A Cognitive Future Internet Architecture

Marco Castrucci[1], Francesco Delli Priscoli[1], Antonio Pietrabissa[1], and Vincenzo Suraci[2]

[1] University of Rome "La Sapienza", Computer and System Sciences Department
Via Ariosto 25, 00185 Rome, Italy
{castrucci, dellipriscoli, pietrabissa}@dis.uniroma1.it
[2] Università degli studi e-Campus
Via Isimbardi 10, 22060 Novedrate (CO), Italy
vincenzo.suraci@uniecampus.it

Abstract. This Chapter proposes a novel Cognitive Framework as reference architecture for the Future Internet (FI), which is based on so-called *Cognitive Managers*. The objective of the proposed architecture is twofold. On one hand, it aims at achieving a full interoperation among the different entities constituting the ICT environment, by means of the introduction of *Semantic Virtualization Enablers*, in charge of virtualizing the heterogeneous entities interfacing the FI framework. On the other hand, it aims at achieving an inter-network and inter-layer cross-optimization by means of a set of so-called *Cognitive Enablers*, which are in charge of taking consistent and coordinated decisions according to a fully cognitive approach, availing of information coming from both the transport and the service/content layers of all networks. Preliminary test studies, realized in a home environment, confirm the potentialities of the proposed solution.

Keywords: Future Internet architecture, Cognitive networks, Virtualization, Interoperation.

1 Introduction

Already in 2005, there was the feeling that the architecture and protocols of the Internet needed to be rethought to avoid Internet collapse [1]. However, the research on Future Internet became a priority only in the last five years, when the exponential growth of small and/or mobile devices and sensors, of services and of security requirements began to show that current Internet is becoming itself a bottleneck. Two main approach have been suggested and investigated: the radical approach [2], aimed at completely re-design the Internet architecture, and the evolutionary approach [3], trying to smoothly add new functionalities to the current Internet towards.

Right now, the technology evolution managed to cover the lacks of current Internet architecture, but, probably, the growth in Internet-aware devices and the always more demanding requirements of new services and applications will require radical architecture enhancements very soon. This statement is proved by the number of financed projects both in the USA and in Europe.

In the USA, there are significant initiatives. NeTS [4] (Networking Technology and Systems) was a program of the National Science Foundation (NSF) on networking research with the objectives of developing the technology advances required to build next generation networks and improve the understanding of large, complex and heterogeneous networks. The follow-up of NeTS, NetSE [5] proposes a clean-state approach to properly meet new requirements in security, privacy and economic sustainability. GENI [6] (Global Environment for Network Innovations) is a virtual laboratory for at scale experimentation of network science, based on a 40 Gbps real infrastructure. Stanford Clean Slate [7] proposes a disruptive approach by creating service platforms available to the research and user communities.

In Europe, Future Internet research has been included as one of the topics in FP6 and FP7. European initiatives appear less prone to a completely clean-state approach with respect of USA ones, and tries to develop platforms which support services and applications by utilizing the current Internet infrastructure. For instance, G-Lab [8] (Design and experiment the network of the future, Germany), is the German national platform for Future Internet studies, includes both research studies of Future Internet technologies and the design and setup of experimental facilities. GRIF [9] (Research Group for the Future Internet, France) and Internet del Futuro [10] (Spain) promotes cooperation based on several application areas (e.g., health) and technology platforms. FIRE [11] is an EU initiative aimed at the creation of an European Experimental Facility, which is constructed by progressively connecting existing and upcoming testbeds for Future Internet technologies.

The contribution of this Chapter is the proposal of a Future Internet architecture which seamlessly cope with the evolutionary approach but is also open to innovative technologies and services. The main idea is to collect and elaborate all the information coming from the whole environment (i.e., users, contents, services, network resources, computing resources, device characteristics) via virtualization and data mining functionalities; the metadata produced in this way are then input of intelligent cognitive modules which provide the applications/services with the required functionalities in order to maximize the user Quality of Experience with the available resources.

The Chapter is organized as follows: Section 2 is devoted to the description of the concepts underlying the proposed architecture; Section 3 describes the Future Internet platform in detail; experimental results showing the potential of the platform are described in Section 4; finally, Section 5 draws the conclusions.

2   Architecture Concept

A more specific definition of the entities involved in the Future Internet, as well as of the Future Internet target, can be as follows:

- Actors represent the entities whose requirement fulfillment is the goal of the Future Internet; for instance, Actors include users, developers, network providers, service providers, content providers, etc.;

- Resources represent the entities that can be exploited for fulfilling the Actors' requirements; example of Resources include services, contents, terminals, devices, middleware functionalities, storage, computational, connectivity and networking capabilities, etc.;
- Applications are utilized by the Actors to fulfill their requirements and needs exploiting the available resources.

In the authors' vision, the Future Internet target is to allow Applications to transparently, efficiently and flexibly exploit the available Resources, thus allowing the Actors, by using such Applications, to fulfill their requirements and needs. In order to achieve this target, the Future Internet should overcome the following main limitations.

(i) A first limitation is inherent in the traditional layering architecture which forces to keep algorithms and procedures, laying at different layers, independent one another. In addition, even in the framework of a given layer, algorithms and procedures dealing with different tasks are often designed independently one another. These issues greatly simplify the overall design of the telecommunication networks and greatly reduce processing capabilities, since the overall problem of controlling the telecommunication network is decoupled in a certain number of much simpler sub-problems. Nevertheless, a major limitation of this approach derives from the fact that algorithms and procedures are poorly coordinated one another, impairing the efficiency of the overall telecommunication network control. The issues above claim for a stronger coordination between algorithms and procedures dealing with different tasks.

(ii) A second limitation derives from the fact that, at present, most of the algorithms and procedures embedded in the telecommunication networks are open-loop, i.e. they are based on off-line "reasonable" estimation of network variables (e.g. offered traffic), rather than on real-time measurements of such variables. This limitation is becoming harder and harder, since the telecommunication network behaviours, due to the large variety of supported services and the rapid evolution of the service characteristics, are becoming more and more unpredictable. This claims for an evolution towards closed-loop algorithms and procedures which are able to properly exploit appropriate real-time network measurements. In this respect, the current technology developments, which assure cheap and powerful sensing capabilities, favours this kind of evolution.

(iii) The third limitation derives from the large variety of existing heterogeneous Resources which have been developed according to different heterogeneous technologies and hence embedding technology-dependent algorithms and procedures, as well as from the large variety of heterogeneous Actors who are playing in the ICT arena. In this respect, the requirement of integrating and virtualizing these Resources and Actors so that they can be dealt with in an homogeneous and virtual way by the Applications, claims for the design of a technology-independent, virtualized framework; this framework, on the one hand, is expected to embed algorithms and procedures which, leaving out of consideration the specificity of the various networks, can be based on abstract advanced methodologies and, on the other hand, is expected to be provided with proper virtualizing interfaces which allow all Applications to benefit from the functionalities offered by the framework itself.

The concept behind the proposed Future Internet architecture, which aims at overcoming the three above-mentioned limitations, is sketched in Fig. 1. As shown in the figure, the proposed architecture is based on a so-called "Cognitive Future Internet Framework" (in the following, for the sake of brevity, simply referred to as "Cognitive Framework") adopting a modular design based on middleware "enablers". The enablers can be grouped into two categories: the *Semantic Virtualization Enablers* and the *Cognitive Enablers*. The Cognitive Enablers represent the core of the Cognitive Framework and are in charge of providing the Future Internet control and management functionalities. They interact with Actors, Resources and Applications through *Semantic Virtualization Enablers.*

The Semantic Virtualization Enablers are in charge of virtualizing the heterogeneous Actors, Resources and Applications by describing them by means of properly selected, dynamic, homogeneous, context-aware and semantic aggregated metadata.

The Cognitive Enablers consist of a set of modular, technology-independent, interoperating enablers which, on the basis of the aggregated metadata provided by the Semantic Virtualization Enablers, take consistent control and management decisions concerning the best way to exploit and configure the available Resources in order to efficiently and flexibly satisfy Application requirements and, consequently, the Actors' needs. For instance, the Cognitive Enablers can reserve network resources, compose atomic services to provide a specific application, maximize the energy efficiency, guarantee a reliable connection, satisfy the user perceived quality of experience and so on.

The control and management decisions taken by the Cognitive Enablers are handled by the Semantic Virtualization Enablers, in order to be actuated involving the proper Resources, Applications and Actors.
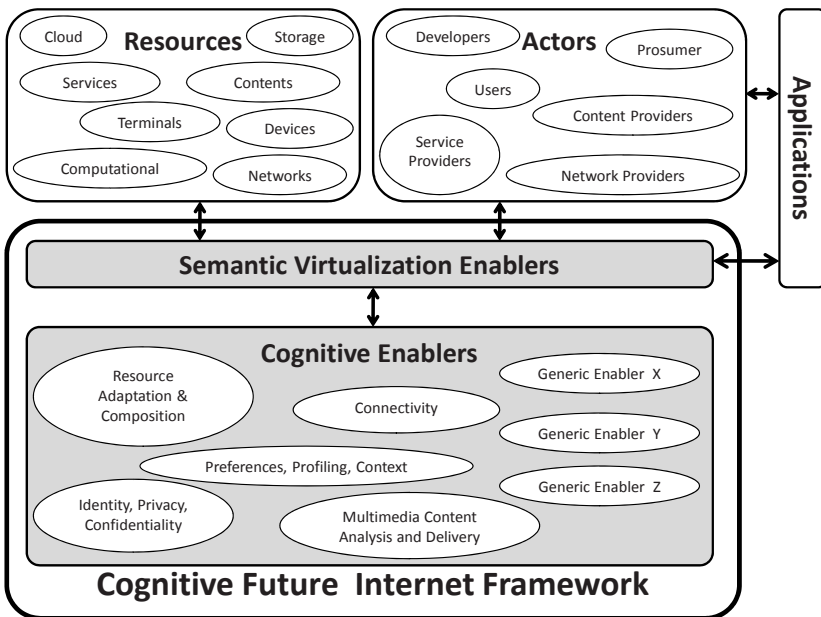

Fig. 1. Proposed Cognitive Future Internet Framework conceptual architecture

Note that, thanks to the aggregated semantic metadata provided by the Semantic Virtualization Enablers, the control and management functionalities included in the Cognitive Enablers have a technology-neutral, multi-layer, multi-network vision of the surrounding Actors, Resources and Applications. Therefore, the information enriched (fully cognitive) nature of the aggregated metadata, which serve as Cognitive Enabler input, coupled with a proper design of Cognitive Enabler algorithms (e.g., multi-objective advanced control and optimization algorithms), lead to cross-layer and cross-network optimization.

The Cognitive Framework can exploit one or more of the Cognitive Enablers in a dynamic fashion: so, depending on the present context, the Cognitive Framework activates and properly configures the needed Enablers.

Furthermore, in each specific environment, the Cognitive Framework functionalities have to be properly distributed in the various network entities (e.g. Mobile Terminals, Base Stations, Backhaul network entities, Core network entities). The selection and the mapping of the Cognitive Framework functionalities in the network entities is a critical task which has to be performed case by case by adopting a transparent approach with respect to the already existing protocols, in order to favour a smooth migration.

It should be evident that the proposed approach allows to overcome the three above-mentioned limitations:

(i) Concentrating control and management functionalities in a single Cognitive Framework makes much easier to take consistent and coordinated decisions. In particular, the concentration of control functionalities in a single framework allows the adoption of algorithms and procedures coordinated one another and even *jointly* addressing in a one-shot way, problems traditionally dealt with in separate and uncoordinated fashion.

(ii) The fact that control decisions can be based on properly selected, aggregated metadata describing, in real time, Resources, Actors and Applications allows closed-loop control, i.e. networks become *cognitive*. In particular, the Cognitive Enablers can, potentially, perform control elaborations availing of information coming from all the layers of the protocol stack and from all networks. Oversimplifying, according to the proposed *fully cognitive* approach, potentially, all layers benefit from information coming from all layers of all networks, thus allowing to perform a full cross-layer, cross-network optimization.

(iii) Control decisions, relevant to the best exploitation of the available Resources, can be made in a technology independent and virtual fashion, i.e. the specific technologies and the physical location behind Resources, Actors and Applications can be left out of consideration. In particular, the decoupling of the Cognitive Framework from the underlying technology transport layers on the one hand, and from the specific service/content layers on the other hand, allows to take control decisions at an abstract layer, thus favouring the adoption of advanced control methodologies (e.g. constrained optimization, adaptive control, robust control, game theory...) which can be closed-loop thanks to the previous issue. In addition, interoperation procedures among heterogeneous Resources, Actors and Applications become easier and more natural.

3 Cognitive Future Internet Framework Architecture

The Cognitive Framework introduced in the previous section consists of a conceptual framework that can be deployed as a distributed functional framework. It can be realized through the implementation of appropriate Cognitive Middleware-based Agents (in the following referred to as Cognitive Managers) which will be transparently embedded in appropriate network entities (e.g. Mobile Terminals, Base Stations, Backhaul Network entities, Core Network entities). There not exist a unique mapping between the proposed conceptual framework over an existing telecommunication network. However we proposed a proof-of-concept concrete scenario in section 4, where the conceptual framework has been deployed in a real home area network test case. Indeed the software nature of the Cognitive Manager allows a transparent integration in the network nodes. It can be deployed installing a new firmware or a driver update in each network element. Once the Cognitive Manager is executed, that network node is enhanced with the Future Internet functionalities and become part of the Future Internet assets.

Fig. 2 outlines the high-level architecture of a generic Cognitive Manager, showing its interfacing with Resources, Actors and Applications.

Fig. 2 highlights that a Cognitive Manager will encompass five high-level modular functionalities, namely the Sensing, Metadata Handling, Elaboration, Actuation and API (Application Protocol Interface) functionalities. The Sensing, Actuation and API functionalities are embedded in the equipment which interfaces the Cognitive Manager with the Resources (*Resource Interface*), with the Actors (*Actor Interface*) and with the Applications (*Application Interface*); these interfaces must be tailored to the peculiarities of the interfaced Resources, Actors and Applications.
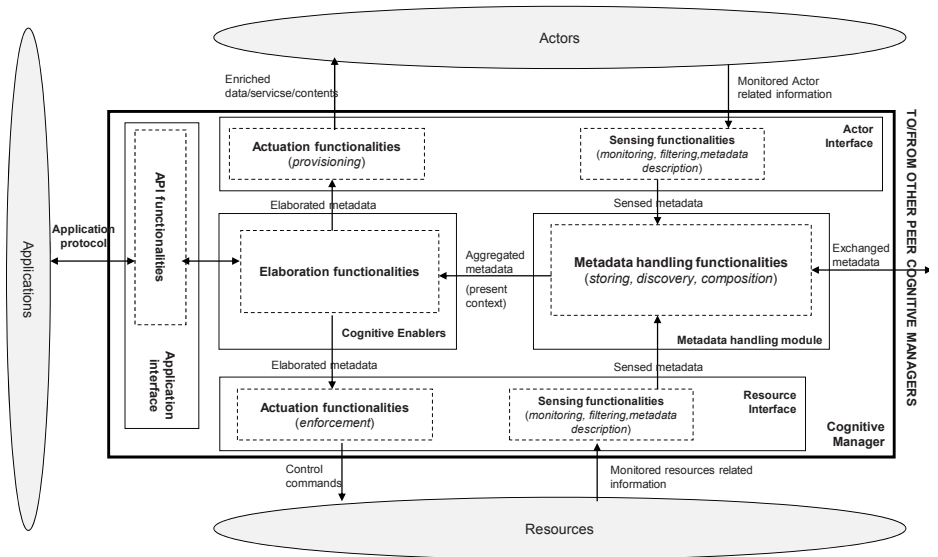

Fig. 2. Cognitive Manager architecture

The Metadata Handling functionalities are embedded in the so-called *Metadata Handling module*, whilst the *Elaboration functionalities* are distributed among a set of *Cognitive Enablers*. The Metadata Handling and the Elaboration functionalities (and in particular, the Cognitive Enablers which are the core of the proposed architecture) are independent of the peculiarities of the surrounding Resources, Actors and Applications.

With reference to Fig. 2, the Sensing, Metadata Handling, Actuation and API functionalities are embedded in the Semantic Virtualization Enablers, while the Elaboration functionalities are embedded in the Cognitive Enablers. The roles of the above-mentioned functionalities are the following.

Sensing functionalities are in charge of (i) the *monitoring* and preliminary *filtering* of both Actor related information coming from service/content layer (Sensing functionalities embedded in the Actor Interface) and of Resource related information (Sensing functionalities embedded in the Resource Interface); this monitoring has to take place according to transparent techniques, (ii) the formal description of the above-mentioned heterogeneous parameters/data/services/contents in homogeneous metadata according to proper ontology based languages (such as OWL – Web Ontology Language).

Metadata Handling functionalities are in charge of the storing, discovery and composition of the metadata coming from the sensing functionalities and/or from metadata exchanged among peer Cognitive Managers, in order to dynamically derive the aggregated metadata which can serve as inputs for the Cognitive Enablers; these aggregated metadata form the so-called *Present Context*; it is worth stressing that such Present Context has an highly dynamic nature.

Elaboration functionalities are embedded in a set of Cognitive Enablers which, following the specific application protocols and having as key inputs the aggregated metadata forming the Present Context, produce *elaborated metadata* aiming at (i) controlling the Resources, (ii) providing enriched data/services/contents to the Actors. In addition, these enablers control the sensing, metadata handling, actuation and API functionalities (these control actions, for clarity reasons, are not represented in Fig. 2).

Actuation functionalities are in charge of (i) actuation of the Cognitive Enabler control decisions over the Resources (*Enforcement functionalities* embedded in the Resource Interface; see Fig. 2); the decision enforcement has to take place according to transparent techniques, (ii) provisioning to the appropriate Actors the enriched data/contents/services produced by the Cognitive Enablers (*Provisioning functionalities* embedded in the Actor Interface; see Fig. 2).

Finally, *API functionalities* are in charge of interfacing the protocols of the Applications managed by the Actors with the Cognitive Enablers.

A so-called *Supervisor and Security Module* (not shown for clarity reason in Fig. 2) is embedded in each Cognitive Manager supervising the whole Cognitive Manager and, at the same time, assuring the overall security of the Cognitive Manager itself (e.g., including end-to-end encryption, Authentication, Authorization and Accounting (AAA) at user and device level, Service Security, Intrusion Detection, etc.). Another key role of this module is to dynamically decide, consistently with the application protocols, the Cognitive Manager functionalities which have to be activated to handle

the applications, as well as their proper configuration and activation/deactivation timing.

The proposed approach and architecture have the following key efficiency and flexibility advantages which are hereinafter outlined in a qualitative way:

Advantages Related to Efficiency

(1) The Present Context, which is the key input to the Cognitive Enablers, includes multi-Actor, multi-Resource information, thus potentially allowing to perform the Elaboration functionalities availing of a very "rich" feedback information.

(2) The proposed architecture (in particular, the technology independence of the Elaboration functionalities, as well as the valuable input provided by the Present Context) allows to take all decisions in a cognitive, abstract, coordinated and co-operative fashion within a set of strictly cooperative Cognitive Enablers. The concentration of the control functionalities in such Cognitive Enablers allows the adoption of multi-object algorithms and procedures which *jointly* address problems traditionally dealt with in a separate and uncoordinated fashion at different layers. So, the proposed architecture allows to pass from the traditional layering approach (where each layer of each network takes uncoordinated decisions) to a scenario in which, potentially, all layers of all networks benefit from information coming from all layers of all networks, thus, potentially, allowing a full cross-layer, cross-network optimization.

(3) The rich feedback information mentioned in the issue (1), together with the technology independence mentioned in the issue (2), allow the adoption of inno-vative and abstract *closed-loop* methodologies (e.g. constrained optimization, data mining, adaptive control, robust control, game theory, operation research, etc.) for the algorithms and rules embedded in the Cognitive Enablers, which are expected to remarkably improve efficiency.

Advantages Related to Flexibility

(4) Thanks to the fact that the Cognitive Managers have the same architecture and work according to the same approach regardless of the interfaced heterogeneous Applications/Resources/Actors, interoperation procedures become easier and more natural.

(5) The transparency and the middleware (firmware based) nature of the proposed Cognitive Manger architecture makes relatively easy its embedding in any fixed/mobile network entity (e.g. Mobile Terminals, Base Station, Backhaul network entities, Core network entities): the most appropriate network entities for hosting the Cognitive Managers have to be selected environment by envi-ronment. Moreover, the Cognitive Managers functionalities (and, in particular, the Cognitive Enabler software) can be added/upgraded/deleted through remote (wired and/or wireless) control.

(6) The modularity of the Cognitive Manager functionalities allows their ranging from very simple SW/HW/computing implementations, even specialized on a single-layer/single-network specific monitoring/elaboration/actuation task, to

complex multi-layer/multi-network/multi-task implementations. In particular, for each Cognitive Manger, the relevant Actuation/Sensing functionalities, the aggregated information which form the Present Context, as well as the relevant Elaboration functionalities have to be carefully selected environment-by-environment, trading-off the advantages achieved in terms of efficiency with the entailed additional SW/HW/computation complexity.

(7) Thanks to the flexibility degrees offered by issues (4)-(6), the Cognitive Managers could have the same architecture regardless of the interfaced Actors, Resources and Applications. So, provided that an appropriate tailoring to the considered environment is performed, the proposed architecture can actually scale from environments characterized by few network entities provided with high processing capabilities, to ones with plenty of network entities provided with low processing (e.g. Internet of Things).

(8) The above-mentioned flexibility issues favours a *smooth migration* towards the proposed approach. As a matter of fact, it is expected that Cognitive Manager functionalities will be gradually inserted starting from the most critical network nodes, and that control functionalities will be gradually delegated to the Cognitive Modules.

Summarizing the above-mentioned advantages, we propose to achieve Future Internet revolution through a smooth evolution. In this evolution, Cognitive Managers provided with properly selected functionalities are gradually embedded in properly selected network entities, aiming at gradually replacing the existing open-loop control (mostly based on traditional uncoordinated single-layer/single-network approachs), with a cognitive closed-loop control trying to achieve cross-optimization among heterogeneous Actors, Applications and Resources. Of course, careful, environment-by-environment selection of the Cognitive Manager functionalities and of the network entities in which such functionalities have to be embedded, is essential in order to allow scalability and to achieve efficiency advantages which are worthwhile with respect to the increased SW/HW/computing complexity.

The following section shows an example of application of the above-mentioned concepts. Much more comprehensive developments are being financed in various frameworks (EU and national projects), which are expected to tailor the presented approach to different environments aiming at assessing, in a quantitative way, the actual achieved advantages in terms of flexibilty (scalability) and efficiency; nevertheless, in the authors' vision such advantages are already evident, in a qualitative way, in the concepts and discussions presented in this section.

4   Experimental Results

The proposed framework has been tested in a home scenario for a preliminary proof-of-concept, but the same results can be obtained even in wider scenarios involving also Access Networks and/or Wide Area Networks. We consider a hybrid home network, where connectivity among devices is provided using heterogeneous wireless (e.g., WiFi, UWB) and wired (e.g., Ethernet, PLC) communication technologies. For

testing purposes, only a simplified version of the Cognitive Manager has been implemented in each node of the network, which includes the following functionalities:

- the *Service and Content adapter*: a QoS adapter module has been implemented, able to acquire information about the characteristics of the flow that has to be transmitted over the network, in terms of Traffic Specifications (TSpec) and QoS requirements, and to map them into pre-defined flow identifiers;
- the *Command and measurement adapter*: a Monitoring Engine has been implemented in order to acquire information about the topology of the network and the status of the links, while an Actuator module has been used to enforce elaboration decision over the transport network, in particular in order to modify the forwarding table used by the node to decide the network interface to be used for the transmission of the packets;
- the *Metadata handling storing functionality*: all the heterogeneous information collected by the Service/content adapter and by the Command and measurement adapter are translated using a common semantic and stored in proper database, ready to be used by elaboration functionalities;
- a *Cognitive connectivity enabler*: it has been implemented to perform technology independent resource management algorithms (e.g., layer 2 path selection), in order to guarantee that flow's QoS requirements are satisfied during the transmission of its packets over the network. In particular, a Connection Admission Control algorithm, a Path selection algorithm and a Load Balancing algorithm has been considered in our tests.

The framework has been implemented as a Linux Kernel Module and it has been installed in test-bed machines and in a legacy router[1] for performance evaluation. Fig. 3 shows three nodes connected together by means of a IEEE 802.11n link at 300 Mbit/s, and two IEEE 802.3u links at 100 Mbit/s.
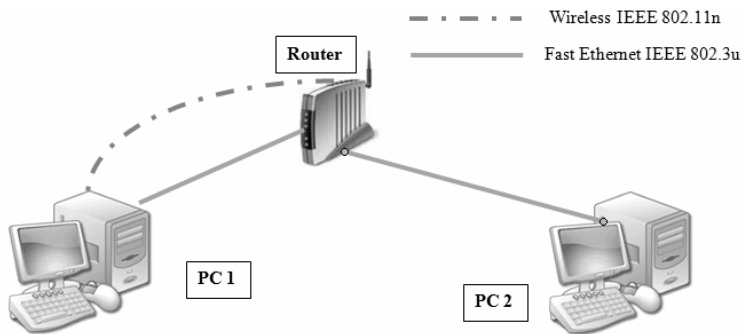

Fig. 3. Test scenario

To test the technology handover performances a FTP download session (file size 175 MB) has been conducted on the Ethernet link. After approximately 10s, one extremity of the Ethernet cable has been physically disconnected from its socket and the flow has been automatically redirected onto the wireless link thanks a context-aware decision taken by the Cognitive connectivity enabler. Switching on the Wi-Fi link causes more TCP retransmissions and an increased transfer time. This is natural, since Ethernet and Wi-Fi have different throughputs. Without the cognitive framework, it is evident that the FTP session would not be terminated at all. As shown in Fig. 4, the experimented handover time is around 240 ms, during which no packet is received. The delay is influenced by the processing time that the framework module spends in triggering and enforcing the solutions evaluated by the path selection routines implemented in the cognitive connection enabler.
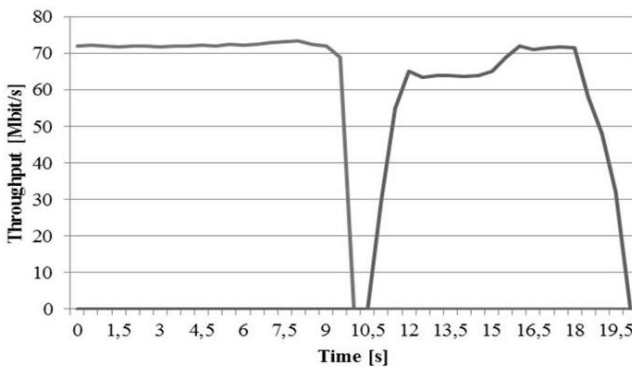

Fig. 4. Technology handover

5   Conclusions

This paper proposes a novel reference architecture for the Future Internet, with the aim to provide a solution to overcome current Internet limitations. The proposed architecture is based on Cognitive Modules which can be transparently embedded in selected network entities. These Cognitive Modules have a modular organization which is claimed to be flexible and scalable, thus allowing a smooth migration towards the Future Internet and, at the same time, allowing to implement only the needed functionalities in a give scenario. Interoperation among heterogeneous entities is achieved by means of their virtualization, obtained thanks to the introduction of Semantic Virtualization Enablers. At the same time, the Cognitive Enablers, which are the core of the Cognitive Managers, can *potentially* benefit from information coming from all layers of all networks and can take consistent and coordinated context-aware decisions impacting on all layers. Clearly, which Cognitive Enabler have to be activated, which input information has to be provided to the Cognitive Enabler, the algorithms the Cognitive Enabler will be based on, have all to be carefully selected case by case; nevertheless, the proposed architecture has an inherent formidable point of strength in the concentration of all management and control tasks in a single technology/service/content independent layer, opening the way, in a natural fashion, to inter-network, inter-layer cross-optimizations.

References

1. Talbot, D.: The Internet is broken, Technology Review, December 2005-January 2006 (2006), `http://www.technologyreview.com/article/16356/`
2. FISS: Introduction to content centric networking, Bremen, Germany, 22 June (2009), `http://www.comnets.uni-bremen.de/typo3site/uploads/media/vjCCN-FISS09.pdf`
3. Miller., R.: Vint Cerf on the Future of the Internet. The Internet Today, The Singularity University (2009), `http://www.datacenterknowledge.com/archives/2009/10/12/vint-cerf-on-the-future-of-the-internet/`
4. National Science Foundation: Networking Technology and Systems, NeTS (2008), `http://www.nsf.gov/pubs/2008/nsf08524/nsf08524.htm`
5. National Science Foundation: Network Science and Engineering, NetSE (2010), `http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503325`
6. BNN Technologies: GENI: Exploring networks of the future, `http://www.geni.net/` (2010)
7. Stanford Clean Slate: `http://cleanslate.stanford.edu/` (2011)
8. G-Lab: `http://www.german-lab.de/home/` (2008)
9. Jutand, F.: National Future Internet Initiatives - GRIF (France), `http://www.francenumerique2012.fr/` (2010)
10. AETIC: Internet del Futuro, `http://www.idi.aetic.es/esInternet/` (2008)
11. ICT FP7 Research: Future Internet Research & Experimentation (FIRE), `http://cordis.europa.eu/fp7/ict/fire/` (2010)