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/D12.3
Contractual Date of Deliverable:	30/11/10
Actual Date of Delivery:	30/11/10
Title of Deliverable:	D12.3: Final Report on Integration 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, Barbara Martini,
IBBT/ Chris Develder,
UESSEX/ Chinwe Abosi, Eduard Escalona
RACTI / Kyriakos Vlachos

Abstract:

This document both overviews the activities conducted during the whole duration of the project and details the third year of activities within the VCE-WP12 Virtual Center of Excellence on Services and Applications (VCE S&A). The objective of the VCE S&A are: investigating architectures for the Service Plane (SP) and asses relevant performance; defining solutions for partitioning of functions between Data, Control, Management Plane and SP in all network segments; defining roadmaps for the evolution of the telecommunication business; conforming conceptually similar but not-interoperable current solutions.
There are fourteen partners involved in this workpackage and seven joint activities have been defined over the three years. Most of them has run for the entire duration of the project. During the third year partners have finalized their studies within the defined joint activities by also setting-up joint testbeds for experimental evaluation of service-oriented platforms for optical networks.
Dissemination of the joint activities in international conferences and journals and in the Open Grid Forum standardization body has been actively pursued.

Keyword list:

Application-network integration, resource virtualization, service engineering.

Clarification:

Nature of the Deliverable

R	Report
P	Prototype
D	Demonstrator
O	Other

Dissemination level of Deliverable:

PU	Public
PP	Restricted to other programme participants (including the Commission Services)
RE	Restricted to a group specified by the consortium (including the Commission Services)
CO	Confidential, only for members of the consortium (including the Commission Services)

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 deliverable of the work package "Virtual Center of Excellence on Services and Applications (VCE S&A)", summarizing the whole activity carried out during the project and reporting the activity from 1/1/2008 to 31/12/2010 with special emphasis on the third year.

During Y2, the ongoing activity in WP21 Topical Project on "Service Oriented Optical Networks" were moved to WP12, as they address similar aspects of application and services in optical networks: (i) Service definition, architectures and implementation, (ii) Application definition, architectures, requirements, (iii) Services and application in an integrated view.

Therefore, this document is reporting the work, the progress, and the final findings of the seven (7) JAs set up during the whole duration of the project as well as some statistics about the integration results.

2. Integration results

2.1 Summary of Joint Activities

The current Joint Activities (JAs) defined within the WP12 are indicated in the table below. Please note that JA5, JA6, and JA7 were inherited from WP21 on Service Oriented Optical Networks.

No	Joint Activity Title	Responsible person	Participants	Mobility Action in 2010	Time frame
1	Service Interconnection Fault-Tolerance	Luca Valcarenghi (SSSUP)	SSSUP, AGH, IBBT, KTH	YES	Whole Project
2	Cloud Computing	Dimitra Simeonidou (UESSEX)	UESSEX, RACTI	NO	Whole Project
3	Service Plane functionalities and demonstration	Fabio Baroncelli (SSSUP), Barbara Martini (SSSUP)	SSSUP, UESSEX, UNIBO, FUB	YES	Whole Project
4	Joint Optimisation of Grid and Network Resources	Tibor Cinkler (BME)	BME, SSSUP, PoliTO	NO	Y1 and Y2
5	Programmable Service Composition Algorithms for Service Oriented Optical Networks	Chinwe Abosi (UEssex)	UESSEX, UoP, RACTI	NO	Whole Project
6	UNI extensions for Service Oriented Optical Networks	Eduard Escalona (UEssex)	UESSEX, RACTI, AIT	NO	Whole Project
7	Photonic Grid Dimensioning & Resilience	Chris Develder (IBBT)	IBBT, AIT, RACTI	NO	Whole Project

2.2 Mobility Actions

The following mobility actions took place during the whole project within WP12. The ones that took place in the third year of the project are labeled.

1. A formal QoS Class-of-Service characterization for deploying advanced services in an heterogeneous network scenario, Valerio Martini, Ph.D. Student at SSSUP, hosted by UEssex from 15/05/2008 to 20/10/2008
2. *Resource Reservation in a Multi-Service Environment*, Pawel Korus, PhD student at AGH, hosted by SSSUP from 15/10/2008 to 31/10/2008
3. *Development of QoS-aware routing algorithms*, Miroslaw Kantor, Research Assistant at AGH, hosted by SSSUP from 23/10/2008 to 25/10/2008
4. *Service interconnection availability in Mobile WiMAX*, Luca Valcarenghi, Assistant Professor at SSSUP, hosted by KTH from 07/02/2009 to 15/02/2009
5. *SIP signalling for user-oriented optical networks*, Walter Cerroni, Researcher at UNIBO, hosted by SSSUP from 14/05/2009 to 14/05/2009
6. *Economic viability of resource infrastructure abstractions in Grid networks (II)*, Chinwe Abosi, Research Student at UEssex, hosted by GET from 01/06/2009 to 25/06/2009
7. *SIP-based Service Platform for Transport Networks*, Barbara Martini, Research Scientist at SSSUP, hosted by UNIBO from 25/11/2009 to 25/11/2009
8. *SIP-based Service Platform for Transport Networks*, Molka Gharbaoui, PhD student at SSSUP, hosted by UNIBO from 25/11/2009 to 25/11/2009
9. *SIP-based Service Platform for Transport Networks*, Walter Cerroni, Researcher at UNIBO, hosted by SSSUP from 15/02/2010 to 15/02/2010 (Y3)
10. *Generalizing the PCE Concept for Load Distribution and Fault Tolerance by Using a Distributed Tree-Aggregation Protocol in a PCE-based WDM Network*, Jawwad Ahmed, PhD student at KTH, hosted by SSSUP from 04/05/2010 to 25/05/2010 (Y3)

2.3 Dissemination & Papers

The various JAs achieved several interesting scientific results. As at the third periodic report of project year 3, the total number of joint papers sums up to 50. The complete list is attached later in this section.

Table 1 reports the tren about the number of papers per year and per institution.

	Year 1	Year 2	Year 3	Total
# papers	10	25	15	50
# joint papers	4	10	7	21
# single papers	7	14	8	29
# institutions	6	14	9	16

Table 1 – Global statistics of paper within WP12

Complete list of papers

1 F. Baroncelli (SSSUP), B. Martini (SSSUP), P. Castoldi (SSSUP), Network Virtualization for Cloud Computing, Annals of Telecommunications, Vol. 65, No. 11-12, pp. 713-721, November 2010. (WP12)

2 S. Pompei (FUB), M. Teodori (FUB), A. Valenti (FUB), S. Di Bartolo (ISCOM), G. Incerti (ISCOM), D. Del Buono (ISCOM), Experimental implementation of an IPTV architecture based on Content Delivery Network managed by VPLS Technique, 2nd IEEE International Workshop on Reliable Network Design and Modeling (RNDM 2010), Moscow, Russia, October 2010. (WP12)

3 G. Zervas (UEssex), V. Martini (SSSUP), Y. Qin (UEssex), E. Escalona (UEssex), R. Nejabati (UEssex), D. Simeonidou (UEssex), F. Baroncelli (SSSUP), B. Martini (SSSUP), K. Torkmen (SSSUP), P. Castoldi (SSSUP), A service oriented multi-granular Optical Network Architecture for the Clouds, IEEE/OSA Journal of Optical Communications and Networking , Vol. 2, No. 10, pp. 883-891, October 2010. (WP12)

4 R. Cafini (UNIBO), W. Cerroni (UNIBO), C. Raffaelli (UNIBO), M. Savi (UNIBO), Enforcing Security in Multi-Service Programmable Routers for Future Internet, Proc. of 21st International Tyrrhenian Workshop on Digital Communications (ITWDC 2010), Island of Ponza, Italy, September 2010. (WP25 WP12)

5 F. Paolucci (SSSUP), F. Cugini (SSSUP), B. Martini (SSSUP), M. Gharbaoui (SSSUP), L. Valcarenghi (SSSUP), P. Castoldi (SSSUP), Preserving Confidentilaity in PCEP-based Inter-Domain Path Computation, ECOC 2010, Turin, Italy, September 2010. (WP22 WP12)

6 B. Statovci-Halimi (TUW), G. Franzl (TUW), QoS differentiation and Internet neutrality: a controversial issue, International Conference on Telecommunications and Signal Processing, Vol. 33, pp. 6, Bonus Parkhotel Baden, Baden bei Wien, Austria, August 2010. (WP12)

7 L. Wosinska (KTH), A. Jirattigalachote (KTH), P. MONTI (KTH), A. Tzanakaki (AIT), K. Katrinis (AIT), Energy Efficient Approach for Survivable WDM Optical Networks, IEEE International Conference on Transparent Optical Networks (ICTON), Munich, July 2010. (WP11 WP12 WP22)

8 B. Statovci-Halimi (TUW), Adaptive admission control for supporting class-based QoS, 6th Euro-NF Conference on Next Generation Internet (NGI 2010), pp. 8, Paris, France, June 2010. (WP12)

9 F. Callegati (UNIBO), W. Cerroni (UNIBO), B. Martini (SSSUP), M. Gharbaoui (SSSUP), A. Campi (UNIBO), P. Castoldi (SSSUP), Configuration of Network Resources for Future Internet Application Services, Proc. of 2010 Future Network and Mobile Summit, Florence, Italy, June 2010. (WP12)

10 L. Wosinska (KTH), J. Chen (KTH), P. Monti (KTH), What Photonics Can Do for Switching in Transparent Optical Networks, European Conference on Networks and Optical Communications (NOC), June 2010. (WP11 WP12 WP22)

11 L. Velasco (UPC), A. Jirattigalachote (KTH), P. Monti (KTH), L. Wosinska (KTH), S. Spadaro (UPC), G. Junyent (UPC), Probabilistic-based Approach for Fast Impairments-aware RWA in All-Optical Networks, Proceedings of OFC, San Diego, CA, USA, March 2010. (WP12 WP11)

12 W. Cerroni (UNIBO), B. Martini (SSSUP), M. Gharbaoui (SSSUP), A. Campi (UNIBO), F. Baroncelli (SSSUP), P. Castoldi (SSSUP), F. Callegati (UNIBO), Experimental validation of a SIP-based Platform for Service Oriented Optical Network, Optical Fiber Communication (OFC/NFOEC 2010) Conference, San Diego, CA, USA, March 2010. (WP12)

13 J. Ahmed, (KTH), P. Monti (KTH), L. Wosinska (KTH), Concurrent Processing of Multiple LSP Request Bundles on a PCE in a WDM Network, Proceedings of OFC, San Diego, CA, USA, March 2010. (WP11 WP12)

14 F. Cugini (SSSUP), S. Xu (NICT), H. Harai (NICT), F. Paolucci (SSSUP), L. Valcarenghi (SSSUP), P. Castoldi (SSSUP), Optical Grid Networking Exploiting Path Computation Element (PCE) Architecture, International Journal of Communication Networks and Distributed Systems, Vol. 5, No. 3, pp. 246-262, March 2010. (WP12)

15 F. Baroncelli (SSSUP), B. Martini (SSSUP), V. Martini (SSSUP), P. Castoldi (SSSUP), A cooperative approach for the automatic configuration of MPLS-based VPNs, Internation Journal of Grid Computing and Multi Agent Systems (GCMAS), January 2010. (WP12)

16 L. Valcarenghi (SSSUP), P. Korus (AGH), F. Paolucci (SSSUP), F. Cugini (SSSUP), M. Kantor (AGH), K. Wajda (AGH), P. Castoldi (SSSUP), Experimental Evaluation of PCE-Based Batch Provisioning of Grid Service Interconnections, Globecom 2009, Vol. 1, No. 1, Honolulu, HI, December 2009. (WP12)

17 R. Aoun (GET), M. Gagnaire (GET), Service differentiation based on flexible time constraints in market-oriented Grids, IEEE Xplore - IEEE Globecom 2009, Piscattaway-NJ, USA, December 2009. (WP12)

18 F. Z. Khan (TUW), M. F. Hayat (TUW), G. Franzl (TUW), Service differentiation via preemptive wavelength conversion in optical burst switched networks, International Conference on Frontiers of Information Technology (FIT2009), Vol. 7, No. paper 410, pp. 6, Abbottabad, Pakistan, December 2009. (WP12 WP27)

19 B. Statovci-Halimi (TUW), V. Chimaobi Emeakaroha (TUW), Framework for Multi-class and Dynamic Flow-aware Admission Control, 34th IEEE Conference on Local Computer Networks (LCN 2009), pp. 344-347, Zürich, Switzerland, October 2009. (WP12)

20 J. Buysse (IBBT), M. De Leenheer (IBBT), C. Develder (IBBT), B. Dhoedt (IBBT), Exploiting relocation to reduce network dimensions of resilient optical grids, Design of Reliable Communication Networks (DRCN 09), Washington DC (USA), October 2009. (WP12)

21 C. Abosi (UEssex), R. Nejabati (UEssex), D. Simeonidou (UEssex), A Novel Service Composition Mechanism for the Future Optical Internet, IEEE Journal of Optical Communications and Networking, Vol. 1, No. 2, pp. A106–A120, September 2009. (WP12 WP24)

22 E. Zouganeli (TELENOR), I. E. Svinnset (TELENOR), Connected Objects and the Internet of Things - a Paradigm Shift, Photonics in Switching 2009, Pisa, Italy, September 2009. (WP11 WP12)

23 M. De Leenheer (IBBT), J. Buysse (IBBT), C. Develder (IBBT), B. Dhoedt (IBBT), P. Demeester (IBBT), Deflection Routing in anycast-based OBS grids, Int. Workshop on Optical Burst/Packet Switching (WOBS), Madrid (Spain), September 2009. (WP12)

24 G. Zervas (UEssex), R. Nejabati (UEssex), D. Simeonidou (UEssex), C. Rafaelli (UNIBO), M. Savi (UNIBO), C. Develder (IBBT), M. De Leenheer (IBBT), D. Colle (IBBT), N. Ciulli (Nextworks), G. Carozzo (Nextworks), M. Schiano (Telecom Italia), Programmable Multi-granular Optical Networks: Requirements and Architecture, 6th Int. Conf. on Broadband Communications, Networks and Systems (Broadnets 09), Madrid (Spain), September 2009. (WP12)

25 B. Martini (SSSUP), V. Martini (SSSUP), F. Baroncelli (SSSUP), K. Torkman (SSSUP), P. Castoldi (SSSUP), Application-driven Control of Network Resources in Multi-Service Optical Networks, Journal of Optical Communications and Networking, Vol. 1, No. 2, pp. A247-A257, July 2009. (WP12)

26 C. Develder (IBBT), J. Buysse (IBBT), M. De Leenheer (IBBT), B. Dhoedt (IBBT), Dimensioning Resilient Optical Grids, Int. Conf. on Transparent Optical Networks (ICTON), Sao Miguel (Portugal), July 2009. (WP12)

27 R. Nejabati (UEssex), J. Aracil (UAM), P. Castoldi (SSSUP), M. De Leenheer (IBBT), D. Simeonidou (UEssex), L. Valcarenghi (SSSUP), G. Zervas (UEssex), J. Wu (BUPT), Advanced Optical Burst Switched Network Concepts, Enabling Optical Internet with Advanced Network Technologies, J. Aracil and F. Callegati editors, pp. 131-154, Springer, UK, July 2009. (WP11 WP12)

28 K. Vlachos (RACTI), L. Raptis (Attica Telecom), A. Teixeira (IT), G. M. Tosi Beleffi (ISCOM), K. Yiannopoulos (RACTI), Optical Switch Fabrics (OSFs) and Thier Application, Enabling Optical Internet with Advanced Network Technologies, J. Aracil and F. Callegati editors, pp. 155-190, Springer, UK, July 2009. (WP14 WP11 WP12)

29 B. Martini (SSSUP), F. Baroncelli (SSSUP), V. Martini (SSSUP), K. Torkman (SSSUP), P. Castoldi (SSSUP), ITU-T RACF implementation for application-driven QoS control in MPLS networks, IM 2009, New York - NY(USA), June 2009. (WP12)

30 K. Vlachos (RACTI), A. Siokis (RACTI), A Service-Transparent and Self-Organized Optical Network Architecture, ICC 2009, Vol. 1, No. 1, pp. 1-7, Dresden, Germany, June 2009. (WP12 WP21)

31 L. Valcarenghi (SSSUP), P. Monti (KTH), I. Cerutti (SSSUP), P. Castoldi (SSSUP), L. Wosinska (KTH), Issues and solutions in mobile WiMAX and wired backhaul network integration, Proceedings of ICTON 2009, June 2009. (WP11 WP12)

32 P. Korus (AGH), F. Paolucci (SSSUP), L. Valcarenghi (SSSUP), F. Cugini (SSSUP), P. Castoldi (SSSUP), M. Kantor (AGH), K. Wajda (AGH), Experimental Evaluation of Batch versus Per-Request Service Interconnection Activation in PCE-based Grid Networking, HPSR 2009, Paris, June 2009. (WP12)

33 G. Zervas (UEssex), M. DeLeenHeer (IBBT), L. Sadeghioon (UEssex), D. Klonidis (AIT), Y. Qin (UEssex), R. Nejabati (UEssex), D. Simeonidou (UEssex), C. Develder (IBBT), B. Dhoert (IBBT), P. Demeester (IBBT), Multi-Granular Optical Cross-Connect: Design, Analysis and Demonstration, IEEE Journal of Optical Communications and Networking, Vol. 1, No. 1, pp. 69-84, June 2009. (WP12 WP21 WP24)

34 C. Develder (IBBT), B. Dhoedt (IBBT), B. Mukherjee (), P. Demeester (IBBT), On Dimensioning Optical Grids and the Impact of Scheduling, Photonic Network Communications, Vol. 17, No. 3, pp. 255-265, June 2009. (WP12)

35 F. Callegati (UNIBO), A. Campi (UNIBO), G. Corazza (UNIBO), D. Simeonidou (UEssex), G. Zervas (UEssex), Y. Qin (UEssex), R. Nejabati (UEssex), Sip-empowered optical networks for future it services and applications, IEEE Communication Magazine, Vol. 47, No. 5, pp. 48-54, May 2009. (WP12 WP21 WP24)

36 A. Campi (UNIBO), W. Cerroni (UNIBO), G. Corazza (UNIBO), F. Callegati (UNIBO), B. Martini (SSSUP), F. Baroncelli (SSSUP), V. Martini (SSSUP), P. Castoldi (SSSUP), SIP-based Service Architecture for Application-aware Optical Network, Proc. of 2nd International Conference on ICT and Accessibility (ICTA 2009), Hammamet, Tunisia, May 2009. (WP12)

37 R. Aoun (GET), M. Gagnaire (GET), An exact optimization tool for market-oriented grid middleware, IEEE Xplore - IEEE CQR conference, Naples-FL, USA, May 12-14, 2009, Piscattaway-USA, May 2009. (WP12)

38 Y. Qin (UEssex), G. Zervas (UEssex), V. Martini (SSSUP), M. Ghandour (UEssex), M. Savi (UNIBO), F. Baroncelli (SSSUP), B. Martini (SSSUP), P. Castoldi (SSSUP), C. Raffaelli (UNIBO), M. Reed (UEssex), D. Hunter (UEssex), R. Nejabati (UEssex), D. Simeonidou (UEssex), Service-Oriented Multi-Granular Optical Network Testbed, OFC 2009, March 2009. (WP21 WP12)

39 B. Martini (SSSUP), A. Campi (UNIBO), F. Baroncelli (SSSUP), V. Martini (SSSUP), K. Torkman (SSSUP), F. Zangheri (UNIBO), W. Cerroni (UNIBO), P. Castoldi (SSSUP), F. Callegati (UNIBO), SIP-based Service Platform for On-demand Optical Network Services, OFC 2009, San Diego - CA (USA), March 2009. (WP12 WP11)

40 F. Baroncelli (SSSUP), B. Martini (SSSUP), V. Martini (SSSUP), P. Castoldi (SSSUP), Service Oriented Optical Network Architecture, Proc. of OFC/NFOEC 2009, San Diego, CA, USA, March 2009. (WP12 WP21)

41 B. Statovci-Halimi (TUW), V. C. Emeakaroha (TUW), Enabling Class-based QoS Commitments through Flow-Aware Admission Control, FITraMEn08 - International Workshop on Traffic Management and Traffic Engineering for the Future Internet, Porto, Portugal, December 2008. (WP12)

42 F. Cugini (SSSUP), S. Xu (NICT), H. Harai (NICT), F. Paolucci (SSSUP), L. Valcarenghi (SSSUP), P. Castoldi (SSSUP), Introducing Path Computation Element (PCE) in Optical Grid Networking, ECOC 2008, September 2008. (WP11 WP12)

43 J. Gál (BME), N. Szabó (BME), A. Ladányi (BME), T. Cinkler (BME), Cost and Time Trade-off of Scheduling Grid Tasks over Grooming Capable Networks, NETWORKS 2008, 13th International Telecommunications Network Strategy and Planning Symposium, Budapest, Hungary, September 2008. (WP12)

44 N. Andriolli (SSSUP), A. Giorgetti (SSSUP), S. Ruepp (COM), J. Buron (COM), L. Valcarenghi (SSSUP), P. Castoldi (SSSUP), Bidirectional Lightpath Provisioning in GMPLS-Controlled Optical Networks, Photonics in Switching (PS) 2008, Japan, August 2008. (WP12 WP22)

45 B. Statovci-Halimi (TUW), G. Franzl (TUW), S. Aleksic (TUW), Performance Comparison of Measurement-Based Admission Control, Proceedings of NOC2008, pp. 230 - 237, Krems, Austria, July 2008. (WP12 WP21)

46 B. Statovci-Halimi (TUW), H. R. van As (TUW), Performance Considerations on Admission Control for Multimedia Services, International Conference on Signal Processing and Multimedia Applications (SIGMAP 2008), Porto, Portugal, July 2008. (WP12)

47 M. Kantor (AGH), K. Wajda (AGH), Inter-domain traffic optimization in resilient Next Generation Network environment, NOC 2008, Krems, July 2008. (WP12 WP21 WP22)

48 L. Valcarenghi (SSSUP), M. Kantor (AGH), P. Cholda (AGH), K. Wajda (AGH), Guaranteeing high availability to client-server communications, ICTON 2008, Vol. 3, Athens, June 2008. (WP12)

49 B. Martini (SSSUP), V. Martini (SSSUP), F. Baroncelli (SSSUP), P. Castoldi (SSSUP), L. Rea (FUB), A. Valenti (FUB), F. Matera (FUB), Dynamic QoS control based on VPLS in Service Oriented Transport Networks, ICTON 2008, June 2008. (WP12)

50 B. Statovci-Halimi (TUW), Support of IP Multi-Services Through Admission Control, ITU-T "Innovations in NGN" Kaleidoscope Conference, pp. 407 - 414, Geneva, Switzerland, May 2008. (WP12)

2.4 Standardization

A standardization effort has been carried out by JA6 partners within the in the Open Grid Forum (OGF). A unofficial draft version of Network Service Interface (NSI) user-case deliverable (work in progress) has been submitted.

2.5 Joint Testbeds

Within the various JAs, the joint testbeds indicated in the table below were set-up.

JA	Involved Partners	Testbed Name	Testbed Type	Joint Collaboration
JA1	AGH, SSSUP	PCE-based batch provisioning of grid service interconnections	Network testbed with commercial MPLS routers	Testbed set-up at SSSUP, remotely operated from AGH
JA3	SSSUP, UEssex	Service-oriented multi-granular optical network	OBS network (at UEssex) combined with service-plane software (implemented by SSSUP)	Testbed set-up and operated at UEssex during a mobility action
JA3	SSSUP, UniBO	SIP-based service platform for GMPLS-enabled transport network	MPLS network infrastructure enhanced with service plane (at SSSUP) integrated with enhanced SIP-based signaling (developed by UniBO)	Network set-up at SSSUP and SIP software located at UniBO → tested remotely run from UniBO

3. JAs Description

WP12 merges the activities on services and applications with the activities that started in WP21 in 2008 on service-oriented optical networks. Thus, the main focus of WP12 is on application and services for optical networks.

This section both overviews the activities conducted during the whole duration of the project and details the technical progress achieved by the joint effort of the partners in Y3.

3.1 JA1 - Service Interconnection Fault-Tolerance

Participants: SSSUP, AGH, IBBT, KTH

Responsible person: Luca Valcarenghi (SSSUP)

3.1.1 JA1 Objectives

The JA1 objective is to investigate how to assure fault tolerance to Quality of Service (QoS)-guaranteed service interconnections

One issue that this JA will investigate is 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 issue that will be investigated how service interconnections can jointly satisfy both reliability and QoS-related requirements (e.g., service set up time). For example, within this JA, the impact on service set up time of different methods for provisioning reliable connections between services will be evaluated.

This JA will therefore design and evaluate algorithms for the efficient provisioning of reliable and QoS-guaranteed connections between services. Distributed and centralized schemes will be considered.

3.1.2 Objectives for Y3

Within Y3 this activity aims at devising and evaluating centralized schemes for the provisioning of QoS guaranteed connection between services. In particular the utilization of centralized schemes based on the Path Computation Element (PCE) concept applied to WDM network will be considered for the bulk service based provisioning optical connection between services. A specific issues will be considered: the irresponsiveness to connection provisioning requests of some commercial devices during connection set up. The activity during Y3 complements the activities carried on Y2 that considered the experimental evaluation of PCE-Based batch provisioning of grid service interconnections in an experimental testbed.

The purpose of Y3 activities of JA1 is to study the batch provisioning problem in PCE-based Grid networks in a scenario where routers are modeled with realistic lockdown times (i.e., time intervals during which routers are irresponsive to any connection request). For this specific work we consider a WDM transport network with wavelength continuity constraint (WCC). One of the key objectives is to demonstrate the benefits of PCE-based batch provisioning to reduce connection set up times in a scenario where routers have large lock-down times.

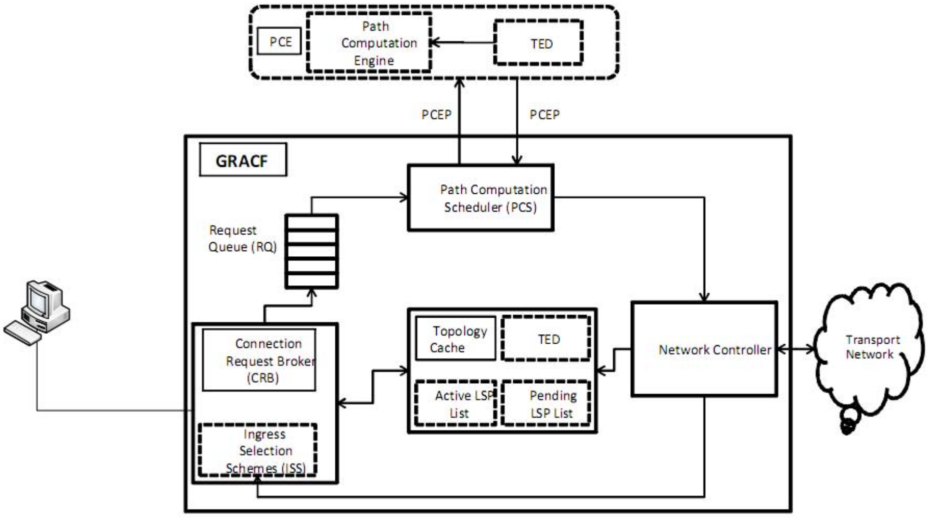


Figure 1 System Architecture

The proposed system model is depicted in Figure 1. We consider a network scenario where a number of routers are connected based on a given network topology to provide transport connectivity among a set of client nodes or terminals. Connection requests for Grid services, in the form of Label Switched Paths (LSPs), are collected at the respective ingress routers and forwarded for path computation to the PCE via a RACF (Transport Stratum, Resource and Admission Control Functions) for Grid services (GRACF) interface. The connection requests are not sent directly to the PCE but they are first sent to a Connection Request Broket (CRB) that puts them into a Request Queue (RQ). Then the Path Computation Scheduler (PCS) organizes these connection requests in batches and feeds them to the PCE. Size and batch formation may vary depending on the batch provisioning policy.

When the PCE finishes the path computation phase for a specific batch, the computed paths are returned to the PCS. They are then sorted based on their ingress routers by the Network Controller (NC) and provisioned by the ingress routers after a signaling phase (i.e., using a Resource Reservation Protocol (RSVP-TE)). The signaling and the provisioning phase are regulated and monitored by the NC. It is assumed that when an ingress router is in the process of provisioning a single or a batch of multiple connection requests, the router is in a lockdown state. In such a state the router does not respond to any further configuration requests until the lockdown time is over. Te lockdown time is a function of: (i) the number of connection requests that have to be processed, and (ii) the router’s CPU load at that specific time instance. During this study we modeled the lockdown time behavior based on the actual experimental data that was NC Controller is also responsible for keeping track of the routers which are currently in lockdown state. For these routers, the NC saves the batches of connection requests that cannot be currently processed in its internal queue. Once a specific ingress router exits the lockdown, the NC assigns the next batch of connection requests.

In our proposed architecture terminals may be connected to their respective ingress/egress routers in a single or in a multi-homed fashion. In a multi-homing scenario, the selection of the ingress and egress routers, for a given connection request, is made at the Connection Request Broker (CRB) using the ISS (Ingress Selection Scheme) module (see Figure 1). Currently in the ISS there two router selection schemes are implemented: ISS0, where the ingress router for an incoming connection request is selected randomly, and ISS1 which selects the ingress router with the minimum number of backlogged traffic batches. Note that for both selection schemes the egress router for a connection requests are selected randomly.

3.1.3 Achieved Results in Y3

This section presents some numerical results for a single and a dual-homed scenario. The topology under consideration is the NSF network with 14 and 20 bidirectional fiber links with 16 wavelengths each. The routing algorithm used at the PCE is a shortest path algorithm that tries to balance the number of resources used in the network, i.e., Least Loaded Routing algorithm. Wavelengths are assigned in a distributed way during the signaling phase, using a random fit approach. In the proposed scenario we assume to have 14 terminals for the single-homed case (i.e., one per each router), and 7 terminal for the dual-homed, i.e., each terminal is assigned two routers. Connection requests arrivals are assumed to follow a Poisson distribution while the connection requests service time is exponentially distributed. The considered value of the load is equal to 160 Erlangs. In addition, deterministic arrivals of connection requests with an arrival rate equal to the Poisson average arrival rate have been considered. Router lockdown time is deterministic and it (slightly) increases linearly as the number of connection requests increases (typical values are in-between 5s and 7s).

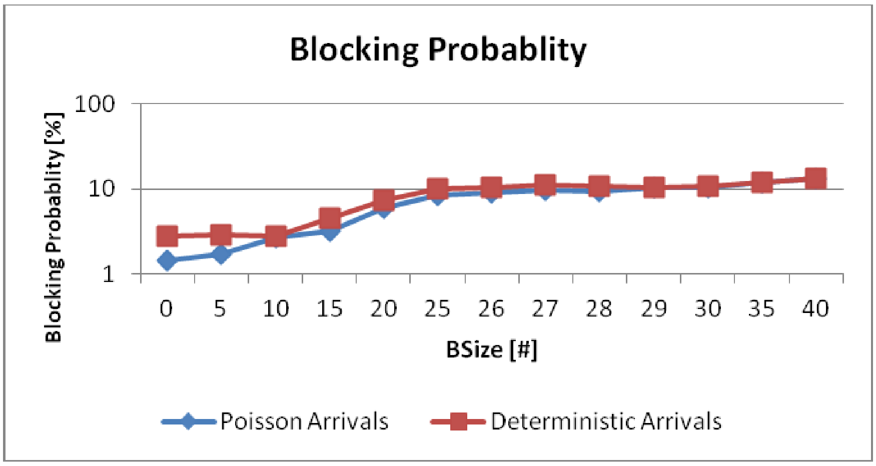


Figure 2: Blocking Probability [%] vs. BSize [#]

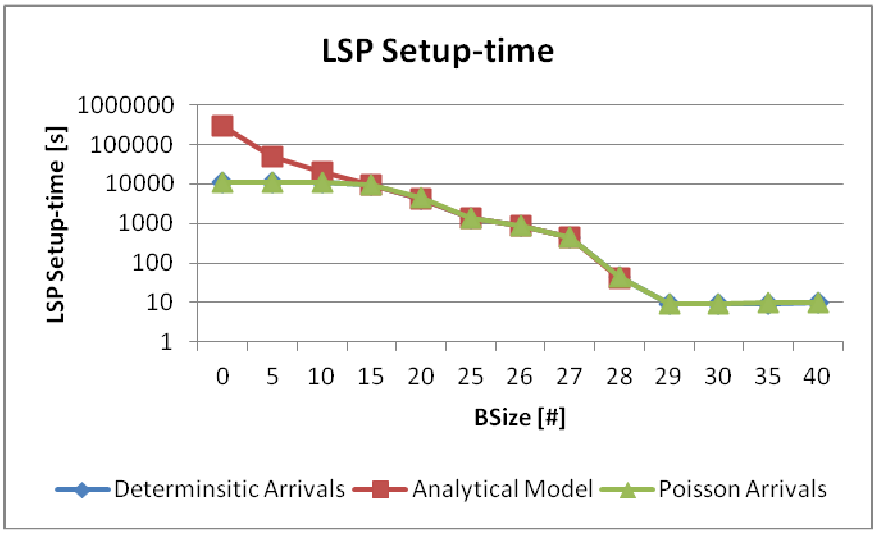


Figure 3: LSP Setup-time [s] vs. BSize [#]

Figures 2 and 3 present the value of the blocking probability and the LSP set up time as a function of the batch size, respectively. The scenario considered is the single-homed case. The plots confirm the intuition that by grouping the connection requests before sending them to the PCE for path computation has a beneficial effect on the connection setup time. This reduction comes with a slight increase of the blocking probability.

Results for the set up time are also compared with the analytical model. The model is based on a $D/D^B/1$ queue and models the service performed by the router to which the terminals are connected. Connection requests arrive at each router from the NC. The arrival rate is modelled with a deterministic arrival rate even though the bulk service implemented by the PCS is based on the batch size and not on a timer. In the model, connection request arrivals are not single connection request arrivals but are batches of connection requests that have the considered router as ingress router. Therefore, batches are actually a set of batches of connection requests coming from the PCS.

In general, the expression of the average delay for a $D/D^B/1$ queue that has a deterministic and fixed T_{lock} time and the service time is negligible with respect to T_{lock} is:

$$\overline{d}(n) = -\frac{B \cdot (n-1) - \lambda \cdot n \cdot T_{lock} + 1}{2\lambda} \qquad (1)$$

where B is the bacth size, λ is the connection interarrival time, and n is the number of considered batches (i.e., the number of batches served during the observation time of the system). By imposing $\overline{d}(n)=0$ and taking the limit for $n \rightarrow \infty$ it is possible to derive the value of the batch size that optimizes the delay as:

$$B^* \geq \lambda T_{lock} \qquad (2)$$

The model can also be valid for modeling the PCS behavior, considering, however an interarrival time scaled (i.e., divided) by the possible combinations of ingress-egress routers.

Similar results are obtained also in the case of the dual-homed scenario (Figure 3 and Figure 4). In this multi-homed scenario it can also be seen that there is an advantages (in terms of reduction of the set up time) in considering the backlogged traffic when selecting the ingress router (i.e., ISS1 scheme). This advantage comes with a price in terms of blocking probability. This is mainly due to the fact that the ingress router with the least backlogged requests may not necessarily be the router on the shortest path from the source and the destination terminal.

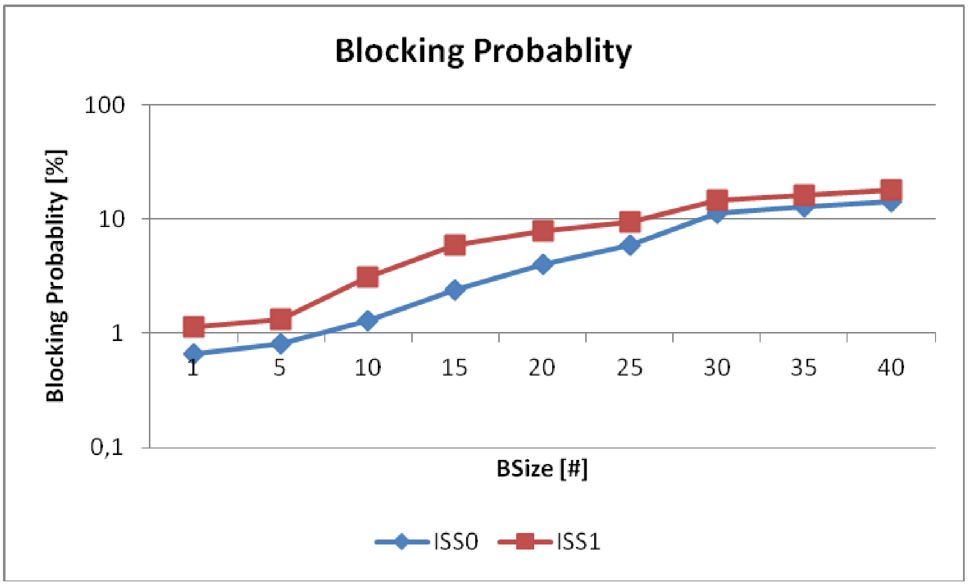


Figure 3 Blocking Probability [%] vs. BSize [#]

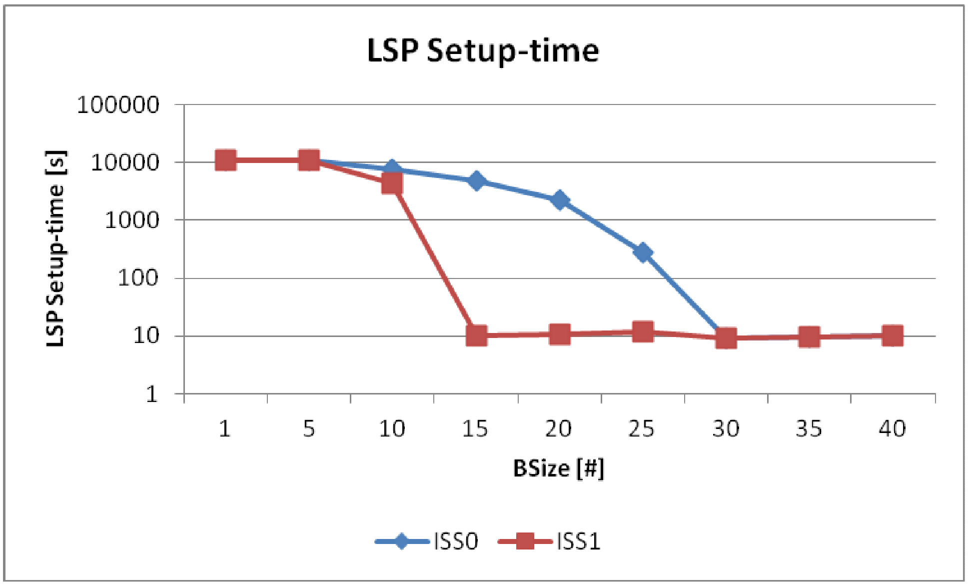


Figure 4 LSP Setup-time [s] vs. BSize [#]

3.1.4 Achieved Results in JA1

JA1 carried on studies on service interconnection fault tolerance by implementing the following activities: the evaluation of PCE-Based batch provisioning of grid service interconnections through experiments, analytical models, and simulations and the evaluation of the reliability of service interconnections in integrated wired and wireless networks.

The first activity showed that feasibility of the integration between PCE (proposed by the IETF) and RACF (part of the ITU-T NGN). This feasibility was proven, in particular, by the set up of an experimental testbed in the SSSUP lab that was remotely controlled by personnel at AGH. In addition the performance of a service interconnection provisioning policy based on bulk service of connection requests was evaluated. The policy resulted effective in minimizing the service interconnection provisioning time when routers are lock down (i.e., they are irresponsive to UNI set up command) when they are setting up connections. In addition, it was shown that, in case of end-user dual homing, a policy that selects the working connection ingress router based on the backlogged connection requests is advantageous from the viewpoint of connection set up time but it my suffer from a higher blocking probability.

In the second activity related to the issue of providing reliable service interconnections in integrated wired and wireless networks, three Mobility Anchor Point (MAP) selection schemes have been evaluated analytically in two sample fixed networks: ring topology and tree topology. Current MAP selection schemes can be efficient in minimizing handover latency but not as efficient in terms of availability. In this activity the impact of current MAP selection schemes on the end-to-end connection availability between mobile client (MS) and fixed server has been evaluated. The considered MAP selection schemes are three distance-based selection schemes. In the first scheme, the farthest MAP selection scheme, the MS selects the farthest MAP in terms of hops among the ones whose advertisements were received by the MS. The Farthest MAP selection (FMS) scheme has been already proposed in the literature and it is utilized by fast moving MS to minimize the number of handovers. In the second scheme, the Closest MAP selection (CMS), the advertised MAP closest to the MS is chosen. CMS can potentially improve the availability of the connection between MS and MAP but it implies a higher number of handovers. In the third scheme, the Half-Way MAP selection (HMS) scheme, the advertised MAP at a distance closest to the average of all the distances advertised by the MAP to the MS is chosen by the MS. The HMS scheme represents a compromise between availability and number of handovers. The results obtained for the ring topology showed that HMS availability is 4% better than the one obtained by FMS while requiring the double of handovers. The average availability achieved by CMS is 12% better than the one achieved by FMS but the number of handovers required is 20 times larger. Therefore, HMS is a good compromise between the obtained availability and the number of required handovers. They confirm that placing the MAP at a higher level of tree improves the linear availability but it needs an exponentially increasing number of handovers.

3.1.5 Conclusions

JA1 focused on the issue of provisioning fault tolerant and QoS-guaranteed service interconnections. Several aspects of the issues have been explored in both fixed network and integrated fixed-mobile networks. In fixed networks it was shown that a centralized service interconnection provisioning based on the integration of PCE and RACF can guarantee fault-tolerance through dual homing and low service set up time. If ingress routers are lock down during connection set up, the latter can be achieved through a bulk service policy.

In fixed-mobile networks it has been shown that there is a trade-off between number of handovers and connection availability when a mobile station (MS) chooses its mobility anchor point (MAP) for accessing services. In general it has been shown that a scheme that chooses a MAP that is half-way along its trajectory achieves a good compromise between availability and the number of required handovers.

A3.1.5.1 References

[1] K. Knightson, N. Morita, and T. Towle, “NGN Architecture: Generic Principles, Functional Architecture, and Implementation,” IEEE Communications Magazine, 2005.

[2] “Homepage of hessian,” http://hessian.caucho.com

[3] “Spring framework,” http://www.springsource.com

[4] WiMAX Forum End-to-End Network Systems Architecture Stage 2–3, Release 1.0.0, March 2007 (http://www.wimaxforum.org/resources/documents/technical/release1) [accessed on May 1, 2009]

[5] R. J. Koodli and C. E. Perkins, “Mobile Inter-Networking with IPv6: Concepts, Principles, and Practice”, John Wiley & Sons, 2007

[6] H. Soliman, C. Castelluccia, K.E. Malki, L. Bellier, “Hierarchical Mobile IPv6 Mobility Management (HMIPv6)”, IETF RFC 4140, August 2005

[7] S. Pack, M. Nam, T. Kwon, Y. Choi, “An Adaptive Mobility Anchor Point Selection Scheme in Hierarchical Mobile IPv6 Networks”, Computer Communications, vol. 29, n. 16, 2006, pp 3066 – 3078.

3.1.6 Published Papers (joint and single partner)

L. Valcarenghi, P. Korus, F. Paolucci, F. Cugini, M. Kantor, K. Wajda, P. Castoldi, “Experimental Evaluation of PCE-Based Batch Provisioning of Grid Service Interconnections”, to appear in Proceedings of Globecom 2009, Nov. 30-Dec. 4, 2009, Honolulu, HI, USA

L. Valcarenghi, P. Monti, I. Cerutti, P. Castoldi, L. Wosinska, “Issues and Solutions in Mobile WiMAX and Wired Backhaul Network Integration”, ICTON 2009, Jun. 28-Jul. 2, 2009, Island of Sao Miguel, Azores, Portugal [INVITED PAPER]

3.2 JA 2 - Cloud Computing

Participants: UESSEX, RACTI, SSSUP

Responsible person: Dimitra Simeonodou (UEssex)

3.2.1 JA2 Objectives

A new architecture of a service aware optical network has been proposed, suitable for cloud computing applications. The networking paradigm introduces service awareness in the core, by creating self-organized islands of service transparency. A service island consists of a single (non-network) resource and a group of networking nodes that constitute the shortest path towards that resource. This group of nodes is service transparent and thus upon a service request, end-users' data are transparently forwarded, to the island's resource and not outside it. The proposed architecture and particularly the service islands are self managed entities in the sense that core nodes are self-organized in an ad-hoc fashion, based on multi-criteria path selection algorithms, thus adapting themselves to updated networking or non- networking conditions.

3.2.2 Objectives of Y3

Propose suitable GMPLS extensions to implement the proposed service-aware network architecture.

3.2.3 Achieved Results in Y3

The proposed concept can be implemented with currently proposed GMPLS extensions. The concept of forwarding adjacency proposed by K. Kompella et al. [0] can be used here. The concept of forwarding adjacency was proposed to aggregate label switched paths (LSPs) by creating a hierarchy of such LSPs. Routing protocols like open-shortest-path-first (OSPF) floods information for FA links as for normal TE links and thus each Label Switching Router (LSR) may keep FA LSPs in its link state database along with its TE information. A forward adjacent LSP is treated as normal TE links by LSRs. To make use of FA notion in the proposed service aware networking scheme, FA-LSPs must be setup dynamically between more than two GMPLS nodes. In that case, each FA-LSP is initiated from the first node (egress router), adjacent to a resource. The newborn FA-LSPs (single link), can be created statically whereas we may assume that there exists an external mechanism (software) to pass performance metrics (i.e CPU availability etc) of that resource. These non-network metrics, together with networking ones must be communicated upstream to extend the service specific FA-LSPs. This is done by network layer protocols like OSPF-TE that keep and flood link state information. To this end and according to 0, the attributes of a FA-LSP are inherited from the LSP that induced its creation. Thus, TE information may propagate with the extension of the FA-LSP. During that process, LSR treats FA-LSPs similar to other TE links and compute which forwarding adjacency to join in a similar way with TE links for path computation. To this end, each LSR node may autonomously decide to join (or not) a FA-LSPs, thus extending a "service island".

It must be noted here that in our case, an LSR has only FA-LSPs in his database on a per service basis and this set of FA-LSPs correspond to the number of resources of that specific service in the network. Each upstream LSR can then perform path computation and decides which FA-LSP to join and thus associate the destination IP address (non-network resource or egress router) of the path message to the computed next hop for that path message. When the service island splits towards different ingress nodes, then the FA-LSP also split to sub-FA-LSPs. In other words, if there exists more than one FA-LSP that originate from different LSR and end to the same LSR, then that common LSP multiplex the two FA-LSPs into a single FA using the concept of Link Bundling. Upon island re-organization and path re-computation, the node may join other FA-LSP, thus directing new service requests to other resources. In that case, ongoing flows continue to be forwarded to the old destination via normal OSPF routing. The old FA-LSPs are not torn down but may start decreasing in size (detaching links), thus ending with a single link between the resource and egress router. Information may continue to be upstream advertised with no upstream node joining it.

3.2.4 Achieved Results in JA2

A new architecture of a service aware optical network has been proposed, suitable for cloud computing applications. The networking paradigm introduces service awareness in the core, by creating self-organized islands of service transparency. A service island consists of a single (non-network) resource and a group of networking nodes that constitute the shortest path towards that resource. This group of nodes is service transparent and thus upon a service request, end-users' data are transparently forwarded to the island's resource and not outside it. The proposed architecture and particularly the service islands are self managed entities in the sense that core nodes are self-organized in an ad-hoc fashion, based on multi-criteria path selection algorithms, thus adapting themselves to updated networking or non- networking conditions.

During the 1st year, the proposed architecture was detailed; its basic notations and metrics were defined; as well as how the networking peers (nodes) interact to each other to form self-managed ad-hoc entities. Node interaction is based on resource discovery algorithms that "discover" information from both network and non-network resource providers and facilitates end-to-end service creation. In general, each service is composed of service attributes that can be different for different services and user requests. Potential implementation using currently proposed GMPLS extensions were also defined. Figure 5 shows the concept of self-organization. In this example, two kinds of services are offered, denoted as service A and B. Each one possesses two set of resources denoted as resource #1 and resource #2, connected to specific egress nodes. For each service, there exists a virtual service plane (replica of the control plane), where nodes are self-organized per resource in such a way that each service request within that network domain is transparently routed to that resource for execution. A service proxy is responsible for service addressing and is placed at the entry points of the network.

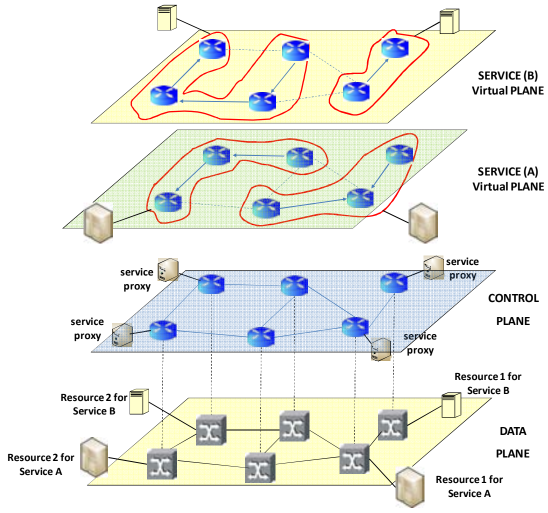


Figure 5: network paradigm for programmable service composition employing service transparent optical islands.

Figure 6 and Figure 7 show an example of a grid network and how network nodes are organized after routing table establishment. For the self-organization process, an optimization function Fs F_s has been used for path (island selection) as follows:

$$F_{s_f^k} = a_1 \frac{B_{R_i} - \mathrm{B_{min}}}{\mathrm{B_{max}} - \mathrm{B_{min}}} + a_2 \left(\frac{h_{R_i \to s_f^k} - h_{\min}}{h_{\max} - h_{\min}} \right)^{-1} + a_3 \left(\frac{D_{R_i \to s_f^k} - D_{\min}}{D_{\max} - D_{\min}} \right)^{-1}, \ldots, + a_k \frac{CPU_{s_f^k} - \mathrm{CPU_{min}}}{CPU_{\max} - CPU_{\min}}$$

where the min, max refer to the corresponding min, max values among the set of all candidate islands in the vicinity of node R_i. In the specific example of node organization, shown in Figure 7, the optimization function was

$$F_s = 50\% \frac{B_{R_i} - \mathrm{B_{min}}}{\mathrm{B_{max}} - \mathrm{B_{min}}} + 50\% \frac{CPU_i - \mathrm{CPU_{min}}}{CPU_{\max} - CPU_{\min}}, \text{ subject to } \quad D \le 80 \cdot h, h \le 3, ST \ge 40\ (a.u)$$

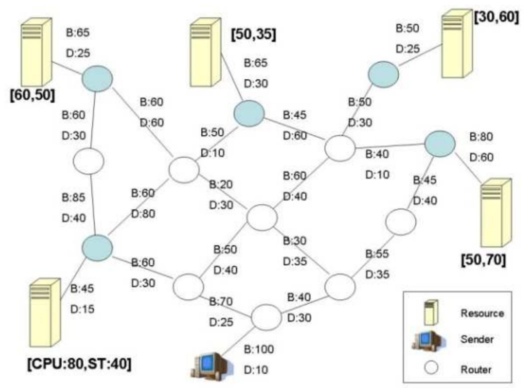


Figure 6: Grid network example consisting of five (non-network) resources and thirteen networking nodes. Numbers on links and resources denote the available bandwidth, delay, CPU and storage respectively.

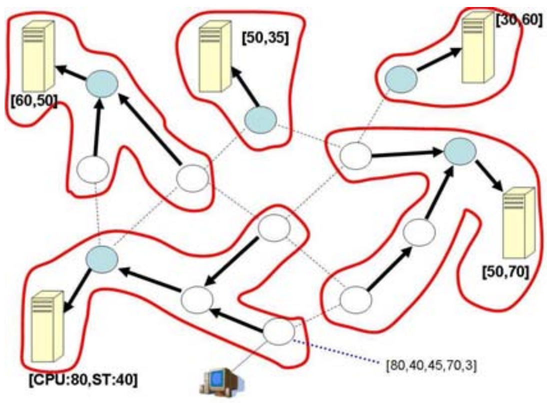


Figure 7: Network example after routing table establishment.

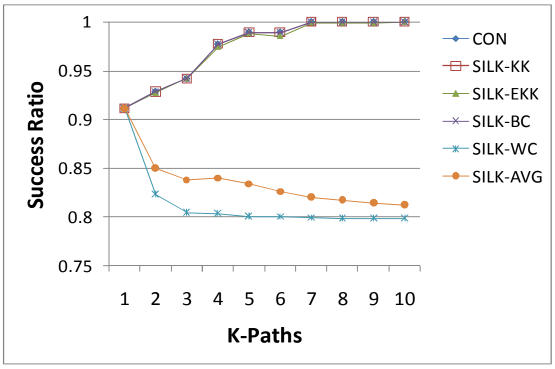
Resource discovery algorithms

Resource discovery algorithms enable resource state information discovery and dissemination by maintaining a resource state database of a resource domain (Figure 6) that facilitates service creation. A mechanism that provides an accurate representation of network and IT resources in an abstracted form with low signalling overheads under dynamic conditions was implemented. The abstracted representation is to allow resource providers to hide internal details of their resource (network and IT) while exposing enough information for successful service compositions. This gives resource providers the opportunity to participate in service creation at whatever level of abstraction they desire. At the highest level of abstraction, service oriented resource discovery requires information about service-end nodes and the connectivity between them. Service end-nodes are nodes that make and/or execute requests for services.

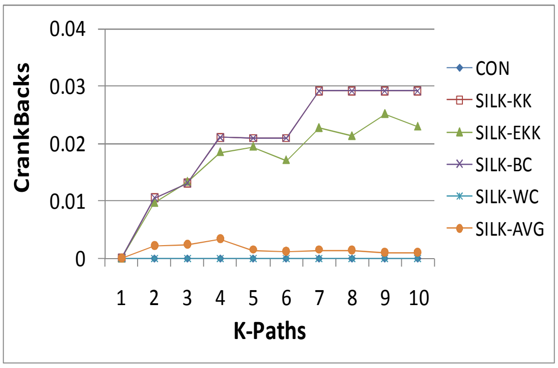
The challenge is to find possible ways to accurately represent resources without exposing too much information about the resource details so that the following objectives are achieved: a) maximise the number of requests accepted, b) minimise the number of possible requests accepted by the abstracted topology, which cannot be satisfied on the physical topology, called crankbacks, c) minimise the number of resource-state update messages propagated to the service plane. To meet the above conditions (a, b and c) the following issues were addressed: (i) Highly accurate abstract structural representation of the resources (ii) Accurate approximation of resource metrics (iii) Adequate monitoring of resources to provide accurate information to the service plane (iv) Implementation of a mechanism to trigger resource updates.

Four network abstraction algorithms were compared: (i) Best Case (SILK-BC) (ii) Worst Case (SILK-WC) (iii) Average Case (SILK-AVG) (iv) Modified Korkmaz-Krunz (SILK-KK) [1], against two novel algorithms (v) CON (vi) Extended Korkmaz-Krunz (SILK-EKK). Three levels of abstraction for IT resources were identified and compred: (i) Detailed (D), (ii) Multiple aggregate (MA), (iii) Single aggregate (SA). Resource-state update algorithms adequately monitor resource-state to provide accurate information to the service plane. A novel hybrid based (HB) update algorithm computes updates resource-state using a combination of event based and timer based update algorithms.

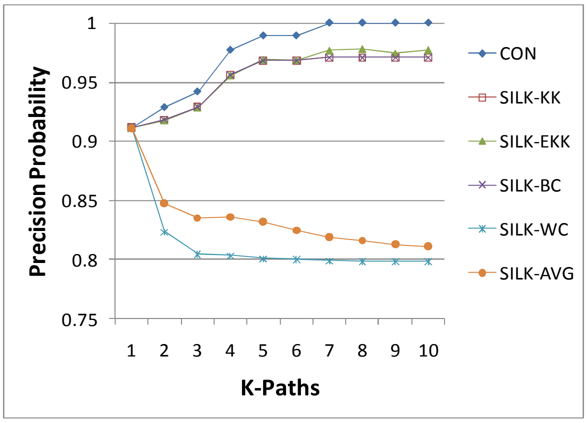
The algorithms were implemented on a variety of network topologies including GEANT European Topology and the US NSFNET under static and dynamic conditions.



a)



b)



c)

Figure 8 – Static Case: The effect of the number of paths used to calculate the connectivity between service-end nodes (a) success ratio (b) crankbacks (c) precision probability

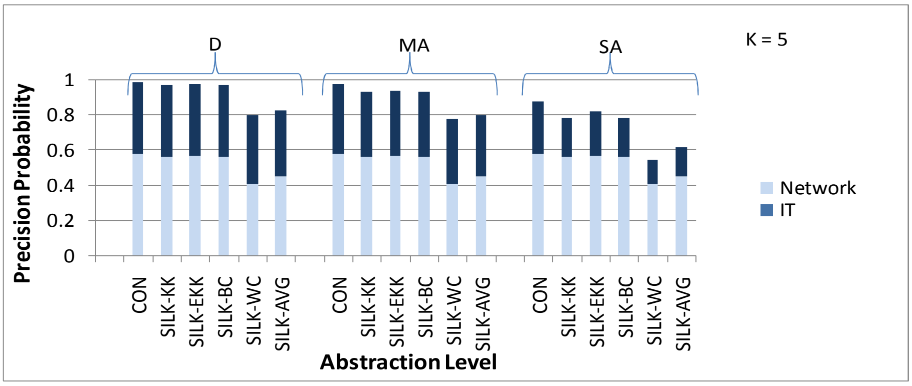


Figure 9 – Static Case: Impact of IT resource abstraction and network abstractions on the precision probability

Under static conditions, SILK-EKK and CON performed best on all evaluation metrics. SILK-BC and SILK-KK perform comparatively. This is because these abstraction models overestimate the network state information by advertising the best delay and bandwidth values. This results in a high probability of accepting feasible paths in the abstracted models. Contrarily, SILK-AVG and SILK-WC have the worst performance in all evaluation metrics. On the interactions between network resource abstractions and IT resource abstractions, it is observed that at lower abstraction (D) levels, the overall performance of IT resources is better than higher levels. This decrease in performance is due to the nature of constraints placed by clients.

Under dynamic conditions, SILK-WC and SILK-AVG are excluded from the results because of their poor performance. SILK-BC and SILK-KK are excluded because they exhibit analogous success ratio performance, which is slightly worse than SILK-EKK. The threshold values chosen are shown in Table 1. Update schemes A – D use the Hybrid algorithm (HB), E - F use the Time-Based algorithm (TB), and G - H use the Event-Based algorithm (EB) [2]. The range (max and min values) over all abstraction models was observed.

	A	B	C	D	E	F	G	H
Timer	3000	1500	3000	1500	3000	1500		
UpperThreshold1	16	16	14	14				
UpperThreshold2	14	14	12	12				
LowerThreshold1	4	4	6	6				
LowerThreshold2	2	2	4	4				
Number λ status change							2	16

Table 1 - Threshold Parameters for Update Algorithm

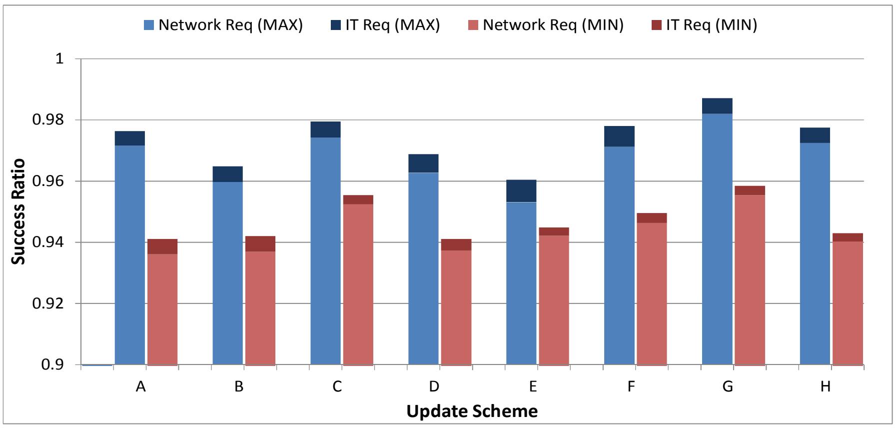


Figure 10 – Dynamic Case: Measure of success ratio on abstraction models using the Hybrid Update Algorithm

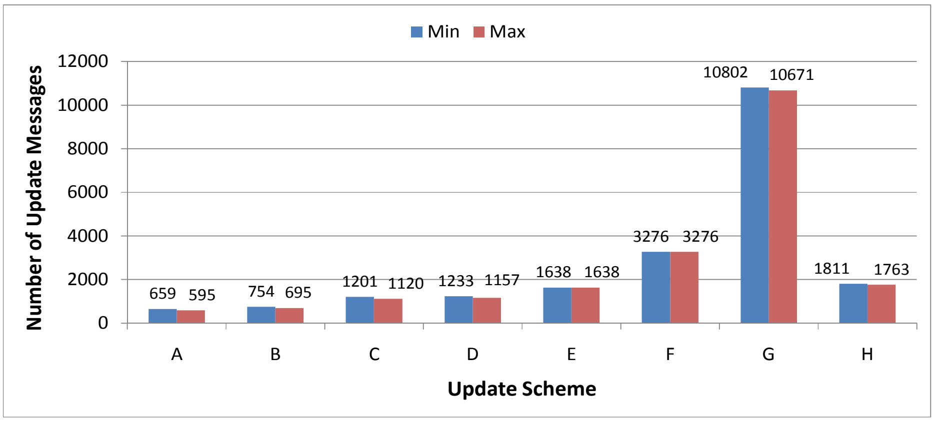


Figure 11 –Dynamic Case: Number of Update Messages

The results obtained from the simulation show that our resource abstraction and update policy algorithms disseminate resources-status information with high accuracy and precision and low signalling overheads. This is an important consideration in the introduction of a service plane since the resource discovery mechanism is responsible for the retrieval of information about the resources that are used to compose services. This in turn determines the efficiency and scalability of the service creation process and the service plane.

A3.2.4.1 References

[1] K. Kompella, Y. Rekhter, “Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)” RFC 4206, Oct. 2005.

3.2.5 Published Papers (joint and single partner)

[1] Kyriakos Vlachos and Apostolos Siokis, “A Service-Transparent and Self-Organized Optical Network Architecture”, in proceed. of ICC 2009, Page(s):1 – 6, 14-18 June 2009.

[2] Kyriakos Vlachos and Apostolos Siokis, “Performance Evaluation of a Service-aware Optical Network Architecture”, submitted to Elsevier Optical Switching and Networks

3.3 JA 3: Service Plane Functionalities and Demonstration

Participants: SSSUP, UEssex, UNIBO

Responsible persons: B. Martini (SSSUP)

3.3.1 JA3 Objectives

The evolution of IT applications (e.g., grid and cloud computing, on-demand multimedia services) towards often requiring effective high-speed connectivity and delay-constrained delivery services across the network, calls for the deployment of new service-related facilities in the optical network infrastructure (e.g,, support of end-to-end Quality of Service (QoS) and Service Level Agreement (SLA)). Currently, such facilities are provided at application layer by the service infrastructure (e.g. web-services in Grid middleware [ogf], Content Delivery Network (CDN)) that off-line interacts with the network operator management system for the actual resource reservation across the network. This is a typical overlay approach, that may results in inefficiencies in the use of the network resources and in limited functionalities, whereas a more integrated and effective approach would be preferable.

This scenario suggests an optical network infrastructure with service-awareness capabilities in order to address connectivity demands issued by IT applications, while providing needed network resources according to SLA and QoS requirements. Moreover a new generation of IT services and applications are desirable that are provided with network-awareness capabilities in order that service and network resources can be uniformly managed in an integrated way while optimizing their usage [mana].

In this JA interactions between IT systems and the optical network are investigated in order to address reciprocal awareness of capabilities while optimizing respective operation. The objective is to theoretically evaluate and experimentally demonstrate the possible exploitation of service-oriented capabilities integrated within the optical network addressing the overall resource usage optimization and service reliability requirements. Specifically the main objective is the design of a service platform for optical networks enabling:

- the exposition of an abstract level of information regarding network resources to IT applications, i.e, technology-independent resource status and logical view of topology
- the automatic and self-consistent translation of service request parameters issued by IT applications into technology-specific and self-consistent directives used by network devices
- advanced network capabilities to improve reliability and preserve quality of service in case of congestions or failure events thereby maximizing the Quality of Experience (QoE) at user side

The proposed approach is based on capabilities provided by the Service Oriented Optical Network (SOON) architecture [jocn-soon]. SOON is a service platform that implements on-demand service provisioning functionalities for optical transport networks. Through a distributed signaling among service elements, i.e., Distributed Service Elements (DSE), SOON fulfills connectivity set-up requests, issued by applications in terms of end-host addresses and perceived quality of service, while performing self-consistent and network-wide device settings across the network. SOON allows for the network resource reservation for the benefit of IT application during the service set-up phase, i.e, in the context of application signaling, triggered by user request. Therefore resource usage can be optimized because only needed network resources are allocated and only at the moment of user request. Moreover, service reliability is improved thanks to network resilience mechanisms that directly support data exchanges among IT systems. Finally, in networks that involve different network technologies, SOON realizes a consistent end-to-end network configuration through the concatenation of domain-by-domain mapping between service request parameters and technology-specific network directives.

In order to enable this approach in an heterogeneous optical network scenario, three sub-activities have being carried out.

Two sub-activities consider two different core network technologies as reference scenario, i.e., Optical Burst Switching (OBS) and (Generalized)Multi Protocol Label Switching ((G)MPLS). Specifically, the first sub-activity is jointly performed by SSSUP and UEssex and consists in the implementation and validation of a Service-Oriented Multi-Granular Optical Network (SO-MGON) architecture that consists in a SOON platform operating on top of a OBS control and data planes. The second sub-activities 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 that enables applications to exchange semantically-rich messages with the network aiming at the reservation of needed resources across the network.

The third sub-activity is performed by FUB and aims at evaluating the service quality perceived by CDN systems, and hence by users, whenever MPLS-enabled network connectivity services are enabled and how such quality is preserved in case of failures or congestion within the network.

3.3.2 Objectives for Y3

During Y1 and Y2 a full implementation and demonstration of a service-oriented OBS network has been achieved [jocn-obs-soon]. During the same period, a full implementation of SIP-based service-oriented GMPLS network has been carried out [ofc2009sip-soon].

The objectives for Y3 are the follows:

- Demonstration of a SIP-based service-oriented GMPLS optical network as follow up of system set-up carried out during Y2
- Demonstration of a resilient CDN operating on top of a MPLS-enabled transport network and evaluation of impact on QoE in case of congestions or failures.

3.3.3 Achieved Results in Y3

Demonstration of SIP-based service-oriented (G)MPLS network

A complete validation of the SIP-based service-oriented GMPLS optical network that has been carried out. In particular the system performance has been evaluated in terms of its capacity to address multiple requests issued at a different rate, each requiring a single media flow channel (scalability analysis).

The testbed set-up used for validating the SIP-based service-oriented GMPLS optical network is shown in Figure 12. It consists of a core network interconnecting N_{ERC} client-side Edge

Routers (ERs) with N_{ERS} server-side ERs. Client-side ERs are logically connected to server-side ERs through a complete mesh of $N_{ERC} \times N_{ERS}$ LSPs across the core network. This ensures that a given overall bandwidth B is pre-allocated between each pair of ERs, whereas routing policies are dynamically applied to allocate portions of the bandwidth B (i.e. the required network resource) to different service requests.

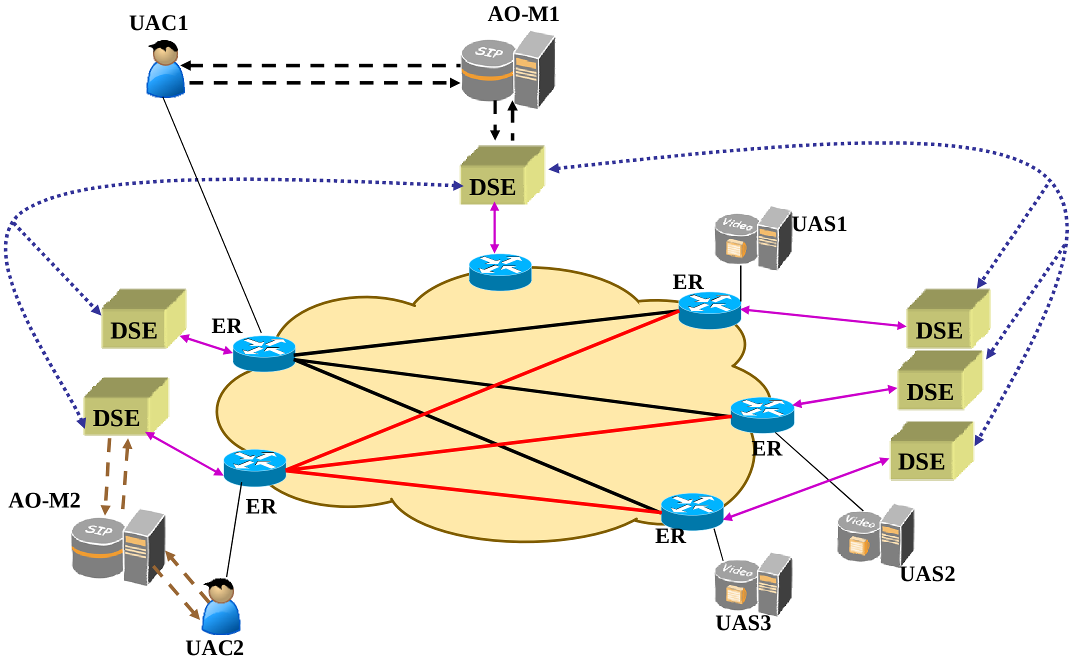


Figure 12 – **Generic topology used in the SIP-based service-oriented GMPLS optical network**

User Agent Clients (UACs) send service requests (i.e. SIP INVITE messages) to the most suitable DSE, which is in charge of coordinating all the DSEs and ERs involved in the resource setup. In Figure 12 this is done by the DSE co-located with AO-M1 for UAC1's requests, as shown by the blue dotted lines.

This topology has been emulated on top of the joint network test-bed available at SSSUP and UNIBO premises to perform the aforementioned scalability analysis. Two sets of experiments have been carried out: the first set was aimed at understanding the performance of the signalling mechanism for edge router configuration and, for this purpose, it has been carried out assuming LSPs with unlimited bandwidth; the second set was a complete characterization of the whole network solution including all the bottlenecks and, therefore, it has been carried out assuming LSPs with realistic available bandwidth values.

- First set of experiments

In the first set of experiments, a number of service requests has been generated from different UACs according to a stochastic Poisson process assuming different values of the average inter-arrival time (IAT). The number of client-side ERs was assumed equal to the number of clients: this assumption was motivated by the intention of considering the bottlenecks only in the server-side ERs, whereas client requests coming from different ERs is considered a realistic scenario. The parameters assumed in the first set of experiments are the following:

- LSP bandwidth: $B = \infty$
- Number of client-side ERs: $N_{ERC} = 100$
- Number of UACs: $N_C = 100$

- Number of server-side ERs: N_{ERS} = 2 / 4
- Number of User Agent Servers (UASs): N_S = 10 / 12, uniformly distributed behind the server-side ERs

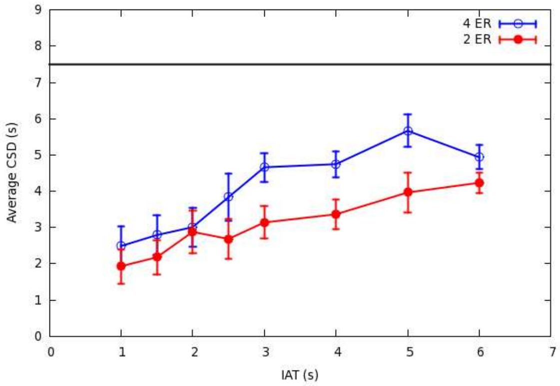


Figure 13 – **Average CSD vs. average request IAT**

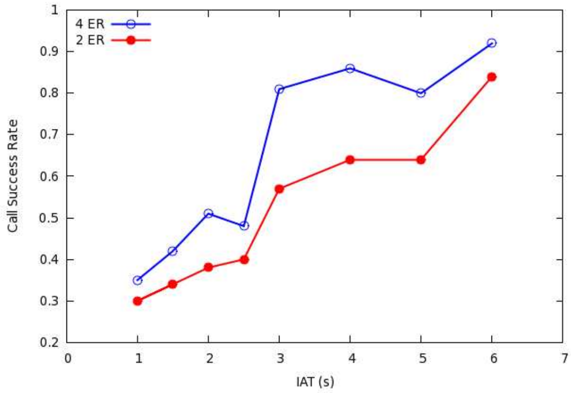


Figure 14 – **Average request success rate vs. average request IAT**

Figure 13 shows the average CSD [Y1531] (with 95% confidence interval) as a function of the average request IAT for the cases of NERS = 2 (NS = 10) and NERS = 4 (NS = 12). The CSD stays well below the 7.5s threshold recommended by ITU-T in Y.1530, even in case of high request rates (e.g., one per second). This demonstrates that the proposed signalling mechanism is quite scalable as the DSEs are capable of coping with multiple simultaneous requests.

The average CSD increases when the IAT or the number of server-side ERs increases. This might seem a bit contradictory, since it appears that the CSD gets better when the chance of contention on the ER configuration increase due to more frequent requests or less ER available. However, the measured CSD includes both the call success and failure cases and this affects the CSD average value. In fact, when a request is successful, that particular CSD includes also the router configuration latency, whereas this is not true in case of request failures caused by the unavailability of the server-side ER, which may be currently busy due to an on-going configuration for a previous request. Therefore, failed requests typically get quicker responses than successful ones. Considering the request success rate, shown in Figure 14 as a function of the average IAT, it is clear that the reduced CSD for small IAT values is mainly due to the majority of requests that fails.

A more accurate measure of the average CSD considering only the successful requests is shown in Figure 15. Here the CSD values are still below the 7.5s threshold and are quite stable also for different IAT values, demonstrating again the scalability of the proposed solution.

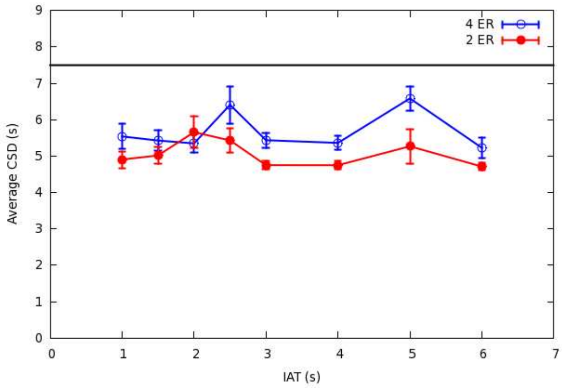


Figure 15 – **Average CSD vs. average request IAT for successful requests only**

- Second set of experiments

In the second set of experiments a limited amount of bandwidth is assumed for each LSP. The parameters chosen are as follows:

- LSP bandwidth: B = 1 Gbps
- Requested bandwidth for service: B_R = 20 Mbps
- Number of client-side ERs: N_{ERC} = 2
- Number of UACs: N_C = 10, uniformly distributed behind the client-side ERs
- Number of server-side ERs: N_{ERS} = 2
- Number of UASs: N_S = 4, uniformly distributed behind the server-side ERs
- Number of requests: N_R = 100

The experiments have been run assuming a deterministic IAT, with average value equal to 2s. The service duration time D (i.e. the time interval when the requested bandwidth is reserved) is either deterministic or exponentially distributed. The following table shows the results of the experiments.

Table 2 Experiment Result

Service duration time	Average CSD (95% confidence interval)	Average request success rate	Average CSD for successful requests only
150s (deterministic)	2.041867 (±0.609775)	0.29	4.869051 (±0.722349)
200s (deterministic)	4.531607 (±1.013995)	0.53	6.320941(±1.188086)
250s (deterministic)	3.280876 (±0.474260)	0.59	4.932275 (±0.087892)
200s (exponential)	3.568297 (±0.656136)	0.53	4.228167 (±0.833040)

The measured performances show a relatively limited influence of the service duration time, both in terms of value and distribution. The random behaviour shown here is due to the random choice of the UAS chosen for each request: this is the result of the emulation of a resource location mechanism that finds the most suitable UAS for the incoming request while trying to reduce the chances to find busy network resources (e.g. lack of LSP bandwidth and/or unavailable ERs being under reconfiguration due to a previous request). The CSD for successful requests is always below the 7.5s threshold. Extensive experiments will be performed in the future to obtain a more accurate characterization of the proposed service oriented network. Finally, analytical models based on queuing theory are currently under development, with the aim of providing a design tool for such a network which will be useful for resource planning.

Demonstration of a resilient CDN operating on top of MPLS-enabled transport network

A resilient CDN delivering VoD services has been achieved across a MPLS-enabled transport network by adopting a VPLS (Virtual Protocol Lan Service) techniques because they allows the CDN to take advantages from features of both restoration and traffic engineering capabilities. In fact, the problem for this type of video service is the reliability in case of increasing of the number of clients with a consequent wider bandwidth occupation on the backbone. At the moment, considering the low levels of the current video traffic, and in particular of the IPTV users, and the capacity of the current Telecom networks, CDN services would not manifest degradations due to traffic congestions. Conversely assuming huge video traffic increasing in the next years we aspect that CDN architecture could be strongly degraded by video traffic peaks. Therefore we believe that problems of admission control on the video servers will be necessary.

To validate the idea a test-bed has been set-up as shown in Figure 16.

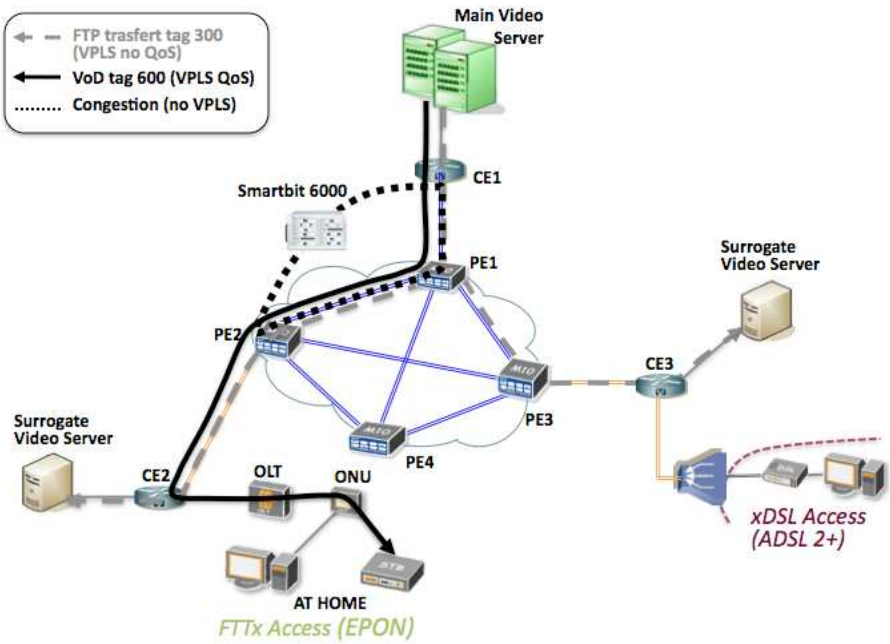


Figure 16. Test Bed Set up.

The test-bed consists in a complete metro network composed by four backbone routers Juniper M10i interconnected with the Roma – Pomezia - Roma single mode optical link at 1550 nm (50 Km). As Customer Edge (CE) routers three Cisco 3845 have been deployed and interconnected to the Provider Edge (PE) routers by Juniper with Multi-Mode fiber at 850 nm.

The access network consists of copper and fiber sections; in the first one, xDSL technique has been used, in particular ADSL2+ by Alcatel-Lucent 7300, ADSL2+ IP-DSLAM. In the fiber section, a FTTH solution has been adopted employing an Ethernet PON.

A CDN is a collection of network elements arranged for more effective delivery of content to end-users. CDNs can take various forms and structures. They can be centralized, hierarchical infrastructure under certain administrative control, or completely decentralized systems. There can also be various forms of internetworking and control sharing among different CDN entities. A CDN provides better performance through caching or replicating content over some mirrored servers strategically placed at various locations in order to deal with the sudden spike in content requests. The users are redirected to the surrogate server nearest to them. This approach allows to reduce network impact on the response time of user requests. In the context of CDNs, content refers to any digital data resources and it consists of two main parts: the encoded media and metadata. The encoded media includes static, dynamic and continuous media data (e.g. audio, video, documents, images and Web pages). Metadata is the content description that allows identification, discovery, and management of multimedia data, and also facilitates the interpretation of multimedia data. Content can be pre-recorded or retrieved from live sources; it can be persistent or transient data within the system. CDNs can be seen as a new virtual overlay to the Open Systems Interconnection (OSI) basic reference model. This layer provides overlay network services relying on application layer protocols such as HyperText Transfer Protocol (HTTP) or Real Time Streaming Protocol (RTSP) for transport. Client requests are redirected to the nearby surrogate, and a selected surrogate server delivers requested content to the end-users. Thus, transparency for users is achieved. Additionally, surrogates send accounting information for the delivered content to the accounting system of the CDN provider.

The goal of CDN proposed in this work is to deliver video content to subscribers of a VoD service, enabling (by RTSP) all the functionalities of a common video player (e.g., Play, Pause, Stop, Fast Forward, etc.). Each server has Video Lan Client (VLC) configured by Video Lan Management (VLM) console. Only the main server has all the VoD channels configured, whereas the surrogate servers have a clean configuration. Each VoD channel is defined in a configuration file that the software VLC reads every time it starts.

Each file is a Full HD Video (1920x1080p) encoded in H.264 video and multi-channel audio with a total bit-rate of about 9 Mbit/s. We use a mpeg2 TS transport stream, i.e., 188 bytes for MPEG packet, 7 MPEG packet for each IP packet. Each video server uses Microsoft Windows 2003 Server with Double Dual core Xeon Processor, 4GB of RAM, and 10K rpm SCSI Hard Drive. Servers are connected to CEs with Gigabit Ethernet copper interface and CAT5E cable.

In this work, one main and two surrogates servers have been employed, assuming that the structure can be replicated more times on the territory, even more complex than just two levels. As shown in Figure 17, the main server is where the new contents are charged. After a defined numbers of requests for the same content, the main server starts to copy that content in the nearest surrogate server to the site where the requests are originated. During the copy, the main server still answers the next requests for the same content. When the copy ends, the management layer of this infrastructure creates the new VoD channel on the surrogate server of the moved content. When the VoD channel is created, the new requests from the same site of the surrogate server are redirected to that surrogate server.

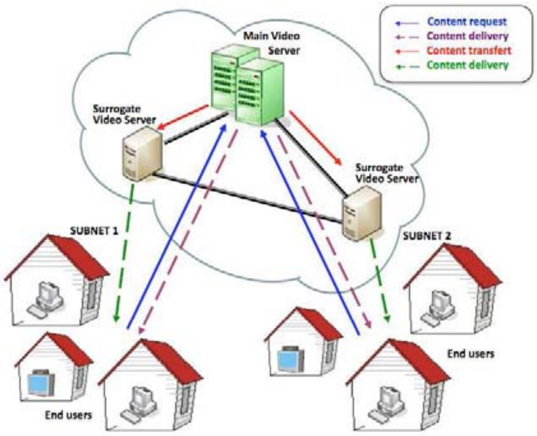


Figure 17. Logical CDN features.

To have a running-like real service, a web portal has been implemented to enable end users to choose which video enjoy. All the details can be found in [cdn-iptv-rndm].

Two kinds of possible impairments have been considered: link congestion and link failure. The first one is recovered by using different class of services for different flows, while the second one by using recovery procedures of VPLS: MPLS Fast ReRoute (FFR) and Standby Secondary Path technique (SSP).

As shown in Figure 17, there are two home access:

- FTTh with two devices connected to the ONU, a standard PC and a Set Top Box (STB) emulator;
- ADSL2+ with a PC connected to the modem.

This test aims at demonstrate how the network works in a situation of overloaded backbone link and how the service is degraded in this condition. Traffic generator, Smartbit 6000, overloads the links between CE1, PE1 and PE2, sending 1 Gb/s load traffic through these three routers.

The VoD traffic among the servers and the STBs, and the FTP traffic among the three servers are delivered using two different VPLS instances. First one, with tag 600, is for VoD and second one, with tag 300, is for FTP traffic. Within the VPLS, two different CoSs are implemented: tag 600 has priority over tag 300 (no priority).

The first test is to verify that in this condition no impairment was seen in the VoD video requested by the user during congestion and during FTP transfer.

The second test aims at verifying how the service works in case of no priority option set on the other traffic in the network. To do this we have modify one of the video to obtain a constant bitrate of about 20 Mbit/s and thus to get a good throughput graphic. We sent the same video to one computer with two logical Ethernet interfaces, one on tag 600 and one on tag 300. The traffic with tag 300 was sent without priority as the Smartbit traffic. On the computer that works as a double STB we record the two videos and the relative .cap file with Wireshark (i.e., software for packet sniffing). The result of this test is shown in Figure 18. During the congestion the throughput graphic presents a bitrate decrement for no quality VoD service that can be noted on the video due to the heavy packet loss.

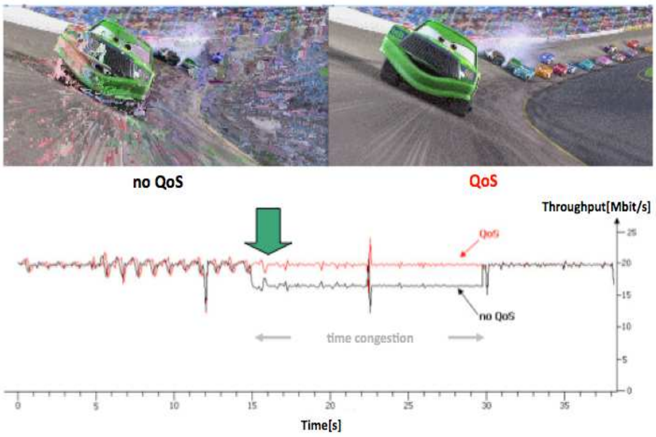


Figure 18. Link Congestion Tests.

This tests demonstrated that by means of VPLS technique a full control of the QoS for VoD services is achieved that can be extended to real-time services. In case of link failure VPLS Fast Reroute is able to recovery the service in less than 50 ms (SDH recovery time) with a little impact on the service. Using CDN architecture over a MPLS-enabled network is not necessary at the moment but we believe that it will be highly required when the number of subscriber of VoD services will remarkably increase as we expect in the next years.

3.3.4 Achieved Results in JA3

- Design and experimental validation of service-oriented OBS-based optical network enabling on-demand bandwidth reservation and switching capability at wavelength and sub-wavelength granularities to IT applications. Service-aware connection establishment and QoS provisioning has been demonstrated while guaranteeing fast provisioning and recovery [jocn-obs-soon].
- Design and experimental validation of a service-oriented GMPLS-based optical network exploiting the concept of session by integrating SIP-based procedures within network resource reservation activity, which enables a stateful approach. Performance characterization of the proposed signalling scheme has been achieved, in terms of service provisioning delay that address ITU-T requirements and in terms of scalability [ofc2009-sip-soon][ofc2010-sip-soon].
- Demonstration of full QoS control enabled by VPLS technique for the benefit of CDN operating on top MPLS networks. Real-time services can benefit of negligible service degradation thanks to service recovery in less than 50 ms enabled by VPLS Fast Reroute.

3.3.5 Conclusions

The validation of service-oriented approach for network service set-up has been achieved through the employment of SOON architecture, thus demonstrating the feasibility of the proposed approach in the context of optical transport networks. The proposed service platform has been evaluated considering different network technology while a comprehensive demonstration of interworking between IT systems, i.e., CDN, and a MPLS-enabled network has been carried demonstrating benefit for service reliability.

The employment of service platforms for optical networks enables higher level of resource usage optimization as well as self-configuration capabilities with saving in operational costs. Moreover this service platform would enable new business models with advantage for both service and network providers. In fact integrated and coordinated provisioning of network and service resources allows service provider to tailor service offerings to users (i.e., different SLA) and network provider to profit from different level of QoS provided.

A3.3.5.1 References

[ogf] I. Foster, H. Kishimoto, and A. S. (eds), "Open grid service architecture v1.0," OGF, Tech. Rep. GFD-I.030, January 2004. [Online]. Available: http://forge.gridforum.org/projects/ogsa-wg

[mana] A. Galis, H. Abramowicz, and M. B. (eds.), "Management and serviceaware networking architectures for future internet," MANA Future Internet group, Position paper, 2009.

[Y.1531] ITU-T Rec. Y.1531 "SIP-based call processing performance", Nov. 2007

3.3.6 Published Papers (joint and single partner)

A3.3.6.1 JA3

[ofc2009-soon] F. Baroncelli, B. Martini, V. Martini, P. Castoldi, "Service Oriented Optical Network architecture", Conference on Optical Fiber Communication 2009 (OFC 2009), San Diego, USA, March 2009

[ofc2009-obs-soon] Y. Qin, G. Zervas, V. Martini, M. Ghandour, M. Savi, F.Baroncelli, B.Martini, P.Castoldi, C. Raffaelli, M. Reed, D. Hunter, R. Nejabati, D. Simeonidou "Service-Oriented Multi-Granular Optical Network Testbed", Conference on Optical Fiber Communication 2009 (OFC 2009), San Diego, USA, March 2009

[ofc2009-mgon] M. Savi, G. Zervas, Y. Qin;, V. Martini, C. Raffaelli, F. Baroncelli, B. Martini, P. Castoldi, R. Nejabati, D. Simeonidou "Data-plane architectures for multi-granular OBS network", Conference on Optical Fiber Communication 2009 (OFC 2009) , vol., no., pp.1-3, 22-26 March 2009

[jocn-obs-soon] B. Martini, V. Martini, F. Baroncelli, K. Torkman, and P. Castoldi "Application-driven Control of Network Resources in Multi-Service Optical Networks", IEEE Journal of Optical Communications and Networking (JOCN), July 2009

[ofc2009-sip-soon]B. Martini, A. Campi, F. Baroncelli, V. Martini, K. Torkman, F. Zangheri, W. Cerroni, P. Castoldi, F. Callegati "SIP-based Service Platform for On-demand Optical Network Services", Conference on Optical Fiber Communication 2009 (OFC 2009), San Diego (USA), March 2009

[im2009-racf] B. Martini, F. Baroncelli V. Martini, K. Torkman, P. Castoldi, "ITU-T RACF implementation for application-driven QoS control in MPLS networks", 11th IFIP/IEEE International Symposium on Integrated Network Management (IM 2009) New York, USA, June 2009

[jocn-soon]B. Martini, V. Martini, F. Baroncelli, K. Torkman, and P. Castoldi "Application-driven Control of Network Resources in Multi-Service Optical Networks", IEEE Journal of Optical Communications and Networking (JOCN), July 2009

[icton2009-vpls-soon]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" 10th Anniversary International Conference on Transparent Optical Networks 2008 (ICTON 2008), Athens, Greece, June 2008

A3.3.6.2 Y3

[ofc2010-sip-soon] W. Cerroni, B. Martini, , M. Gharboui, A. Campi, F.Baroncelli, P.Castoldi, F.Callegati "**Experimental validation of a SIP-based Platform for Service Oriented Optical Network**", Conference on Optical Fiber Communication 2010 (OFC 2010), San Diego, USA, March 2010

[jocn-obs-soon] G. S. Zervas, V. Martini, Y. Qin, E. Escalona, R. Nejabati, D. Simeonidou, F. Baroncelli, B. Martini, K. Torkmen, P. Castoldi "**A Service-Oriented Multi-Granular Optical Network Architecture for the Clouds**" submitted to Journal of Communications and Networking (JOCN)

[soon-cloud] F. Baroncelli, B. Martini, P. Castoldi “**Network Virtualization for Cloud Computing**”, Annals of Telecommunications, published by Springer, July 2010

[comcom-soon-ngn] F. Baroncelli, B. Martini, V. Martini, P. Castoldi “**Extending Next Generation Network (NGN) architecture for connection-oriented transport** ” Special Issue of Elsevier Computer Communications on "Next Generation Networks Service Management", February 2010

[gcmas-soon]F. Baroncelli, B. Martini, V. Martini, P. Castoldi “**A cooperative approach for the automatic configuration of MPLS-based VPNs**”, International Journal of Grid Computing and Multi Agent Systems (GCMAS), January 2010

[cdn-iptv-rndm] S. Pompei, M. Teodori, A. Valenti, S. Di Bartolo, G. Incerti, D. Del Buono, “Experimental implementation of an IPTV architecture based on Content Delivery Network managed by VPLS technique” Proc. of IEEE Reliable Networks Design and Modelling, Moskow, Russia, October 19-20, 2010

[summit-futint] F. Callegati (UNIBO), W. Cerroni (UNIBO), B. Martini (SSSUP), M. Gharbaoui (SSSUP), A. Campi (UNIBO), P. Castoldi (SSSUP) “**Configuration of Network Resources for Future Internet Application Services**” Proc. of 2010 Future Network and Mobile Summit, Florence, Italy, June 2010.

[glob-futint]F. Callegati, A. Campi, W. Cerroni, Transport Service for the Future Internet: Concepts and Operations, Proc. of 4th IEEE Workshop on Enabling the Future Service-Oriented Internet, in conjunction with IEEE Globecom 2010, Miami, FL, December 2010.

3.4 JA 4: Joint Optimisation of Grid and Network Resources

Participants: BME, SSSUP, PoliTO

Responsible person: Tibor Cinkler (BME)

3.4.1 JA4 Objectives

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.

3.4.2 Achieved Results in JA4

This activity has ended in year 2. The following summary refers to the achieved results in Y1 and Y2 and reported for completeness as described in deliverable D12.2.

To achieve JA4 objectives, the work was organized in sub-activities. The progress of the sub-activity is reported next.

Sub-activity 1 of JA4

Heuristic methods for reducing the state-space, thus reducing the time required for optimization, have been proposed. Furthermore, a simple greedy algorithm has been proposed that is based on recursive use of Dijkstra’s shortest path algorithm.

Sub-activity 2 of JA4

The model of PoliTO for optimizing storage resources over a single layer network is being extended with the multilayer optical beared network model of BME.

Sub-activity 3 of JA4

A task distributed between multiple distant GRID computers is considered executed if all subtasks are executed and all the sub-results are collected. Either a network or GRID-computer failure can hinder a task to be executed. Therefore, we consider the following strategies:

1. If network resources fail, in either direction at any time, a new connection will be used or set up either via protection or via restoration, respectively.
2. If any of the GRID computing resources fail the task will be sent out to closest fastest GRID unit to execute it as fast as possible not optimizing for the cost any more, since the result is obtained only if all subtasks are all completed.
3. All the sub-tasks are sent to two geographically distinct nodes connected via physically diverse network resources, the sub-tasks executed twice, and sub-results collected. This is the best from the resilience point of view, since multiple network and computing unit failures can be easily overcame with no delay at all, the accuracy of sub-results can be verified, unfortunately it waists both, computing and transmission resources.

Co-operation with JA7 was discussed. Particularly sub-activity 4 of JA4 that deals with resilience is closely related to JA7. Merging of JA4 and JA7 was also proposed, not yet decided.

Sub-activity 4 of JA4

To meet the QoS requirements of global grid-enabled applications, a dynamic and deterministic joint optimisation of both computational and network resources is necessary. In wavelength-switched optical network (WSON), the joint optimisation is needed when bandwidth-greedy applications require the provisioning of lightpaths among grid resources. A way of achieving this joint optimisation is the cooperation between the grid resource manager (GRM) and the network resource manager (NRM). The typical objective of GRM is to achieve the maximisation of the overall amount of grid services successfully established and the minimisation of delivery time subject to the required QoS constraints. However, the lack of network information at the GRM does not guarantee the efficient network resource utilisation and, in turn, the maximisation of the established grid services.

The introduction of the path computation element (PCE) within the control plane of optical networks is an appealing solution to provide efficient utilisation of network resources. The PCEP protocol is proposed here as the new standard interface between the GRM (behaving as PCC) and the NRM (embedding a PCE implementation). Two novel schemes for grid resource selection specifically designed to exploit the standard PCE protocol (PCEP) features are proposed. The schemes resort to the currently defined PCE architecture and PCEP implementation to provide a feedback to the GRM on the expected network resources utilised by alternative choices of computational resources. In this way, the GRM can evaluate in advance the impact on the network performance of different alternatives and, among them, select the one that minimises the grid service delivery time and the overall network resource utilisation. The proposed schemes are suitable in case of grid services requiring large connection bandwidth, high level of connectivity among multiple grid resources (e.g., full mesh of lightpaths) and long duration. These schemes do not require additional control plane extensions or interfaces specific for grid purposes. In addition, the proposed schemes preserve the requested level of confidentiality between GRM and NRM, concerning the detailed network information (i.e., link bandwidth availability and strict routes information are not disclosed by NRM). Therefore, the proposed PCEP-based solution can be implemented at NRMs, e.g., independent internet service providers (ISPs) not belonging to any grid virtual organizations.

Different grid-networking schemes with specific relationship between GRM and NRM are considered. The former two schemes are generally applied, the latter two are specifically proposed for joint optimization:

- *Dg scheme*: GRM selects the least loaded *g* computational resources. NRM is in charge of setting up network connections among them. Lightpaths are computed and provisioned by GMPLS protocols in a distributed way. The PCE is not utilized. Joint optimisation is not achieved.
- *Pg scheme:* GRM selects the least loaded *g* computational resources. NRM acts as PCE having full and detailed visibility of the network and may perform multiple and concurrent path computation. Network resources are better utilized, however joint optimization is not achieved.
- *Pkg scheme*: GRM considers the combination of all the possible least loaded *g* computational resources (assuming *k* equally loaded resources). NRM acts as PCE and is requested to perform multiple concurrent path computation of combinations furnished by GRM. NRM replies providing the GRM with path computation solutions including a TE metric representative of network QoS level. GRM selects the solution minimizing the combined grid-network metric.
- *PGkg scheme***:** in an offline PCEP-based communication, GRM requests a non-synchronised independent path computation between all node pairs connecting the grid resources potentially involved. PCE replies with a related static metric value representing a lower bound of the actual metric (i.e., optimistic value, considering the shortest path which is not always available). Such metrics are used by GRM upon new grid requests occurrences, identifying the *k* resources that minimise the considered metrics. Random selection performed in *Pkg* scheme is thus avoided. Successive steps are the same as *Pkg* scheme.

Scheme	*Description*	*GRM*	*NRM*	*Complexity (# path computations)*
Dg	Grid networking scheme not utilising the PCE	g	Distrib.	$p = g(g-1)/2$
Pg	Grid networking scheme utilising the PCE	g	PCE	$p = g(g-1)/2$
Pkg	PCEP-based scheme with feedback information	$k\ (>g)$ g	PCE	$cp = k!/(g!(k-g)!\ g(g-1)/2$ $p = g(g-1)/2$
PGkg	PCEP-based scheme with feedback information and a-priori communication	G $k\ (>g)$ g	PCE	$cp' = k!/(g!(k-g)!\ G(G-1)/2$ $cp = k!/(g!(k-g)!\ g(g-1)/2$ $p = g(g-1)/2$

Table 3: Considered grid networking schemes

The four schemes are summarized in Table 3, together with their associated complexity (i.e., number of required path computations), where *G* is the total number of grid resources in the network.

The considered schemes are evaluated through simulations considering the NSF network topology where each link carries 8 wavelengths and one grid resource per node is considered. The case with $g = 3$, and thus $p = 3$ bidirectional lightpaths is reported. Services are sequentially established and never torn down. The path computation is based on integer linear programming (ILP) formulation. The implemented objective function minimises the overall amount of used network resources and, as secondary objective function, minimises the load of the most loaded link.

Figure 19.1 and Figure 19.2 show the blocking rate versus the load. Results confirm that *Dg* and *Pg* achieve similar results. The two curves are almost completely overlapped, due to the low number of equal cost shortest routes of the topology. Scheme *Pkg* with $k = 4$ and $k = 5$ significantly improves the blocking rate compared to *Dg* and *Pg* schemes. The *Pkg* scheme with $k = 5$ achieves a higher blocking rate compared to the case with $k = 14$. The *PGkg* scheme, which performs upon service request 10 path computations, further improves the *Pkg* performance, closely approximating the optimal performance achieved by *Pkg* or *PGkg* with $k = 14$ (requiring, however, 364 path computations).

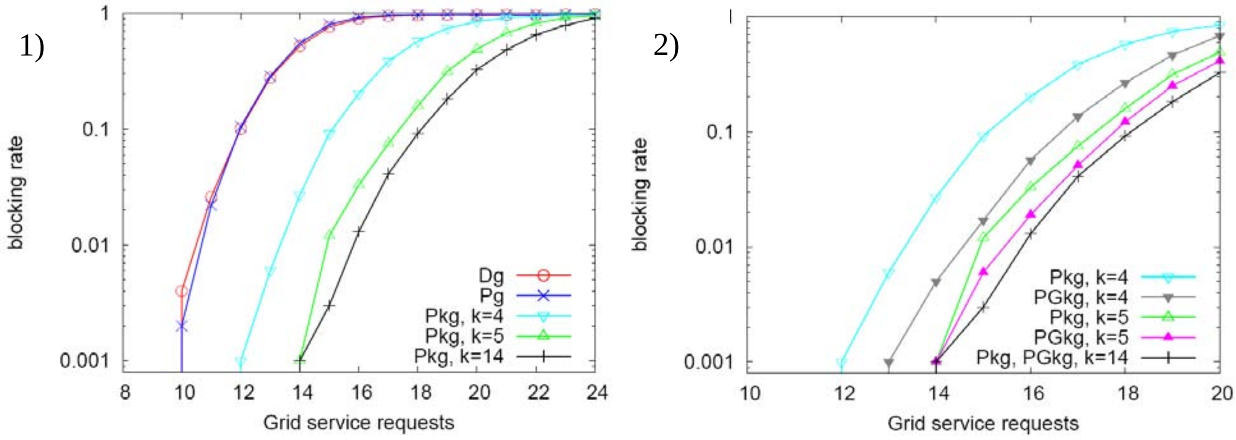


Figure 19: Simulation results

The proposed schemes are experimentally evaluated on the 'JGN2plus's metropolitan area WSON testbed (Figure 20). The GMPLS control plane runs on control PCs exploiting generic routing encapsulation (GRE) connections. The data plane consists of 4 wavelengths generated by gigabit Ethernet interfaces equipped with DWDM media converters. The NRM consists of two sub-blocks: the PCE and the lightpath activator. The PCE is implemented in a Linux PC and is based on C code and ILP formulations. The PCEP module is based on C++ and TCP socket libraries. The five edge nodes are connected to equally least loaded grid resources (i.e., $G = 5$). The GRM triggers one grid service *si*, requiring bidirectional lightpaths connecting in a full mesh $g = 4$ grid resources.

When the *Dg* scheme is applied, if *g* grid resources randomly selected includes router R1, the service is rejected since only one wavelength is available on the link connecting router R1 to OXC X1 (i.e., full mesh of lightpaths cannot be supported). If the selection excludes R1, the service may also be rejected because of lack of network resources on the other links (e.g, unless the lightpaths are assigned wavelength *w*0 for lightpath R2-R5 and wavelength *w*3 for lightpath R3-R4). The *Dg* scheme experiences a service rejection equal to 97.7%. When the *Pg* scheme is adopted, the latter critical condition affecting *Dg* is avoided thanks to the concurrent path computation. However, the former critical condition related to the random choice of R1 persists. Service rejection is reduced to 80%. By applying the *Pkg* and the *PGkg* schemes with $k = 5$, the service is successfully established since both previous critical conditions are avoided. The five combinations of $p = 6$ lightpath computations exclude router R1 and allocate the wavelengths that avoid service rejection.

The figure indicates also the sequence of PCEP message exchanged between the grid user and the NRM. *Pkg* and *PGkg* scheme requires an additional PCEP message. However, such additional PCEP does not significantly increase the overall grid service delivery since the overall lightpath activation latency is dominated by the high OXC switching time (e.g., some seconds). For this reason, grid services exploiting the current optical network technologies can be provisioned dynamically but at a limited rate.

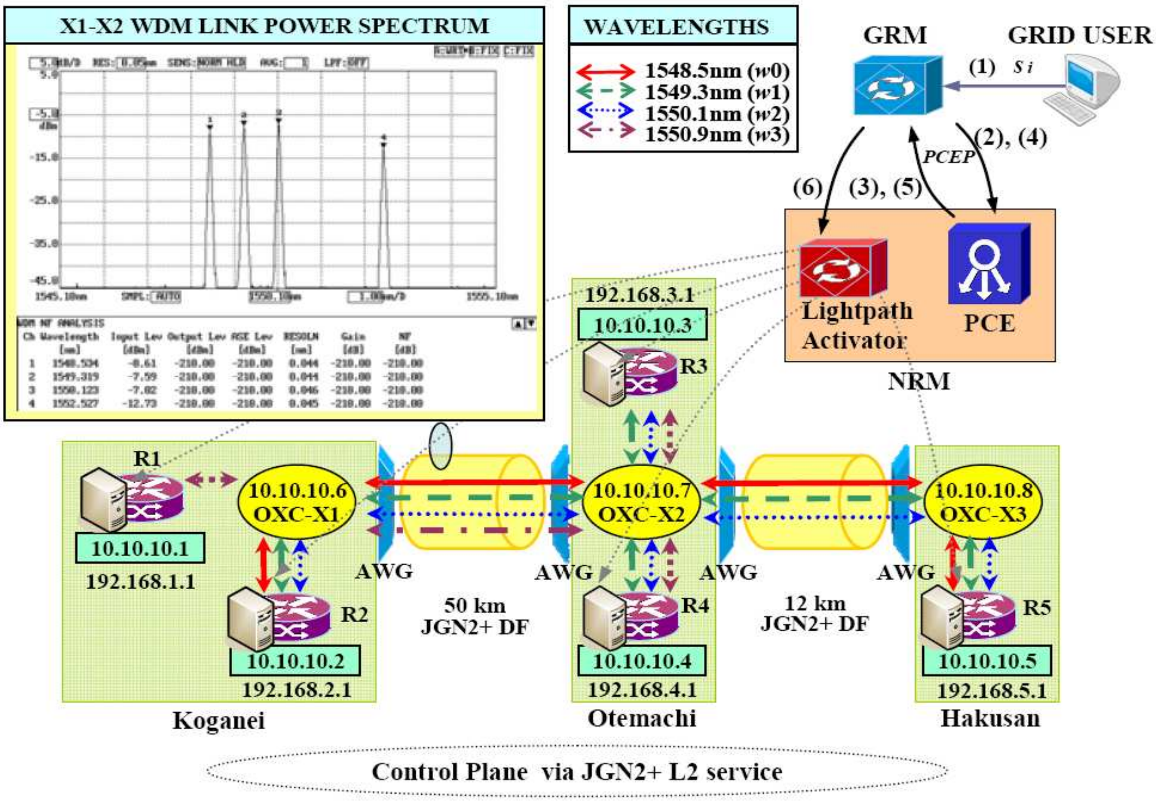


Figure 20: Experimental implementation on the JGN2plus testplan

3.4.3 Published Papers (joint and single partner)

1. J. Gál (BME), N. Szabó (BME), A. Ladányi (BME), T. Cinkler (BME), *Cost and Time Trade-off of Scheduling Grid Tasks over Grooming Capable Networks*, NETWORKS 2008, 13th International Telecommunications Network Strategy and Planning Symposium, Budapest, Hungary, September 2008.

2. F. Cugini (SSSUP), S. Xu (NICT), H. Harai (NICT), F. Paolucci (SSSUP), L. Valcarenghi (SSSUP), P. Castoldi (SSSUP), *Introducing Path Computation Element (PCE) in Optical Grid Networking*, ECOC 2008, September 2008.

3. F. Cugini (SSSUP), S. Xu (NICT), H. Harai (NICT), F. Paolucci (SSSUP), L. Valcarenghi (SSSUP), P. Castoldi (SSSUP), *Optical Grid Networking Exploiting Path Computation Element (PCE) Architecture*, International Journal of Communication Networks and Distributed Systems, Vol. 5, No. 3, pp. 246-262, March 2010.

3.5 JA5 - Programmable Service Composition Algorithms for Service Oriented Optical Networks

Participants: IT/Telecom-ParisTech, UESSEX, UoP, RACTI, SSSUP

Responsible person: Chinwe Abosi (UEssex)

3.5.1 JA5 Objectives

In the context of the research in this JA, service composition is the process of assembling independent, reusable service components across multiple providers to construct end-to-end services. It relies heavily on the orchestration of resources from several infrastructure providers, which provide the building blocks of basic services. Providers' resource information may be passed to the service provider in an abstract format to accommodate heterogeneity, scalability and confidentiality. Our focus is on high-end current and emerging next generation Internet applications, thus resource orchestration combines both network and non-network IT resources based on local resource-state information from individual, independent providers. Specifically, each service composition request includes service specific parameters which can be categorised into network specific and IT resource specific service parameter. For each service request, the user perceived QoS parameters need to be mapped into technology dependent network directives. Thereafter, service composition algorithms need to find an appropriate provider to provide each resource component of the request, which together form an end-to-end service. In cases where there are several providers able to provide the basic service components, the service composition algorithm needs to select one such provider to use taking into account fair management that defends the interests of both the end-users and the providers.

The objective of JA5 is to design programmable algorithms that support effective and efficient implementation of service oriented optical network functionality. The objective of JA5 was to

- Propose an architecture for programmable service composition.
- Implement a set of algorithms that enable dynamic coordination of resources to deliver application functionality as a service through:
    - Resource/Request Information Mapping,
    - Infrastructure Resource Abstraction,
    - Service Composition
    - Resource Provisioning,
    - Fair Management in Composing Services.
- Propose a set of algorithms that enable efficient collaboration across heterogeneous Infrastructure providers (IT and network providers),

3.5.2 Objectives for Y3

In Y3, JA5 focused on:

1) Enhancing the Self-Organising, Service Oriented Optical Network to create end-to-end services

2) Service oriented resource orchestration algorithms to optimize resource selection and allocation for dynamic service requests.

3) Resource provisioning for enhanced services for Cloud Environments.

3.5.3 *Achieved Results in Y3*

Enhancement to the Self-Organising, Service Oriented Optical Network

RACTI detailed a new networking paradigm, termed as SO-SOON ("*Self-Organizing, Service Oriented Optical Network*"), to introduce service awareness in the core optical network, by creating *self-organized islands of service transparency.* A *service island* consists of a single (non-network) resource and a group of networking nodes that constitute the shortest path towards that resource. This group of nodes is service transparent and thus upon a service request, end-users' data are transparently forwarded, to the island's resource and not outside it. The proposed architecture and particularly the *service islands* are self managed entities in the sense that core nodes are self-organized in an ad-hoc fashion, based on multi-criteria path selection algorithms, thus adapting themselves to updated networking or non- networking conditions.

With respect to services, a specific algorithm is proposed for service composition. This combines both network and non-networks resources and is called distance vector cost. The fields of the vector denote local network state information that are the values of the parameters taken into account to reach a destination address (service aware routing). These parameters are service specifics and constitute attributes to be requested, when the user request is submitted. These attributes can be categorized in network (such as bandwidth (B), end-to-end delay (D) and hop-count (h)) and non-network specific ones (such as CPU cycles or storage capacity for grid services).

Suppose that a user located at the ingress node R_i requests service S_r with the specific attributes 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:

$$\begin{bmatrix} B_{R_i \to S_r^k} \\ h_{R_i \to S_r^k} \\ D_{R_i \to S_r^k} \\ CPU_{R_i \to S_r^k} \\ ST_{R_i \to S_r^k} \end{bmatrix} \leq \begin{bmatrix} B_{job} \\ h_{job} \\ D_{job} \\ CPU_{job} \\ ST_{job} \end{bmatrix}$$

Thus, in order to create end-to-end services out of programmable multi-dimensional vectors, 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, 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 i.e. routers, switches, etc). The edges of the graphs are inter-connected with a unique *service composition algorithm as the one described above.*

Service-oriented resource orchestration

Demanding requirements of next-generation applications require simultaneous access to multiple heterogeneous IT resources interconnected by high-speed optical network. These applications demand a resource orchestration model that views and optimizes multiple heterogeneous resources in a holistic manner. UEssex proposed a new on-demand algorithm that views heterogeneous resource-types in a holistic manner. The proposed service-oriented resource orchestration model was based on optimised co-scheduling of heterogeneous resources owned by different infrastructure providers. The optimised orchestration algorithm aimed to serve a higher number of end-users while balancing the load over the multiple resource-types. For this purpose, a weighted scheduling cost was introduced and the adaptive cost-based algorithm was proposed. The adaptive cost-based algorithm decides which resources to assign jobs to. As it is an online scheduling algorithm, it has no knowledge of future jobs. Thus it tries to assign jobs in such a way as to improve the chances of accepting future requests. It achieves this by assigning jobs according to a weighted cost function. The weighted cost relies on the simultaneous identification of available resources of each resource-type. It ensures that jobs which require IT resources are assigned to the closest IT resource node with the most-free resource capacity. Jobs that do not require IT resources are assigned to resources with a trade-off between distance and available link bandwidth capacity. The weighted cost-based algorithm introduces weights based on the utilisation of resources to adapt to the state of underlying infrastructure. The proposed model aims to optimise the efficiency and performance from both the end-users' and infrastructure providers' point of view. An Integer Linear Programming (ILP) was presented to benchmark the performance of the algorithm. Simulations were run to confirm the scalability of the algorithm. The algorithm is also compared against existing resource allocation schemes.

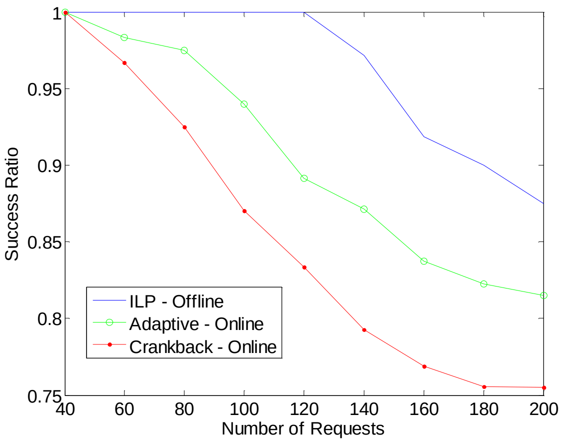
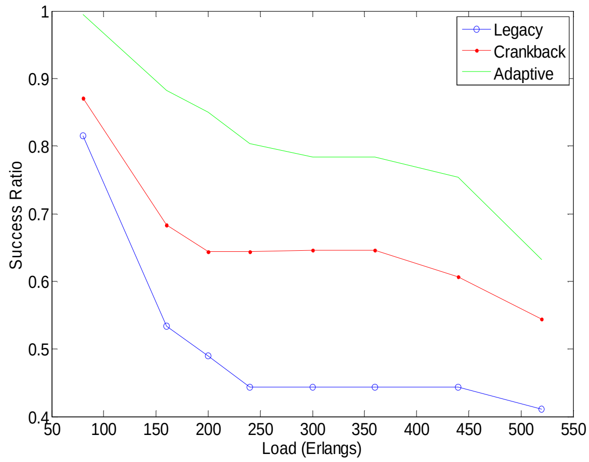


Figure 21: Success Ratio vs Number of Requests (6-Node Topology)



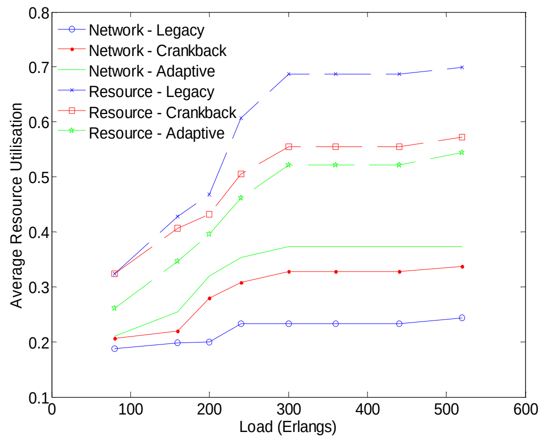


Figure 23: Success Ratio vs Load NSFNET 19-Node Topology

Figure 23: AvgResUtil vs Load 19Node NSFNET

Resource provisioning for enriched services in Cloud environment

Cloud services are based on the provisioning of computing, storage and networking resources able to satisfy requests generated by remote end-users. High speed Internet access and the emergence of multi-core Virtual Machines (VMs) enable today to enrich the types of services to be considered in Clouds. IT/Telecom Paris Tech has investigated the impact of two innovative services on resource allocation carried out by Cloud Service Providers (CSPs). These services correspond to distributed data storage and to multicast streaming respectively. In the former case, instead of storing each block of data onto a unique site, we consider the possibility to distribute this storage onto different locations. The latter case aims to distribute a same video sequence in streaming-mode from one server to multiple remote nodes assuming a restricted number of network nodes are able to manage Point-to-Multipoint (P2MP) connections. Two variants are then possible. The first one consists in simply duplicating from the server to multiple destinations the same block of data by means of multiple point-to-point (P2P) connections. The second variant exploits one or several multicast nodes to send the same data stream towards multiple destinations. IT/Telecom Paris Tech has proposed a Mixed Integer Linear Programming (MILP) formulation of this extended version of the resource allocation problem. It is assumed that the underlying infrastructure is typically based on a mix of end-to-end connections with guaranteed sustainable bandwidth (such as Carrier-Grade Ethernet (CGE) circuits). Our numerical results show that an average gain of 3% is achieved in terms of accepted requests when the multicast feature is enabled. In addition, an average gain of 4% is achieved in terms of accepted requests when the distributed storage feature is enabled. When both features are enabled, an average gain of 7% is obtained. It is worth noting that the observed simulation runtimes exceed the considered scheduling period. For this purpose, a heuristic approach might be preferable.

3.5.4 Achieved Results in JA5

- A new architecture for programmable service composition (RACTI) was proposed and designed. The networking paradigm introduces service awareness in the core optical network, by creating self-organized islands of service transparency. A service island consists of a single IT resource and a group of networking nodes that constitute the shortest path towards that resource. During the 1st year, the proposed architecture was detailed; its basic notations and metrics were defined, as well as how the networking peers (nodes) interact to each other to form self-managed ad-hoc entities. Node interaction is based on resource discovery algorithms that "discover" information from both network and non-network resource providers and facilitates end-to-end service creation. In the consequent years 2 and 3, the architecture was enhanced to include service composition algorithms.

- An Infrastructure resource abstraction mechanism (UEssex-IT/TelecomParisTech) was proposed and implemented by UEssex. Resource abstraction allows for resource-state information to be discovered at different levels of detail. The abstraction mechanism provides an accurate representation of network and IT resources in an abstracted form with low signaling overheads. The aim of having an abstracted representation is to allow resource providers to hide internal details of their resource (network and IT) while exposing enough information for successful service compositions. This gives resource providers the opportunity to participate in service creation at whatever level of abstraction they desire. Performance evaluations were also performed to measure the effectiveness of proposed update algorithms.
  Improvements to the resource abstraction mechanisms were conducted as part as a joint mobility action with IT/TelecomParisTech.
- Service differentiation in grid marketplaces (IT/TelecomParisTech) through fairer management algorithm that defends the interests of both the end-users and the providers was investigated by IT/TelecomParisTech. Management algorithms were implemented that specified Grid Service Provider (GSP) business models taking into account the expectations of both end-users and service providers (network providers, storage providers and computing resources providers). The motivation was twofold. Firstly to serve a higher number of end-users while respecting each user's job utility. Secondly, to increase the benefit of service providers as much as possible.
- Framework to map user perceived QoS parameters to technology dependent network directives (SSSUP) was a sub-activity carried out within JA5. It defined a mapping framework between connectivity requests issued by applications in terms of end-host addresses and perceived QoS parameters and technology-dependent network directives. The objective was to enable an automatic and self-consistent node configuration across the network for on-demand connectivity set-up while avoiding error-prone correlation activities carried out by human operators. Specifically a Multi Protocol Label Switching (MPLS) network scenario was considered. The MPLS is a packet-based technology extensively used in transport networks to provide traffic engineering and efficient aggregation of multi-service traffic. MPLS is also becoming the preferred solution used by Internet Service Providers (ISPs) for the provisioning of secure, Quality of Service (QoS)-enabled, scalable and cost efficient Virtual Private Networks (VPNs) among corporate sites. MPLS networks are provided with built-in control plane functionality and comprises routing and signaling capabilities that require to be configured in each network element and need to be coordinated before the provisioning of MPLS-based network services. Configuration of MPLS network elements require to specify protocols instances and relevant parameters, and Label Switched Paths (LSPs) route and characteristics. SSSUP proposed a signaling architecture for network configuration based on distributed entities that cooperate by means of a dedicated signaling scheme to correlate network configuration parameters for a self-consistent and automatic configuration of MPLS-based VPN. In addition,

SSSUP aimed at further advancing the state of the art toward a fully distributed configuration approach aligned with the current recommendations in network management that pushes for a cooperative management architecture. The proposed solution follows the principles of the service signaling architecture proposed by the authors in [Martini], already applied in [Baroncelli] to enable on-demand provisioning of (G)MPLS-based network services. The same approach has also been proposed by the authors for supporting GMPLS capabilities within the ITU-T Next Generation Network (NGN) architecture [Baroncelli2]. The proposed architecture allows the processing of a VPN service request coming from management entities, e.g., a Service Broker of an Operation Support System (OSS), its validation according to the current network status, and finally its mapping in self-consistent vendor-specific instructions over the set of involved network elements. In particular, the processing and validation operations enable OSSs to trigger and automate the VPN provisioning eliminating any need of human intervention. The mapping operations exempt OSSs from knowing the network access topology (reachability) and relevant implementation details, and allow to express a VPN request in terms of client network addresses and QoS parameters (e.g., packet delay and jitter, bandwidth) as perceived by the clients.

- Service Oriented Resource Orchestration (UEssex) proposed a novel service-oriented resource orchestration model based on the optimisation of heterogeneous IT and network resources owned by different infrastructure providers. The resource orchestration model is a high level scheduling algorithm that uses abstract information to select suitable infrastructure providers and respective IT resource sites and network paths suitable to satisfy end-to-end services. UEssex presented ILP (Integer Linear Programming) formulations for the optimization problem. Algorithms for the problem was also proposed and analyzed taking into account the interests of both the end-users (accepted requests) and the providers (resource utilisation).
- Resource Provisioning for Enriched Services (IT/TelecomParisTech) proposed a novel resource provisioning model for provisioning innovative services on resource allocation carried out by Cloud Service Providers. IT/TelecomParisTech proposed a MILP (Mixed Integer Linear Programming) formulation for the optimization problem.

3.5.5 Conclusions

The work planned for JA5 was to design frameworks and algorithms that support effective and efficient implementation of service oriented optical network functionality. This included an architectural framework and algorithms to facilitate the creation of end-to-end services. Six different research areas, detailed in Section 3.5.4, were identified and studied within this framework. The work carried out in this JA5 led to a couple of collaborations, with two mobility action performed between UEssex and IT/TelecomParisTech. The mobility action saw two PhD students, one from each institution, spent a month each at the other’s institution. This mobility action resulted in a joint journal paper submission, in addition to the numerous publications that were published from work carried out in the JA. We believe that the work we present for this JA5 match the objectives planned for this joint activity.

A3.5.5.1 References

1. [Martini] B.Martini, F.Baroncelli, and P.Castoldi, “A novel service oriented framework for automatically switched transport network,” in *9th IFIP-IEEE International Symposium on Integrated Network Management (IM)*, Nice, France, 15-19 May 2005.

2. [Baroncelli] F.Baroncelli, B.Martini, V.Martini, and P.Castoldi, “A distributed signaling for the provisioning of on-demand vpn services in transport networks,” in *10th IFIP-IEEE International Symposium on Integrated Network Management (IM)*, Munich, Germany, May 2007.
3. [Baroncelli2] F.Baroncelli, B.Martini, V. Martini, and P.Castoldi, “Supporting control plane-enabled networks within itu-t next generation network (ngn) architecture,” in *IEEE/IFIP Network Operations and Managment Symposium NOMS 2008*, Salvador de Bahia, Brasil, 7 - 11 April 2008.

3.5.6 Published Joint Papers (joint and single partner)

A3.5.6.1 JA5

1. Kyriakos Vlachos and Apostolos Siokis, “A Service-Transparent and Self-Organized Optical Network Architecture”, ICC 2009
2. Chinwe Abosi, Reza Nejabati and Dimitra Simeonidou, “A Service Plane Architecture for Future Optical Internet” ONDM 2009.
3. Chinwe Abosi, Reza Nejabati and Dimitra Simeonidou, “A Service Plane Architecture for Future Optical Internet” JOCN 2009
4. R. Aoun, M. Gagnaire, “Service differentiation based on flexible time constraints in market-oriented grids”, IEEE Globecom conference, Honnolulu-USA, November 30- December 4, 2009.
5. R. Aoun, and M. Gagnaire, “An exact optimization tool for market-oriented grid middleware”, in Proceedings of the 14th international IEEE Communication Quality & Reliability (CQR) Workshop, Florida, USA, May 2009.
6. R. Aoun, and M. Gagnaire, “Towards a fairer benefit distribution in grid environments”, in Proceedings of the first international workshop on grid Computing (GRIDCOM), Rabat, Morocco, May 2009.
7. C. E. Abosi, R. Nejabati and D. Simeonidou, “Performance Evaluation of a Novel Service Provisioning Mechanism for Future Optical Internet Infrastructure” OFC 2010.
8. F. Baroncelli, B. Martini, V. Martini, P. Castoldi “A cooperative approach for the automatic configuration of MPLS-based VPNs”, International Journal of Grid Computing and Multi Agent Systems (GCMAS), January 2010.
9. C.E. Abosi, R. Nejabati, D. Simeonidou, "A Novel Service-Oriented Resource Allocation Model for Future Optical Internet" Invited, ICTON 2010.
10. R. Aoun, E.A. Doumith, M. Gagnaire, “Resource provisioning for Cloud environment: an exact approach”, accepted paper for the 2nd IEEE CloudCom conference, Indianapolis-USA, Nov. 30-Dec.3, 2010.
11. R. Aoun, C. E. Abosi, E. A. Doumith, R. Nejabati, M. Gagnaire, D. Simoneidou, “Towards an optimized abstracted topology design in Cloud environment ”, submitted to IEEE Transactions on Software Computing journal.

A3.5.6.2 Y3

1. C.E. Abosi, R. Nejabati, D. Simeonidou, "A Novel Service-Oriented Resource Allocation Model for Future Optical Internet" Invited, ICTON 2010.
2. R. Aoun, E.A. Doumith, M. Gagnaire, “Resource provisioning for Cloud environment: an exact approach”, accepted paper for the 2nd IEEE CloudCom conference, Indianapolis-USA, Nov. 30-Dec.3, 2010.
3. R. Aoun, C. E. Abosi, E. A. Doumith, R. Nejabati, M. Gagnaire, D. Simoneidou, “Towards an optimized abstracted topology design in Cloud environment ”, submitted to IEEE Transactions on Software Computing journal.

3.6 JA6 - UNI Extensions for Service Oriented Optical Networks

Participants: UESSEX, RACTI, AIT

Responsible person: Eduard Escalona (UEssex)

3.6.1 JA6 Objectives

The main goal of this activity has been to study, analyze and propose an interface between the service (Grid) layer and the optical transport network layer considering the requirements imposed by next generation services both from an application and a network perspective. This implies the development of interoperable procedures for requesting and establishing dynamic network services between clients and application servers connected by the transport network. The development of such procedures requires the definition of an enhanced UNI to support connectivity services and auto discovery procedures to aid signalling, which facilitate on demand and in advance network services using a network provisioning system such as GMPLS control plane.

Most of the work carried out in this joint activity has been performed within the framework of the OGF NSI Working Group (WG). This workgroup aims to facilitate interoperation between Grid users, applications and network infrastructures spanning different service domains, via the development of abstract messaging and protocols. The NSI WG must provide a general and open definition independent of implementation of provisioning systems or control planes. It should be flexible, modular and scalable to facilitate future enhancements.

3.6.2 Objectives of Y3

The planned activities for Y3 were to consolidate the work performed in the first two years and extend it with the definition of abstract messages to support an efficient communication for requesting network services. These abstract messages are independent of specific protocols but a study of a possible architecture can be considered.

3.6.3 Achieved Results in Y3

The NSI provides reliable and secure sessions for service related communication between a Requestor NSA and a Provider NSA as shown on Figure 24

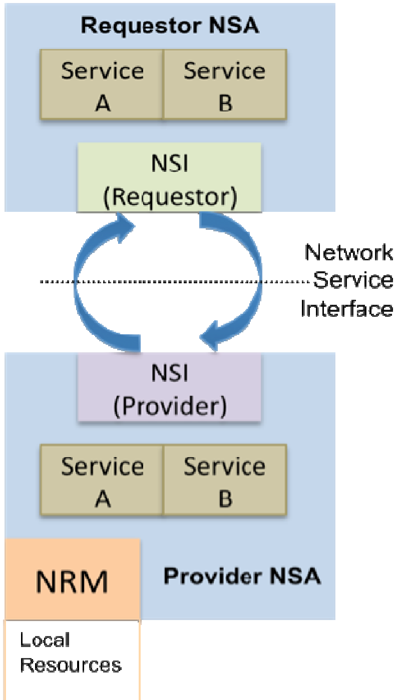


Figure 24: NSA and NSI

The Requestor NSA requests network resources and the Provider delivers these network resources to create a service. The NSA incorporates several functional elements, some of them defined as NSI network services and others as NSA internal functions. For example connection service and topology service are representing the former, whereas path-finding the latter. A network resource manager also exists in the Provider NSA to manage the part of the Network Service implemented locally. It’s important to say here that NSI protocol exchanges abstracted information of transport services, which reduces complexities of delivering a particular transport service.

The connection service supports the creation, management and removal of connections. It is message based command-response protocol that operates between a requester NSA and a provider NSA. The protocol includes a set of defined primitives such as Reserve, Provision, Cancel, Query and Notify. These command primitives allow developing the connection service lifecycle as in Figure 25.

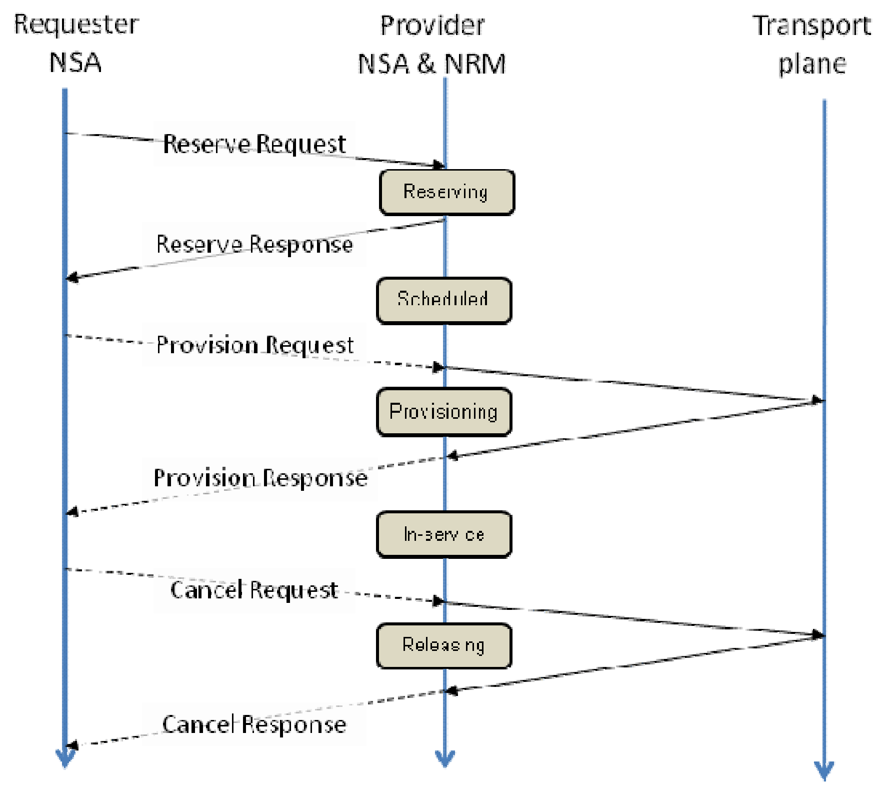


Figure 25: Connection lifecycle

NSI v1.0 will just support single channel point-to-point connections, which can be unidirectional or bidirectional. Connections are basically composed by an ingress (access) section, a transport section and an egress (access) section (Figure 26). The transport section can have any type (one or many) of framing protocol as long as the end-to-end preservation (from ingress to egress points of transport section) of user payload data stream is honoured. Thus, the Connection requester can specify the transport connection end points, access framing, and other parameters, and the Network Service Provider can then select the transport protocol and associated transport path to allow for a discovery and reservation of a valid, available and optimal path. In this case, it is clear that “connection” service has a different perception for the requestor and the provider due to the abstraction. However, the omniscient requestor is able to participate on connection planning decisions through the NSI architecture. They can also be concatenated to form longer connections.

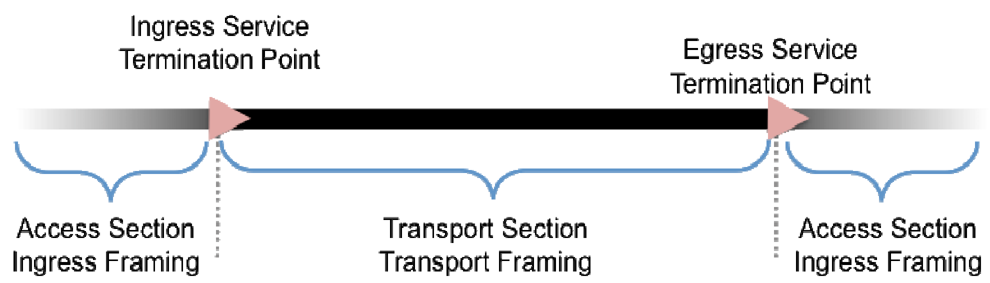


Figure 26: NSI connection

The connection lifecycle starts with a connection request. The first phase is the Reserving phase, which includes a path selection and resource reservation. Resource reservation can also be immediate or in advance. Once the provider NSA has completed processing the request it sends a success confirmation or an error response accordingly. The second phase performs the actual provisioning of the connection. Provisioning is the process where the connection is physically instantiated by configuring each device along the path to reflect the path plan developed and reserved in the Reservation phase. Once provisioning is complete, the connection then moves into an “In-Service” state and the user are notified that the connection is ready for use. Finally, the Releasing phase ends the connection provisioning and make sthe resources available for future connections. The Releasing phase is triggered by a release event which can be instantiated by the requester NSA or because the scheduled time has finished.

NSI attributes

Each NSI message includes a set of attributes exchanged between NSAs to manage a connection. The NSI message attributes are divided into 3 groups:

- Message attributes: These include attributes such as NSI version, NSA addressing, security, type of service, transaction identifier.
- Service attributes: These include attributes such as connection service version, connection identifier, service primitive, service transaction identifier.
- Primitive attributes: These include specific primitive attributes such as start time, end time, service parameters or path objects.

NSI implementation

The NSI Agent is the interface driver for the NSI requestor and the NSI provider. A common and effective approach to its implementation is to use SOAP Web Services. The main reasons for this choice are:

- SOAP is widely known by the programmers community
- There are Open Source suites like AXIS2 that have support for Java and C++.
- The user can ask to the NSI server the information needed to write clients, (asking the wsdl files to the server directly).

An NSI Agent needs to have a specific function to bind the queries submitted by the NSI requestor. Appropriate conversions among the NSI and the network provider parameters need to be implemented to allow a coherent operation.

NSI proxy

In some cases the implementation of a proxy could be useful to translate NSI messages when one of the actors does not support the NSI protocols (e.g., using a GMPLS control plane). This has been a use case study performed at University of Essex. An overlay control plane model allows the user to invocate a connection request using an interface (UNI) directly connected to an edge node. The architecture scenario proposed deploys an NSI proxy which implements a Connection Web Service towards the client and the standard UNI interface towards the GMPLS core network. In a typical optical connection request, a user specifies source, destination and bandwidth. With the implemented Connection Web Service the user also indicates the service class that will use the connection. It can be either: Real Time, Streaming or Transactional. This will be used at the NSI proxy to determine the expected QoS by mapping the acceptable Packet Loss Rate (PLR) per application into QoT (Q Factor) values according to Table 4.

	PLR	BER	Q
Real Time	< 5 E-5	< 1.1 E-6	~13.4
Streaming	< 1 E-3	< 2.8 E-5	~12.0
Transactional	< 1 E-2	< 2.3 E-4	~10.9

Table 4: QoS for Service class

The PLR classification and the Q Factor relation with respect to the Bit Error Rate (BER) have been taken from the literature [1][2], while the correlation between packet loss rate and BER has been experimentally obtained by monitoring continuous data traffic with variable packet size. It is of great importance to fulfil these requirements to guarantee the QoS to the user application. For instance, video quality is directly affected by packet loss, above all in real time applications such as HD videoconferencing.

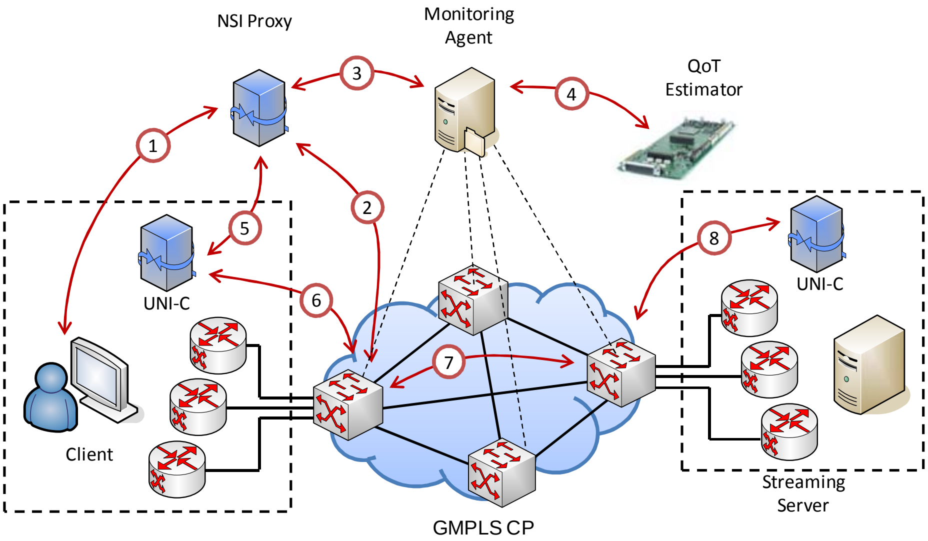


Figure 27: NSI Proxy implementation

As an example for a proof of concept we have set up the scenario shown in Figure 27 with a GMPLS controlled core network, two clients and a Quality of Transmission (QoT) estimator. Upon a connection request from the user (Step 1) the NSI proxy interacts with the Path Computation Engine (PCE) to obtain a route (Step 2). At the same time it gathers from GMPLS the list of the already established LSPs that share any link with the calculated route and sends the information, including the desired transmission rate, to the Monitoring Agent (MA) (Step 3). The MA adds physical layer parameters obtained through actual monitoring and preconfigured static values that characterize each link, to form a complete picture of the network. This information is transmitted to the hardware-accelerated QoT estimator, which calculates the impact of the new LSP over the existing LSPs in the form of Q factor values (Step 4). These values are sent back to the NSI proxy, which accepts or rejects the route according to the previously specified service class and requested QoS. If the route is found appropriate, standard GMPLS mechanisms are triggered to establish the connection (Steps 5-8). Otherwise, a disjoint route is calculated by the PCE and the QoT is estimated again. If a disjoint acceptable route is not possible, the NSI proxy retries Step (3) with the next possible transmission rate specified by the client. If no possible combination rate/route is found the connection request is blocked.

This experiment shows how using an extended UNI (in this case NSI proxy) can facilitate the incorporation of advanced functionalities that allow a better cooperation of the service and network layers.

3.6.4 Achieved Results in JA6

This joint activity has developed extensions to current standard definition of the User to Network Interface (UNI). These extensions have been defined in terms of abstract messages described in a modular way, allowing different protocol implementations.

The main achieved results include a study of the requirements for such extensions, description of the identified functionalities and abstract messages and a basic implementation of a proxy that uses the defined interface extensions in a real test-bed experiment.

3.6.5 Conclusions

Overall, the results achieved in this joint activity demonstrate the necessity of having an extended interface between the service and network layers. This interface will be extensive enough to allow the inclusion of future advanced functionalities and simple enough to be adopted by both Grid and networking communities. Also, a common definition of such interface will ensure interoperability between different applications and network providers regardless of proprietary implementation of provisioning systems. This interface is being defined and standardized under the OGF NSI working group.

A3.6.5.1 References

[1] A. Stavdas, Core and Metro Networks, Wiley (2010)

[2] P. Pavon, Offline Impaiment Aware RWA Algorithms for Cross-Layer Planning of Optical Networks, JLT vol.27, no.12, June (2009)

[3] OGF NSI WG, http://www.gridforum.org/gf/group_info/view.php?group=nsi-wg

3.6.6 Published Joint Papers (joint and single partner)

Part of the work carried out in this joint activity has been done in the context of the GF NSI Working Group and included in a use-case draft document and an architecture definition draft document. Both documents are still in draft version and available at the workgroup website [3].

3.7 JA7 - Photonic Grid Dimensioning and Resilience

Participants: IBBT, AIT, RACTI, AGH, KTH

Responsible person: Chris Develder (IBBT)

3.7.1 JA7 Objectives

This JA focuses on photonic Grids, which aim at providing cost and resource efficient delivery of network services with possibly high data rate, processing and storage demands, for a geographically widely distributed user base. In such a context the unicast routing is required that enable users to transmit data for processing and delivery, without assigning and explicit destination. Specifically the objectives of this JA are:

- Design and evaluation of network dimensioning algorithms capable of optimizing network topology, resource capacity and resource placement. Processing and storage resources dimensioning strategies will be also examined. In judging the relative optimality of the different approaches, simulations will be used. Various optical switching paradigms can be considered.
- To study and propose algorithms and strategies for providing reliable Grid operation. Based on a general analysis to provide resiliency in optical Grid environments, we will define protection and restoration strategies, and present specific case studies.

3.7.2 Objectives for Y3

The objectives of the research work planned for Y3 are the follows.

- **IBBT**: Dimensioning of resilient optical grids based on large-scale studies, using an extension of ILP methodology as well as an heuristic approach. Then resiliency strategies will be expanded to include protection against resource failures in addition to network (link) failures.
- AIT & KTH: Study of differentiated path provisioning in resilient WDM optical networks supporting a variety of services.
- **RACTI**: Investigation on various parameters that affect the quality of transmission of the selected primary and backup light paths.
- **AGH:** Test of resilience differentiation methods and study of control plane issues for supporting enhanced resilience in OBS-based grids will be carried out using OBS simulator. Then the implementation of GridSim tool is planned..

3.7.3 Achieved Results in Y3

The results obtained in Y3 are detailed below:

- Scalable algorithms for dimensioning resilient optical grids in case of network link failures (IBBT)

In [Buysse2009DRCN], we have considered the impact of using the relocation mechanism for the case of shared path protection, i.e. we allow the wavelengths used for backup paths to be shared among multiple backup paths as long as the corresponding primary paths are link disjoint. However, we had to confide in small-scale studies, since the proposed classical Integer Linear Programming (ILP) is not scalable to larger scale scenarios.

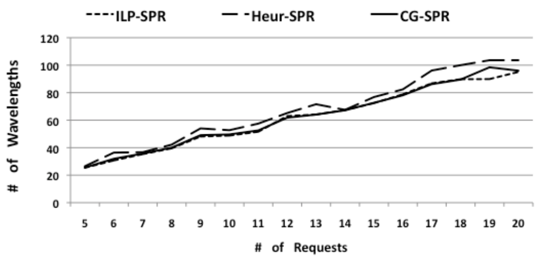
To address this scalability issue, we proposed to use a Column Generation (CG) approach for modeling and solving the problem. The philosophy of CG is to limit the number of variables explicitly included in the ILP problem, thus reducing the problem size. We refer to [Jaumard2010] for details on the CG model and CG solution techniques. As an alternative to linear programming solutions, we also devised a heuristic to solve the dimensioning problem. We developed one in [Buysse2010ONDM] for the case of an anycast routing problem, where only traffic sources are given, but destinations are to be chosen. In [Jaumard2010], the destinations are assumed to be given a priori.

We evaluated the above approaches in a case study that consider a European topology with 28 nodes and 40 bidirectional links. We have chosen 5 nodes as grid server resource sites. We assessed both the optimality and running time of the considered algorithms. Each algorithm was used to calculate the capacity requirements (i.e. wavelengths summed over all links) to fulfill the traffic demand while providing resiliency against single link failures. Two resiliency strategies were compared, each time using either one of the ILP/CG/heuristic algorithms: (i) CSP is the classical shared protection case, providing backup paths to the original destination, while (ii) SPR is the case where we allow relocation in case of failures: instead of providing a backup path to the original location, grid jobs can be relocated to a different secondary server site, since grid users typically do not care what server performs their processing.

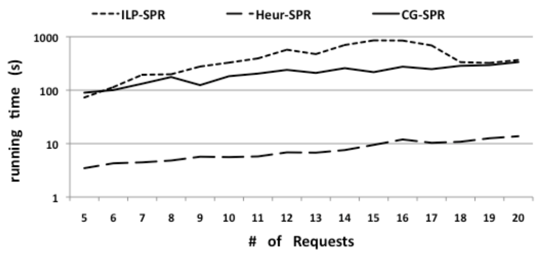
With respect to the solution methods, we found that the heuristic performs quite well, but with comparable computing times, and solutions of slightly inferior quality than CG, when the number of requests remains small. Comparing the results generated by the CG method and the heuristic on small demand instances (Figure 28a), we conclude that the gap between their optimized solutions remains fairly constant, where CG matches the optimal output as found by an exact ILP solution very well and the heuristic has suboptimal solutions, which are of satisfactory quality. The trend is fairly similar for larger demands, as illustrated Figure 28c, where we plotted the total number of wavelengths for the demand sets with 50 to 300 requested connections. We ascertain that in this particular study [Jaumard2010], the difference between the total number of wavelengths for the heuristic and CG averages to 4.99% for the CSP case and 6.92% for SPR.

Scalability is a known issue in planning problems solved by ILPs. This is our main motivation for using other techniques such as heuristics and decomposition column generation techniques. Results on larger traffic demand instances (Figure 28c-d) show that the heuristic generates very good solutions, at least up to 200 requests. However, it is with a computational cost that is comparable with the solution of the CG model. It therefore shows that, on the one hand, CG model offers a fairly scalable tool even for large ILP models, and on the other hand that there is a trade off in heuristics between computational times and qualities of the solutions: Accurate solutions may be costly to reach with a heuristic when the problem to be solved is quite combinatorial in nature.

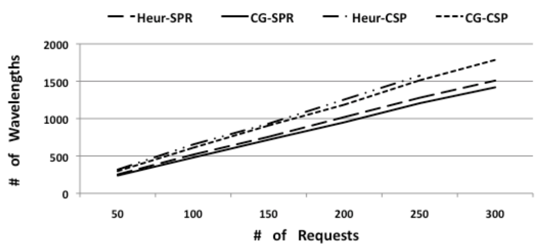
With respect to the impact of exploiting relocation, our work in [Jaumard2010] confirmed the conclusions from our earlier works [Buysse2009DRCN] and [Buysse2010ONDM]: relocation impacts the network dimension by introducing a network load reduction (NLR). Here, it amounts to 22%, independently of the requested number of connections.



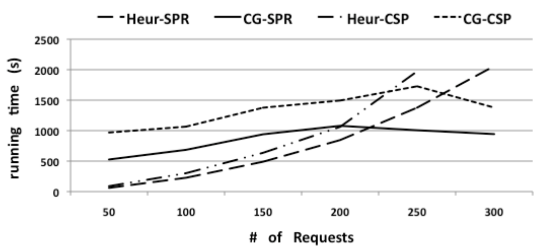
(a) Number of wavelengths for small instances



(b) Running time for small instances



(c) Number of wavelengths for large instances



(d) Running time for large instances

Figure 28 – Sample results

- Dimensioning of resilient optical grids against both network and server site failures (IBBT)

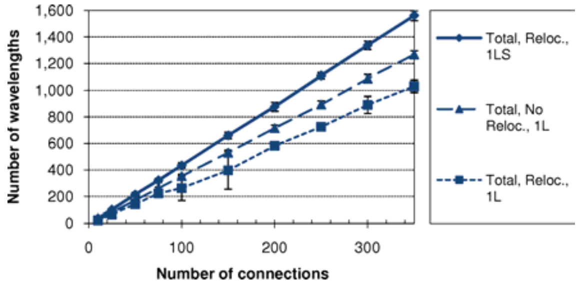
In [DevelderICC], we expanded the approach discussed above and we (i) consider not just network but also server site failures to protect against, (ii) simultaneously optimize the amount of server resources (instead of just network), and (iii) do not fix the destination server site (under failure free conditions) a priori. We expanded on the relocation idea to reduce the amount of overall backup server capacity. We assess this reduction quantitatively by dimensioning the network to survive both single link and server failures. The dimensioning model is generic and can address any failure scenario that can be expressed as a set of jointly failing link resources, i.e., a so-called shared risk link group (SRLG). Note that node failure can be represented as joint failure of its incident links (hence leading to an SRLG comprising all of them). For the column generation model, using this SRLG concept, we refer to [DevelderICC].

We evaluated this methodology on a European network topology comprising 28 nodes and 41 links. To assess the cost of resilience against both link and server failures, as well as the benefit of relocation, we compared the following three scenarios: (i) 1L, No Reloc.: Single link failures only, no relocation; (ii) 1L, Reloc.: Single link failures only, with relocation; (iii) ILS, Reloc.: Single link or server site failures, with relocation. To model the failures, the set S of SRLGs was constructed as follows: (i) for single link failures (1L), S has elements s each containing a pair formed by a link and its reverse; (ii) for single link or server failures (1LS), we additionally include in S the singletons formed by each single link towards a server site.

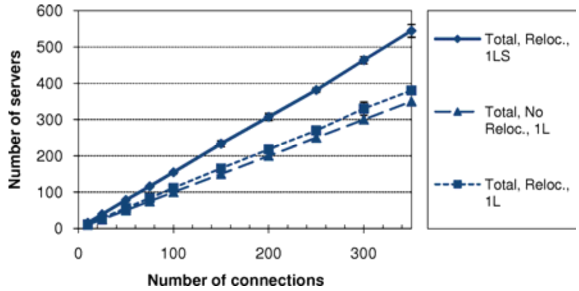
In our results below, we consider the case of K = 3 server sites, however we did not fix the locations a priori (as opposed to [Jaumard2010]). Instead, we use the server location ILP of [Develder2009] to choose them, based on the demand. For the latter, we varied the total number of unit demands between 10 and 350. For each D, we created 10 random instances. The measures plotted below are averages over those 10 random instances per demand case.

Looking at Figure 29a, we confirm our earlier findings [Buysse2009DRCN, Jaumard2010] that exploiting relocation enables substantial savings in the number of network resources (wavelengths); in this case in the order of 19% (comparing 1L, Reloc. vs. 1LS, No Reloc., averaged over the data points plotted in the figure). Using our model incorporating server capacity dimensioning, we are also able to quantitatively assess that relocation to protect against single link faults incurs an increase in server capacity around 11% (averaged over the 10-350 demands), as can be derived from Figure 29b. Hence, relocation indeed has a considerable net benefit in resource requirements. (Note that our model jointly minimizes total server and network cost.)

Considering now the protection against both network and server site failures (1LS, Reloc.), we note that in terms of server capacity we do better than a 1+1 protection strategy at the server sites, since we only need around 55% extra server capacity (compared to 1L, No Reloc., and averaged over the considered 10-350 demands). The additional network capacity amounts to 26% (avg. over the 10-350 demands). Note that the difference in network capacity mainly stems from difference in backup wavelengths (cf. we noted that the maximal deviation in network resources in the relocation cases only amount to around 10%).



(a) Wavelengths



(b) Servers

Figure 29 – Results for EU network with 3 server sites. Error bars (which are largely hidden behind data point markers) indicate standard deviation among the 10 random instances per data point.

- Differentiated Path Provisioning in resilient WDM Optical Networks supporting a variety of services (KTH+AIT)

This activity is a follow up of the activity started in Y1. In Y1 an Impairment Constrained Based Routing algorithm supporting differentiation of services at the BER level, referred to as ICBR-Diff [1] was presented. This year's activity extends this work by including the effect of optical impairments on the quality of transmission for the protection path. As the protection paths are typically much longer than the primary paths, they are in general more susceptible to physical layer impairments. This in turn: (i) degrades the quality of the signal traversing the protection paths and (ii) has a direct impact on the working path, since the working-protection path-pair might not able to guarantee the required signal quality. The objective of the activity is to provide comprehensive results, in terms of blocking probability, of the ICBR-Diff algorithm applied to protected connection requests, based on both shared path protection (SPP) and dedicated path protection (DPP). Detailed description of algorithm steps is provided in [Wosinska2009].

The ICBR-Diff algorithm with protection has been evaluated through simulation, using the Pan-European network topology defined by COST 239 [1]. In order to differentiate between various QoS requirements, two distinct classes of grid service requests (Class-1 and Class-2) were chosen based on available IP traffic measurements [2, 3]. Class-1 and Class-2 grid service requests are assumed to require BER less than 10^{-15} and 10^{-9} respectively. For benchmarking purposes, we also evaluated two additional provisioning algorithms, namely shortest path (SP) and Impairments Aware Best Path (IABP). Essentially these two approaches do not employ service differentiation, but instead handle both Class-1 and Class-2 with a single transmission quality threshold, i.e., they block all the connection requests (regardless of service class) with BER greater than 10^{-15}. In our experiments we assumed that the traffic comprises of 30% of Class-1 grid service requests and 70% of Class-2 grid service requests.

Figure 30 and Figure 31 compare the blocking probability performance of SP, IABP and ICBR-Diff when all grid service requests are assigned dedicated path protection (DPP) and shared path protection (SPP), respectively. Great improvement of the network blocking probability can be achieved when impairment aware routing is applied to both the primary and the protection path, when compared to SP routing. The figures also show that ICBR-Diff is able to outperform IABP due to its inherent better utilization of high quality links. It can be noticed that under high load conditions, blocking due to insufficient resources (case C) is dominating in the DPP scheme. On the other hand, when using the SPP scheme blocking probability due to violation of signal quality requirement plays the most important role when SP and IABP algorithms are used (compare case B and case C in Figure 30 and Figure 31 respectively). This is because in the SPP case protection paths tends to share resources as much as possible and, as a result, they tend to be longer than the protection paths in the DPP case. This has a negative effect in their signal quality as longer paths are typically more impaired. Both figures show that, on the other hand, the total blocking probability is higher in the case of DPP, a result that was expected.

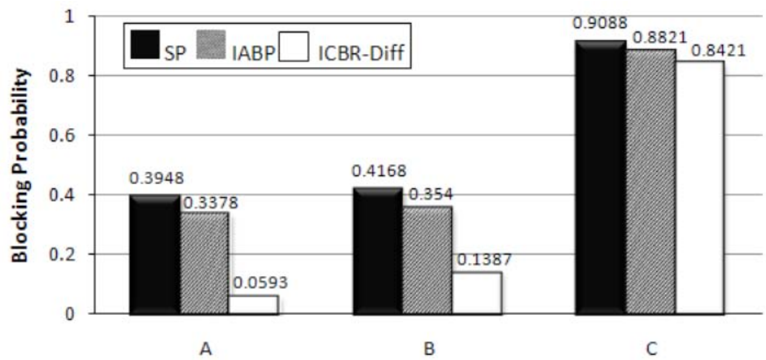


Figure 30 – Dedicated path protection (DPP). Case A: blocking probability (due to insufficient signal quality) in low load conditions (60 Erlangs). Case B: blocking probability (due to insufficient signal quality) in high load conditions (300 Erlangs). Case C: total blocking probability (due to insufficient signal quality and insufficient network resources) in high load conditions (300 Erlangs).

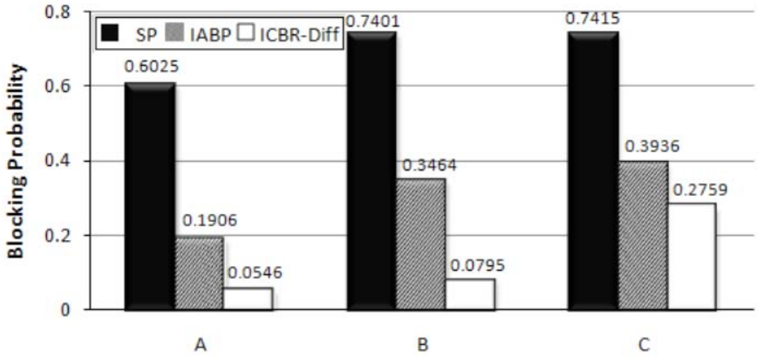


Figure 31 – Shared path protection (SPP). Case A: blocking probability (due to insufficient signal quality) in low load conditions (60 Erlangs). Case B: blocking probability (due to insufficient signal quality) in high load conditions (300 Erlangs). Case C: total blocking probability (due to insufficient signal quality and insufficient network resources) in high load conditions (300 Erlangs).

- Performance Evaluation of Impairment-Aware Routing Under Single- and Double-Link Failures (AIT)

In the framework of this work [Georgakilas2010a], the performance of a WDM optical networks in the presence of single failures was studied, while the path computation takes into account the physical layer impairments. Resilience was provided through a Shared Backup Path Protection (SBPP) [4] mechanism compared to an extended version where sharing was reinforced. When considering dual-link failures, the metric of connection loss rate due to double failures was defined and the dual-failure restorability was evaluated through a path restoration mechanism.

During simulations, we modeled the network as a set of nodes and directional links interconnecting them. Wavelength continuity is a strict constraint across the network and incoming requests adhere to a Poisson arrival process with exponentially distributed time duration. Full single link failure recovery was provided, since each incoming demand is accepted only after a primary and a link-disjoint backup path are found, otherwise it is blocked. The Routing and Wavelength Assignment problems are solved in two separate steps. Routing for both primary and backup paths is based on Dijkstra's shortest path algorithm, assigning the appropriate link weights. For primary path computation, physical impairments are considered, thereby implementing the Impairment-Aware Routing (IAR) scheme. During backup path computation, the IAR scheme is compared against a classical minimum hop (MH) routing approach (unity link cost). In both case, BER computed for each path is then compared to a predefined threshold (10-15) to decide whether the path quality is accepted. Wavelength assignment is solved with the First Fit algorithm both for primary and backup path. [Markidis2008b].

Network resilience is provided through the SBPP scheme, where backup wavelengths can be shared between protection paths whose primary paths are link-disjoint. This scheme provides better resource utilization, whereas 100% survivability against single link failures is still provided. Moreover, the impact of a reinforced sharing scheme is studied where during the backup path routing process, preference is given to links that are already used in order to reinforce sharing of links when this is possible.

Four different scenarios of RWA and resilience combinations are defined and evaluated, namely IAR_IAR_SBPP, IAR_IAR_SBPPreinf, IAR_MH_SBPP, IAR_MH_SBPPreinf (denoting [primary path routing]_[backup path routing]_[resilience scheme]). The metrics used to evaluate the network performance under these four scenarios are: (i) total blocking probability (single-link failures), which is the result of blocking due to unavailable bandwidth on either primary or backup paths and blocking due to impaired primary or backup paths, (ii) resource utilization, connection loss rate due to dual-link failures (the number of affected established connections (by a dual link failure) over the number of active connections) and (iii) dual-failure restorability (the percentage of affected connections (after a dual link failure) that are successfully restored with the remaining spare capacity).

The case study was based on two network topologies, the 11-node and 26-link COST 239 and the 16-node and 24-link NSFnet networks, assuming that each directional link employs one fiber with a capacity of 32 wavelengths. Some indicative results are presented in the following figures:

Figure 32 illustrates the total blocking probability for the COST 239 network in the presence of only single-link failures. We observe that that IAR scheme is preferable for both primary and backup path computation, since the total blocking probability is much lower (50% reduction) compared to minimum hop routing in the backup path. In addition, the main contributor to the total blocking probability was the highly impaired backup paths, an expected result due to the commonly longer length of protection paths compared to working paths. Figure 33 depicts the network resource utilization of the different schemes. A counter-intuitive observation is related to the better resource utilization of the IAR scheme compared to min-hop, although the utilization graph shows the opposite. This explanation is based on the previous results of blocking probability, where the IAR scheme performs much better, thus utilizing more of the network resources. Finally, Figure 34 shows the average connection loss rate (right y-axis) and dual-failure restorability (left y-axis) for the NSFNET network. Both IAR and min-hop in the backup path demonstrate similar performance in terms of both metrics and actually worse in general than the COST network (not shown here) because of the lower node degree of the topology. The impact of dual link failures is very low (under 3,5%), whereas 20-60% dual failure restorability figures are achieved. Although these values seem low and the dual-link failures do not occur as frequently as single link failures, the results indicate that spare capacity placement or efficient restoration schemes have to be seriously considered for networks requiring high availability figures of the order of 99,999 %.

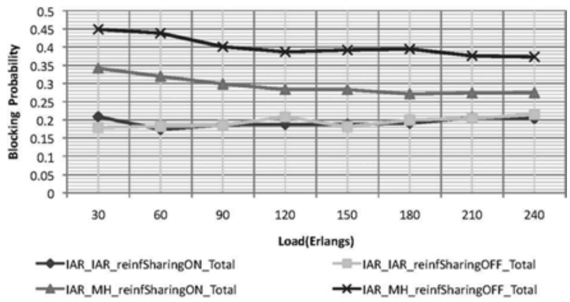


Figure 32 – Blocking Probability (COST 239)

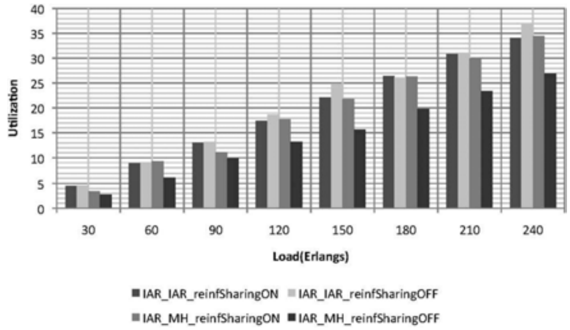


Figure 33 – Utilisation (COST 239)

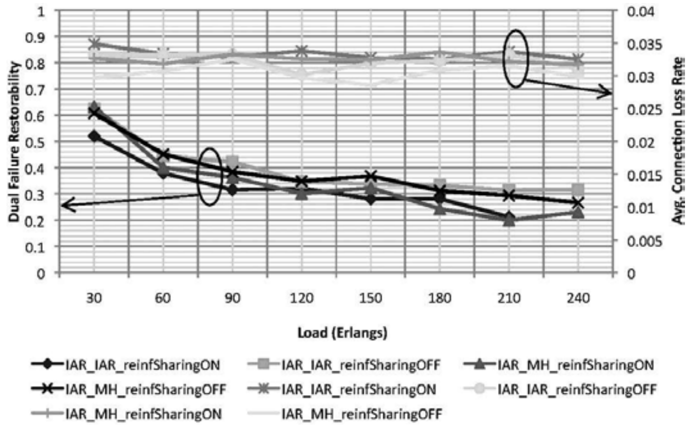


Figure 34 – Dual Failure Restorability and Avg Connection Loss Rate (NSFNET)

- Path Protection in WDM Networks with Quality of Transmission Limitations (RACTI)

In case of Quality of Transmission (QoT) limitations, the issue of providing degradation-resistant WDM networks is important as fault tolerance, since under some scenarios the probability of lightpath reaching an unacceptable level of QoT may be higher than the probability of a lightpath failure, due to a broken link. Both events make the lightpath useless and the establishment of a new lightpath the only solution. In case of protection if the backup lightpath does not have acceptable QoT when the primary lightpath fails, then the backup lightpath is useless and a new lightpath needs to be calculated (restoration). In this way several resource are wasted for a long period of time. Therefore mechanisms that are efficient for impairment constrained optical networks are required that are able to both serve as many connection requests as possible, but also should be able to ensure the long survivability, in terms of QoT, of both the primary and the backup lightpaths.

To address such issue, a multicost approach is proposed [Christo2009] for the selection of the primary and backup lightpaths taking into account QoT performances. According to this approach, the primary and the backup lightpaths should then be selected so as: i) to have valid QoT, ii) to do not damage the QoT of the established lightpaths, iii) to result in low future connection blocking, iv) to result in increased survivability of the backup lightpath. In case of 1:1, the survivability of backup path is the probability that it will have valid QoT for the duration of the primary's lightpath lifetime. In case of 1+1 the survivability of the backup lightpath (but also of the primary) is preserved by checking whether the QoT of the already established lightpaths (primary and backup) is affected by the establishment of a new primary lightpath and of the corresponding backup.

The multicost approach provides a set of candidate lightpaths (instead of single one) for each source destination pairs, together with their QoT parameters. Candidate lightpaths are calculates using Impairment-Aware (IA)-RWA algorithms that can be viewed as a generalization of Dijkstra's algorithm considering a vector, instead of scalar, link costs. In this way more than on paths between two nodes are calculated, since the optimization parameters are more than one. Moreover, in order to increase the duration of the validity of the backup lightpaths, in terms of its QoT, we propose the differentiation of the way the primary and the backup lightpaths are chosen. In particular, for the selection of the primary lightpath we use actual current network utilization, while for the backup lightpath we use worst-case interference assumption. Specifically the following approaches are proposed: (i)

- bestQ-bestQ: First, the primary lightpath is selected using the maximum Q-factor value as criterion. Next, the disjoint backup lightpath is selected from the ones available, using again the maximum Q-factor criterion.
- sumQ: In contrast to the previous policy, it is possible to select the primary and the backup jointly. One such possible scheme is to select the lightpaths (primary and backup) with the largest sum of Q-factor values.

The criteria used for selecting the primary and the backup lightpaths, in these previous approaches, do not directly consider the survivability of the lightpaths. The approach that we propose and evaluate in this work, in the case of 1:1 protection, is to consider the actual current interference among lightpaths so as to select the primary lightpath (with the best Q-factor), while using the worst case interference assumption for the backup lightpath (selecting the lightpath with the best Q-factor). In the latter assumption the wavelengths on all links are considered as fully utilized and the transmission quality of the candidate lightpaths is calculated under this assumption. A lightpath that is chosen in this way is bound to have acceptable transmission quality during its entire duration, even if future connections that interfere with it are established. However, such an approach reduces the candidate path space, leading to the selection of backup lightpaths that are not the best possible (in terms of hops, QoT etc), considering the actual interference in the network.

In our simulation experiments, we consider an all-optical transparent network, where connections arrive dynamically and have to be served as they come. The experiments were performed assuming the DT network topology, which is a transparent candidate network, as identified by the DICONET project [5]. Connection requests (each requiring bandwidth equal to 10Gbps) are generated according to a Poisson process with rate λ (requests/time unit). The source and destination of a connection are uniformly chosen among the nodes of the network. The duration of a connection is given by an exponential random variable with average $1/\mu$ (time units). Thus, λ/μ gives the total network load in Erlangs.

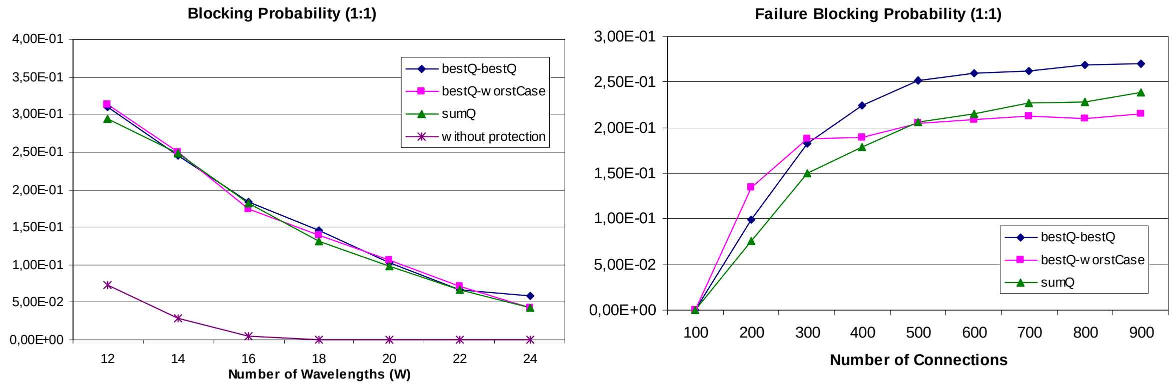


Figure 35 – (a) The blocking probability versus the number of available wavelengths and (b) the failure blocking probability as a function of the number of connection requests, for 1:1, protection, using the bestQ-bestQ, sumQ, bestQ-worstCase and no protection (bestQ) policies, for fixed network load (equal to 100 Erlangs).

Figure 35a shows the blocking probability of several multicost based IA-RWA algorithms as a function of the number of available wavelengths, performing 1:1 protection. We observe that in case of 1:1 protection the differences between the different protection policies are small. It seems that since the backup lightpath is not activated, its QoT related selection has small impact on the network performance. Moreover, as expected, the blocking probability is increased when using 1:1 protection, since more lightpaths are reserved in total, reducing in this way the possibility of future connection requests finding a valid lightpath to use. Also, we observe that the blocking probability decreases as the number of available wavelengths increases. Figure 35b shows the failure blocking probability as a function of the number of connection requests, for 1:1 protection. We use the failure blocking probability as an indicator of the survivability of the backup lightpaths. In our experiments we measure this parameter as follows. We assume that the connections' durations are infinite. After all the connections have been served we assume that a particular link has a failure and measure the number of backup lightpaths used (in the place of affected/failed primary lightpaths) that do not have valid QoT over all the established lightpaths. Next, we re-establish this link and perform the same operation/measurement for all the other links, sequentially. In the end we average the measured values. In our experiments, we observe that when the network is lightly loaded the sumQ metric produces the best results, increasing the survivability of the backup lightpaths (decreasing the failure blocking probability), while the bestQ-worstCase policy the worst. On the other hand as the network load increases the performance of the sumQ deteriorates, while the bestQ-bestQ policy results in a large number of backup lightpaths with invalid QoT. On the other hand the bestQ-worstCase policy seems to behave better under this heavy load scenario. This is due to the fact that the worst case interference assumption leads to the selection of backup lightpaths, which are not affected by future established lightpaths. This characteristic becomes more evident when the number of affected by a failure primary lightpaths is large.

- Service class differentiation for OBS-based grid networks (AGH)

We have used EON network topology in order to check how OBS network handles with some 2 traffic classes distinguished according to QoS requirements. In our studies we focused on most popular method of QoS provisioning inside OBS network which employs offset-time differentiation. Using ns2 simulator [6] expanded with OBS-ns [7] and EVALVID [8] packages we have performed simulations to check influence of offset-time differentiation for quality of video transmission. Each of performed simulation assumed to sent video file coded in H.264 format between two dedicated nodes using different offset-time. Inside network we defined two classes of traffic:

- class 1 (high priority) – traffic between video source and destination, considered as foreground traffic
- class 0 (low priority) – data traffic, emulating grid service

At the beginning we assumed, that offset-time for both high priority traffic and normal data is the same. Next simulations were performed with increased value of offset-time parameter, but only for high priority traffic. Using reported tools we were able to measure QoS parameters of high priority traffic transmission as well as to compare quality of received video with reference one. For video sent in every single scenario, packet and frame loss probability were measured, also SSIM (Structural Similarity) parameter was calculated. For calculation two models were used:

– frame loss mode – where damaged frames were removed from received video
– packet loss mode – where damaged frames were included in received video.

Obtained results are presented in Figure 36 and Figure 37.

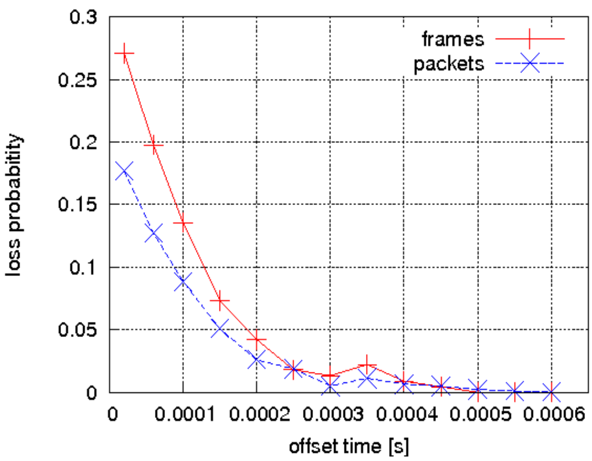


Figure 36 – Frame and packet loss probability for different values of offset-time parameter.

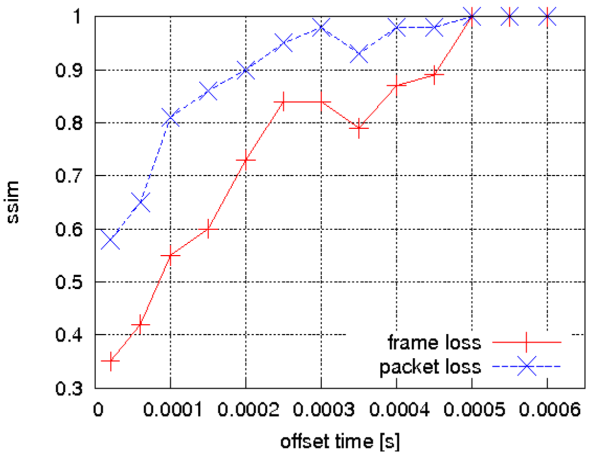


Figure 37 – SSIM parameter for different values of offset-time parameter.

As Figure 36 shows, increasing value of offset-time parameter for high priority traffic makes it more resistant to data loss. For offset-time 25 times longer than for normal data, loss probability is negligible. Proper offset-time assignment gave us solid isolation of traffic classes.

- Usage of FDL for improvement burst loss probability in OBS-based grid networks with failures (AGH)

We also considered OBS-based grid network (EON topology) with a single link failure. To perform simulations we used simulation tool written in Universitat Politecnica de Catalunya in Barcelona. This simulator allows to study version of OBS network called Offset-time Emulated OBS (E-OBS). Two restoration algorithms were under investigation: global routing update and local deflection. For each mechanism we used {2,4,8,16} Fiber Delay Lines (FDL).

First we examined global routing update scheme. As we expected, link failure deteriorates network performance – measured as Burst Loss Rate (BLR), but the employment of FDL can significantly compensate this undesirable effect (Figure 38).

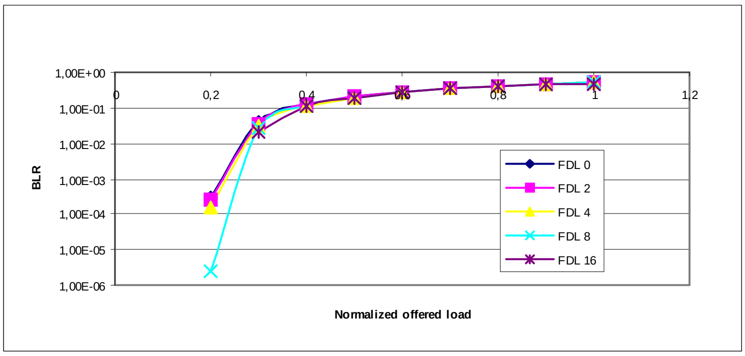


Figure 38 – Results for EON topology - Global routing update

We can observe, that BLR decreases with increasing number of delay lines. The best results were achieved for low values of offered load e.g. for the normalized load equal to 0,2 and with 16 FDL used we were able to eliminate burst losses. For the values of offered load greater than 0,3, impact of deployment FDL is negligible.

The second examined scheme was local deflection mechanism. Results from this section of our studies shows difference between both algorithms.

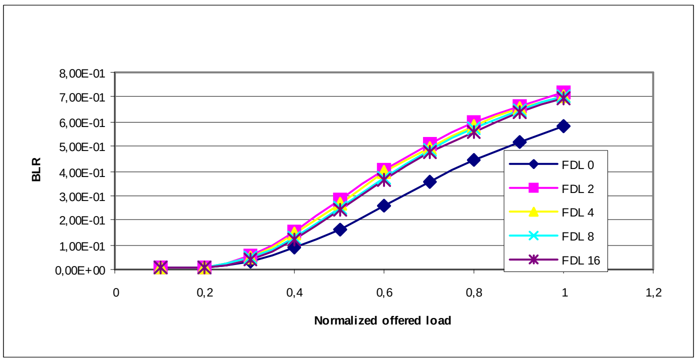


Figure 39 – Results for EON topology - Local deflection

In this case burst loss rate is larger when fiber delay lines are used (Figure 39). This phenomenon can be explained. Node is directing bursts to an alternate link, which is already occupied by bursts, which traverse through paths that are not affected by a faulty link. It increases the probability of congestion on a mentioned link and can lead to more burst losses. To overcome this issue we propose to define the limit on maximum number of bursts that can be redirected to a deflection path.

3.7.4 Achieved Results in JA7

We successfully addressed the grid dimensioning and resiliency problems set out at the start of the JA. Specifically, our results include:

- a general strategy to tackle the grid network planning problem [Develder2009]
- Routing and scheduling algorithms for grid applications, including anycast and advance reservations (eg. [Varvarigos2008, Stevens2009, Buysse2009, etc.])
- Differentiated service approaches for improved survivability, taking into account optical impairments [Markidis2008a, Wosinska2009, Christo2009, etc.]
- OBS-based approaches for optical grids [Buysse2009ICNS]

3.7.5 Conclusions

By the variety of results as outlined above, we are confident to have met our JA objectives, both in terms of research performed and collaborations that have been set-up (and are still ongoing).

A3.7.5.1 References

[1] P. Batchelor et al., "Study on the implementation of optical transparent transport networks in the European environment-Results of the research project COST 239", Photonic Network Communications, vol. 2, no. 1, pp. 15-32, 2000.

[2] Traffic Measurements and Models in Multi-Service Networks project, Celtic project, "TRAMMS IP Traffic report no. 1, April 2008", http://projects.celtic-initiative.org/tramms/.

[3] Traffic Measurements and Models in Multi-Service Networks project, Celtic project, "TRAMMS IP Traffic report no. 3, June 2008", http://projects.celtic-initiative.org/tramms/.

[4] J. Doucette, M. Clouqueur and W. D. Grover, "On the availability and capacity requirements of hared backup path protected mesh networks," Optical Networks Magazine, November/December 2003.

[5] Dynamic Impairment Constraint Network for Transparent Mesh Optical Networks (DICONET), www.diconet.eu

[6] *Network Simulator - ns2 available at http://www.isi.edu/nsnam/ns/*

[7] OIRC OBS-ns Simulator available at *http://wine.icu.ac.kr/~obsns/index.php*

[8] EvalVid - A Video Quality Evaluation Tool-set available at *http://www.tkn.tu-berlin.de/research/evalvid/fw.html*

3.7.6 Published Joint Papers (joint and single partner)

The joint papers are indicated with a "*". Contributing partners are listed at the end of each paper.

A3.7.6.1 JA7

[Varvarigos2008] E. Varvarigos, V. Sourlas, K. Christodoulopoulos, "Routing and Scheduling Connections in Networks that Support Advance reservations", Computer Networks (58), 2008. (RACTI)

[Kokkinos2008] P. Kokkinos, K. Christodoulopoulos, A. Kretsis, E. A. Varvarigos, "Data Consolidation: A Task Scheduling and Data Migration Technique for Grid Networks", Proc. 8th IEEE Int. Symp. on Cluster Computing and the Grid (CCGRID 2008), 19-22 May 2008, pp. 722-727. (RACTI)

[Doulamis2008] N. Doulamis, P. Kokkinos, E. A. Varvarigos, “Spectral Clustering Scheduling Techniques for Tasks with Strict QoS Requirements”, Proc. 14th Int. Conf. on Parallel and Distributed Computing (Euro-Par 2008), Las Palmas de Gran Canaria, Spain, 26-29 Aug. 2008, pp. 478-488. (RACTI)

[Markidis2008a] G. Markidis and A. Tzanakaki, “Network Performance Improvement through Differentiated Survivability Services in WDM networks”, J. of Optical Networking, June 2008, vol. 7, no 6, pp. 564-572. (AIT)

[Markidis2008b] G. Markidis and A. Tzanakaki, “Routing and wavelength assignment algorithms in survivable WDM networks under physical layer constraints,” in 5th Int. Conf. on Broadband Communications, Networks and Systems, 2008, pp. 191–196 (AIT)

***[Jiratti2009]** A. Jirattigalachote, L. Wosinska, P. Monti, K. Katrinis and A. Tzanakaki, "Impairment Constraint Based Routing (ICBR) with Service Differentiation in Survivable WDM Networks", in Proc. 35th European Conf. and Exhibition on Optical Communication (ECOC 2009), Vienna, Austria, September 2009. (AIT+KTH)

[Develder2009] C. Develder, B. Mukherjee, B. Dhoedt, and P. Demeester, “On dimensioning optical grids and the impact of scheduling,” Photonic Network Commun. (PNET), Vol. 17, No. 3, pp. 255-265. (IBBT)

***[Stevens2009]** T. Stevens, M. De Leenheer, C. Develder, B. Dhoedt, K. Christodoulopoulos, P. Kokkinos, E. Varvarigos, “Multi-Cost Job Routing and Scheduling in Grid Networks”, Future Generation Computer Systems, vol. 25, no. 8, Sep. 2009, pp. 912-925. (IBBT+RACTI)

[Buysse2009ICNS] J. Buysse, M. De Leenheer, C. Develder, B. Dhoedt, P. Demeester, "Cost-Effective Burst-Over-Circuit-Switching in a Hybrid Optical Network”, Proc. 5th Int. Conf. on Networking and Services (ICNS 2009), pp.499-504. (IBBT)

[Buysse2009DRCN] J. Buysse, M. De Leenheer, B. Dhoedt, and C. Develder, “Exploiting relocation to reduce network dimensions of resilient optical grids,” in Proc. 7th Int. Workshop Design of Reliable Commun. Netw. (DRCN 2009), Washington, D.C., USA, 25–28 Oct. 2009. (IBBT)

***[Wosinska2009]** L. Wosinska, A. Jirattigalachote, P. Monti, A. Tzanakaki, K. Katrinis, “Lightpath Routing Considering Differentiated Physical Layer Constraints in Transparent WDM Networks,” in Proc. of IEEE/OSA/SPIE Asia Communications and Photonics Conference (ACP), Shanghai, China, 2-6 Nov. 2009 (AIT+KTH)

[Christo2009] K. Christodoulopoulos, K. Manousakis, E. Varvarigos, M. Angelou, I. Tomkos, “A Multicost Approach to Online Impairment-Aware RWA”, Proc. IEEE ICC, 2009.

A3.7.6.2 Y3

[Buysse2010] J. Buysse, M. De Leenheer, B. Dhoedt, and C. Develder, “On the impact of relocation on network dimensions in resilient optical grids,” in Proc. Int. Conf. on Optical Network Design and Modelling (ONDM 2010), Kyoto, Japan, 31 Jan.–3 Feb. 2010. (IBBT)

[Jaumard2010] B. Jaumard, J. Buysse, A. Shaikj, M. De Leenheer, and C. Develder, “Column generation for dimensioning resilient optical grid networks with relocation,” in Proc. IEEE Global Telecommun. Conf. (Globecom 2010), Miami, FL, USA, 6–10 Dec. 2010, to appear. (IBBT)

[VanHoudt2010] B. Van Houdt, C. Develder, J. F. Perez, M. Pickavet, and B. Dhoedt, “Mean field calculation for optical grid dimensioning,” IEEE/OSA J. Opt. Commun. Netw., vol. 2, no. 6, pp. 355–367, Jun. 2010. (IBBT)

[DevelderICC] C. Develder, J. Buysse, A. Shaiki, B. Jaumard, M. De Leenheer, B.Dhoedt. “Survivable optical grid dimensioning: anycast routing with server and network failure protection”, submitted to ICC 2010. (IBBT)

[BuysseCC] J. Buysse, M. De Leenheer, B. Dhoedt, C. Develder, “Providing resiliency for optical Grids by exploiting relocation: A dimensioning study based on ILP”, submitted to Computer Communications (under review). (IBBT)

[Georgakilas2010a] Konstantinos N. Georgakilas, Kostas Katrinis, Anna Tzanakaki, and Ole B. Madsen, "Performance Evaluation of Impairment-Aware Routing Under Single- and Double-Link Failures," J. Opt. Commun. Netw. 2, 633-641 (2010) (AIT)

[Georgakilas2010b] K.N. Georgakilas, K. Katrinis, A. Tzanakaki and O.B. Madsen, "Impact of dual-link failures on impairment-aware WDM routed networks" (Invited paper), in Proceedings of the 12th International Conference on Transparent Optical Networks (ICTON 2010), Munich, Germany, June 2010 (AIT)

[Kokkinos2010] P. Kokkinos, K. Manousakis, E. A. Varvarigos, “Path Protection in WDM Networks with Quality of Transmission Limitations,” IEEE ICC, 2010. (RACTI)

4. Conclusions

This report has summarized the achievements of all joint activities that have been active within VCE-WP12, including a few joint activities that were moved from former WP21 Topical Project on Service Oriented Optical Networks at the end of the first year.

Overall, 16 partners were involved in this workpackage and seven joint activities have been active in the WP. Along the WP development, the collaboration among partners increased and among others, allowed to set-up three joint testbeds for experimental evaluation of service-oriented platforms for optical networks, based on GMPLS or OBS. Dissemination of the joint activities in international conferences and journals and in the Open Grid Forum standardization body has been actively pursued. The overall integration in the WP can be assessed by the observing that nearly half of the WP publications are joint.

Research outcomes were achieved in the field of strategy and approaches to improve and guarantee the service availability. In particular, the relocation of services was proposed as a way to guarantee resilience and capable of reducing to reduce the capacity requirements of the network. Evaluation of the reliability of wireless services on an integrated wireless and wired network was carried out and showed a tradeoff between reliability and number of handovers experienced by the service.

A new architecture of a service aware optical network has been proposed, suitable for cloud computing applications. The networking paradigm introduces service awareness in the core, by creating self-organized islands of service transparency. To allow for service-aware networking, a complete suite of algorithms, strategies, and mechanisms for implementing service abstraction and resource virtualization were proposed and investigated. Mechanisms for the co-selection and co-reservation of the resources in the service plane were shown to improve the performance of service provisioning for future Internet applications. Also, the impact of the service plane as an integrated entity vs. trusted or separate entity was evaluated. Also, the configuration of the service layer and network virtual platforms to support cloud computing have been proposed and discussed.

The experimental activity demonstrated the feasibility of the service plane and service-oriented optical networks. During the experimental assessment, the service set-up latency experienced in a service-plane testbed was evaluated and was shown to meet ITU-T recommendations. PCE-based control plane enabled with functionality for the setup of grid services has been tested on an optical network with commercial equipment and validated the simulative results.

Acronyms

AO-M	Application-Oriented Module
ASN	Access Service Network
ASN-GW	Access Service Network Gateway
B2B	Back-to-Back
BC	Binding Cache
BCH	Burst Control Header
BCP	Burst Control Packet
BER	Bit Error Ratio
BGP	Border Gateway Protocol
BS	Base Station
CE	Customer Edge
CMS	Closest MAP Selection
CN	Correspondent Node
CoA	Care of Address
CSD	Call Set-up Delay
CSE	Centralized Service Element
CSN	Connectivity Service Network
CSPF	Constraint Shortest Path First
DoS	Denial of Service
DSE	Distributed Service Element
DS-TE	DiffServ-aware Traffic Engineering
E-OBS	Offset-Time Emulated OBS
ER	Edge Router
FA	Forward Adjacent
FER	Frame Error Ratio
FMS	Farthest MAP Selection
FPGA	Field Programmable Gate Array
(G)MPLS	(Generalized) Multi-Protocol Label Switched Path
GRACF	Admission Control Function for Grid Services
GRM	Grid Resource Manager
HA	Home Agent
HMIPv6	Hierarchical Mobile IPv6
HMS	Half-Way MAP Selection
HoA	Home Address
IaaS	Infrastructure as a Service
IAT	Inter-Arrival Time
ILP	Integer Linear Programming
ICBR	Impairment Constraint Based Routing
JA	Joint Activity
LCoA	Link Care of Address
LSP	Label Switched Path
LSR	Label Switched Router
NaaS	Network as a Service

NVP	Network Virtualization Platform
MAP	Mobility Anchor Point
MEMS	Micro-Electro-Mechanical System
MG-OXC	Multi-Granular Optical cross-connect
MIPv6	Mobile IPv6
MS	Mobile Station
NE	Network Element
NR-DB	Network Resource Data Base
NRDL	Network Resource Specifications
NRM	Network Resource Manager
NSA	Network Service Agent
NSI	Network Service Interface
OGF	Open Grid Forum
OBS	Optical Burst Switching
OPS	Optical Packet Switching
OBT	Optical Burst Transport
PCE	Path Computation Element
PCEP	Path Computation Element Protocol
PCC	Path Computation Client
PCS	Path Computation Scheduler
PCReq	Path Computation Request
PS	Proxy Server
PSN	Packet Switched Network
QoS	Quality of Service
RA	Requestor Agent
RCoA	Regional Care of Address
SaaA	Software as a Service
SCF	Service Control Functions
SIP	Session Initiated Protocol
SOON	Service Oriented Optical Network
SO-MGON	Service-Oriented Multi-Granular Optical Network
SP	Service Plane
SR	Service Request
SRLG	Shared Risk Link Group
TED	Traffic Engineering Database
UAC	User Agent Client
UAS	User Agent Server
VPLS	Virtual Private LAN Service
WG	Working Group
WP	Work Package
WSON	Wavelength-Switched Optical Network
Y1	First Year
Y2	Second Year
Y3	Third Year