SEVENTH FRAMEWORK PROGRAMME

Project Number: FP7-ICT-2007-1 216863

Project Title: Building the Future Optical Network in Europe (BONE)

CEC Deliverable Number: FP7-ICT-216863/SSSUP/R/PU/D12.1

Contractual Date of Deliverable: 30/11/08

Actual Date of Delivery: 30/11/08

Title of Deliverable: D12.1: Report on inventory of available expertise, planned strategy towards integration, and activated joint activities

Workpackage contributing to the Deliverable: WP12 : Virtual Center of Excellence on Services and Applications (VCE S&A)

Nature of the Deliverable Make your choice from: R (Report)

Dissemination level of Deliverable Make your choice from: PU (Public)

Editors: Partner / Names
SSSUP / Piero Castoldi, Luca Valcarenghi

Abstract:

This document is the report of the first year of activities in the VCE-WP12 Virtual Center of Excellence on Services and Applications (VCE S&A). It includes the achievement of milestone M12.1 due at month 3 that report highlights the participants' expertise. This reports includes also the status of the joint activities proposed within VCE-WP12.
There are sixteen partners involved in this workpackage and four joint activites have been proposed so far, each of them involving at least one mobility action.

Keyword list:
Application-network integration, resource virtualization, service engineering.

Disclaimer

*The information, documentation and figures available in this deliverable, is written by the BONE ("*Building the Future Optical Network in Europe*) – project consortium under EC co-financing contract FP7-ICT-216863 and does not necessarily reflect the views of the European Commission*

Table of Contents

1. Executive Summary :

This document is the first deliverable of the work package "Virtual Center of Excellence on Services and Applications (VCE S&A)". The main objectives of this deliverable is collect the partner expertise and describe the status of the planned joint activities in the framework of BONE project. Based on common research topics four (4) joint activities have been proposed and planned so far. The topics covered and planned by this work package try to address the three main aspects of application and services: (i) Service definition, architectures and implementation, (ii) Application definition, architectures, requirements, (iii) Services and application in an integrated view.

2. Introduction

The Virtual Center of Excellence on Services and Applications is focused on the definition of a roadmap

(i) for conceptually characterizing the service access problem both at the customer application side and at the network side (through all segments)

(ii) for identifying and structuring network services in terms of VPN to fulfil the bandwidth and level of connectivity needed by applications

(iii) for identifying and structuring non-network services as perceived by a customer application

(iv) for restructuring the business chain of telecommunications between transport providers (i.e., network service providers) dealing with network resources and service providers dealing with non-network resources.

With this aim, specific objectives of this WP are:

- To integrate the research efforts on applications and services in Europe with special emphasis on those based on optical networks.

- To collect inputs and research outcomes for preparing guidelines of the most appropriate approaches for the application-to-network interaction to be offered to European providers and vendors.

- To investigate on service platform architectures that apply to the various network segments handled by telecom operators.

- To define roadmaps for the evolution of the telecommunication business in terms of services based on network and non-network resources.

Section 3 presents the list of partners that have declared man power on this WP.

Section 4 lists a summary of the most relevant bibliography of those partners that have declared commitment so far.

Section 5 reports the outcome of a poll realized among contributing partners on which key issues they feel of importance with respect to the scope of this WP.

Section 6 reports the methodology adopted for integration and reports a provisional list of Joint activities that have been defined so far

Section 7 povides an update on the activities carried in the JAs so far.

…

3. VCE S&A Participants

The following member organisations have allocated manpower in VCE S&A as of today.

Participant	Total Effort Spent	Total Effort Planned
1 - IBBT	0,2	1,5
2 - TUW	0,572	0
12 - UC3M	0,1	0,25
15 - UPVLC	0,25	0,75
21 - RACTI	0,7	1,75
24 - BME	0,5	1
27 - FUB	0,3	0,75
30 - PoliTO	0,5	0
31 - SSSUP	1,5	3,75
32 - UNIBO	0,02	2
35 - TELENOR	0	1,5
38 - AGH	0,2	1
41 - KTH	0,05	1
47 - UEssex	1,1	1,5
49 - Ericsson	0,5	0,5
	6,49	**17,25**

The table above show the status so far as reported in the BONE Directory Service. It only includes efforts up to the third reporting period (included) since at time of preparation of this deliverable the fourth period is not terminated yet.

4. Selected essential recent bibliography of partners

Scuola Superiore Sant'Anna

L. Valcarenghi, F. Paolucci, F. Cugini, and P. Castoldi, "Quality of service aware fault tolerance for grid-enabled applications", Optical Switching and Networking, Volume 5, Issues 2-3, June 2008, pp. 150-158

IBBT (Ghent University)

M. Pickavet, P. Demeester, D. Colle, D. Staessens, B. Puype, L. Depré, I. Lievens, "Recovery in multilayer optical networks", (invited paper), Journal of Lightwave Technology, ISSN 0733-8724, Vol. 24, Nr. 1, January 2006, pp. 122-134.

Technische Universitaet Wien

B. Statovci-Halimi: "A Resource Provisioning Scheme for Multiservice IP Networks"; Broadband Europe 2007, Antwerp, Belgium; 03.12.2007 - 06.12.2007; in: "BroadBand Europe 2007", (2007), ISBN: 9789076546094.

Akademia Gorniczo-Hutnicza

Piotr CHOŁDA, Andrzej JAJSZCZYK, Krzysztof WAJDA, "A unified framework for the assessment of recovery procedures", HPSR 2005, Hong Kong, 12–14 May 2005.

Universidad Carlos III de Madrid

Isaias Martinez-Yelmo, David Larrabeiti, Ignacio Soto, and Piotr Pacyna, "Multicast traffic aggregation in mpls-based vpn networks", Communications Magazine, IEEE, 45(10):78-85, Oct. 2007.

DTU Fotonik

S. Ruepp, N. Andriolli, J. Buron, L. Dittmann and L. Ellegaard, "Restoration in All-Optical GMPLS Networks with Limited Wavelength Conversion", accepted for publication in Computer Networks Journal

PoliTo

A. Bianco, J. Finochietto, L. Giraudo, M. Modesti, F. Neri, "Network Planning for Disaster Recovery", LANMAN 2008, The 16th IEEE Workshop on Local and Metropolitan Area Networks, September 3-6, 2008, Transylvania

BME

J. Gál, N. Szabó, Á. Ladányi, T. Cinkler: "Cost and Time Trafe-off of Scheduling GRID Tasks over Grooming Capable Networks", NETWORKS 2008, 13th International Telecommunications Network Strategy and Planning Symposium, Budapest, Hungary, Sept 28 - Oct 2, 2008, www.networks2008.org

5. Key issues and inventory of expertise

A number of key issues have been identified by partners. These key issues have been grouped and consolidated in three main sections that reflect the scope of this Virtual Center of Excellence that aims at integrating services (suppliers of connectivity and resources) and applications (consumer of connectivity and resources).

The three main areas are (i) service definition, architectures and implementation; (ii) application definition, architectures, requirements; (iii) services and applications in an integrated view.

For each of the three areas a table is reported where each partner has declared interest in passive or active mode. "X" indicates that a partner is interested and skilled (active), "#" denotes interest but not available skills. The last column indicates whether each key issues is covered by a joint activity (as later detailed) or the relevant research activity has to be decided (TBD).

All areas of research have already been covered by joint activities.

5.1 *Service definition, architectures and implementation*

This area covers the main aspect of service definition when user and network are in consumer/supplier relationship. The services offered by a network may include connectivity services based on network resources and higher level services such as storage or computation based on network resources.

Topic name / Partner interest	SSSUP	Uessex	UC3M	UNIBO	TUW	AGH	UPVLC	RACTI	FUB	KTH	POLITO	IBBT	in JA
Service definition and abstraction	X	X		X	X	X		#					**all**
Resource virtualization	X	X		X				#					**1, 2, 3**
Service plane architectures	X	X		X				#		#			**3**
GMPLS-based complex services	X	X	#								X		**3**
L1, L2, L3 Virtual Private Networks	X		X#						X				**2, 3**
Multi-domain services		X	X	X						#			**3**
Service routing, signaling, resilience	X		X			X		X		X		X	**1**
Automatic Network Configuration	X		#	X						X			**3**
Paradigms for end-to-end QoS					X		#		X	X			**3**

5.2 *Application definition, architectures, requirements*

This area covers the main aspect of application definition as a consumer of services, i.e. what type of intelligence and functional interface it will realize for the final user and what kind of interaction with the network it will have.

Voice over IP has to be covered by a joint activity.

Topic name/Partner interest	SSSUP	Uessex	UC3M	UNIBO	TUW	AGH	UPVLC	RACTI	FUB	KTH	POLITO	IBBT	in JA
Voice over IP (VoIP)	#		#	X	X		#						**TBD**
Grid over optical networks	X	X	#	X			#					X	**3**
Peer-to-peer over optical networks	#		#					#	X	X			**3**

5.3 Services and application in an integrated view

This area covers the integration of application and services, i.e. achieving the final goal of matching consumer and suppliers and all the necessary functional blocks to realize the proper coupling of the two functional abstracted entities.

A few research topics need still to be covered by joint activities.

Topic name/Partner interest	SSSUP	Uessex	UC3M	UNIBO	TUW	AGH	UPVLC	RACTI	in JA
ITU-T Next Generation Network		X			X				**TBD**
Service Control Functions	X	X		#				X	**3**
Resource and Admission control Functions	X		#	#	X			X	**3**
Service platforms	X	X			#				**3**
Service composition	X			X				X	**3**
Service Plane - Control Plane interworking	X	X						#	**3**

6. Plans for integration in view of Joint activities definition

6.1 Methodology

WP12 is a relatively small cluster of partners within the BONE project. Nevertheless there is a core group of partners already active within e-photon/ONE and e-photon/ONE+ with a solid background on application and services.

The integration methodology of this WP envisions sessions of cross-dissemination based on meetings and exchange of results and the creation of joint activities where more and less experienced partners with common interests decide to work together to reach a specific goal.

At time of preparation of this report four JAs have been defined within WP12 and are detailed in the following.

Since the aggregation is new and some partners have declared interest, there in expectation that the number of JAs will increase along the project evolution.

6.2 List of JAs

The JAs summarized in Table 1 have been defined so far.

Table 1 JA activities

No	Joint Activity Title	Responsible person	Participants	Mobility Action	Time frame
1	Service Interconnection Fault-Tolerance	Luca Valcarenghi (SSSUP)	SSSUP, AGH, IBBT, KTH	Yes	(Apr. 2008- Sep. 2009)
2	Cloud Computing	Dimitra Simeonidou (UESSEX)	UESSEX, RACTI	Yes	Jun. 2008- Sep. 2009
3	Service Plane functionalities and demonstration	Fabio Baroncelli (SSSUP), Barbara Martini (SSSUP)	SSSUP, UESSEX, UNIBO, FUB	Yes	Apr.2008- Sep. 2009
4	Joint Optimisation of Grid and Network Resources	Tibor Cinckler (BME)	BME, SSSUP, PoliTO	Yes	Apr. 2008- Dec. 2010

6.2.1 Joint Activity n. 1: Service Interconnection Fault-Tolerance

Participants: SSSUP, AGH, IBBT, KTH

Responsible person: Luca Valcarenghi (valcarenghi@sssup.it)

Description:

An important role in service delivery is played by service interconnection resilience. Indeed, interconnection between services and between services and clients is of paramount importance for the successful service delivery. So far, schemes for service fault tolerance and network fault tolerance have evolved independently. Service layer fault-tolerant schemes mainly focused in guaranteeing resilience against application failures (e.g., server crash, server failure). Network layer fault-tolerant schemes focused on providing a reliable network service to upper layer applications. However the integration of such schemes have the potential of increasing the overall reliability while lowering the required costs. Indeed, for instance, replicating a service may contribute in increasing service reliability without requiring more expensive network reliability increase.

One of the objectives of this joint activity is to investigate how to combine application layer and network layer fault-tolerant schemes to assure the required reliability to service clients. Up to a certain extent, integrated schemes could guarantee the reliability required by clients that could not be reached by implementing fault-tolerant schemes just at a single layer.

Another important issue to be investigated is to guarantee service interconnection Quality of Service (QoS)-aware fault tolerance. That is, interconnections will need to jointly satisfy both reliability and minimum guaranteed bandwidth requirements.

The joint task will be therefore organized as follows. First of all a survey of current fault-tolerant schemes both at the application and a the network layer will be conducted. Then they will be analyzed based on their capability of guaranteeing a specific reliability level. Furthermore novel integrated fault tolerant schemes will be proposed. They will be evaluated based on their capability of overcoming current schemes based on specific criteria, such as their cost based on the reliability they can guarantee. Finally, QoS-aware reliability schemes will be proposed and evaluated.

Expected outcome of the joint research activity:

The partners involved in this joint activity will benefit of their complementary expertise. IBBT and SSSUP have a long lasting experience network layer fault tolerant schemes. KTH and AGH have a thorough knowledge about availability and reliability studies together with facilities for scheme testing. The first outcome of the joint activity will be a fault tolerant scheme survey focused on service interconnection. The second main output will be the design and evaluation of a QoS-aware fault tolerant scheme integrating both application layer and network layer fault tolerance.

Timescale: Apr. 2008 – Sep. 2009

Targeted call for papers (please include deadline – month and year):

ICTON 2008, OFC 2009, Globecom 2009, JON

Published Joint Papers:

L. Valcarenghi, M. Kantor, P. Cholda, K. Wajda, "**Guaranteeing high availability to client-server communications",** ICTON 2008, 22-26 June 2008, Athens, Greece

Current Activities:

The first sub-activity focused on developing a *network-aware server placement heuristic* to guarantee high reliability to client-server (two-terminal) communications. The heuristic rationale is based on the following observations: i) link disjoint shortest paths form parallel end-to-end connections, thus the higher is the number of parallel connections between client and server the higher is the two-terminal availability; ii) a set of disjoint paths with a low number of links is more reliable; iii) a path with few links is more reliable. The aforementioned consideration are exemplified in Figure 1 where the reliability between pair s-d R_{sd}^A in the network A is higher than the s-d pair reliability R_{sd}^B in the network B.

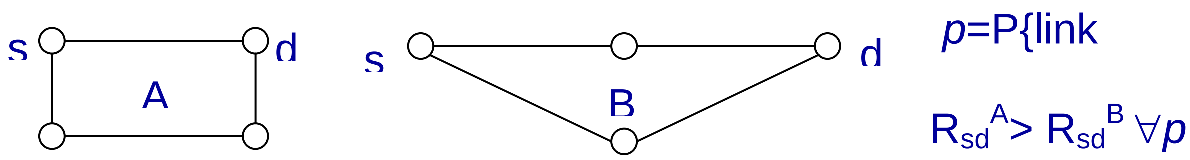

Figure 1 Network-aware server placement heuristic example

study a heuristic algorithm is proposed to obtain a sub-optimal server choice.

The pseudo-code of the algorithm utilized to select each client-server pair is summarized in Figure 2. The first selection criterion in based on the fact that, if paths are link disjoint, they form parallel connections between client and server. Thus, the more are the parallel connections (i.e., the more are the paths) between source and destination the higher is the reliability.

```
Find the set S1 of servers for which the number of link disjoint paths to client c is the maximum
    if |S1|=1
        select the only server in S1
    else
        find the sub-set of servers S2 ⊆ S1 for which the sum of all the links belonging to the
        link disjoint paths is the lowest
        if |S2|=1
            select the only server in S2
        else
            find the sub-set of servers S3 ⊆ S2 for which the path to the client is the
            shortest
            if |S3|=1
                select the only server in S3
            else
                select one server in S3 with uniform probability
```

Figure 2 Server selection algorithm pseudo-code.

The second selection criterion is the number of hops in the edge disjoint paths. Indeed, intuitively, given the same number of edge-disjoint paths, the paths with the lowest sum of hops are the most reliable. This is also in accordance with reliability theory which states that the less are the elements (in our case – hops, which represent links) connected in series in the path the lower is the failure probability. The third criterion is to choose the server that is reachable by means of the shortest path.

The extensive analysis of the proposed server selection algorithm is performed by considering the prism network depicted in Figure 3. The prism network consists of 6 nodes and 9 bidirectional links. Each link is assumed to have the same reliability value p and links are assumed to fail independently.

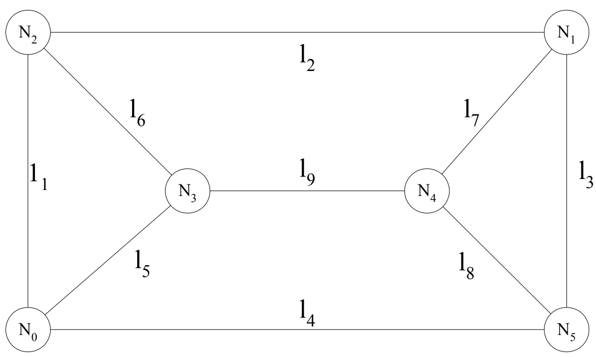

Figure 3 Prism network

Since the prism network consists of a limited number of nodes and links, the proposed server selection algorithm has been compared against an exact (optimal) calculation of the two terminal reliability obtained by means of graph transformation. In addition two-terminal availabilities obtained by a shortest path- (SP-) based server selection algorithm are reported. Only two client placements are considered but, due to the prism network symmetry, they are representative of all the possible client placements.

Table 2 Algorithm comparison results for the prism network

Client location	algorithm	p=0.95		p=0.99	
		Selected server	reliability	Selected server	reliability
N0	optimal	N2, N3	0.99973	N2, N3	0.99999796
	SP	N2,N3,N5	0.99968	N2,N3,N5	0.99999763
	proposed	N2,N3	0.99973	N2,N3	0.99999796
N3	optimal	N0,N2	0.99972	N0,N2	0.99999796
	SP	N0,N2,N4	0.99968	N0,N2,N4	0.99999763
	proposed	N0,N2	0.99972	N0,N2	0.99999796

The reliability values reported in Table 2 represents the average relaibility values for all the servers that can be potentially selected by the considered algorithms, assuming that their selection is equiprobable. In particular, for the optimal algorithm all the servers that provide the maximum reliability are considered, for the SP algorithm all the servers whose path to the client is the shortest, in terms of hops, are considered, and for the proposed algorithm all the servers belonging to the set S_3 are considered. The results show that, for both the considered link reliability values (i.e., p=0.95 and p=0.99) and for both the client placements, the proposed server selection algorithm matches the results obtained by the exhaustive optimal analytical computation. Indeed the same servers are selected. In addition the proposed algorithm outperform the SP-based algorithm.

The first sub-activity involved also a short visit of Miroslaw Kantor (AGH) to SSSUP for the Development of QoS-aware routing algorithms” from 23/10/2008 to 25/10/2008.

The second sub-activity dealt with the implementation of a resource reservation service in a multi-service environment. The resource reservation service utilizes the heuristic developed in the first sub-activity the perform path computation. Some of the current results obtained by this sub-activity stem from the visit of Pawel Korus (AGH) to SSSUP for the implementation of “Resource Reservation in a Multi-Service Environment” from 15/10/2008 to 31/10/2008.

To serve service interconnection requests an architecture based on a Network Management Service (NMS) that utilizes Path Computation Element (PCE) for route computation and then directly interacts with Network Elements (NEs) to commit and reserve the paths is considered. Two serving policies are experimentally evaluated. The first one, *batch provisioning*, serves and commits accepted connection requests in a batch based on a time-out. The second policy, *per-request provisioning*, serves and commits requests as soon as they arrive. An architecture overview of the proposed system that utilizes adaptive queueing is depicted in Figure 4. The NMS is decomposed into two interoperating modules: NMS-client and NMS-agent.

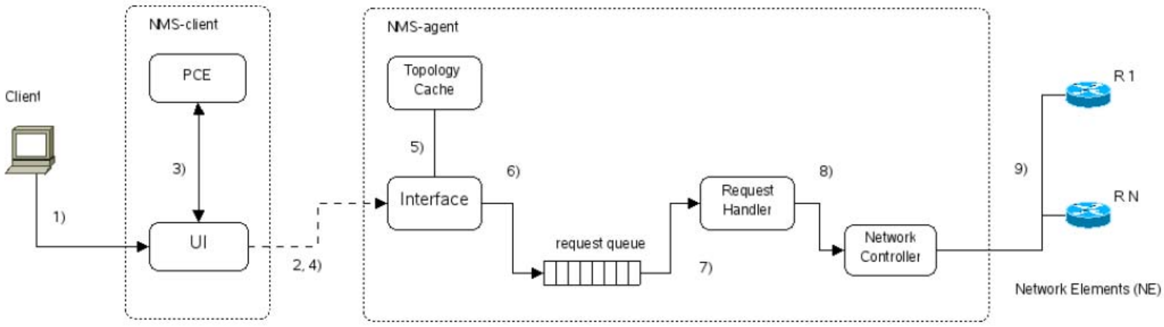

Figure 4 Network Management System Architecture (NMS)

The NMS-client is a client application that provides a command line interface for the users. It also contains the PCE module responsible for routing services. Multiple instances might be running at once. The NMS-agent is the central server for NMS. It implements features related to request queueing, providing current network topology snapshots and interaction with NEs.

The two service policies implemented in the NMS-agent are the following. In the *per-request provisioning* policy requests are processed as soon as they arrive. In the *batch provisioning* policy an adaptive queueing mechanism is utilized to minimize delay caused by Label Switch Path (LSP) activation. The main idea is to perform periodic queue monitoring, then serve the requests in a batch to configure LSP simultaneously.

The two policies have been implemented and tested in a test-bed at SSSUP depicted in Figure 5. LSP requests arrive with an exponential inter-arrival time whose average value is λ. Each experiment considers 15 independent requests. The requested LSP source is chosen with uniform distribution from the set *r7, r8, r9* and the remote node is picked by a routing algorithm from *r1, r3*. The considered performance parameter is the *service request service time* defined as the time elapsing between the client request arrival and time in which traffic can be injected in the set up connection. It includes routing, network device configuration, commit time and all necessary delays in queues.

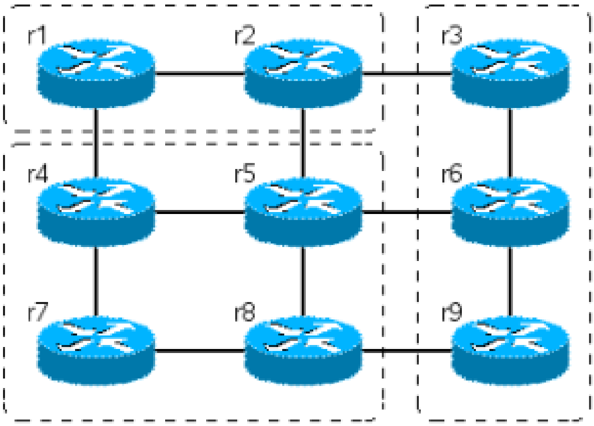

Figure 5 Test-bed

Figure 6 shows the average service time of successive requests served by means of the per-request provisioning policy. If λ is low the service elements are capable of concluding the LSP set up process before the arrival of the successive request. However when λ increases, the successive request arrive before the earlier request set up process has been concluded. Beacause during the set up process network elements cannot serve any other request the service time increases.

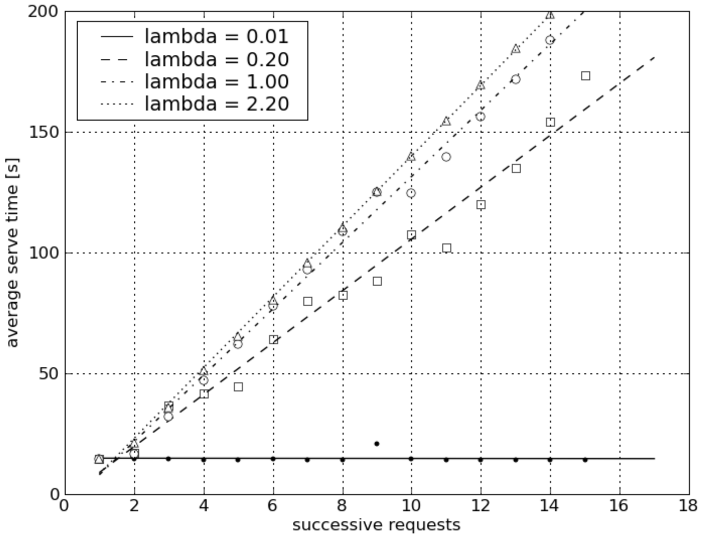

Figure 6 Per-request provisioning policy performance

On the other hand, by using batch provisioning, the service time increase is low if not null as shown in Figure 7. This is due to the fact that by using the adaptive queue and accumulating LSP requests from the same source they can be committed all at once without additional delay.

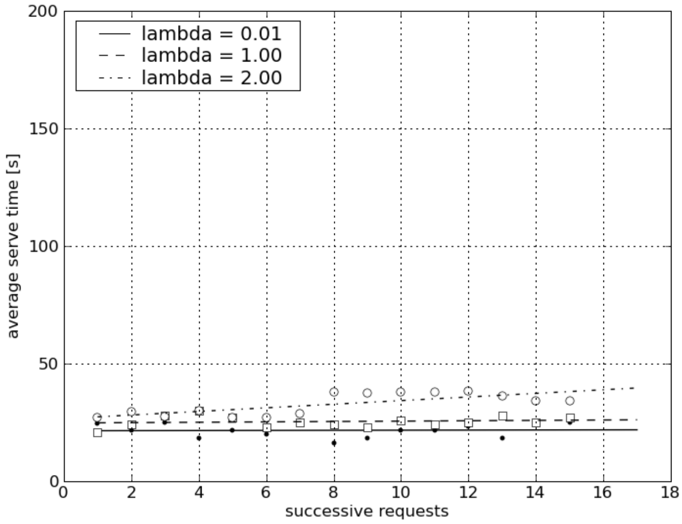

Figure 7 Batch-provisioning policy performance

The main conclusion of this study is that the utilization of batch-provisioning policy and adaptive queueing mechanisms, despite much more complicated module interactions, leads to significant improvement of the average service time with respect to a per-request service policy.

Planned Activities:

Develop server selection heuristics taking into account bandwidth availability jointly with link reliability.

Develop server selection heuristics for client in motion.

Further study the resource reservation server and implementing in it the proposed server selection heuristics.

6.2.2 Joint Activity n. 2: Cloud Computing

- Objectives
  - Clouds provide utility computing as a commercial service.
  - Technologies: virtualization and Web service interfaces.
  - Clouds are provided by commercial infrastructure/service providers (Google, Amazon, eBay, Microsoft, IBM
  - Cloud computing providers will use high bandwidth optical networks in order to co-locate computing resources in remote locations (i.e. secure sites, renewable energy sites)
  - Define optimised end-to-end service offering scenarios (routing algorithms, NML, virtualisation)

Participants: UEssex, RACTI

Responsible person: Dimitra Simeonodou (UEssex)

Current Activities:

RACTI has been working on service aware routing algorithms for cloud computing architectures. Within a cloud, there exist different services each with different resources that are located in remote locations. User requests are transparently forwarded to these resources, ensuring QoS with the specific attributes requested. In a cloud computing architecture, we denote, as $S = (S_1, S_2, \dots, S_N)$ the set of services offered. The provision of each service to end-user means that there exist service specific resources that can be computing elements, storage, VoD servers etc. Each resource can be replicated more than one time in the network either for survivability issues or for serving more or different customers. We denote the k^{th} resource for i service as S_i^k and relying on the overlay approach, there exist a virtual service plane for each individual service in the network.

Considering that abstract concept of cloud computing, the formal definition is *G(V,E)*, is extended to *G(V,E,S)* with the addition of the overlaying services, running on top of the network infrastructure. *G* denotes the network graph, *V* is its set of vertices and *E* is its set of edges.

Problem definition*:* S*ervice aware routing problem formulation in cloud computing architectures.* Suppose that a user located at the ingress node R_i requests service $S_r \in S$, with the specific attributes (service specifics) represented by the following constraint vector: $J = [B_{job}, h_{job}, D_{job}, CPU_{job}, ST_{job}]^T$

with entries being the *bandwidth*, *hop count*, *delay*, *CPU cycles* and *storage capacity*. Assuming that $S_r = (S_r^1, S_r^2, \dots, S_r^K)$ is the set of resources for service *r*, the problem is to find a path (or paths) $P(R_i, S_r^k)$ with $k \in [1, \dots, K], S_r \in S$, $S_r \in S$, such that:

$$D\left(P(R_i, S_r^k)\right) \le J \Rightarrow \begin{bmatrix} B_{R_i \to S_r^k} \\ h_{R_i \to S_r^k} \\ D_{R_i \to S_r^k} \\ CPU_{R_i \to S_r^k} \\ ST_{R_i \to S_r^k} \end{bmatrix} \le \begin{bmatrix} B_{job} \\ h_{job} \\ D_{job} \\ CPU_{job} \\ ST_{job} \end{bmatrix} \qquad \text{Eq. 1}$$

For comparing multi-dimensional metrics put on paths, we refer to the notion of lattices 0. The comparison with ≤ between the two vectors of path information follows that:

$$\left(B_{R_i \to S_r^k} \ge B_{job}\right) \cap \left(h_{R_i \to S_r^k} \le h_{job}\right) \cap \left(D_{R_i \to S_r^k} \le D_{job}\right) \cap \left(CPU_{R_i \to S_r^k} \ge CPU_{job}\right) \cap \left(ST_{R_i \to S_r^k} \ge ST_{job}\right)$$

.In order to create end-to-end path vectors for service delivery, we rely on using additive, multiplicative or limited vector operators. For example hops and delay are cumulative costs in the sense that going from node R_k to R_{k+1}, the relevant cost of the path vector is $P_{R_1,R_2} = \begin{bmatrix} D_{R_1} \\ h_{R_1} \end{bmatrix} + \begin{bmatrix} D_{R_2} \\ h_{R_2} \end{bmatrix} = \begin{bmatrix} D_{R_1} + D_{R_2} \\ h_{R_1} + h_{R_2} \end{bmatrix}$, while for limiting parameters like storage capacity or bandwidth, it is: $P_{R_1,R_2} = min[R_1, R_2]$.. Note that although *hop count* is a parameter that has to do with the number of hops traversed, it could be seen as parameter associated to links by assigning $h[R_1,R_2] = 1$ for every link $[R_1,R_2]$.

Having all the above in mind, the network could be seen as two graphs: one graph containing the non-network resources (computing elements, storage area, VoD servers etc), and the other one the network resources (edge and core network equipment ie routers, switches, etc). The edges of the graphs are inter-connected with a unique *score function* F_s. This score function is calculated using the aforementioned parameters that constitute the distance vector cost. Thus, every router is linked with every resource with a score. If a router does not lead to a destination, then the score towards that resource is zero. For example, in Figure 9, node R_1 points to only a single resource denoted as S_r^1 with a score F_{S_1}. Similarly node R_2 points to two resources (S_r^3 and S_r^n) with different scores but finally points to only one (S_r^3). Each node stores its distance cost vector and communicates it, upstream to be updated until it reaches an ingress node. Figure 9 graphically shows the end-to-end shortest paths/tunnels formed starting from the egress nodes of Figure 9. The distance vector cost for node R_i that is adjacent to egress node R_k is defined as follows:

$$\left\{D\left(P(R_i,S_r^k)\right) = D\left(P(R_i,R_k)\right)\circ D\left(P(R_k,S_r^k)\right)\right\}$$

$$= \begin{bmatrix} B_{R_i\to R_k} \\ h_{R_i\to R_k} \\ D_{R_i\to R_k} \\ - \\ - \end{bmatrix} \circ \begin{bmatrix} B_{R_k\to S_r^k} \\ h_{R_k\to S_r^k} \\ D_{R_k\to S_r^k} \\ CPU_{S_r^k} \\ ST_{S_r^k} \end{bmatrix} = \begin{bmatrix} min\left\{B_{R_i\to R_k}, B_{R_k\to S_r^k}\right\} \\ h_{R_i\to R_k}+1 \\ D_{R_i\to R_k}+D_{R_k\to S_r^k} \\ CPU_{S_r^k} \\ ST_{S_r^k} \end{bmatrix}$$

However, node R_i does not necessarily point to resource S_r^k. It may point to any resource of service S_r, through any neighbor node (R_k') that maximizes score function, F_s, as follows:

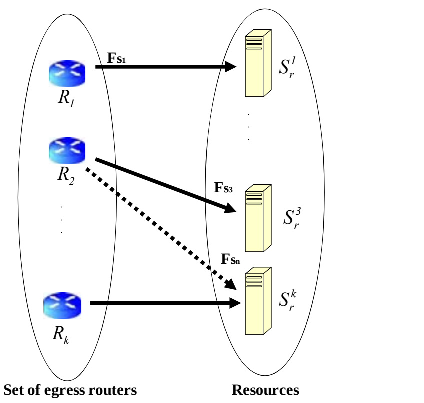

Figure 8: Representation of the network as two sets of node

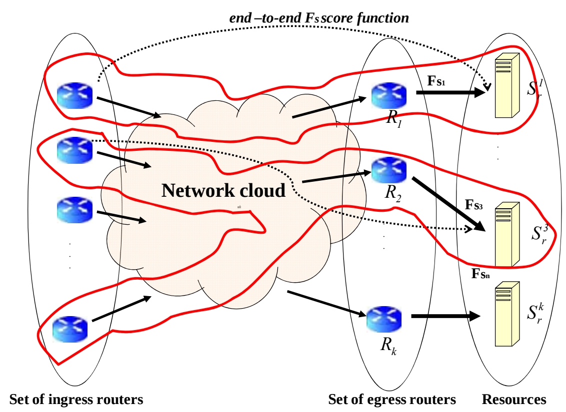

Figure 9: End-to-end shortest path formed between the set of ingress nodes to the available set of resources.

$maxF_s\left\{\overline{D}\left(P(R_i, R_i^j)\right)\circ\overline{D}\left(P(R_i^j, S_r^k)\right)\right\} = maxF_s\left\{\overline{D}\left(P(R_i, R_i^j)\right)\circ\overline{D}(R_i^j)\right\}, \quad \forall R_i^j \in N_i$. In the above equation, R_i^j is any of the N_i set of neighboring nodes of R_i. Therefore, node R_i selects that adjacent node that maximizes its score value, thus pointing to a specific resource. For calculating the distance vector cost of the next to R_k node, we used the operator ° for handling multi dimensional vectors. That operator defines the logical functions (*MIN, MAX, ADD,* etc) between the vector entries.

References

A. Jukan and G. Franzl, “Path Selection Methods with Multiple Constraints in Service-Guaranteed WDM Networks”, IEEE/ACM Transactions on Networking, 12(1):59–72, Feb. 2004.

Planned Activities:

Future activities of include performance simulations using ns-3 in order to validate end-to-end service interconnection under failures of either network resources (fiber cuts, node failures etc) or non-network resources (i.e. server failure). It will be investigated, how these failures affect the end-to-end service offering for both new requests that may be blocked (Denial of Service – DoS) and ongoing services (rerouting and reallocation).

6.2.3 Joint Activity n. 3: Service Plane functionalities and demonstration

Participants: SSSUP, UEssex, UNIBO, FUB

Responsible person: B. Martini (SSSUP), F. Baroncelli (SSSUP),

Description:

A new generation of IT applications (e.g. grids computing, on-demand multimedia services) is emerging which demands access to remote IT systems and services (e.g., computing resources, distributed data storage facilities, media servers, content repositories and scientific instruments) often via high-speed network infrastructures. Initially developed by collaborative virtual research communities, these applications are steering the development of new service-oriented optical transport network technologies and architectures. The aim is enabling IT applications to directly request a given IT resource (i.e. remote storage space, computation capability or media stream) as well as the communication facilities required to connect to it with the proper QoS fully relying on the network infrastructure.

The reciprocal awareness between applications and networks push for the disintegration of numerous barriers that normally separate computing IT systems and application services from optical transport network infrastructures and technologies. Although there has been substantial progress in the technology of management and delivery IT systems, there is a big semantic-gap between parameters handled by network devices, such as commercial routers, and parameters used by the end-host applications, such as SIP-based Proxy.

Application-to-network mapping is the key-point needed to fill this semantic-gap and dynamically provision high-bandwidth and QoS-enabled network services to qualified IT applications. This “mapping” capability enables the translation of service requests generated by application session signalling, e.g. SIP, into technology-specific and self-consistent directives used to configure network devices. Moreover, in networks that involve different network technologies, the fulfilment of service required issued by IT application require the concatenation of different mapping to realize a real end-to-end network configuration.

The objective of this activity is to theoretically evaluate and experimentally demonstrate the possible exploitation of service platform architectures for optical networks. Such architecture are required to expose to IT applications abstract levels of information regarding offered network and non-network services while enabling the consistent network node configuration.

The proposed platform is based on Service Oriented Optical Network (SOON) architecture. SOON is a service architecture for optical transport networks conceived to fulfill connectivity service requests issued by applications, i.e. in terms of end-host addresses and perceived quality of service, while performing network-wide and self-consistent device settings based on a distributed signaling among “service nodes” called Distributed Service Elements (DSEs). We consider an heterogeneous network scenario comprising OBS and IP/MPLS as core network technologies and Ethernet as aggregation network technology.

The partners involved in this joint activity will benefit of the complementary experience in this field. SSSUP will focus mainly on architectures and the result of its testbed built on top of commercial routers, UNIBO will bring his competence on SIP protocol, FUB has a multivendor testbed, while UEssex has an optical testbed already developed and available since the two editions of e-photon/ONE.

This activity has been conceived to operate in collaboration with WP21. The theoretical activities will be carried out in WP12 while most experimental results will be realized within the framework of WP21.
The following two sub-activities have being carried out. The first one is jointly performed by SSSUP and UEssex and consists in the implementation and validation of a service-oriented multi-granular Optical Bust Switching (OBS) network. The second one is jointly performed by SSSUP and UNIBO and consists in the functional specification and experimental validation of a SIP-based service platform for (G)MPLS-enabled optical network

Current activities:

1. Sub-JA: SSSUP-Uessex

This work presents results of a novel service-oriented multi-granular Optical Bust Switching (OBS) network demonstrator for the support of future Internet applications. The experimental validation of the concept relies on the demonstration of Quad-HD video-on-demand (VoD) services over the OBS Testbed. The network architecture is based on the integration of Service Oriented Optical Network (SOON) framework functionalities with OBS control and data plane technologies.

We worked on the demonstration that the proposed SOON-enabled OBS architecture can be used to effectively bridge the informational gap between the Application layer and the Network layer by introducing a suitable formalism that facilitates a mapping process between Application requests and the Network services. The Service Oriented Optical Network (SOON) framework is based on a distributed approach conceived for overseeing a new paradigm of application-to-network interaction. The aim is to disjoint the parameters perceived by an end user from the technology-specific directives needed by Network devices. It also enables automatic network configuration for establishing on-demand connection with different classes of service.

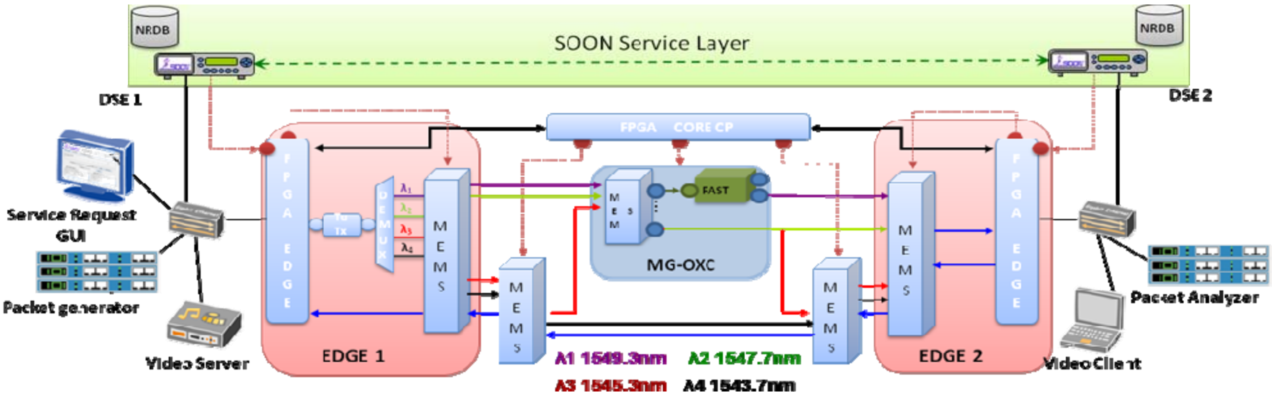

Figure 1. Architectural block diagram of the SOON-enabled E-OBS-T Test-bed and network demonstrator

In this implementation, the SOON framework is able to map a set of parameters that an application (e.g. multimedia) specifies within a network service request, into a set of specific parameters used by the network (e.g. edge aggregation buffer thresholds, offset, lambda/sub-lambda light-paths) while hiding the network resource technology and topology details from the application. SOON also translates this mapping into a direct request for dynamic configuration of that service. The E-OBS-T network is then able to parse the SOON request and on-the-fly configure aggregation scheduler (buffer size, time), select the appropriate granularity and establish light-paths (slow-fast path) to support different services.

The main experiment carried out in ESSEX University validates the SOON service and connection establishment as well as high definition video over OBS transmission. The SOON-JIT messages encapsulated in BCHs are sent over OBS control plane, and the generated variable optical bursts over Ethernet-type data plane. In order to study the effect of OBS on the real-time transmission of high performance media, four pre-recorded videos, with different qualities, were used in different streaming media scenarios across the OBS test-bed Network. These videos varied from High Definition with resolutions of 1280x720 and 1440x1080 and bit-rates of 27 Mbps and 46 Mbps respectively, to Quad-HD of 2560x1600 and bit-rates of 106 and 156 Mbps. TCP background traffic of around 200Mbps was also generated through traffic generator in order to emulate the current internet traffic behaviour (between TCP and UDP data). The aggregation developed is hybrid and combines both size and time threshold. Size and time thresholds are also dynamically changed per SOON service with maximum size threshold of 5000 bytes and time limit of 2ms.

The Fig. 2 shows that for service 1 (OBT) more than 95% of the UDP packets have a delay of less than 3 ms with a maximum delay less than 4ms which is well within the acceptable level and for service 2 the value is less than 1.8ms. It also shows that for service 2 (OBS) the jitter also remains below 1.4 ms for 100% of the traffic and below 0.9ms for service 2, also a well accepted value. The packet loss of the OBS network is zero the whole amount of data.

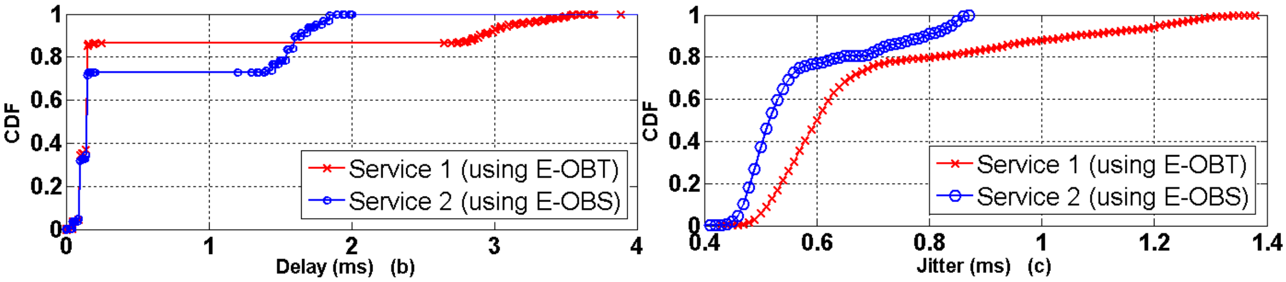

Figure 2 a) Delay and b) Jitter of receiving 46Mbps High Definition Video over the E-OBS-T test-bed which has 200Mbps background TCP traffic. Service 1 is defined and implemented in the edge node (FPGA) as best effort class. Service 2 is defined and implemented in the edge node (FPGA) as timing critical class.

2. Sub-JA: SSSUP-UniBo

In this sub-activity we carried out the experimental validation of a service platform for optical transport networks based on the Session Initiation Protocol (SIP) that enables applications to exchange semantically rich messages with the network to issue service requests for the reservation of needed resources. The platform subsequently translates such requests into technology-dependent directives for the network nodes.

The functional architecture of the service platform is presented in Fig.3. Three functional layers are in evidence: the Application-Oriented Network Layer which manages application signaling and generates the service requests; the Service Plane (SP) which manages the mapping between service requests and technology-specific directives; the (G)MPLS-enabled transport network infrastructure which is responsible of data forwarding. The coupling of the transport network with the SP constitutes the Service Oriented Optical Network (SOON).

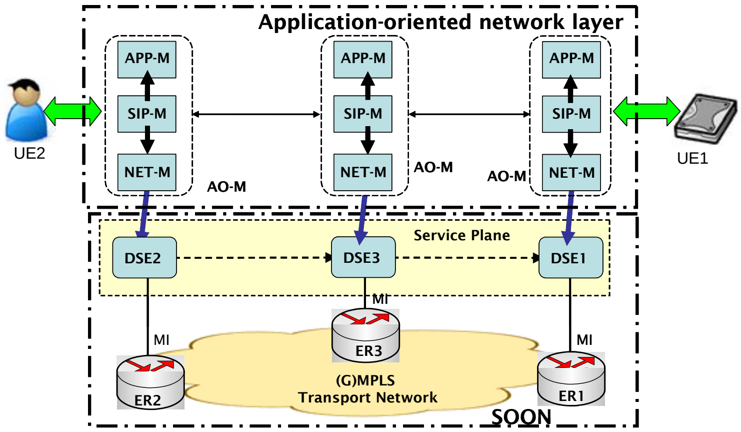

Fig.3 The application-aware service platform

The Application-Oriented Network Layer handles communication among end-hosts and maps generic communications instances into network service sessions that are managed using the SIP protocol. The Application-Oriented Network Layer is in charge of the resource management and deals with both application resources (i.e., remote storage space, remote multimedia content) and network resources (i.e., connectivity with the needed QoS assurance). It resides in nodes called Application-Oriented Modules (AO-M). The AO-M talks with the applications using a Resource Description Framework (RDF) vocabulary called Network Resource Description Language (NRDL). Specifically the AO-M includes three sub-modules.

- The APP-M maps application communication requests into network service sessions by parsing and elaborating the NRDL messages carrying the communication needs required by the application.
- The SIP-M is an enhanced SIP proxy server, in charge of establishing and managing dialog sessions by relaying messages between applications and maintaining the current state using session attributes.
- The NET-M handles the requests for the communications facilities of the various sessions and sends the requests for network resources to the SOON to enable the required data transfer across the network.

In the proposed architecture resources are published and can be searched and reserved by applications through the AO-M using SIP session management messages, which carry resource information in the user-oriented NRDL language. The SIP proxy in the SIP-M has been enhanced with the capability to read the NRDL from the message body and pass it to the APP-M that parses the message, interprets the communication needs and triggers the NET-M to send the service request to SOON. In turn, SOON translates such a request into the proper set of technology-specific directives for an effective resource allocation across the network.

A test-bed has been set-up to validate the proposed architecture. The experiment we carried out reproduced the issue of a service request, i.e., channel establishment specified in terms of end-host addresses and perceived quality of service, from the AO-M to the SOON during an enhanced SIP signaling between applications and the AO-M. Upon the arrival of the service request, then SOON process such a request and configures the traffic policies in the edge routers to enable a data flow with the required Quality of Service (QoS) assurances between end-hosts. Without lack of generality, a single AO-M was included coupled with DSE3.

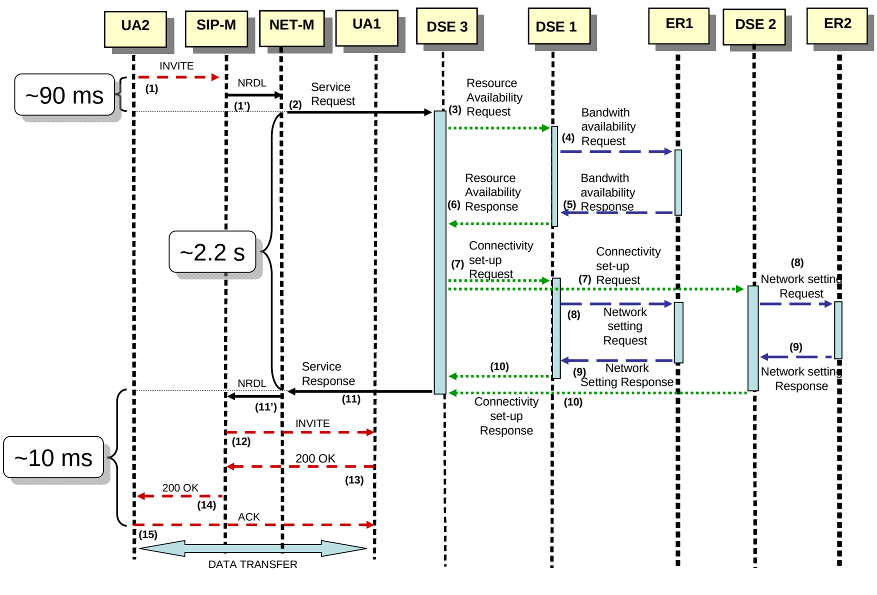

Fig. 4 – Example of message flow for application driven service configuration

The message flow of the experiment is shown in Fig. 4. The whole procedure is started with an INVITE message from UA2 to the AO-M carrying a NRDL document that requests the HD video service. The SIP-M receives the INVITE and passes the NRDL to the APP-M. The APP-M parses it and triggers the NET-M to ask the DSE3 to arrange the network resources reservation across the MPLS network domain. The service request consists of an XML file including the UA IP addresses and the service characteristics, i.e., service type and quality level. From the IP addresses the DSE3 deduces the edge routers terminating the required network path, i.e., ER1 and ER2. From the service type and quality level, the DSE determines the traffic category and the bandwidth requirements, that in turn determines the the proper LSP between ER1 and ER2 to carry out video traffic across the MPLS network (i.e., LSP-video) and the amount of bandwidth (e.g., 20 Mbit/s for HD video) to be reserved within such LSP. The DSE3 verifies whether such bandwidth is available along the LSP by interrogating one of the edge routers (e.g., DSE1). In case of positive response, the DSE3 triggers the bandwidth reservation in the LSP by asking the DSE1 and the DSE2 to issue the related network-specific directives to ER1 and ER2. Such directives enforce traffic policies on both ER1 and ER2 to assure the proper forwarding of video packets at the ingress and egress of LSP-video, the consistent packet marking as they enter such LSP and the needed transmit rate at the output transmission queue. When the reservation is fulfilled, the DSE3 confirms to the NET-M that the requested resources are available. The NET-M triggers the SIP-M to complete the session set-up. An INVITE message is sent to UA1 that replies with an OK message as required by the SIP three-way handshake. The ACK by UA2 concludes the session setup and data can now flow in the LSP-video experiencing the required QoS.

The average timing obtained from the test-bed experiment is shown in Fig. 4. The total signaling delay needed to reserve network resources is in the order of 2-3 seconds and is fully compatible with the dynamics of typical applications.

Planned activities:

Enlarge the set of network services provisioned to applications, e.g., bidirectional channel, compound channel supporting contemporary video, voice, chat sessions. Possibly include the Ethernet-based domains in the network scenario.

Dissemination process:

- 1 paper published

B. Martini, V. Martini, F. Baroncelli, P. Castoldi, L. Rea, A. Valenti, F. Matera “Dynamic QoS control based on VPLS in Service Oriented Transport Networks” Proc. of ICTON 2008, Athens, June 22-26, 2008.

- 2 papers submitted at OFC 2009 and 1 at ICC 2009

6.2.4 *Joint Activity n. 4: Joint Optimisation of Grid and Network Resources*

Participants: BME, SSSUP, PoliTO

Responsible person: Tibor Cinkler (BME)

Description:

The objective of Joint Activity 4 is to consider and optimise ***jointly***, on the one hand, the Services and Applications (S&A), and on the other hand, the optical network resources.

The services and applications are assumed to be GRID services, namely network-wide distributed storage and computing services. Nowadays it seems a preferred strategy to virtualise and share own resources than building huge private supercomputers and data storage centers. This approach is much more efficient for bursty computing requirements, and can be much cheaper for end users to cover slightly higher power-costs, however, having access to practically unlimited parallel computing resourecs.

The networks considered are assumed to be controlled by the GMPLS protocol family. In different studies we assume different data planes ranging from a general packet-switching capable one, to more heterogeneous wavelength-switched and packet-switched grooming-capable one.

The composition of partners is as follows: SSSUP having expertise in CP issues, PoliTo having expertise in Storage issues, BME having competence in Computing GRID issues, and all three partners experienced in related optimization issues. Our on-going and future activities are described in more details in Sections “Current Activities” and “Planned Activities”.

Timescale:

- April 2008 – Dec. 2010

Targeted calls for papers:

1. ICTON 2009, www.itl.waw.pl/konf/icton/2009, submission: March 31, 2009
2. DRCN 2009, submission: ~May 30, 2009
3. ONDM 2010, submission: ~September 30, 2009
4. OSN journal, 2010

Current Activities:

1. Control Plane issues (GMPLS) for GRID services

   GRID services need strong support from the Control plane. Current Control Plane standardisation does not include this support. In this activity it is investigated weather the current GMPLS “standard” can efficiently support GRID sevices or some extensions to it are needed?

2. ILP optimisation for GRID storage resources

   Disaster Recovery and Business Continuity issues are becoming fundamental in networks since the importance and social value of digital data is continuously increasing. On the one hand, there is an obvious need of backing up data for resilience against major failures; in many situations the process of storing backup data is also enforced by the law. On the other hand, providing services that allow the migration of applications in realtime through virtualization techniques is becoming a mandatory feature in several business situations.

Here we analyze the problems and the challenges of off-site data replication and virtual machine migration. In particular, we discuss the issue of optimizing network planning to support disaster recovery and business continuity.

ILP (*Integer Linear Programming*) formulations for the optimization problem are presented with different objective functions. Heuristics are also proposed and analyzed taking into account both network cost minimization and fault recovery efficiency.

3. ILP optimisation for GRID computing resources
   Optical networks are a promising inter-site connectivity solution for data intensive GRID applications. Possible applications include e.g. data-intensive scientific computations and high-resolution video processing. In such applications a job assignment involves in the transfer of a considerable amount of data over the network. Therefore the job scheduling can not solely rely on the reliability of the processing resources. Both these and the available network resources need to be taken into account. The network is assumed to be a two-layer optical –beared grooming-capable one.
   We present an ILP formulation for task assignment assuming the above scenario, and in addition some heuristics to improve its scalability. We use simulations to compare different variations of the algorithm, e.g., weather the cost or the execution time is the critical objective.

Planned Activities:

1. Using simpler and faster yet efficient heuristics based on greedy algorithm and shortest path searches instead of ILP for optimising GRID storage resources.
2. Using simpler and faster yet efficient heuristics based on greedy algorithm and shortest path searches instead of ILP for optimising GRID computing resources.
3. Applying joint optimisation of network resources and GRID storage resources assuming that the network consists of two layers, a WDM wavelength switched layer and a TDM capable layer over it.
4. Considering resilience issues in joint optimisation of network resources and GRID computing resources. Namely, if a network link fails or a computing node fails a part of results is not received by the initiator. Therefore different scenarios are investigated ranging from those that optimise the speed to those that optimise the cost.
5. Evaluating the Control Plane support for all of the above methods.

7. Conclusions

This WP represents a relative small cluster of partners with special interest in Services and Applications. Partners who declared interest in this WP are both experienced and less expert partners who can get benefits from the integration/dissemination actions of the network of excellence. The three main areas have been defined to reflect partners expertise or interest (i) service definition, architectures and implementation; (ii) application definition, architectures, requirements; (iii) services and applications in an integrated view.

The currently activated four joint activities has been outlined, covering core topics of this WP. More activities are expected to come.