also in networking equipment to provide, as an example, queuing and forwarding of different classes of packet traffic in separate memory buffers.
Reproducibility represents a feature of the quality of the services offered based on virtual resources. Providing assurances on the behaviour of a single resource is possible through a careful engineering of the hardware and software, but there is the lack of a general set of rules and procedures. As an example, computing elements in FEDERICA have been chosen to provide specific functionalities in hardware, like virtualization aware CPUs. Some circuits are connected to network QoS capable line cards in the Juniper MX. In other cases, where hardware was not available, the physical resources have been adequately dimensioned, to avoid any overbooking and minimize contention. While it is then possible to create a slice with a set of resources, each with reproducibility guarantees, providing reproducibility to the global behaviour of the resources connected in a slices is much more complex and, the complexity increases with the number of resources involved. The classic problem of guaranteeing the end-to-end QoS of an IP flow exemplifies the issue. In case of virtual infrastructures, as in the case of Internet traffic, the project's experience is that most of the virtual infrastructures are not required to have strict reproducibility guarantees, and rather a "best effort" behaviour is sufficient for the researcher's scope.

- *Virtualization services definition in an infrastructure capable of both wide area network and computing virtualization.* In both the cloud computing environment and in the network environment the developments of services is proceeding. Effort on standardization of resource representation, management and control is needed and also a standard infrastructure description language. Such language can be an evolution, as an example, of the network description language effort [17]. Such developments open the possibility to federate infrastructures and to develop new business models featuring a user 's virtual infrastructure extending in more than one domain.
- *Complexity*. The complexity of the systems based on virtualization, in particular when coupling network and computing resources, increases fast with the increase of number of resources. In addition, the possibility of having recursive virtualization resources on the same physical resource generates additional challenges to the monitoring system, privacy and security model, and management and control architecture. The complexity may actually reduce the reliability and the quality of the system, increasing its operational cost for management and problem resolution. This challenge requires a focused effort on standardization in a short timeframe, to avoid un-compatible implementation of services, applications and architectures. It also requires a broader set of competences in selected experts to plan and operate such services.

6. Infrastructure relation with "cloud computing"

The generic term "clouds" is a rapidly developing paradigm shift. It is often associated to "computing" [4][5] and refers to a set of "services" of type application (e.g. Google apps), platform (e.g. Windows Azure) and infrastructure (e.g. Amazon EC2). These services are offered through Internet with a simple user interface and leverage virtualization technologies.

Those services can be offered by an infrastructure like FEDERICA fully based on virtualization. The more natural comparison is with an Infrastructure as a Service (IaaS) offering with which it shares a simple user interface, neutrality to user's operating system and application, scalability and parallelism in users' utilization.
There are also some noticeable differences in respect to current cloud computing offerings. In FEDERICA virtualization is available in all elements (including network ones, like switches and routers) and specific importance and careful engineering is placed on the reproducibility (QoS in brief) of the single elements behaviour. This will enhance the possible service offering and the overall capabilities of a slice as research infrastructure when and where needed. A strong coordination and collaboration for the definition of standards is also envisaged.

7. Conclusion

The use of virtualization technologies in Internet is increasing and expanding. The advantages of adding computing capabilities to the network infrastructure are evident when coupled with virtualization. Such an infrastructure based on virtualization in all of its components and built with both computing and network resources may represent a new architecture for parts of Internet, in particular it can provide an ideal environment for innovative research and rapid service development.
The virtual infrastructures created, or slices, may contain any combination of the basic, "raw" fundamental virtual resources in arbitrary topologies and hosting any operating system and application type. Such virtual infrastructures allow full control to their owner, may easily connect to other infrastructures and are configurable in a short amount of time. The physical infrastructure in the NRENs and the telecommunication environment are already capable of supporting virtualization in many of their components. The FEDERICA project is an example of the capabilities of current hardware and is researching on the many challenges of virtualization. The project will continue to support researchers and to develop the architecture of virtualization based infrastructures. A strong connection to the research and standardization in cloud computing is expected.

References

[1] Paul Barham et al, Xen and the Art of Virtualization, *SOSP'03*, October 19–22, 2003, Bolton Landing, New York. ACM 1-58113-757-5/03/001. http://www.cl.cam.ac.uk/research/srg/netos/papers/2003-xensosp.pdf

[2] VMware Inc. is a provider of virtualization software founded in 1998. http://www.vmware.com

[3] Kernel-based Virtual Machine (KVM), see http://www.linux-kvm.org/page/Documents

[4] For a definition of Cloud Computing, see NIST Definition of Cloud Computing, P. Mell, T. Grance, Version 15, 10-7-09, http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc

[5] *IEEE Computing now*, October 2009: System Virtualization, September 2009: Cloud Computing: Opportunities and Challenges, http://www.computer.org/portal/web/computingnow/articles

[6] Chowdhury, N. M. M. K., Boutaba, R.,: Network Virtualization: State of the Art and Research Challenges. *IEEE Communications Magazine*, July 2009, 20-26 (2009)

[7] GÉANT, the European Research and Education backbone, http://www.GÉANT.net

[8] FEDERICA, "Federated E-infrastructure DEdicated to Researchers Innovating in Computing Architectures", European Commission co-funded in the 7th Framework Work Programme, project n. RI-213107, http://www.fp7-federica.eu/
[9] P. Szegedi, S. Figuerola, M. Campanella, V. Maglaris, C. Cervello-Pastor: "With Evolution for Revolution: Managing FEDERICA for Future Internet Research", *IEEE Communications Magazine* Vol.47 No.7 pp. 34-39, July 2009
[10] M. Campanella, "The FEDERICA Project: creating cloud infrastructures", In *Proceedings of Cloudcomp 2009* (CloudComp), October 19-21, 2009, Munich, Germany. ISBN 978-963-9799-77-6, to be published in LNICST
[11] For a detailed analysis of NRENs in Europe and their role, see the documents (in particular the TERENA compendium) available at http://www.terena.org/publications/
[12] Internet2 Dynamic Circuit Network, at http://www.internet2.edu/network/dc/
[13] M. Campanella, R. Krzywania et al, Bandwidth on Demand Services for European Research and Education Networks, Proceeding of the 1st IEEE International Workshop on Bandwidth on Demand, Nov. 2006, San Francisco, USA, Volume 1, Nov. 2006 Page(s): 65 – 72
[14] B. Tierney, J. Boote, E. Boyd, A. Brown, M. Grigoriev, J. Metzger, M. Swany, M. Zekauskas, Yee-Ting Li, and J. Zurawski, "Instantiating a Global Network Measurement Framework", LBNL Technical Report LBNL-1452E, January 2009. See also http://www.perfsonar.net/publications.html
[15] Global Environment for Network Innovation, http://www.geni.net
[16] Future Internet Research and Experimentation, http://cordis.europa.eu/fp7/ict/fire/
[17] Freek Dijkstra, Bert Andree, Karst Koyman, Jeroen van der Ham and Cees de Laat. A Multi-Layer Network Model Based on ITU-T G.805 Computer Networks, Vol. 52, Issue 10, pp. 1927-1937, July 2008 and the web site http://www.science.uva.nl/research/sne/ndl/