A Network Control Plane architecture for on-demand co-provisioning of optical network and IT services

Giada LANDI[1], Nicola CIULLI[1], Jens BUYSSE[2], Konstantinos GEORGAKILAS[3], Markos ANASTASOPOULOS[3], Anna TZANAKAKI[3], Chris DEVELDER[2], Eduard ESCALONA[4], Damian PARNIEWICZ[5], Artur BINCZEWSKI[5], Maciej STROIÑSKI[5]

[1]*Nextworks, via Livornese 1027, Pisa, 56122, Italy*

Tel: +39 050 3871600, Fax: +39 050 3871601, [g.landi, n.ciulli]@nextworks.it

[2]*IBBT-Ghent University (IBCN), Gaston Crommenlaan 8, B-9050 Gent, Belgium*

Tel: +32 093314900, Fax: +32 093314899, [jens.buysse,chris.develder]@intec.ugent.be

[3]*Athens Information Technology, Markopoulou Avenue, Athens, 19002, Greece*

Tel: +302106682766, Fax: +302106682703, [koge, manast, atza]@ait.gr

[4]*University of Essex, Wivenhoe Park, CO4 3SQ, United Kingdom*

Tel: +44 1206874233, Fax: + 44 1206872900, eescal@essex.ac.uk

[5]*PSNC, ul. Noskowskiego 10, Poznan, 61-704, Poland*

Tel: +48 618582001, Fax: +48618525954, [damianp, artur]@man.poznan.pl

Abstract: The convergence between IT and optical network services is a fundamental step to support new emerging applications, typically distributed and with strict requirements in terms of performance and service reliability. The FP7 GEYSERS project has designed a new network architecture capable of joint and on-demand provisioning of 'Optical Network + Any-IT' resources for end-to-end service delivery. This service, referred to as Network + IT Provisioning Service (NIPS), is enabled through the cross-layer cooperation between the Service Middleware Layer (SML) and an enhanced Network Control Plane (NCP+) along the overall service lifecycle. This paper describes the GEYSERS NCP+ architecture that extends standard ASON/GMPLS and PCE solutions to operate over a virtual optical network infrastructure and offer on-demand provisioning of enhanced network transport services between virtual IT resources connected to the network edges. In the proposed architecture, some of the functionalities traditionally provided by the SML, like the selection of the IT end-points, are moved to the NCP+ side, allowing global optimization of the utilization of network and IT resources and of energy consumption. Simulation results are provided to evaluate the benefits of the combined computation of IT end-points and network paths, both in terms of reduced service blocking probability and utilization of the infrastructure resources.

Keywords: Optical network and IT convergence, Network Control Plane, end-to-end network + IT provisioning, dynamic infrastructures, Future Internet, PCE

1. Introduction

Next generation applications require direct access to IT resources with high storage and computation capabilities. These applications are typically data intensive and introduce huge amounts of data traffic that need to be processed. However, the required computational and storage resources cannot always be located where the data is generated or where the application user needs it. The concept of clouds and data centres arose to cope with these application requirements, but to transport the huge amounts of data in efficient way, a dynamic underlying optical network, equipped with advanced control mechanisms, is required.

Typically, IT services and network services are disassociated and provisioned separately. The lack of converged service provisioning and management leads to unexpected incompatibilities, bottlenecks and costs. Integrated procedures to jointly establish IT and network services allow performing the end-to-end provisioning seamlessly, in a single step upon a single request from the service consumer. This approach creates the potential to adopt unified algorithms for discovering the best combination of IT and network resources, where the metrics and constraints to be applied are determined by the dynamic requirements of the applications. Moreover, having a single step end-to-end provisioning, significantly reduces the setup time and the probability of finding outdated or inaccurate resource availability information.

In this paper we describe the GEYSERS [1] approach, which takes on the challenge of defining converged advanced provisioning mechanisms to maintain the Quality of Service (QoS) associated with aggregated services composed of networking and IT resources. In particular, we present the architecture of the GEYSERS Network Control Plane (NCP+), which operates over a Virtual Infrastructure (VI) and allows network service providers to offer customized connectivity services automatically tailored to the application requirements. The network and IT convergence is enabled through the cooperation between the NCP+ and the Service Middleware Layer (SML), where in the most advanced scenarios of the anycast services, the NCP+ is responsible for the joint selection of network and IT resources. In this paper we evaluate the advantages of this combined computation through a set of simulations that compare the performances of unicast and anycast services in terms of service blocking probability and resource utilization.

The paper is structured as follows: section 2 describes the architecture of the NCP+ and its positioning in the GEYSERS architecture. Section 3 analyzes the performance of the anycast services through simulations. Section 4 describes the experimental validation of the NCP+ prototype in the GEYSERS test-bed, while section 5 presents the conclusions.

2. The GEYSERS Network Control Plane architecture (NCP+)

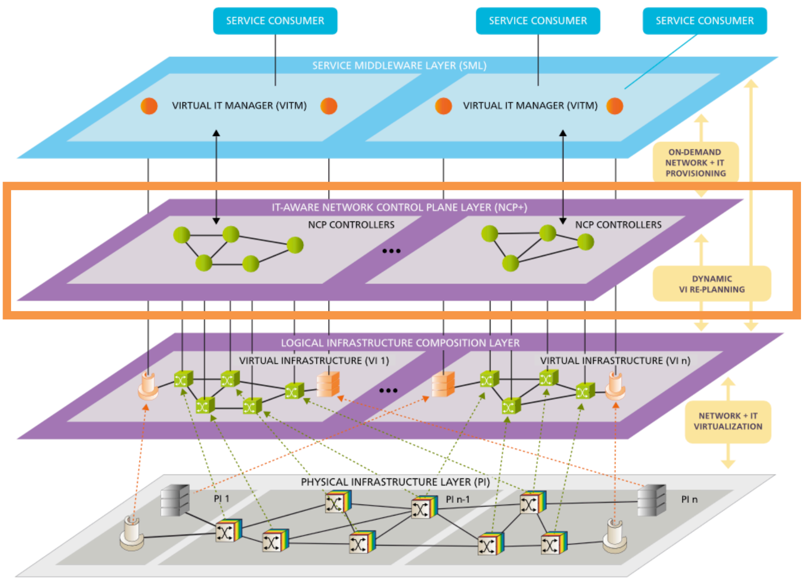

Figure 1 GEYSERS architecture

The GEYSERS NCP+ is an ASON/GMPLS (Automatically Switched Optical Network / Generalized Multiprotocol Label Switching) and PCE-based (Path Computation Element) control plane that operates over a virtual optical network infrastructure, accessed through the Logical Infrastructure Composition Layer (LICL). The LICL performs an abstraction of the physical resources, allowing the NCP+ to control the virtual nodes (VNs) through a homogeneous interface that hides specificities and constraints of the heterogeneous multi-vendor devices of the physical infrastructure. On the other hand, the information model at NCP-to-LICL interface offers a powerful description of the VNs, as required by the NCP+ to efficiently control the VI. For example, the LICL exposes the switching technology details of the VNs, but following a uniform model. Moreover, abstract "green" parameters describing the power consumption of the VI are provided to the NCP+ and used to provide energy efficient services.

The NCP+ is extended to:

- offer on-demand provisioning of enhanced network transport services connecting IT virtual resources located at the network edges, in support of joint Network + IT Provisioning Services (NIPS);
- trigger the automatic re-planning of the VI to optimize the resource utilization in the medium and long term and adapt the VI size to the actual traffic load.

The NCP+ positioning and key functions in the overall GEYSERS architecture [2] are depicted in Figure 1 and further elaborated in the next sections.

2.1 – Cross-layer cooperation for network and IT provisioning: the NIPS UNI

The NIPS is key for the convergence between IT and optical networks and requires the active interaction between NCP+ and SML, that controls the whole IT services and the IT virtual resources. This interaction is enabled through a Service-to-Network interface, called NIPS UNI (Figure 2). It allows joint optimization of resource provisioning in IT and optical segments improving the entire VI utilization in many perspectives, e.g. in terms of energy efficiency, service availability and resiliency or combined allocation of network and IT resources. The cross-layer, on-demand approach for end-to-end service delivery adapts resource provisioning to the application dynamicity, guaranteeing QoS and reliability.

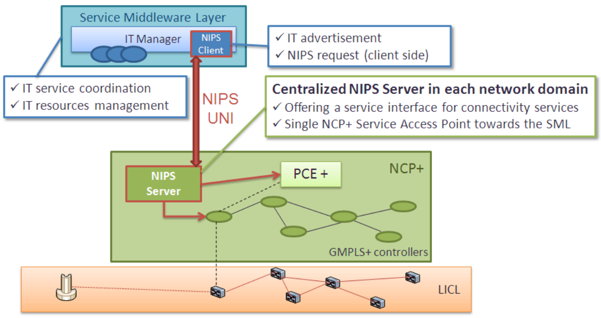

Figure 2 NIPS UNI

In traditional approaches the SML takes autonomous decisions on the IT end-points to be used for a given service, considering only the capabilities of the IT resources available in the controlled data centres. In GEYSERS, these decisions are improved taking into account also network related information received from the NCP+ or are fully delegated to the NCP+ itself. The following types of connectivity services are supported:

- ***Assisted unicast services***, where the NCP+ provides network service quotations for possible combinations of IT end-points in terms of network-related prices or performances, and the SML is responsible for the final end-points selection.
- ***Restricted anycast services***, where the SML specifies a set of candidate destination end-points, equivalent from the IT perspective, and the NCP+ takes the final decision about the destination, according to network constraints specified in the request.

- ***Full anycast services***, where the SML specifies the end-points only in terms of IT characteristics required for the complete service (e.g. amount of storage), while the NCP+ discovers potential end-points and selects the optimal set of resources.

The NIPS UNI is a REST (Representational State Transfer) interface, where the SML acts as client. The NCP+ exposes a centralized access point for the NIPS UNI services through a NIPS Server. The main services offered over the NIPS UNI are the following:

- ***IT resource advertisement,*** that allows the NCP+ to know the capabilities and availabilities of the IT resources. This information is used in anycast services, when the selection of the IT end-points is partially or fully delegated to the NCP+.
- ***NIPS request,*** used by the SML to dynamically request the setup, tear-down or modification of transport network connectivity services associated with an IT service.
- ***NIPS monitoring,*** with synchronous or asynchronous notifications for the SML about status or performance of the underlying network services. This data is used to assess the status of the service and its compliancy with the given Service Level Agreement (SLA). Recovery procedures with cross-layer escalation strategies are also supported.

2.2 – NCP+ core functionalities

The main NCP+ functionalities are the NIPS control procedures for service setup, tear-down, monitoring and recovery, and the re-planning of the virtual network infrastructure.

NIPS control. The NIPS is managed at the NCP+ through the *NIPS call* concept, which is an extension of the ITU-T G.8080 ASON call [3] which involves two IT end-points that can be explicitly defined or declared implicitly through the description of the IT resources (anycast service). The NIPS call consists of multiple call segments with connections in each domain traversed by the NIPS. The NCP+ adopts a logical separation between call and connection control. The call is managed by Network Call Controllers (NCC) located on the GMPLS+ controllers at the domain boundaries, while connections are managed by all the GMPLS+ controllers within a single domain. NCP+ signalling adopts the RSVP-TE protocol, extended to carry the specification of the IT resources in full anycast NIPS calls. The end-to-end call path is determined at the first NCC, through a path computation request to the domain's local PCE+.

Advance reservations. A NIPS with advance reservation allows the scheduling of a service and the associated resource reservation in a given future time-slot. During the setup phase, the resources are reserved in advance for their usage in the specified time interval and they are automatically activated at the service start-time. Advance reservations are managed through a set of calendars, maintained and updated at the PCE+, that describe the resource bookings in the future time-slots.

Virtual Infrastructure re-planning. A new concept introduced by GEYSERS is the possibility to resize a VI any point during its lifecycle according to the evolving dynamics in the traffic load or in support of a growing business. The NCP+ is capable of detecting inefficient conditions in the infrastructure provisioning taking into account forecasts and statistics about service requests and network resource availability. Consequently it can act as an automatic trigger for the dynamic re-planning procedures in order to optimize the resource utilization in the medium and long period.

Re-planning detection/triggering and management are split between PCE+ and VRM (VI Re-planning Manager). The PCE+ is responsible to detect potential re-planning conditions and compute the up-scaling of the resources that could be applied on the VI. The VRM processes this information applying further policies (e.g. related to re-planning cost), interacts with the VI for re-planning procedures and reconfigures the NCP+.

2.3 – The enhanced PCE for network and IT path computation (PCE+)

Path computation functionalities in the NCP+ are provided by dedicated Path Computation Elements (PCEs), centralized for each routing domain. These elements, called PCE+, are compliant with the standard PCE architecture [4] and are extended to provide energy-efficient and IT-aware end-to-end path computation in support of NIPS services.

The routing algorithms implemented in the PCE+ apply specific constraints like the minimization of power consumption or network/IT cost or the optimization of IT resource utilization [5]. The Traffic Engineering Databases (TE-DBs) are enhanced to include not only the traditional network TE information, but also additional "green" parameters describing the power consumption of the network entities and information describing capabilities, availabilities, power parameters and cost of the IT resources at the end-points. The TE-DBs are fed with the network information flooded through the OSPF-TE routing protocol, extended to carry "green" TE parameters. On the IT side, the information collected at the NIPS server (through the IT advertisement service of the NIPS UNI) are regularly pushed using an extended version of the PCEP protocol, called PCEP+ (Figure 3).

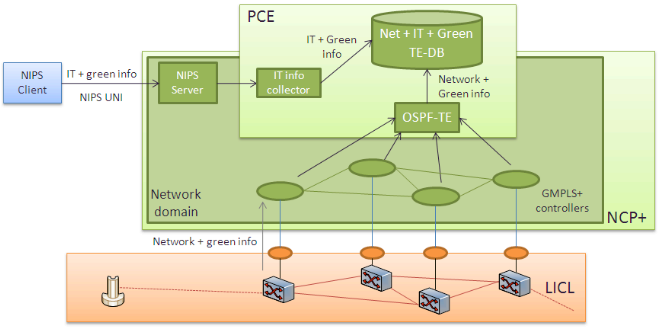

Figure 3 Sources for network and IT information

Path computation in multi-domain scenarios is based on the hierarchical PCE cooperation model, where a parent PCE is in charge of coordinating the end-to-end path computation through multiple edge-to-edge intra-domain requests to its child-PCEs located along the candidate domain paths. In full anycast scenarios, IT information collected on a per-domain basis through the NIPS UNI are reported by the child PCE+ to the parent PCE+ using PCEP+ Notify messages. Consequently, the parent PCE+ is aware of the IT resources availabilities in its child domains and can perform a preliminary selection of the candidate IT end-points, followed by the computation of the associated candidate domain paths.

3. Simulation of unicast and anycast services

3.1 – Simulation scenarios

In order to evaluate the benefits of the NCP+ anycast services, we compare unicast and anycast routing using a simulator [6] specially developed to simulate the functionalities of the NCP+ for both unicast and anycast requests. A NIPS request consists of (i) a single network unit demand (i.e. a wavelength) and (ii) several units of IT capacity (e.g. a number of servers). We assume a Poisson process for the arrival of the requests [7]. The VI is described through a topology graph representing an optical network with nodes inter-connected by optical fibres with a given number of wavelengths, and data centre nodes located at the edges. For each request the objective is to find an appropriate data centre for the destination and a route between the source site and the destination data centre. Both the data centre destination and the route to it must be compliant with the request specifications. Furthermore, we assume that, whenever a request cannot be routed or scheduled because of resource unavailability, it is blocked, leading to a so called blocking probability.

In this work we use three scheduling mechanisms to choose a data centre that has enough IT capacity available to accept the request, where the first two are network-unaware:

- *MaxLoad*: chooses the available data centre that currently has the highest IT resources load.
- *MinLoad*: chooses the available data centre that currently has the lowest IT resources load.
- *Closest*: chooses the available data centre that is closest in terms of available network path.

The path from the source node to the data centre is selected in all cases based on a length-based shortest path routing algorithm. The combination of MaxLoad/MinLoad with shortest path routing conforms to network-unaware scheduling with unicast routing (first IT scheduling, network routing afterwards; once the IT resources is chosen, unicast routing can be used to find the network path) while Closest with shortest path routing corresponds to anycast routing as network and IT resources are treated in the same time instead of sequentially. All shown figures are averages over 5 simulations with the same parameter set but with different seeds in order to produce statistically more reliable results. All simulations stop after having served 200.000 requests and have a warm-up period which stops after sending 10% of the total number of requests.

3.2 – Simulation results

We plot the total blocking probability in **Error! Reference source not found.**-[a, b, c], where blocking can be attributed to either (i) not finding a path from the requesting source to one of the destinations (network blocking) or (ii) not finding a data centre which is able to process the given request (IT blocking). Secondly we plot the average network load and average data centre load in **Error! Reference source not found.**-[a, b, c] and Figure 6-[a, b, c].

For low and medium load values (i.e. low $\frac{\#erlangs}{source\ site}$ and low $\frac{\#server}{request}$), the *Closest* scheme achieves lower blocking probability compared to the other two schemes. The selection of the IT site for each request is based on the length of the paths from the source to the IT sites with available capacity. This way, shorter path and thus more efficient utilization of the network resources is achieved, leading to higher acceptance rate for the given load range.

When considering high load conditions (i.e. high $\frac{\#erlangs}{source\ site}$ and high $\frac{\#server}{request}$), the IT blocking increases and the network blocking decreases. If we would only focus on IT blocking, we notice that *MaxDCUtil/MinDCUtil* have 8% and 4% less IT blocking than *Closest* scheduling respectively. On the other hand, *Closest* still tries to minimize network blocking and consequently the total blocking is still less: there is a difference between *MaxDCUtil/MinDCUtil* and Closest in network blocking of about 9% and 6% respectively. So the performance gain in IT blocking from *MaxDCUtil/MinDCUtil* is neutralized by the difference in network blocking from *Closest*.

The same ideas can be applied on the average network and IT load.

- Although *MinDCUtil* tries to balance the IT load among all data centres, their geographical spread impacts the network blocking because longer paths need to be taken. Consequently, although IT capacity is still available, some requests cannot be processed by the data centres as there is no route available to some of them.
- *MaxDCUtil* tries to compact all the requests to one data centre and when this data centre no longer has enough capacity, it will schedule requests to other data centres. This also means that a lot of the same network resources need to be used to reach the highest loaded data centre and this again influences the network blocking probability:

some sources are isolated from the other data centres because links around the source sites are saturated and incur a penalty in network blocking.

These statements are also reflected in the network load figures. *MaxDCUtil* and *MinDCUtil* on average have 23% and 10% more network load to process then *Closest*. More network load means less free network capacity and leads to higher network blocking.

Concluding, we see that treating network resources and IT resources separately (and thus using unicast routing after IT scheduling) has a drastic impact on overall performance of the infrastructure, while some optimizations can be achieved when we consider both resources at the same time (IT+network-aware anycast).

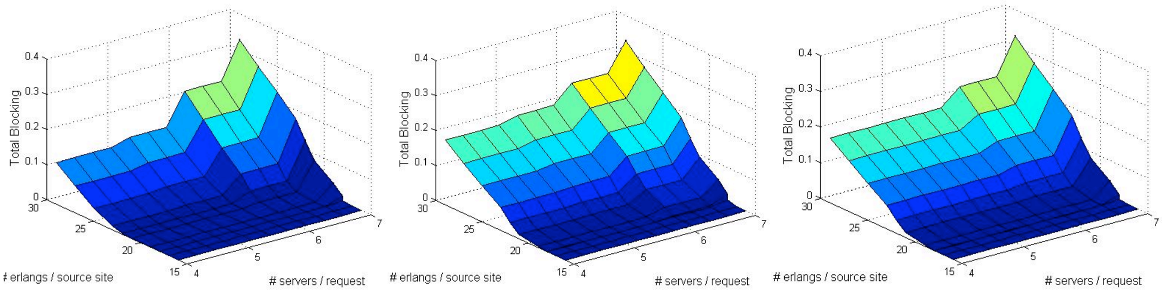

Figure 4a-b-c Total Blocking figures for Closest, MaxDCUtil, and MinDCUtil scheduling policies

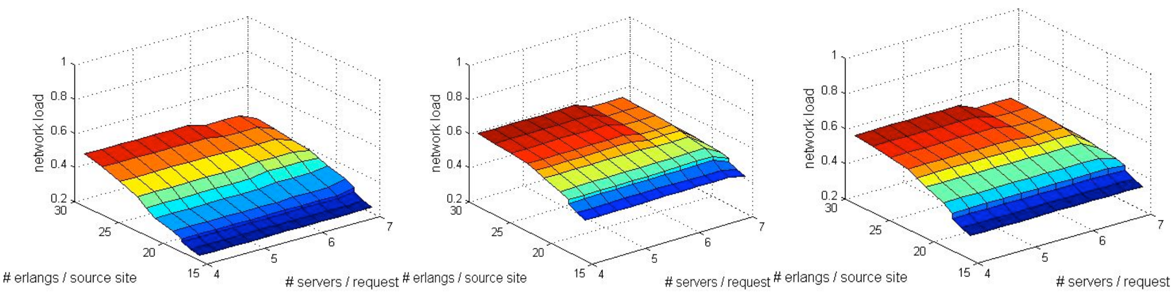

Figure 5a-b-c Network load for Closest, MaxDCUtil, and MinDCUtil scheduling policies

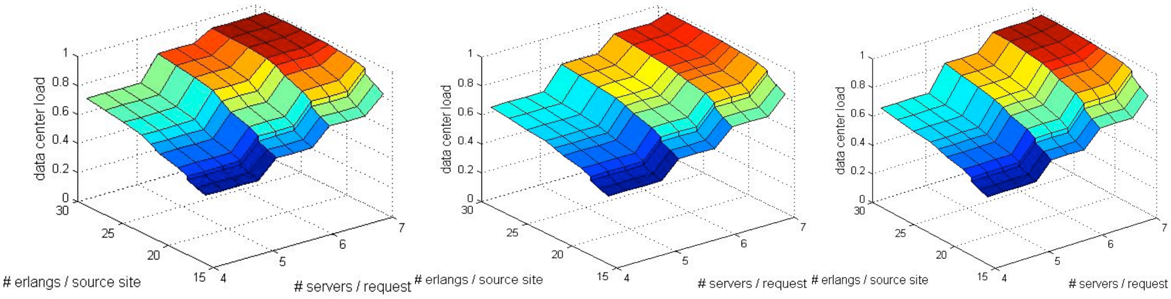

Figure 6a-b-c IT load for Closest, MaxDCUtil, and MinDCUtil scheduling policies

4. NCP+ experimental validation

The NCP+ prototype is deployed in the GEYSERS test-bed, composed by network and IT resources provided by project partners (i2CAT, IBBT, Interoute, Lyatiss, PSNC, Telefonica, Univ. of Essex, Univ. of Amsterdam, Tech. Univ. of Braunschweig) and installed at their premises. These sites are interconnected, as shown in Figure 7. Some examples of network equipment used in the test-bed are Alcatel-Lucent 1850 TSS-160, ADVA FSP 3000 R7, ADVA FSP 3000 RE-II, Calient FibreConnect 320X, W-Onesys Proteus System, Alcatel-Lucent 1850 TSS-320 and LambdaOpticalSystems LambdaNode2000.

Within the GEYSERS test-bed, the optical network equipment and optical links are virtualized and offered as logical resources; NCP+ controllers control logical resources within a single virtual network. Depending on the technological aspects of the underling transport network and organizational boundaries, several domains of NCP+ instances are configured. IT resources are also virtualized and managed by Virtual IT Managers in each IT centre connected to the test-bed. The first step of the NCP+ deployment in the test-bed is the software integration of the NCP+ elements with networks resources (e.g. network virtualization systems) and Virtual IT Managers together with service middleware applications. The validation of the NCP+ is carried out by internal tests and public demonstrations of the provided NIPS services and the internal re-planning functionalities. Results of early integration activities have preliminarily validated the intercommunications among the NCP+ components and the related protocol extensions.

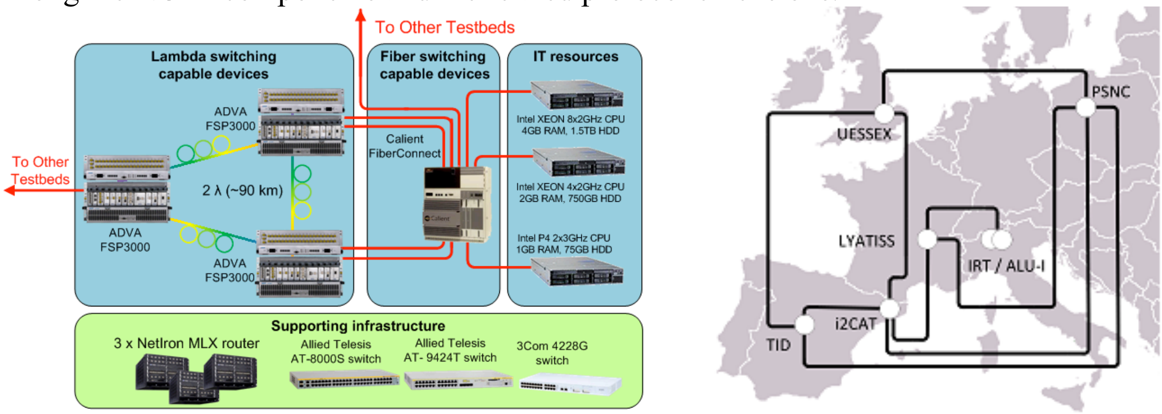

Figure 7 An example of GEYSERS test-bed site (on the left) and the pan-European Geysers test-bed for integration, tests, validation and demonstrations (on the right) [1].

5. Conclusions

This paper has described an enhanced NCP+ operating over a virtual optical infrastructure to provide IT-aware and energy-efficient connectivity services between data centres. Simulation results have demonstrated the advantages of joint procedures for selection of IT resources integrated with network path computation. Future work will include refining anycast scheduling and routing schemes (e.g. the schemes in [5]) which will be included as algorithms in the proposed NCP+ architecture. Furthermore the performance of routing algorithms applied over specially designed virtual infrastructures will be investigated.

The prototype of the NCP+ will be further validated through experimental activities on the GEYSERS test-bed and public demonstrations are expected starting from the second quarter of 2012.

Acknowledgement: This work was carried out with the support of the GEYSERS (FP7-ICT-248657) project funded by the European Commission through the 7th ICT Framework Program.

References

[1] GEYSERS web site, http://www.geysers.eu

[2] Vicat-Blanc P., Figuerola, S., Chen, X., Landi, G., Escalona, E., et al. *Bringing Optical Networks to the Cloud: an Architecture for a Sustainable future Internet*. Springer Lecture Notes in Computer Science, 2011

[3] ITU-T G.8080/Y.1304 Recommendations, *Architecture for the Automatic Switched Optical Network (ASON)*, 2001

[4] Farrel A., Vasseur J.P., Ash J., *A Path Computation Element (PCE) – Based Architecture*, IETF RFC 4655, August 2006

[5] Tzanakaki, A., Anastasopoulos, et al., *Energy Efficiency in integrated IT and Optical Network Infrastructures: The GEYSERS approach.* INFOCOM 2011 workshop on "Green Communications and Networking", Shanghai, China, 2011

[6] De Leenheer M., Buysse J., Mets K., Dhoedt B., Develder C., *Design and Implementation of a Simulation Environment for Network Virtualization*, Proc. 16th IEEE Int. Workshop Computer Aided Modeling, Analysis and Design of Commun. Links and Netw. (CAMAD 2011), 2011

[7] Christodoulopoulos K., Varvarigos E., Develder C., De Leenheer M., Dhoedt B., *"Job demand models for optical grid research"*, Proc. 11th Int. IFIP TC6 Conf. on Optical Netw. Design and Modeling (ONDM2007), Athens, Greece; Lecture Notes in Computer Science, vol. 4534, pp. 127–136, May 2007