Novel Network Architecture for Optical Burst Transport

Von der Fakultät für Informatik, Elektrotechnik und Informationstechnik
der Universität Stuttgart zur Erlangung der Würde
eines Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

vorgelegt von

Christoph Martin Gauger

geb. in Esslingen am Neckar

Hauptberichter:	Prof. Dr.-Ing. Dr. h. c. mult. Paul J. Kühn
Mitberichter:	Prof. Biswanath Mukherjee, Ph.D., UC Davis, USA
Tag der Einreichung:	09. Mai 2006
Tag der mündlichen Prüfung:	09. August 2006

Institut für Kommunikationsnetze und Rechnersysteme
der Universität Stuttgart

2006

To Susanne.

Summary

Transport networks form the backbone of communication networks by cost-efficiently offering huge bandwidth and by guaranteeing a high service quality and availability. These requirements can best be met by using optical communication technologies. Currently, wavelength-switching is the most prominent network technology employing optical fiber communication and wavelength division multiplexing. It transports data in circuit-switched wavelength channels, the so-called lightpaths. While for years progress in optical networks has been defined by ever increasing transmission bit-rates, higher flexibility and manageability as well as multi-service and multi-layer integration are equally important criteria today. Accounting for these trends, optical burst switching (OBS) has been proposed as a highly dynamic optical network architecture. It offers fine-granular transport of different packet-switched services and applies statistical multiplexing directly in the optical layer.

This thesis presents the design, modeling, and evaluation of the optical burst transport network architecture (OBTN). The architecture is motivated by the need for flexible, scalable, and cost-efficient transport in next generation networks. In addition, it is stimulated by the research activities towards highly dynamic optical network infrastructures.

OBTN defines a network architecture to transport and switch optical burst data in a core network. The design objectives for the OBTN architecture are (i) an overall high quality of service, (ii) a network design allowing for cost-efficiency and scalability, and (iii) a network evolution perspective based on the current wavelength-switched networks. These objectives are achieved by combining selected concepts, architectures, and strategies of optical burst and optical packet switching as well as of multi-layer traffic engineering.

In order to provide the background information for the design of OBTN, Chapter 2 introduces the general characteristics, requirements, and trends for next generation transport networks. Also, it discusses the concept of layering in next generation networks and its application in layer networks for the virtualization of transport resources. Consequently, virtual topology design and dimensioning are analyzed to quantify the trade-offs regarding connectivity and resource requirements. Chapter 2 also reviews the fundamental technologies as well as currently emerging data and control plane architectures for optical transport networks. This presentation is then extended towards a long-term perspective. It describes architectural constraints and classification criteria for highly dynamic optical network architectures. These criteria are used to characterize the *fast optical circuit switching*, *optical burst switching*, and *optical packet switching* architectures. Then, hybrid optical network architectures are discussed as a framework to combine wavelength-switched and optical burst/packet-switched networks.

Chapter 3 discusses the state of research and technology for optical burst switching to structure the design space and identify promising approaches. Thus, it presents the requirements for the different functions in an OBS network and classifies the proposed architectures and mechanisms. Particularly, it addresses contention resolution which is necessary to achieve a high QoS in burst-switched networks. Here, wavelength conversion, fiber delay line buffering, alternative/deflection routing, and their combinations are looked at. It is concluded that wavelength conversion is a promising primary contention resolution strategy but should be complemented by FDL buffering and/or alternative routing. Thus, architectures, parameters, and operational strategies for FDL buffers are discussed in detail. This is supported by Appendix A which analyzes the performance of shared FDL buffers for different configurations and traffic characteristics. The review of alternative/deflection routing shows that it can only support other contention resolution schemes if it is closely controlled, i.e., if extensive deflections and route variations are avoided. Finally, architectures and realization aspects for burst-switched core nodes are presented to understand their resource and scalability constraints.

Chapter 4 presents the design rationale for OBTN and explains how OBTN combines a burst-switched client layer network with a wavelength-switched server layer network. Then, it introduces its fundamental concepts, namely the dense virtual topology, constrained alternative routing, and shared overflow capacity. These components are analyzed regarding their consequences for the overall node and network architecture. Further architectural details and variants as well as operational strategies of OBTN are discussed to complete the presentation. Finally, a qualitative discussion of OBTN with respect to optical burst switching and hybrid optical networks concludes this chapter.

Chapter 5 describes a unified resource model which allows to dimension and evaluate burst-switched architectures with different virtual topologies. Also, it details the dimensioning process for OBTN. Then, it addresses the simulation methodology and the reference evaluation scenario used in Chapter 6. It discusses metrics for node and network resources as well as for QoS performance. Finally, it derives QoS objectives for burst-switched core networks.

Chapter 6 evaluates OBTN and compares it with the two burst-switched reference architectures OBS and *Burst-over-Circuit-Switching*. OBS uses a sparse virtual topology while BoCS employs a full-mesh virtual topology. The evaluations in Section 6.1 show that for the same high target QoS, suitable OBTN dimensionings require substantially less resources in burst-switched nodes than OBS and slightly less than BoCS. This improvement comes at the cost of higher resource requirements compared to OBS in the underlying wavelength-switched server layer. However, applying the cost relations for lambda grid networks, in which bandwidth is considered a commodity and client layer resources the major cost driver, OBTN yields an overall cost reduction. The best results for OBTN are obtained when approximately 10 % of the network capacity is assigned as shared overflow capacity.

The comparison of OBTN and OBS is extended towards OBS architectures without an FDL buffer and OBS architectures with alternative routing. It is demonstrated that bufferless OBS with and without alternative routing requires approximately the same amount of server layer resources as OBTN. However, it consumes more client layer resources. For OBS with an FDL buffer, alternative routing does neither impact the client layer nor the server layer resources substantially. Furthermore, the effectiveness of constrained alternative routing and of the shared overflow capacity in OBTN is assessed by isolating them. This is achieved by comparing OBTN

and BoCS which use the same virtual topology but differ in the routing flexibility and network dimensioning. The evaluations show that applying any of the concepts alone to BoCS does not yield any or only limited improvements or even produces severe penalties. However, if applied together as OBTN, they harmonize and effectively improve performance.

Evaluations with different FDL buffer architectures and dimensionings quantify the trade-off of improved contention resolution, thus reduced node and network resources and the additional node resources required for the FDL buffer. The results show that increasing the number of FDL buffer ports up to approx. 16–24 leads to fewer node and network resources. Beyond this dimensioning, node resources increase with diminishing reductions regarding resources in the underlying server layer. Extending the studies to advanced virtual topologies, it is demonstrated that OBTN is not restricted to the full-mesh virtual topology used for all previous evaluations. A traffic demand-based and a path-length-based virtual topology design approach is introduced and investigated. For one example parameterization, the number of virtual links can be more than halved with a small penalty in client layer resources and a small gain in server layer resources.

Finally, the impact of changing the total traffic demands on the performance and on resource requirements of all three architectures is studied. It shows that when increasing the total traffic demand, BoCS becomes increasingly efficient. For small traffic demands OBS becomes more attractive. However, OBTN provides an attractive solution for intermediate traffic demands where neither OBS nor BoCS can offer optimal solutions.

Concluding, OBTN is shown to offer an overall high QoS, to effectively reduce the node resources of the burst-switched client layer, and to perform well in a wavelength-switched network context.

Zusammenfassung

Transportnetze stellen das Rückgrad von Telekommunikationsnetzen dar, für die sie große Bandbreiten kostengünstig, ausfallsicher und mit hoher Dienstgüte bereitstellen. Da diese Anforderungen durch optische Kommunikationsnetze am besten erfüllt werden können, basieren Transportnetze heute meist auf Glasfasernetzen mit Wellenlängenmultiplex-Technik. Die momentan wichtigste optische Netzarchitektur schaltet im Netz Wellenlängenkanäle und baut damit sogenannte Lichtpfade auf. In diesen Lichtpfaden werden die Daten verschiedenster darüber liegender Netzschichten transportiert. Viele Jahre lang wurde Fortschritt in optischen Netzen mit stetig steigenden Übertragungsbitraten gleichgestellt. Inzwischen haben aber hohe Flexibilität und Steuerbarkeit sowie die Möglichkeit verschiedene Transportdienste und Netzschichten auf einer Platform zu integrieren eine ebenso hohe Bedeutung erlangt.

Um diesen Trends Rechnung zu tragen, wurde Optical Burst Switching (OBS) als hochdynamische optische Netzarchitektur vorgeschlagen. OBS ermöglicht feingranularen Transport und statistisches Multiplexen in der optischen Ebene zur Integration paketvermittelter Transportdienste. Durch Verkehrsaggregation und Assemblierung zu Bursts vermeidet es jedoch die Probleme optischer Paketvermittlung, in der einzelne Pakete vermittelt werden müssen.

Diese Dissertation behandelt den Entwurf, die Modellierung sowie die Bewertung der *Optical Burst Transport Network* Architektur (OBTN). Sie ist durch den Bedarf der Next Generation Networks (NGN) an Datentransport mit hoher Flexibilität, Skalierbarkeit und Kosteneffizienz motiviert. Darüberhinaus ordnet sie sich in Forschungsaktivitäten in Richtung auf hochdynamische optische Netze ein.

OBTN beschreibt eine Netzarchitektur, die optische Bursts in einem Transportnetz vermittelt. Dem Entwurf der OBTN Architektur lagen folgende Ziele zugrunde: (i) eine insgesamt hohe Dienstgüte (QoS), (ii) Kosteneffizienz und Skalierbarkeit und (iii) eine Perspektive für die Netzevolution auf der Basis der momentanen wellenlängen-vermittelten Netze. Um diese Entwurfsziele zu erreichen, bedient sich OBTN Konzepten, Architekturen und Strategien der Netztechniken Optical Burst und Optical Packet Switching sowie des Multilayer-Verkehrsmanagements.

Zunächst führt das Kapitel 2 allgemeine Eigenschaften, Anforderungen und Trends der nächsten Generation von Transportnetzen ein, die den Hintergrund des Entwurfs von OBTN darstellen. Dazu wird das Schichtenkonzept (Layering) für NGNs sowie dessen Anwendung als sog. *layer networks* zur Virtualisierung von Transportressourcen diskutiert. Der Entwurf und die Dimensionierung virtueller Topologien werden analysiert, um Abhängigkeiten zwischen Konnektivität und Ressourcenbedarf zu quantifizieren. Anschließend werden die grundlegenden Technologi-

en und momentan in der Einführung befindlichen Architekturen für die Daten- und Steuerebene optischer Netze besprochen. Diese Aufarbeitung wird im Hinblick auf Architekturen mit einer langfristigen Perspektive fortgesetzt. Dazu werden architekturelle Randbedingungen und Klassifizierungskriterien für hochdynamische optische Netzarchitekturen zusammengestellt. Diese Kriterien werden wiederum verwendet, um die Architekturen *Fast Optical Circuit Switching*, *Optical Burst Switching* und *Optical Packet Switching* zu charakterisieren. Schließlich zeigt der Abschnitt über hybride optische Netze Möglichkeiten auf, solche hochdynamischen Netze mit den aktuellen, wellenlängen-vermittelten Netzen zu kombinieren.

Kapitel 3 arbeitet den Stand von Wissenschaft und Technik zu Optical Burst Switching (OBS) auf, um die Entwurfsoptionen zu strukturieren und vielversprechende Lösung zu identifizieren. Deshalb werden für die wesentlichen Funktionsblöcke in OBS jeweils Anforderungen genannt, sowie verfügbare Architekturen und Mechanismen für deren Realisierung klassifiziert. Insbesondere geht das Kapitel dabei auf Verfahren zur Blockierungsauflösung ein. Diese sind in OBS notwendig, um eine hohe Dienstgüte zu erreichen. Dazu werden Wellenlängenkonversion, Pufferung in Faserverzögerungsleitungen (*Fiber Delay Lines*, FDL), sowie alternative Verkehrslenkung (*Deflection Routing*) besprochen. Schließlich beschreibt das Kapitel Technologien und Architekturen für burst-vermittelte Kernetzknoten, um Ressourcenbeschränkungen und Skalierbarkeitsgrenzen aufzuzeigen.

Im 4. Kapitel wird die OBTN-Architektur eingeführt und alle wesentlichen Elemente diskutiert. Dazu werden ihre Kernkonzepte, eine stark vermaschte virtuelle Topologie, alternative Verkehrslenkung mit eingeschränkter Wegewahl und gemeinsam genutzte Überlaufkapazität, erklärt. Deren Auswirkungen auf die darunter liegende, wellenlängen-vermittelte Netzschicht sowie auf die Knotenarchitektur werden analysiert. Außerdem werden weitere architekturelle Details und Varianten sowie mögliche Steuerabläufe beschrieben. Eine qualitative Diskussion von OBTN und Vergleiche mit verwandten Architekturen schließen das Kapitel ab.

Kapitel 5 schlägt eine einheitliche Ressourcen-Modellierung für burst-vermittelte Architekturen mit unterschiedlichen virtuellen Topologien vor. Diese wird anschließend zur Ressourcen-Dimensionierung von OBTN-Netzen angewandt. Darüberhinaus werden die Bewertungsmethodik für Kapitel 6 und das Untersuchungsszenario diskutiert. Schließlich werden Metriken zur Bewertung der Dienstgüte und des Ressourcenbedarfs eingeführt sowie im Fall der Dienstgüte durch entsprechende Zielwerte konkretisiert.

Kapitel 6 untersucht die OBTN-Architektur und vergleicht sie mit den zwei burst-vermittelten Referenzarchitekturen OBS und *Burst-over-Circuit-Switching* (BoCS). Dabei verwendet OBS eine schwach-vermaschte virtuelle Topologie, während BoCS eine stark vermaschte virtuelle Topologie einsetzt. Die Bewertung in Kapitel 6 zeigt, dass OBTN prinzipiell in der Lage ist, eine hohe Dienstgüte bei guter Ressourcenausnutzung bereitzustellen. Geeignete Dimensionierungen von OBTN benötigen in den burst-vermittelten Knoten deutlich weniger Ressourcen als OBS und sogar etwas weniger als BoCS. Diese Verbesserung wird jedoch durch einen höheren Ressourcenbedarf als OBS in der darunter liegenden, wellenlängen-vermittelten Netzschicht erkauft. Da in Transportnetzen typischerweise die Kosten für die Netzelemente höherer Netzschichten diejenigen niederer Netzschichten dominieren, ergeben sich durch OBTN insgesamt Kosteneinsparungen.

Diese Untersuchungen werden bzgl. OBS-Architekturen mit erweiterter Funktionalität zur Auflösung von Blockierungen fortgesetzt. Die Ergebnisse zeigen allerdings, dass auch diese OBS-Varianten höhrere Ressourcenanforderungen an die Burstknoten stellen als OBTN. Anschließend werden die einzelnen Architekturkomponenten von OBTN bezüglich ihrer Effektivität isoliert betrachtet. Es zeigt sich, dass jede einzelne Komponente keine oder nur eine relativ geringe Verbesserung ergibt. Wirken die entsprechenden Komponenten aber wie in OBTN vorgeschlagen zusammen, lassen sich die angestrebten Leistungsmerkmale erreichen.

Anschließend werden die Einflüsse komplexerer Einheiten zur Blockierungsauflösung auf den Ressourcenbedarf burst-vermittelter Knoten analysiert. Während die Installation solcher Einheiten den Ressourcenbedarf meist direkt erhöht, ermöglichen sie aber auch eine höhere Auslastung und damit eine Verringerung benötigter Ressourcen. Die Untersuchung dieser gegenläufigen Effekte für den Fall von FDL-Puffern ergibt, dass bereits sehr einfache Puffer optimal hinsichtlich der benötigten Ressourcen sind. Mit Hilfe von erweiterten virtuellen Topologien, die ihre Konnektivität über Mindestwerte für Verkehrsangebote oder für Pfadlängen definieren, wird gezeigt, dass OBTN auch mit weniger stark vermaschten Topologien vergleichbar leistungsfähig ist. Dadurch kann der Aufwand zur Steuerung und Verwaltung der virtuellen Links reduziert werden. Schließlich thematisieren Studien den Einfluss kleiner und großer Verkehrsangebote auf die Dienstgüte und den Ressourcenbedarf von OBTN, OBS und BoCS.

Zusammenfassend wird festgestellt, dass die OBTN-Architektur die an sie ursprünglich gestellten Anforderungen nach hoher Dienstgüte, nach Ressourceneffizienz sowie nach einer Migrationsperspektive erfüllt.

Contents

List of Figures

List of Tables

Abbreviations and Symbols

Abbreviations

Alt-FDL	Hunting mode preferring an alternate route over FDL buffering
ASON	Automatically switched optical network
ASTN	Automatic switched transport network
ATM	Asynchronous transfer mode
AWG	Arrayed waveguide grating
BER	Bit error ratio
BoCS	Burst-over-circuit switching
BRM	Basic reference model
CapEx	Capital expenditures
CCDF	Complementary cumulative distribution function
CR-LDP	Constraint-based routing label distribution protocol
CWDM	Coarse wavelength division multiplexing
DBORN	Dual bus optical ring network
DPSK	Differential phase-shift keying
DSL	Digital subscriber line
DSLAM	Digital subscriber line access multiplexer
DWDM	Dense wavelength division multiplexing
FBM	Fractional Brownian motion
FDL	Fiber delay line
FDL-Alt	Hunting mode preferring FDL buffering over an alternate route

FEC	Forward equivalence class
FF	First-fit
FRP	Fast reservation protocol
FSK	Frequency-shift keying
FTTX	Fiber-to-the-X (X denotes the location of the optical terminal)
GFP	Generic framing procedure
GMPLS	Generalized multiprotocol label switching
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPDV	IP packet delay variation
IPLR	IP packet loss ratio
IPTD	IP packet transfer delay
ISP	Internet service provider
ITU-T	International Telecommunication Union - Telecommunication Standardization Sector
JET	Just-enough-time
JIT	Just-in-time
LAUC	Latest available unscheduled channel
LAUC-VF	Latest available unscheduled channel with void filling
LCAS	Link capacity adjustment scheme
LDP	Label distribution protocol
LMP	Link management protocol
LSP	Label-switched path
MAN	Metropolitan area network
minConv	Probing wavelength converters and FDL buffers to minimize converter usage
minDelay	Probing wavelength converters and FDL buffers to minimize the delay
MPLS	Multiprotocol label switching
MPλS	Multiprotocol lambda switching
MSN	Manhattan street network

NGN	Next generation network
NNI	Network network interface
OBS	Optical burst switching
OBTN	Optical burst transport network
OCS	Optical circuit switching
OIF	Optical Internet Forum
OpEx	Operational expenditures
OPS	Optical packet switching
OSI	Open Systems Interconnection
OSPF	Open shortest path first
OTH	Optical transport hierarchy
OTN	Optical transport network
PCE	Path computation element
PDH	Plesiochronous digital hierarchy
PEBS	Pre-estimate burst switching
PHY	Physical layer in OSI layer model
PMON	Polymorphic multi-service optical network
PON	Passive optical networks
PostRes	Resource reservation scheme which schedules the output wavelength only after buffering
PreRes	Resource reservation scheme which schedules the output wavelength before buffering
QoS	Quality of service
RFD	Reserve-a-fixed duration
RLD	Reserve-a-limited duration
RSVP	Resource reservation protocol
RSVP-TE	Resource reservation protocol extensions for LSP tunnels
SAN	Storage area network
SCM	Sub-carrier modulation

SDH	Synchronous digital hierarchy
SDL	Switched delay line
SOA	Semiconductor optical amplifier
SONET	Synchronous optical network
TAG	Tell-and-go
TAW	Tell-and-wait
TCP	Transmission control protocol
TDM	Time division multiplexing
TTL	Time-to-live
TWC	Tunable wavelength converter
UNI	User network interface
VCAT	Virtual concatenation
VDSL	Very high speed digital subscriber line
VON	Virtual optical network
VPN	Virtual private network
WAN	Wide area network
WDM	Wavelength division multiplexing
WDS	Wavelength and delay selection

Symbols

$\mathbf{A}$	Traffic demand matrix
A	Total traffic demand
$a_{i,j}$	Traffic demand from node i to node j
A_{thr}	Traffic demand threshold in virtual topology design
$B(\cdot,\cdot)$	Erlang-B formula
B_{FDL}	Effective storage capacity per wavelength channel of a fiber delay line
b	Bit-rate
$\mathbf{C}$	Client layer capacity matrix
$\mathbf{C}_{\text{D}}$	Client layer capacity matrix of the direct end-to-end lightpath
$\mathbf{C}_{\text{S}}$	Client layer capacity matrix of the shared overflow capacity
C	Total capacity of the client layer
C_i	Total capacity of the client layer for the intersected overprovisioning factor α_i
$C_{\text{D,target}}$	Total target number of ports for direct end-to-end lightpaths
$C_{\text{S,target}}$	Total target number of ports for shared overflow capacity
$c(\cdot)$	Capacity dimensioning function (scalar)
c_{T}	Coefficient of variation
D	Delay granularity in buffers with multiple FDLs
$d_{\text{A}}(\cdot,\cdot)$	Mean hop distance of routed traffic demands
$d_{\text{D}}(\cdot,\cdot)$	Mean hop distance for individually dimensioned traffic demands
$d_{\text{G}}(\cdot)$	Mean hop distance for the topology graph
$\mathbf{E}$	Matrix with all elements 1 except for 0 on diagonal
$e_{i,j}$	Element (i,j) of the matrix $\mathbf{E}$
F	Number of FDLs in multi-FDL buffer
H	Number of physical layer hops a virtual link spans
H_{thr}	Physical hop threshold for virtual links in virtual topology design
h	Burst transmission time
M	Number of wavelengths in input or output fiber

N	Number of input and output fibers of a node
N_{D}	Number of ports for direct end-to-end lightpaths
N_{F}	Number of FDL buffer ports, product of W_{F} and F
$N_{\mathrm{F,total}}$	Total number of FDL buffer ports in the network
N_{L}	Number of local add/drop ports
$N_{\mathrm{L,total}}$	Total number of local add/drop ports in the network
N_{S}	Number of ports for shared overflow capacity
n	Number of nodes in the topology
n_{MAN}	Mean number of MANs per point-of-presence
$\mathbf{P}$	Physical topology matrix
P_{F}^{*}	Probability that no suitable FDL can be scheduled in the FDL buffer at the time of a burst blocking event
P_{L}	Burst loss probability
$P_{\mathrm{L,target}}$	Target burst loss probability
P_{O}	Probability that all output wavelength channels are blocked at burst arrival
$P_{\mathrm{O}}^{*}(t)$	Probability that all output wavelength channels are blocked at time t after a burst blocking event
R_{CR}	Maximum number of constrained alternative routes between a pair of network nodes
R_{FM}	Number of routes between a pair of network nodes in a full-mesh topology
r_{C}	Ratio of number of wavelength converters and number of input/output wavelength channels
$\mathbf{S}$	Server layer capacity matrix
$\mathbf{S}_{\mathrm{D}}$	Server layer capacity matrix of the direct end-to-end lightpaths
$\mathbf{S}_{\mathrm{S}}$	Server layer capacity matrix of the shared overflow capacity
S	Total capacity of the server layer
S_i	Total capacity of the server layer for the intersected overprovisioning factor α_i
$\mathbf{T}$	Topology matrix
$\mathbf{T}_{\mathrm{P}}$	Physical topology matrix
$\mathbf{T}_{\mathrm{V}}$	Virtual topology matrix

T_{F}	Delay of FDL
$T_{\mathrm{F,i}}$	Delay of FDL i in a buffer with multiple FDLs
t_{a}	Arrival time of the burst control packet
$t_{i,j}$	Element (i, j) of the topology matrix
W_{C}	Number of wavelength converters in converter pool
W_{F}	Number of wavelengths in FDL
$W_{\mathrm{F,i}}$	Number of wavelengths in FDL i in a buffer with multiple FDLs
W_{V}	Number of wavelengths in a virtual link

α	Overprovisioning factor of the client layer
α_{D}	Overprovisioning factor of the direct end-to-end lightpaths
α_{FM}	Overprovisioning factor in case of a full-mesh virtual topology
α_i	Intersected overprovisioning factor of the client layer for the target QoS
α_{P}	Overprovisioning factor if the virtual and physical topology are identical
β	Share of (server layer) network capacity assigned as shared overflow capacity
β_i	Intersected share of (physical) network capacity assigned as shared overflow capacity for the target QoS
$\gamma(\cdot)$	Capacity dimensioning function (matrix operation)
$\tilde{\gamma}(\cdot)$	Alternative capacity dimensioning function (matrix operation)
Δ_i	Reservation time in core node i
δ	Offset time between burst control packet and data burst
$\phi(\cdot,\cdot)$	Demand routing function (matrix operation)
λ_{FDL}	Wavelength channel in FDL
λ_{in}	Input wavelength channel
λ_{out}	Output wavelength channel
ρ_{MAN}	Mean utilization of resources
τ	Propagation delay of a network link
$\|\| \cdot \|\|$	Sum of all matrix elements

1 Introduction

The rapid and successful emergence of the Internet in the 1990s leveraged the penetration of our society and economy with information and communication technologies. This process significantly accelerated transformations towards an information society and a knowledge-based economy. At the same time, major economic shifts and technical trends materialized. From an economic perspective, competition among operators, market deregulation, and globalization changed the telecommunication industry. From a technical perspective, the explosion of digital media, new high-bandwidth applications, and the support of mobility defined new requirements for communication networks. Within telecommunications, this triggered the break-up of many monolithic, service-specific communication infrastructures and founded a strong convergence trend towards general platforms for services and transport.

Within few years, the Internet has developed from a scientific network for a limited user group, via an era of the *free* public Internet with its non-profit aura, to the most important commercial platform for any type of information and communication service. Shopping and trading, information retrieval, human-to-human communication, and television are only a selection of services that have been shifted to the Internet. However, the Internet suffers from the fact that its basic concepts and infrastructures have neither been designed for wide-spread commercial usage nor for the operation by telecommunication carriers. Next generation networks are thus embraced as a framework for realizations to bridge the gap between the Internet and carrier networks. They provide the convergence towards flexible, scalable, and cost-efficient platforms to offer services and to transport data.

1.1 Photonic Transport Networks

Transport networks form the backbone of communication networks by cost-efficiently offering huge bandwidth and by guaranteeing a high service quality and reliability. These requirements can best be met by using optical communication technologies. Currently, wavelength-switching is the most prominent network technology employing optical fiber communication and wavelength division multiplexing (WDM). It transports data in circuit-switched wavelength channels, the so-called lightpaths. While for years progress in optical networks has been defined by ever increasing transmission bit-rates, higher flexibility and manageability as well as multi-service and multi-layer integration are equally important criteria today.

Accounting for these trends, optical burst switching (OBS) has been proposed as a highly dynamic optical network architecture. It offers fine-granular transport of different packet-switched services and applies statistical multiplexing directly in the optical layer. Over the past years it has become an active field of research with a considerable number of publications and a series of workshops dedicated to OBS. Most contributions focus on the architecture and functionality of nodes and study performance aspects. Far less work is available on realization topics, resource requirements or migration scenarios for OBS and wavelength-switched networks.

However, especially optical network architectures have to comply with several realization and operational constraints. On the one hand, they are restricted by limitations of optical technologies and on the other hand they can be stimulated by new break-through components or systems. Also, they have to fit into the operational environment of a transport network provider regarding quality of service (QoS) and reliability. Finally, new optical network architectures should offer a network evolution perspective and a migration path.

Considering these trends and requirements, this thesis designs, models, and evaluates, the optical burst transport network architecture (OBTN). OBTN combines a burst-switched with a wavelength-switched network layer to minimize the resource requirements for the OBS layer. In addition, OBTN defines advanced routing, dimensioning, and contention resolution schemes to achieve a high QoS.

1.2 Overview of the Thesis

This thesis is structured into the fundamental chapters on transport networks (Chapter 2) and optical burst switching (Chapter 3) as well as the chapters on the architecture (Chapter 4), modeling and dimensioning (Chapter 5), and evaluation (Chapter 6) of OBTN.

Chapter 2 discusses next generation networks (NGN) with a focus on transport networks. It first introduces the NGN basic reference model, its transport stratum, and layering concept. This concept can be used for virtualization of network resources: a transport network can provision an arbitrary virtual topology to a client layer network. Also, it reviews key technical and non-technical characteristics of transport networks and their operation. Then, optical transport networks are surveyed in terms of fundamental technologies as well as current and emerging data and control plane architectures. Furthermore, architectural constraints for future, highly dynamic optical network architectures are presented and related to each other. They are used to define and characterize the fast optical circuit switching, optical burst switching, and optical packet switching architectures. Finally, it treats virtual topology design and dimensioning and derives quantitative relations for their resource requirements.

Chapter 3 takes a detailed look at OBS and its functional components. It classifies mechanisms, architectures, and the available literature for aggregation and assembly, reservation, scheduling, contention resolution, switching in core nodes, and QoS support. This in depth review not only structures the design space but also identifies promising concepts and deficiencies in OBS. The chapter is supported by Appendix A in which new performance results for shared fiber delay line (FDL) WDM buffers are presented. It demonstrates the principal behavior of FDL buffers and allows to derive suitable parameterizations for burst traffic and FDL buffers to be used in Chapter 6.

The new optical burst transport network architecture (OBTN) is introduced in Chapter 4. In a design rationale, QoS requirements, cost-aware network design, and the application scenario for burst transport networks are discussed. Then, the key extensions of the OBTN architecture with respect to OBS are presented, namely a densely-meshed virtual topology, constrained alternative routing, and shared overflow capacity on a subset of the virtual links. These extensions are analyzed regarding their consequences for the node and network elements. Finally, a qualitative discussion and comparison with respect to related reference architectures concludes this chapter.

Chapter 5 describes a unified resource model, which allows one to evaluate burst-switched architectures with different virtual topologies. This model is used to extend the dimensioning process for virtual topologies introduced in Chapter 2 towards OBTN. Then, the methodology, the QoS and resource metrics as well as the scenario of the evaluations in Chapter 6 are discussed.

Chapter 6 evaluates OBTN and compares it with two basic OBS reference architectures using different virtual topologies. It analyzes both the QoS performance of the architectures for a given dimensioning and the required node and network resources for a given target QoS. Further evaluations extend the comparison to other OBS reference architectures with a higher degree of routing flexibility, without an FDL buffer, or using different node control strategies. This also allows to isolate the different architectural components of OBTN and to assess their effectiveness. Then, variants of the node design and of the OBTN virtual topology design are studied to further optimize node resources and network operation. Finally, the impact of the total traffic demand in the network is analyzed.

Chapter 7 summarizes this thesis and presents an outlook to future work.

2 Next Generation Transport Networks

This chapter introduces the fundamental transport network concepts and architectures supporting Next Generation Networks (NGN). The term NGN is frequently but also ambiguously used to refer to new concepts in networking, ranging from first voice-over-IP (VoIP) deployments to entirely new service and transport architectures. In the context of this chapter, NGN, on the one hand, refers to the particular instantiation of the Global Information Infrastructure (GII) [Y.100] as covered by the ITU in its Y-series of recommendations. These documents and the recommended reference models introduce and structure key concepts of NGN and their design. On the other hand, the term *next generation* allows for a more fundamental and possibly disruptive change in networking technology rather than marginal evolutionary developments. In this sense, the chapter not only looks at current and emerging but also at future transport network architectures with a long-term perspective.

The first section of this chapter discusses the role of *transport* in the NGN and concepts for structuring transport networks. The second section focuses on current and emerging optical transport networks and presents main architectures for the data and control plane. Then, the third section introduces future, highly dynamic optical network architectures like fast optical circuit switching, optical burst and optical packet switching. Finally, the design and dimensioning of virtual topologies are discussed in the fourth section.

2.1 Transport Networks in Next Generation Networks

A central characteristic of the NGN is that service and transport should be decoupled in order to allow them to be offered separately and to evolve independently. This is reflected in the NGN Basic Reference Model [Y.2011] and discussed below. As this thesis concentrates on transport networks, the following subsections then take a closer look at the transport stratum and transport networks characteristics.

2.1.1 NGN Basic Reference Model

For several years, the well-known seven layers OSI Basic Reference Model (BRM [ISO94], [X.200]) represented the main model of communication systems. It formulates the principal concept of layered architectures and rigidly assigns functionality and specific characteristics to the seven layers. However, it does not adequately cover service platforms and shows deficiencies in several practical situations. Problems occur, for instance, if the functionality of the different

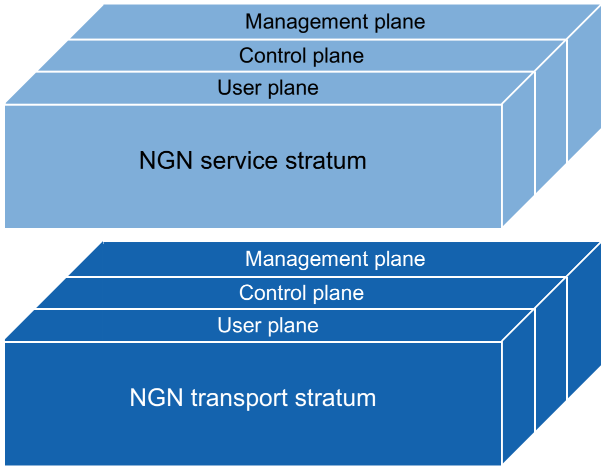


Figure 2.1: Basic reference model for next generation networks [Y.2011]

layers does not correspond to the OSI model or if the ordering of layers significantly differs, e.g., due to tunneling. Therefore, the NGN BRM defines a more general layering concept which offers more flexibility. It should be noted that it does not replace the OSI BRM but complements it with respect to overall network design aspects[1].

Figure 2.1 shows the separation of an NGN into the service stratum at the top and the transport stratum at the bottom. Both strata comprise user, control, and management planes, which are independent across the strata boundary. The **service stratum** is the part of an NGN exchanging service-related data and the functions to produce, control and manage the services as well as their resources. In contrast, the **transport stratum** comprises all functions to transfer data (including user, control and management information) and to control and manage the transport itself as well as the transport resources. This data transport thus provides user-to-user, user-to-services-platform, and services-platform-to-services-platform connectivity. As this thesis focuses on the user plane of transport networks, only the transport stratum will be further discussed.

2.1.2 The Transport Stratum

The NGN transport stratum is implemented by a hierarchy of layer networks as described in the recommendations of the functional architecture of transport networks [G.805, G.809]. These layer networks can use different types of switching technologies (circuit-switched, packet-switched) and different connection concepts (connection-oriented, connectionless). In principle, they also have their own user, control, and management planes. Nevertheless, a unified control and management can be implemented across layers in a single protocol (cf. Section 2.2.5).

The concepts of *layering*, i.e., the decomposition into independent layer networks, and of *partitioning* a layer network into sub-networks are described for transport network in [G.805] and [G.809] for connection-oriented and connectionless networks, respectively. The recommenda-

[1]The NGN BRM document [Y.2011] discusses the relation of the two BRMs in detail. Note that the terms transport stratum and transport network are not related to the layer 4 in the OSI BRM called *transport layer*.

tions are technology-independent and also introduce a graphical and formal textual notation to model layering, partitioning as well as generic functions of layer networks. Obviously, layering supports the concept of virtualization of resources, i.e., it allows to independently design, operate, add, and modify layer networks. In transport networks, it also supports recursive aggregation and multiplexing from higher to lower layers such that coarser granularities with lower cost per bit can be employed. In contrast, partitioning is mainly motivated by boundaries defined by operator or vendor domains as well as by administrative domains within the network of one operator. While layering plays an important role in the context of this thesis and is thus treated with more detail, partitioning is not considered in the following.

A hierarchy of layer networks as depicted in Figure 2.2 is called a multi-layer network. In a multi-layer network, the upper of two adjacent layer networks acts as a transport service requester and the lower layer as a transport service provider respectively. Therefore, [G.805] refers to the upper as the *client layer network* and to the lower as the *server layer network*. The connectivity and the characteristic functionality provided by a server layer is called a transport service which may or may not be offered to external customers depending on the operator's business model. As the links in the client layer network and thus the entire client layer topology is provisioned by the server layer, they are called *virtual links* and *virtual topology*. The terms virtual or alternatively logical topology refer to the fact that the client layer connectivity does not have to match any physical topology but can be changed by the server layer. For instance, Figure 2.2 shows that the virtual link between nodes C1 and C4 in the client layer is provided as a trail of server layer links.

Having multiple layer networks in the transport stratum requires coordination in order to avoid conflicts and inefficiencies. For instance, traffic engineering or failure recovery functions can benefit significantly if information is explicitly exchanged or protocols are suitably parameterized. Therefore, multi-layer or cross-layer design and engineering are actively worked on in several areas of communications. In transport networks, multi-layer traffic engineering and resilience attempt to optimally exploit the specific granularities and dynamics of the different network technologies [Mod01, Cin03, NGB03, VPD04, Köh05b, OM05, ZZM05]. In wireless communications, architectures and protocols on different layers are mostly adapted to match the special characteristics of radio transmission links, e.g., high packet losses or transfer delays [KHFK05, KHFK06]. From a strategic point of view, however, it is of utmost importance that cross-layer design and operation do not neutralize the benefits realized by layer independence.

The very generic definition of the transport stratum covers a wide range of network types, technologies and layers ranging all the way from access to core, from wireline to wireless, and from a physical to a network layer. In practical realizations, the transport stratum typically comprises the layers 1–3 of the OSI BRM. As the NGN explicitly targets packet-based networks [Y.2001] and as the interface to the services layer should be a convergence layer, the top-most layer of the transport stratum will likely be based on the Internet protocol (IP). The attractiveness of the IP layer [RFC791, RFC2460, Com06] is not only based on its packet-based communication or the availability of IP-centric control and management protocols defined by the IETF. It is mostly based on its universal deployment in the Internet and the multitude of applications and services built on top.

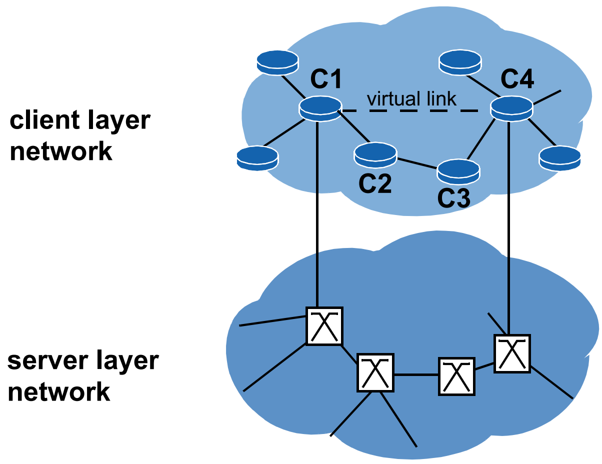


Figure 2.2: Client and server layer network in a layered transport network

2.1.3 Transport Network Characteristics

In contrast to the broad coverage of the term transport stratum, the term *transport network* is commonly used in a more restricted way. It primarily refers to high-capacity aggregation networks, metropolitan area networks (MAN) or wide area networks (WAN). Consequently, most transport networks are wireline except for high-speed point-to-point microwave radio systems. Typical transport networks comprise several layer networks using technologies such as multiprotocol label switching (MPLS) [RFC3031], asynchronous transfer mode (ATM) [I.121, ATM06], the synchronous (SDH) [G.707][2] and plesiochronous (PDH) [G.702] digital hierarchy and WDM. In addition to those technologies, Ethernet is currently extended towards *carrier-grade Ethernet* to satisfy the higher operational requirements in an operator compared to an enterprise network [GPTV04, GPTB05, Med05, ABMR06]. Due to its frame-based transport, its aura as a cost-efficient solution, and the huge market potential for Ethernet-based transport services it can be expected to win a significant market share over the next years. From the point of view of a transport network, the IP layer is usually not considered as a transport network but acts as a client layer network.

Apart from their high capacity, transport networks are also characterized by their role as the backbones of public networks. In order to offer attractive capacity, transport networks have to satisfy requirements towards low capital expenditures (CapEx) and low operational expenditures (OpEx). Although the components of CapEx and OpEx are not clearly specified, the former commonly comprises the cost of switching and transmission equipment. In contrast, the latter focuses on the cost of network operations, which not only includes control and management but also parts of the inbound and outbound logistics of a transport service provider [PKI$^+$05, KIA$^+$05]. Despite their importance, the cost of real estate infrastructure, e.g., central offices or fiber ducts, and the respective utilities like power and cooling are not consistently assigned to either CapEx or OpEx.

[2]In the US and Canada, the respective standard is called synchronous optical network (SONET). For the remainder of this thesis, only SDH will be used to refer to both though.

Regarding CapEx, a transport network technology has to prove cost-efficiency in terms of cost per bit, scalability, upgradability, and multi-service support. These requirements can be satisfied by high bit-rate transmission, aggregate and coarse-grain processing and switching, as well as transparency. Low OpEx can be achieved by effective control and management of the network, both in normal operation mode and in failure situations [PKI$^+$05]. As transport services require specialized, expensive equipment and often installations, provisioning has traditionally taken months and involved paper-based orders, long-term contracts as well as service personnel in the field. Thus, an automated transport service provisioning process is essential to achieve lower OpEx values [KIA$^+$05]. Finally, low power consumption and heat emission as well as a small foot-print are desirable, e.g., using small form factors or ultra-long-haul transmission to avoid inline amplifier locations.

As transport networks form the backbone, onto which operators deploy all other layer networks and all services, their service quality and reliability is of highest technical and strategic importance. The lower a transport network is in the hierarchy of layer networks, the higher are these requirements as they are the aggregation of the individual requirements of the higher layer networks and services. Therefore, network survivability as well as network and traffic engineering are important research areas in transport networks [GLM$^+$00, Cha03a, Cha03b, Cha04, VPD04, Cha05a, Muk06].

Transport networks are not necessarily owned by the service provider, whose traffic it transports. Although this is still the case for the incumbent telecommunications operators, many regional, national or global carriers only own transport infrastructure and market wholesale or value-added transport services, e.g., Ethernet LINE [MEF6] or virtual private networks (VPN) on the OSI layers 1, 2, and 3 [KL04, MEF6, TBPOB05]. Their customers in turn are large enterprises, Internet service providers (ISP), competitive operators with only geographically limited presence or even virtual network operators without any transport network. The increasingly complex business relations also motivate the trends towards higher degrees of flexibility, layer independence, and automation in transport networks.

Summarizing this discussion of non-technical characteristics, transport networks constitute huge capital investments by their operators which only amortize over relatively long periods. In addition, quality of service and reliability are primary operational objectives. Therefore, technology evolution in transport networks proceeds slower than in other business segments and more directly depends on an operator's business model and market position[3]. Therefore, the architecture for optical burst transport proposed in Chapter 4 takes an evolutionary approach to introduce burst switching techniques. It builds on an existing optical server layer instead of a commonly assumed greenfield deployment.

2.2 Optical Transport Networks

After this general introduction of the technical and non-technical characteristics of transport networks, this section discusses current and emerging transport network architectures in more

[3]This conservative tendency explains at least partly why several of the expectations of the Internet bubble period did not materialize at all or not as quickly as promised—certainly apart from technological immaturity of several solutions and wrong estimations of the traffic growths.

detail. It first outlines the scope of optical networking with an emphasis on wavelength-switched networks. Then, the vision of IP-over-WDM networking is presented which stimulated many recent developments in the user and control plane. Finally, multi-layer networks and the control of optical transport networks are looked at more closely.

2.2.1 Optical Networking

The term *optical networking* is neither precisely defined nor does it refer to a single, well-defined network technology. Also, it has undergone a continuous change over the past fifteen years as technology evolved and expectations grew and shrinked. Nevertheless, a fundamental part of optical networking is the transmission of data by means of light in free space or over glass fiber (in low-cost, short-reach applications also plastic fiber). The key benefits of optical compared to electrical transmission are the low attenuation and distortion at very high bit-rates, the immunity to electro-magnetic interference, and the lack of electromagnetic radiation. Consequently, optical communication can achieve very long transmission distances, does not suffer from or produce cross-talk across fibers, and is difficult to eavesdrop.

In fiber-based communication, lasers generate signals with wavelengths in the infrared spectrum around 1550 nm, 1310 nm or 850 nm. As the signal attenuation of glass fibers is minimal around 1550 nm, this spectrum is most attractive for long-reach transmission while the lower wavelength spectra are most frequently used in short-reach optics. Currently, transmission bit-rates of 10–40 Gbps across hundreds of kilometers with carrier-grade equipment and of 1–10 Gbps across up to ten kilometers with standard Ethernet equipment are reached in the field.

This already high capacity is increased even more by wavelength division multiplexing (WDM), i.e., the transmission of many wavelength channels in parallel within a single fiber. Depending on the frequency/wavelength spacing, *dense* WDM (DWDM, 12.5–100 GHz spacing around 1550 nm [G.694.1]) and *coarse* WDM (CWDM, 20 nm →3.3 THz spacing over a much broader spectrum [G.694.2]) are distinguished. The smaller number of wavelength channels for CWDM is tolerated because less complex and thus more cost-efficient equipment can be deployed. In addition to the higher transmission capacity of a fiber using WDM, the joint amplification of all wavelength channels with the same fiber amplifier leads to significant CapEx reductions. These fiber amplifiers *pump* the data wavelengths with an additional laser in erbium-doped (EDFA for 1550 nm) [FML$^+$89] or praseodym-doped (PDFA for 1310 nm) [Wan96] fiber segments.

In contrast to optical transmission, optical switching has been introduced later and is still not always present in optical networks. Traditionally, this has been due to the relative infancy of optical compared to electronic switching matrices and the lack of flexible optical memory. However, several optical switching technologies are available (cf. Table 4.1 in [Buc05]) and currently make their way into operational networks. If switching is not performed in optics there are two main categories of optical networks: (i) Optical transmission links can be used as point-to-point links between client layer nodes. (ii) Alternatively, a switch converts the optical signal from optics to electronics (O/E) at the input and from optics to electronics (O/E) at the output. In-between, this O/E/O or *opaque* switch uses an electronic switching matrix without terminating the end-to-end server layer signals. Nevertheless, both cases suffer from the high number and cost of optical receivers and transmitters, which are called transponders if paired.

As an extreme vision for optical networking, all-optical or photonic networks transport and process data entirely in the optical domain. However, combining the advantages of electronics, e.g., processing, with those of optics, e.g., transmission and (partial) transparency, seems more promising [CHH94]. Transparency describes the capability of a network element to both, perform its task independent of certain signal attributes and not to interfer with certain signal attributes. Transparency is attractive as it allows providers to operate different services, technologies, and types of equipment together in the same optical network. [BHJD96, SSJ97] classify transparency with respect to the carrier type (e.g., carrier frequency), the signal type (analog, digital), the modulation format, the maximum supported bit-rate, the line coding, the transmission type (e.g., continuous, burst-mode), and the transmission format (e.g., SDH, Ethernet). Note that for this classification order, each level of transparency also includes all of the following levels. Transparency is often also used in a more general sense. Then, it refers to an optical signal which is transported from transmitter to receiver entirely in the optical domain without O/E/O conversions [RFC3717].

Today, optical networking has also found its place in many network segments. In core networks, wavelength-switched DWDM networks, which are discussed in the next subsection, are primarily deployed. In metro networks, more cost-efficient and robust CWDM ring networks with SDH or metro Ethernet client layers [MEF] are deployed. Finally, FTTx (fiber-to-the-curb, -premises, -building, -desktop, etc.) [Gre04, Shi05] installations rely on passive optical networks (PON) [UOF$^+$01, KP02] or active point-to-point links and are used in access and distribution networks. In the following, only WSN, WDM multi-layer networks, and future, highly dynamic optical networks will be considered though.

2.2.2 Wavelength-Switched Networks

A wavelength-switched network (WSN) represents the fundamental optical network architecture. It transports data in optical connections, which are called lightpaths and several of which form a *lambda grid*. Lightpaths start in an optical transmitter, span wavelength channels on several fibers, and terminate in an optical receiver. In intermediate nodes lightpaths are switched as circuits either statically using an optical patch panel or dynamically using an optical add-drop multiplexer (OADM) or an optical cross-connect (OXC). Thus, *routing and wavelength assignment* are the essential tasks when setting up lightpaths in WSNs [BM96, Spä02]. Dynamic WSN not only support rapid and automated service provisioning and capacity adaptation but also enable more efficient resilience concepts. Protection resources no longer have to be dedicated but can be shared and assigned dynamically when a failure occurs. Also, networks can be dynamically reconfigured/restored upon a failure accounting for the current network state.

Figure 2.3 depicts a WSN with two lightpaths connecting the terminal nodes T1 and T3 as well as T2 and T4 by lightpaths. The lightpaths span several fiber links and are switched in the optical cross-connects OXC1, OXC2, and OXC3. If a lightpath has to use the same wavelength channel on all fiber links, which is called the wavelength continuity constraint, it is called a wavelength path. Wavelength converters [WPW$^+$02], which can translate an optical signal from one wavelength channel to another, can lift this constraint and yield substantial performance improvements [RM98, Spä02]. While the lightpath T2→T4 in Figure 2.3 is wavelength converted in node OXC2, T1→T3 satisfies the continuity constraint forming a wavelength path.

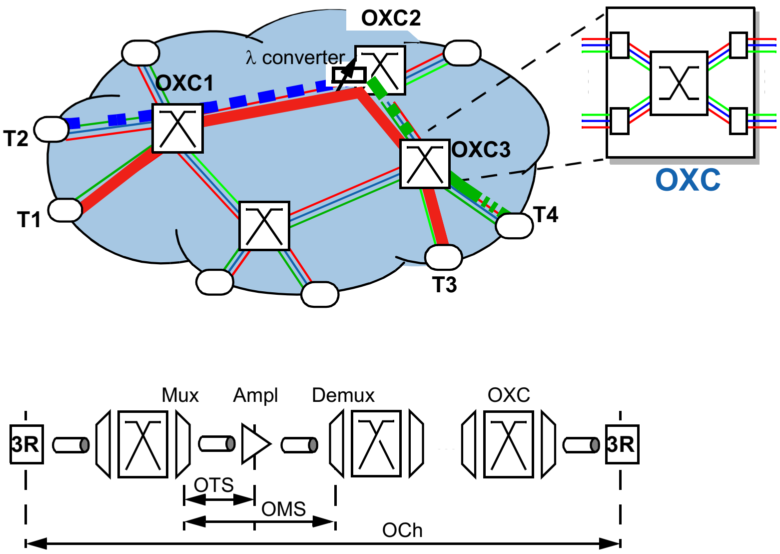


Figure 2.3: Wavelength-switched network and OTN sections

The ITU architecture of a WSN is the optical transport network (OTN) which is defined and modeled as a layered transport network in [G.872] (also cf. [ITG06]). The WSN is decomposed into the optical channel (OCh), the optical multiplex section (OMS), and optical transmission section (OTS) layer networks to account for the different layers and their supervisory overhead. An OCh network connection represents a lightpath as an optical connection between reference points with 3R signal regeneration[4]. It can be member of a group of WDM signals and can span multiple multiplex sections in the physical layer. An OMS transports several optical channels between a WDM multiplexer and a WDM demultiplexer. Finally, the OTS covers the multiplexed channel group across several fibers or amplifier sections. Figure 2.3 illustrates the OTN transport entities OCh, OMS, and OTS in a WSN with transparent optical cross-connects. The optical transport hierarchy (OTH) [G.709] implements the G.872 OCh with a digital framed signal with digital management overhead. Further, it supports the operation and management of OTNs across sub-networks and across domains by complete internal and restricted external interfaces. Despite the digital implementation of G.709, the OTN architecture is more general and also covers implementations with a greater degree of transparency [ITU03] as assumed for the remainder of this thesis.

Wavelength-switched networks allow multi-service integration in WDM by transporting different services in different lightpaths. However, even large customers (enterprises, research centers) are still hardly able to directly connect to and exploit the huge offered bandwidth of lightpaths. Therefore, WSNs are currently mainly applied as server layer for electronic client layer transport networks, which provide sub-wavelength granularity, e.g., SDH/WDM or MPLS/WDM. The task of dimensioning and routing in such networks is defined as grooming [Muk06]. Nevertheless, high-speed customer interfaces and advances in control plane technologies will make lightpath services more common. At the same time, traffic growth expectations and operational requirements motivate the extension of WSN towards coarser granularities, i.e., entire wavebands or fibers, which is called hierarchical optical switching [IGW$^+$02, ZZZM03].

[4] 3R regeneration comprises re-amplification, re-shaping, and re-timing. Despite significant efforts towards optical 3R regeneration, most solutions still comprise O/E and E/O conversion.

2.2.3 IP-over-WDM Networks

The terms IP-over-WDM or IP/WDM integration are synonymous with the vision of a multi-layer network only comprising IP and wavelength-switched layer networks. However, they also name cornerstone activities in the evolution of optical networking basically recognizing two important facts: First, that optics has more to offer than static point-to-point transmission links and thus requires support for dynamic optical networking in the user and control plane. Second, that the traditionally high number of layers stacked in a transport network (e.g., IP/ATM/SDH/WDM) introduce unbearable overhead.

The trend towards dynamic optical networks triggered a huge amount of research and development activities [Dix03], the key resulting architectures of which are discussed in Section 2.3. Also, several of the control plane activities led to new standards (cf. Section 2.2.5) which are currently introduced in carrier networks. In order to leverage the integration with IP, these standards use or extend IP-based protocols for signaling, routing, and resource management.

The trend towards fewer transport network layers initiated detailed considerations on the functionality and effectivity of different overall architectures. While ATM deployments in transport networks are disappearing, SDH has remained in the networks due to its large installed base and new data-centric functionality. It has been extended to map various client layer data services (generic framing procedure, GFP [G.7041]) and to flexibly (virtual concatenation, VCAT [G.707]) and adaptively (link capacity adjustment scheme, LCAS [G.7042]) provision transport capacity. Nevertheless, carrier-grade Ethernet may replace SDH as soon as high-speed transport interfaces as well as traffic engineering and management support become available [GPTV04, GPTB05, Med05, ABMR06].

Unfortunately, IP-over-WDM integration activities also partly over-stimulated the market expectations and thus finally contributed to the *optical* speculation bubble. Therefore, the term IP-over-WDM has the more general meaning of packet-aware WDM transport networks today. IP is no longer assumed to become the only client layer network to be supported. Still, it clearly functions as the convergence network layer and thus remains at the top of many layered transport networks for years to come. However, in addition to IP, network operators market (premium) layer 1 and layer 2 transport services, e.g., FiberChannel for storage area network (SAN) interconnection [TDY05] or Ethernet VPNs [MEF]. Independent of the services evolution, these network layers are built on top of WDM optical networks.

2.2.4 Multi-layer Transport Networks

Multi-layer transport networks are currently the architecture of choice for transport networks. They comprise wavelength-switched and electronic transport layers, e.g., SDH/WDM, MPLS/WDM, etc. As a key advantage, they bridge the bandwidth and cost gap between the packet-switched data world and the circuit-switched transport world. While higher layers offer fine-grain multiplexing but have higher cost per bit, the WDM layer provides a coarse granularity but substantially smaller cost per bit. Optimal multi-layer network designs follow two principal strategies: On the one hand, they apply recursive aggregation and multiplexing to achieve a high utilization. On the other hand, they offload transit traffic to lightpaths, which bypass client layer nodes, and thus decrease the number of client layer switch ports. For instance, in Figure 2.2, the direct virtual link between nodes C1 and C4 offloads traffic to the server layer, which avoids transit traffic in nodes C2 and C3. Section 2.4 quantifies principal trade-offs when applying those two strategies in virtual topology design. In addition to the design related CapEx reductions, multi-layer networks can also improve OpEx by coordinated control and management across layers.

Multi-layer network engineering allocates resources in the server layer to optimally transport static or average client layer traffic demands. This task comprises the design and dimensioning of a virtual topology offered to the client layer (cf. Section 2.4). In contrast, multi-layer traffic engineering, commonly referred to as grooming, routes and assigns client layer connections or traffic flows to this virtual topology. For static traffic demand and network scenarios, multi-layer network and traffic engineering are frequently treated together using optimization or derived heuristics [DR00, Mod01, Cin03, ZZM05, Muk06]. Alternatively, they can be performed sequentially to reduce the design complexity or to account for different layer dynamics. For instance, the WDM layer can be static or only be reconfigured on much higher time-scales than the client layer [SBG04, BSG04, Gal05]. Finally, completely dynamic multi-layer scenarios can again integrate both tasks to react to explicit client layer service requests [Nec02, NGB03, Köh05b, ZZM05] or traffic variability [GG05, YMC$^+$03].

Similar to network and traffic engineering, multi-layer resilience [VPD04, OM05] targets an overall optimum in terms of minimal resource requirements, fast reaction to failures, and high availability. Here, the WDM layer can directly detect and thus quickly react to failures in the physical infrastructure. In contrast, higher client layers experience failures only indirectly or delayed due to their limited, logical topology view, which slows down root cause analysis and reaction. Also, if the WDM layer recovers a coarse transport granularity, the client layer does not have to individually recover a large number of connections or flows. However, recovery on lower layers cannot offer survivability towards client layer failures, e.g., a router interface failure.

Section 2.4 discusses virtual topology design and dimensioning and analyzes the resource requirements of different virtual topologies under dynamic client layer traffic. Client-server hybrid optical networks introduced in Section 2.3.4 constitute multi-layer networks with burst-switched and wavelength-switched technology. The optical burst transport network architecture (OBTN) introduced in Chapter 4 implements such a hybrid optical network for cost-optimal burst transport.

2.2.5 Control of Optical Transport Networks

Section 2.1 and the previous subsections showed the trend towards more flexible and dynamic optical networks in next generation layered transport networks. In order to enable a customer or client network to set-up a transport service on demand, e.g., a lightpath, control planes have to be introduced to overcome the limitations of the existing provisioning schemes using the management plane. Similarly, flexible and fast reaction to failures, e.g., a fiber cut, across layered networks demand for control planes especially if they belong to different administrative domains.

In order to ensure compatibility, the control plane functions and conventions for resource management, routing, signaling, and fault management have to be standardized. ITU-T and IETF both formulated their requirements, key issues and candidate models for the control plane. While the ITU requirements for automatic switched transport networks (ASTN [G.807]) are network technology independent, IETF's IP over optical networks framework [RFC3717] is more centered around the interaction of IP/MPLS and optical networks. Both standard bodies focus on IP-based protocols in the optical control plane—non-IP protocols are not excluded though.

Although layering of completely independent networks offers many benefits through separation of concerns, the lack of any coordination can also lead to conflicts or inefficiencies. Thus, control planes have to interconnect to enable the required coordination while protecting operational independence and business boundaries. The different types of layer interconnection can be classified by three basic models [RFC3717]:

- In the *overlay* model, the layer networks run separate instances of the control plane (cf. Figure 2.4 left) and apply independent overlay routing. Through a user network interface (UNI), they only exchange the information absolutely necessary to set-up and maintain a service but no topology information. Therefore, the overlay model protects internal operational details in business models, in which an operator offers the transport service of a server network to customer client networks [PV05, BDL$^+$01a].
- The *peer* model tightly couples or integrates the control planes (cf. Figure 2.4 right) to share all information and perform integrated routing and unified signaling. This model can be applied in trusted environments such as inside an operator's administrative domain to facilitate traffic engineering and resilience [BDL$^+$01a].
- Finally, the *augmented* model only loosely couples the control planes by separate routing instances and limited exchange of information, e.g., only regarding reachability. In this case, routing would be domain-specific in the sense that a routing protocol is running between the control planes.

While [RFC3717] introduces these models in an unbiased way, the ASTN requirements [G.807] demand that the layer independence of [G.805] be respected and server layer topology cannot be assumed to be available to a client layer.

A further classification criterion of control plane architectures is whether information storage and processing is centralized or distributed. While a centralized approach benefits from consistent information and better justifies complex optimizations, distributed intelligence scales better and avoids single points of failure. Here, ITU-T formulates no preference in [G.807], whereas the IETF favors the distributed approach in [RFC3717].

Based on the ASTN requirements, ITU-T standardized the control plane architecture and functional requirements for the automatically switched optical network (ASON [G.8080]) applicable to SDH and OTN layer networks. The recommendation realizes an overlay model the functional components of which can be implemented either centrally or distributed. Additional recommendations further specify the component architectures and all protocols used. For instance,

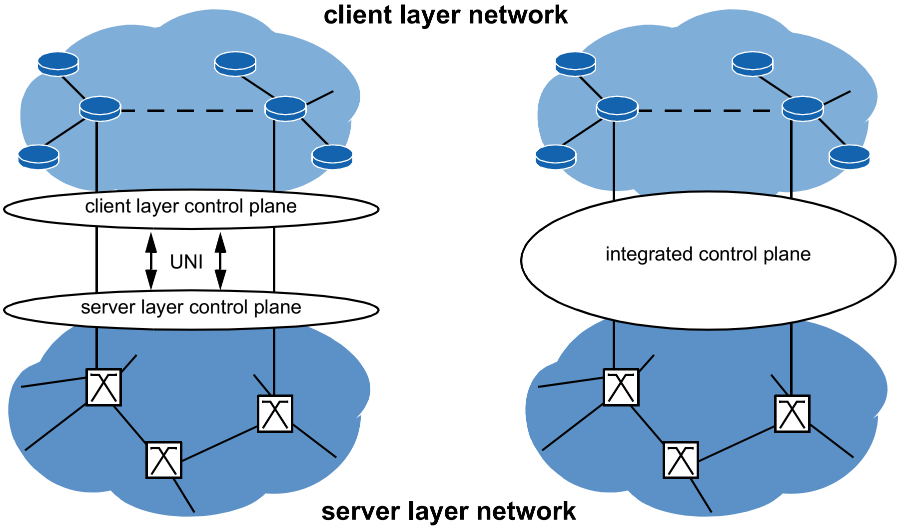


Figure 2.4: Control plane interconnection models: overlay (left) and peer model (right)

in [G.7713.2], ITU-T adopted the UNI 1.0 standardized by the Optical Internet Forum (OIF) [OIF04]. This UNI applies the IP-based signaling protocols CR-LDP [RFC3036, RFC3212] and RSVP-TE [RFC3209, RFC3471, RFC3477].

Generalized multiprotocol label switching (GMPLS) [RFC3945] constitutes the unified control plane architecture of the IETF intended to cover various packet-switched and circuit-switched architectures [BDL$^+$01a, BDL$^+$01b, MTF05]. This is achieved by first defining a hierarchy of switching capabilities for label-switched paths (LSP) ranging from packet-switched, via label-switched, TDM-switched (time-division-multiplexing, i.e., SDH) to wavelength-, waveband-, and fiber-switched[5]. Second, it extends the control protocols for signaling (e.g., RSVP [RFC3209] or LDP [RFC3472]), routing (e.g., OSPF [RFC3630, RFC4203]), and resource/link management (LMP [RFC4204, RFC4209]) accordingly.

In GMPLS, constraint-based routing allows to compute label-switched paths respecting certain topology, traffic engineering, resilience or signal quality requirements. Typical constraints relate to hop count, total delay, underlying protection or in case of optical networks physical signal impairments [RFC4054]. Constraint-based LSPs are signaled by RSVP-TE [RFC3209] or CR-LDP (constraint-based routing LDP) [RFC3212] and are routed, e.g., by the extended OSPF-TE [RFC4203].

Although GMPLS originally targeted a peer model with a distributed implementation, this is only the case for the unified GMPLS approach. In addition, the GMPLS UNI [RFC4208], which offers superior flexibility compared to the OIF UNI [PV05], can realize an overlay model. Also, a path computation element (PCE) [IET06] and the respective protocols for information exchange are being developed to allow for (semi-) centralized path computation supporting more complex constraint-based routing.

[5]GMPLS generalizes the control framework of MPλS, pronounce MPLamdaS, which was proposed as the application of MPLS to optical networks.

While control plane architectures and realizations are not in the center of this thesis, principal control plane interconnection models and functional requirements are discussed for OBTN in Section 4.2.5.

2.3 Architectures for Future Optical Networks

While the previous section discussed current and emerging optical transport network architectures, this section looks farther into the future and prepares the in depth discussion of optical burst switching in Chapter 3. Highly dynamic, switched optical network architectures aim to support the high traffic variability of data services better than wavelength-switched networks by offering finer, sub-wavelength granularities. In the layered transport network hierarchy, these networks act as a server layer to carry IP/MPLS client layer traffic. Mainly because of technological challenges, most of these future IP-over-WDM architectures have a midterm to long-term perspective.

The first subsection introduces key architectural constraints with which highly dynamic optical network architectures have to comply. It classifies fast optical circuit switching as well as optical burst and packet switching and introduces OPS and OBS in greater detail. Finally, architectures for *hybrid* optical switching, i.e., the combination of an optically burst/packet-switched with a wavelength-switched network, are classified and qualitatively discussed.

2.3.1 Architectural Constraints

In order to structure the design space for highly dynamic optical networks, this subsection discusses key architectural constraints and practical design rules. Here, the switched granularity, the resource reservation scheme and the used technology are of particular importance. They are characterized by the transmission time (of the switched granularity), the round-trip time in case of end-to-end signaling, the switching time, and the maximum FDL delay, respectively. Figure 2.5 summarizes typical time ranges for them over a logarithmic time-line ranging from 1 ns to 100 s [San01].

Relevant ranges for transmission times in future optical networks can be obtained by analyzing the switching granularities of circuits and packets. Future dynamic wavelength/circuit-switched networks can be expected to support much higher dynamics with holding (transmission) times as low as seconds. In contrast, the transmission times of typical IP packets ($\approx$ 40–1500 Byte) at 10 Gbps range from few tens of nanoseconds to approx. 1 microsecond. Obviously, they further decrease for higher bit-rates. In order to reduce the requirements towards forwarding and switching performance of core packet switches [Jun06] and to decouple the switched-granularity in the client and server layer, IP packets can be aggregated to form larger transport entities, the so-called bursts. While the aggregation function leads to minimal burst transmission times of a few microseconds, there is no definite maximum. However, reasonable maximum values of several hundreds of milliseconds can be derived from the shortest holding times of dynamic circuits.

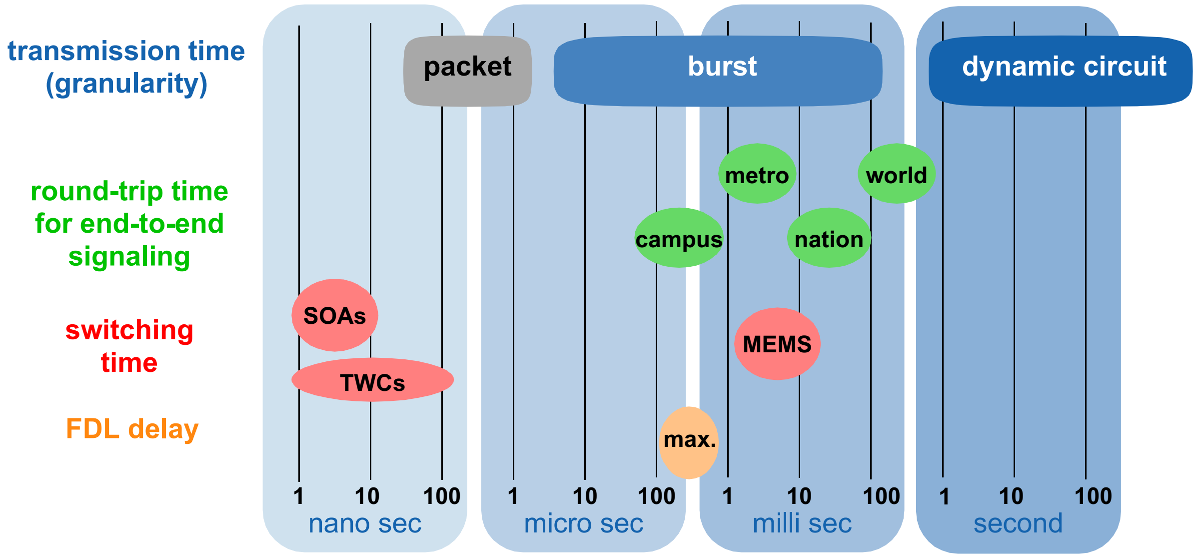


Figure 2.5: Timing constraints in highly dynamic optical networks [San01]

Optical networks, other than electronic networks, cannot apply extensive buffering in network nodes to resolve resource conflicts. Therefore, several architectures seek to avoid conflicts by explicitly reserving network resources prior to data transport, instead. These schemes commonly apply end-to-end signaling, which takes at least one round-trip time, and store the reservations in the respective network elements. Thus, if the round-trip time is large compared to the transmission time, an unbearable amount of state information aggregates in the nodes. Also, if resources are already exclusively reserved from the time the reservation request arrives, a substantial amount of bandwidth is wasted due to the large bandwidth-delay-product (cf. [Dol04] Figure 2.11). In these cases, sending data without advance reservation is advantageous as further discussed in Section 3.2. Figure 2.5 shows approximate round-trip propagation delays for campus, metro, national, and international distances between ingress and egress nodes (20, 200, 2000, and 20,000 km at a speed-of-light in the fiber of 200,000 km/s, delays for processing the reservation requests are not included). From these numbers, it can be concluded that end-to-end resource reservation is only scalable for rather long bursts or dynamic circuits.

The switching technology and thus the switching time determine the minimal guard time between contiguous packets, bursts or circuit connections. Thus, the switching time should be much smaller than the transmission time to achieve a small transmission overhead. Figure 2.5 contains approximate switching times for three representative devices [Buc05]. Micro-electro-mechanical systems (MEMS), which redirect light beams by tilting mirrors, have typical switching times of 10 ms and less. Semiconductor-optical amplifiers (SOA) as on/off switches are much faster and reach switching times of less than 10 ns. Finally, tunable wavelength converters (TWC) can be realized with tuning times well below 1 μs. TWCs affect the guard time as they are used as converters in series with other switching devices or in wavelength switching stages together with static arrayed waveguide gratings (AWG, cf. Section 3.5).

The transmission time also formulates requirements towards contention resolution schemes and the transmission infrastructure. Section 3.4.3 argues that practical fiber delay lines (FDL) can be assumed to have a maximum length of approx. 80 km and thus the maximum delay of approx. 400 μs shown in Figure 2.5. From Section A.1 follows that FDLs require delays of at least 2 mean burst transmission times to effectively resolve contentions. Thus, by combining both arguments it can be concluded that only architectures with transmission times of approx. 200 μs or below can apply FDL buffers. In addition, short transmission times mandate that the fiber amplifiers and receivers are burst-mode capable, i.e., the receivers can synchronize and adapt the decision level to highly variable power levels within a negligible fraction of the burst transmission time. Fiber amplifiers have to be stable in the presence of power fluctuations, e.g., using fast power control as shown for OBS in [FPS02, Buc05]. Similarly, the receivers have to be able to adjust the decision level for the optical signal and to synchronize onto the bitstream within very few leading bytes [EDLW04].

Based on these architectural constraints, only a subset of all combinations of switched granularity, resource reservation, and switching technology comply with practical design rules and are thus practically relevant. Also, due to the wide range of burst transmission times not all architectures for burst transport are able to cover the complete burst granularity spectrum. Therefore, two regimes of the burst transmission time are distinguished in the succeeding list of highly dynamic, switched optical network architectures.

- Optical packet switching (OPS): it transports client layer packets without aggregation leading to minimal transmission times, which require fast switching technologies and prohibit end-to-end resource reservation. FDL buffers can be used for contention resolution.

- Optical burst switching (OBS) with short transmission times: it aggregates traffic to reduce the forwarding and processing effort but requires fast switching technologies and prohibits end-to-end resource reservation. FDL buffers can be used for contention resolution.

- OBS with long transmission times: it aggregates enough traffic to make end-to-end resource reservation feasible in the targeted network context. It is not useful to apply FDLs for contention resolution. If the transmission time is significantly larger than the switching time of MEMS switches, such slower switching technologies can be applied.

- Fast optical circuit switching can be considered as the extrapolation of the current wavelength-switched networks towards shorter set-up and holding times. It can be distinguished from OBS with long transmission times by the fact that it sets-up an end-to-end lightpath and not only reserves resources for a transmission time which in most cases is shorter than the propagation time. Thus, the switching time of a slow optical switching technology, e.g., MEMS, is small enough. Also, traffic aggregation and burst assembly are not considered an integral function here. Instead, the client layer requests dynamic lightpaths as transport circuits. These request can be triggered by machine-to-machine communication, e.g., in grid computing [BBE$^+$05].

As fast optical circuit switching requires fundamentally different architectures than OPS and OBS, only the latter are relevant to this thesis and further characterized and contrasted in the following.

2.3.2 Optical Packet Switching

The success of electronic packet-switched networks based on Ethernet, ATM or IP technology stimulated research on optical packet switches [Kar93, Haa93, GRG+98a, DDC+03, BBR+03]. Research on OPS explored the possibilities and limitations of agile optical switching, optical processing, and contention resolution including wavelength conversion and fiber delay line buffering [OSHT01, YMD00, BBR+03]. As OPS laid the fundamentals for these topics still highly relevant in OBS, many contributions are discussed in the context of OBS in Chapter 3.

Pure OPS definitions demand that the packet header is read optically and controls the switch optically, which is still the focus of long-term technology research. Therefore, more practical definitions allow for an opto-electronic approach. Here, the control header is read and processed electronically to control the node while the data payload is switched transparently in optics. Initial OPS proposals used inline control headers which require to tap and receive all data wavelength channels [GRG+98a]. Alternatively, the control information can be multiplexed more cost-efficiently onto a sub-carrier [BCR+99] or onto the data wavelengths using an orthogonal modulation format [Nor03, VZC+03].

While switches with slotted operation allow for simpler operation and fewer resource conflicts they also mandate synchronization of optical packets to slot-boundaries [Muk06]. Thus, proposals for unslotted also called asynchronous optical packet switches gained importance during the past years. Optical packets can have either fixed or variable length. Depending on the client layer packet format and the optical packet format this can require basic segmentation, aggregation or padding functions. Still, systematic traffic aggregation and assembly is commonly not part of OPS but a defining concept for OBS.

2.3.3 Optical Burst Switching

Optical burst switching has been proposed in the late 1990s as a novel photonic network architecture directed towards efficient transport of IP traffic [WPRT99, QY99, Tur99][6]. While this subsection only introduces the most fundamental concepts and topics, Chapter 3 discusses OBS and its functional components in-depth. Because of the large amount of contributions in the field of OBS and the wide range of burst transmission times, the definitions of OBS vary substantially in literature. Nevertheless, the following core concepts can be regarded as defining for OBS and are thus assumed for the remainder of this thesis. Figure 2.6 illustrates these characteristics in a network scenario.

[6]Already in the 1980s, *burst switching* was proposed for integrated voice and data switching with distributed control [Ams83, Ams89]. It aimed at delivering high bit-rate services (200 kbps at that time) to the end-customer over the twisted-pair copper telephone infrastructure. This application required electronic burst switches close to the subscriber in the field and small enough ($< 1\,\text{ft}^3$) to fit into street cabinets or to be mounted onto poles. Similarities with OBS are basically the aggregation of data and voice bursts (e.g., talkspurts) for transport through the network and the distributed control. Although it has not been deployed as such, the concept of small switching elements close to the subscriber recently materialized: The rollout of very high speed digital subscriber lines (VDSL) mandates that small digital subscriber line access multiplexers (DSLAM) be installed close to the subscriber [NDT02].

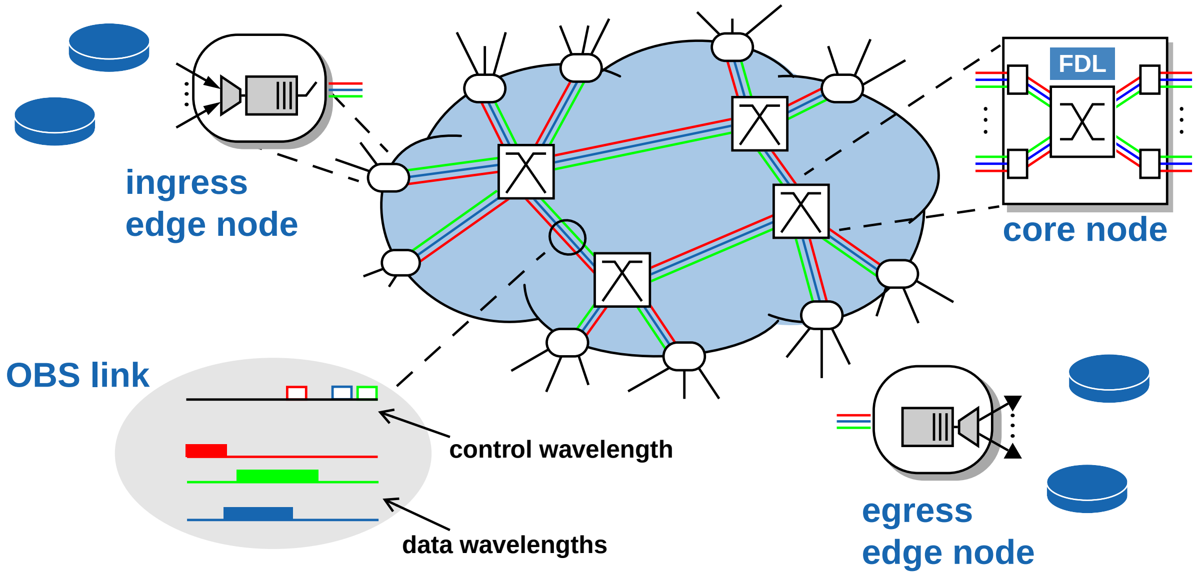


Figure 2.6: Network scenario for optical burst switching

- OBS ingress edge nodes aggregate client layer traffic, e.g., IP packets, and assemble variable length optical bursts. These bursts are transported through the network to the egress OBS edge node, which disassembles the bursts and forwards the client layer traffic, e.g., to an IP router.

- Data bursts are asynchronously and transparently switched in core nodes, i.e., they stay in the optical domain until they reach their egress edge node.

- Control information for routing and resource reservation is signaled in control header packets out-of-band and is processed electronically in core nodes. Burst control headers are transmitted on a separate control wavelength as depicted in Figure 2.6 or are multiplexed onto the data wavelength using an orthogonal modulation scheme [Nor03].

  This separation of control and data allows the core nodes to only terminate the point-to-point control information and to transparently switch the data bursts. The fact that control headers are processed electronically while data bursts are switched in optics is frequently referred to as a hybrid approach. However, the term hybrid is used in a different sense here as introduced in Section 2.3.4.

In the context of this thesis, the previous definitions apply to OBS architectures with longer and shorter transmission times. This is due to the fact that these definitions neither specify resource reservation schemes nor realization technologies.

As a principal qualitative comparison, the bandwidth granularity and the switching complexity of OBS are in-between those of WSN and OPS networks. With respect to WSN, OBS provides more bandwidth flexibility in the optical transport network, i.e., it can better adapt to changes in the traffic. However, it also mandates more agile switching technology and support for burst-mode transmission. Compared to fast optical circuit-switching, OBS aggregates and assembles smaller data transfers inside the OBS layer network, which do not justify a complete circuit set-up and should not or cannot be transported over electronic packet-switched networks. Regarding optical packet switching, OBS requires less complex header processing technology and simpler forwarding control. Because of aggregation, less complex switching technology can be employed as bit-rates further increase. Also, in contrast to slotted OPS architectures, there is no need for synchronization in OBS nodes. Finally, in literature, OBS is frequently distinguished from OPS based on the fact that it does not use fiber delay line (FDL) buffering. However, OBS can greatly benefit even from simple FDL buffers. Thus, FDL buffers are frequently proposed for OBS nodes and an integral part of the discussion on contention resolution schemes for OBS in Section 3.4.

2.3.4 Hybrid Optical Switching

In addition to the individual network architectures outlined so far, several *hybrid* optical network architectures have been proposed. They target an overall improved network design by combining several network technologies as motivated in [HN01]. In [GvBK+06], they are defined to employ at least two of the basic optical network technologies, namely fiber, waveband, wavelength, burst and packet-switching. OBS alone is often called *hybrid*, because it combines electronic control and optical switching. However, it is not hybrid by itself according to the former definition mandating that multiple network technologies be deployed simultaneously.

The following presentation focuses on combinations of optical burst-switched as well as wavelength-switched networks because they are most relevant to this thesis. Also, their combination is particularly instructive because they define entirely different requirements towards optical technology. For instance, OBS requires burst-mode capable receivers and transmission links and mandates faster switching technology than WSN.

Existing proposals for hybrid optical networks are divided into following three classes based on the degree of interaction and integration of the network technologies [GvBK+06]: (i) client-server, (ii) parallel, and (iii) integrated hybrid optical networks. The following paragraphs and figures define these classes in progressing order of integration level and show where current proposals of architectures belong to. In addition, they qualitatively discuss key consequences of these approaches.

Client-server Hybrid Optical Networks

The first class employs a hierarchy of optical layer networks with different network technologies in client layer/server layer relationship. For efficient network design, a hierarchy of bandwidth granularities is established with finer granularities in the higher layers and coarser granularities in the lower layers. Again, the discussion focuses on architectures in which the client layer is an OBS network and the server layer is a WSN. Figure 2.7 shows such a client-server hybrid optical network consisting of an OBS client layer and a WSN server layer, which provides a virtual topology of lightpaths. Optical bursts are only switched in the client layer nodes and transparently flow in lightpaths through the WSN and its circuit-switched OXCs. the As direct lightpaths bypass intermediate OBS nodes, they reduce the amount of transit traffic in the client layer nodes (cf. Section 2.4). Also, they decrease the number of contention situations along a

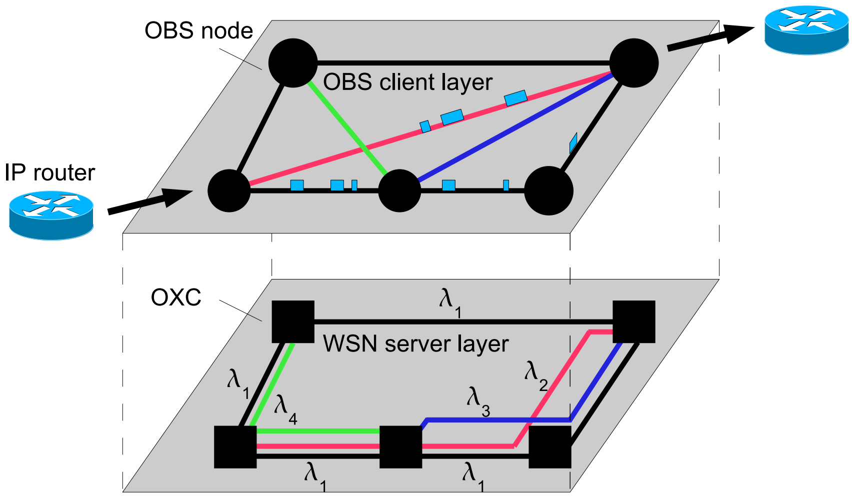


Figure 2.7: Client-server hybrid optical network architecture [GvBK+06]

path from source to destination [CW03][7]. Such virtual topologies support attractive and likely introduction or migration scenarios, in which OBS uses capacity provisioned by an existing WSN. Finally, capacity adaptation and failure recovery schemes for the OBS client layer can benefit from dynamic lightpath setup and tear-down in the server layer.

Client-server hybrid optical networks define several requirements for the WSN and for both control planes. As mentioned above, OBS mandates a burst-mode capable transmission infrastructure and may use orthogonal modulation schemes for control information [Nor03]. Hence, the underlying wavelength-switched transmission infrastructure also has to be burst-mode capable and transparent to the used modulation schemes. If the OBS network and the WSN are operated by different service providers, wavelength services are leased by the OBS operator to set-up the virtual topology links. Then, control planes can chose to implement either an overlay interconnection model with separate state information, an augmented model with partial information sharing or a peer model with complete information sharing (cf. Section 2.2.5).

The migration scenarios outlined in [OSHT01] and [CSBO03] define an extreme type of a client-server hybrid optical network. The client layer OBS nodes basically act as traffic aggregation nodes at the edge of the core network. As these nodes are interconnected in a dense or full-mesh virtual topology of direct lightpaths, the respective architectures are termed *Burst-over-Circuit-Switching* (BoCS) in the following. In BoCS, most or all transit traffic bypasses intermediate nodes, which minimizes the number of OBS switch ports required for transit traffic. However, OBS depends on a high statistical multiplexing gain on network links to achieve a high utilization for a target quality of service (QoS) level. Increasing the connectivity of a network by virtual topology links as in BoCS leads to less traffic per virtual link and thus a reduced multiplexing gain and eventually higher resource requirements (cf. Section 2.4.3). Consequently, a dense virtual topology saves switch ports for transit traffic in intermediate nodes but

[7]From an overall network design point of view, this effect is rather small compared to the effects of transit traffic reductions. If the individual burst loss probabilities in the nodes along the path are small and independent of each other, the end-to-end burst loss probability is approximately proportional to the number of nodes on the path [DG01, ID03]. Thus, the end-to-end burst loss probability only changes marginally when decreasing the path lengths. At the same time, the penalty in burst loss probability due to the lower statistical multiplexing gain on the virtual links is substantial. This will be further discussed in Section 6.1.

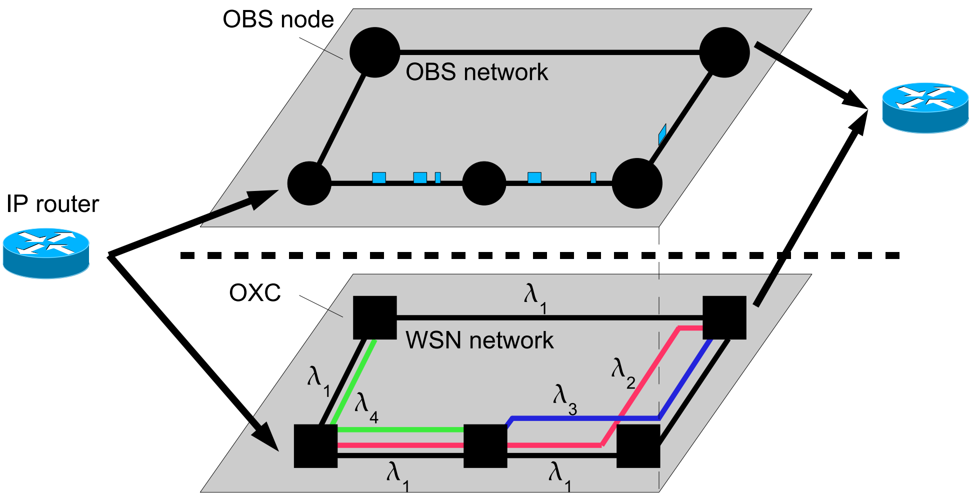


Figure 2.8: Parallel hybrid optical network architecture [GvBK+06]

requires additional resources at end nodes[8]. Therefore, the virtual topology has to be carefully designed to reduce overall network cost. For OBS/WSN client-server hybrid optical networks, Chapter 6 quantitatively evaluates this trade-off of QoS performance and resource requirements.

Parallel Hybrid Optical Networks

In the second class of hybrid optical networks two or more optical server layer networks with different network technologies are accessible in parallel. An intelligent (client layer) edge node selects the respective transport services individually or in combination in order to optimally serve customer service requirements. Virtual optical networks (VON) [QY99] and polymorphic multi-service optical networks (PMON) [dM+04] introduce generic frameworks and describe possible realizations of this class of hybrid networks. They combine WSNs of different dynamics with OBS (either with or without end-to-end resource reservation). [LWZ+03] and [XQYD03] both propose and analyze a parallel hybrid optical network based on OBS without end-to-end resource reservation and on a dynamic WSN. The edge nodes select a network technology based on explicit user request, based on traffic characteristics like bandwidth or expected flow duration or based on QoS requirements. Figure 2.8 illustrates such a parallel hybrid optical network in which an IP edge node can choose from an OBS or WSN transport service. Thus, IP traffic can be either transported in optical bursts or in a continuous byte stream inside a lightpath.

Resources for transmission and switching can be either dedicated to or shared among the different network technologies. The realization presented in Figure 2.8 has dedicated resources for both transmission and switching. Alternatively, resources for transmission could be shared while those for switching would remain separate. Finally, both transmission and switching resources could be completely shared. While sharing of resources improves resource utilization it also mandates that all involved equipment supports the requirements of the most demanding network technology. For instance, an integrated OBS and wavelength-switching node

[8]Note that this argument also applies to an IP client layer with SDH or WDM server layers [BSG04, Gal05]. However, its effect is less pronounced because the IP routers benefit from large and flexible electronic buffers.

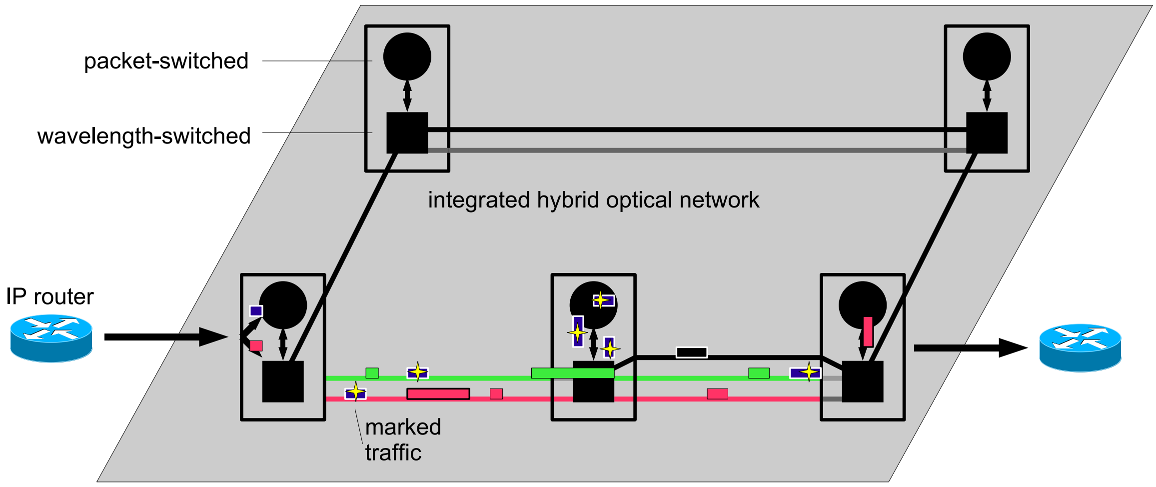


Figure 2.9: Integrated hybrid optical network architecture [GvBK+06]

[LWZ+03, XQYD03, dM+04] employs the faster switching technology and the burst-mode capability of OBS nodes even for the wavelength-switched services. Consequently, the design of such parallel hybrid architectures has to trade-off efficiency and realization complexity. Similar to the arguments on resource sharing, a unified control plane across all parallel optical layer networks facilitates network operation [dM+04]. However, it also requires that all specifics of the different network technologies are considered in its design and implementation.

The discussion of this class implicitly assumed that the selecting edge node is under the control of the service provider. However, the customer could also decide upon the transport service. In CHEETAH [VZL+03], a high capacity customer has two network attachments and chooses between a primary, dynamic SDH and a secondary TCP/IP-based transport service.

Integrated Hybrid Optical Networks

Integrated hybrid optical networks extend the parallel approach by collapsing the two networks and sharing all resources at all times. Figure 2.9 illustrates this approach for nodes comprising a wavelength-switched and a packet-switched part. In this case, an intermediate node has access to a bypassing wavelength channel without terminating it. Thus, if a bypassing lightpath is underutilized, an intermediate node can inject additional traffic destined to a downstream node into that lightpath. Such traffic is marked by an outband label to be in a special burst/packet mode and can be detected and extracted downstream based on this marker. For instance, the center node in Figure 2.9 extracts and injects marked traffic from and to the bypassing green and red lightpaths.

From a pure resource point of view, this approach can be assumed to be optimal. However, it is also the most complex, both from a control and a technology point of view. As explained above for parallel hybrid optical networks, the realization of integrated nodes is substantially more difficult compared to non-integrated nodes. Thus, only OpMiGua [Bjo04] and overspill routing in optical networks (ORION) [vBCC+03] currently propose integrated hybrid optical networks. In both architectures, every node is able to detect the current mode of operation of traffic in all lightpaths and to insert and extract traffic without disturbing existing wavelength-switched traffic. Avoiding collisions is extremely complex so that both approaches rely on advanced optical functionality and components. While OpMiGua encodes the marker by an orthogonal polarization state, ORION uses an orthogonal modulation format, e.g., frequency shift keying (FSK) [VZC+03].

2.4 Virtual Topology Design and Dimensioning in WDM Networks

This section discusses virtual topology design and dimensioning as fundamental tasks in network engineering for multi-layer networks. It surveys fundamental work on virtual topology design, particularly heuristics which are adapted and evaluated for OBTN in Section 6.6. Also, it introduces a formalization of the virtual and physical design process based on a flow analysis and matrix transformations. Then, two limiting virtual topologies are analyzed to quantify key trade-offs regarding traffic offloading.

In this thesis, virtual topology design and dimensioning are only used to quantify principal trade-offs and dimension example virtual topologies for the performance and resource evaluation of OBTN in Chapter 6. Therefore, design and dimensioning apply a flow analysis and are not approached by optimization.

2.4.1 Virtual Topology Design

Virtual topology design comprises the sub-problems of deciding which client layer nodes to connect with a virtual link and of dimensioning the respective virtual links. The dimensioning step usually mandates that both the traffic demands and the routing on the virtual links are known. For small networks, solutions for the virtual topology optimization problem can be found even under additional constraints like wavelength continuity in the server layer [RS96, RR00, BM00, Muk06]. However, for larger networks the problem is often decomposed into its sub-problems and approached with heuristics as, e.g., proposed in [RS96, RR00].

Traffic demands and the lengths of virtual links in the physical topology are fundamental criteria in the sub-problem of deciding which client layer nodes to connect. Regarding traffic demands, the Heuristic Logical topology Design Algorithm (HLDA) [RS96] assigns virtual links starting from the largest end-to-end demands $a_{i,j}$ until all available resources are used, i.e., this approach integrates the dimensioning step. All other traffic demands are then routed onto the set-up virtual links. In order to isolate the selection step and control the absolute size of traffic demands that qualify for virtual links, only traffic demands $a_{i,j}$ between two nodes i and j exceeding a threshold A_{thr} could be selected instead (cf. the demand-based virtual topology design in Section 6.6).

With respect to the route lengths in the physical topology, the Traffic Independent Logical topology Design Algorithm (TILDA) [RS96] selects virtual links starting from the shortest physical hop distance $H_{i,j}$ again until all available resources are used. The combined heuristic in [BM00] selects a virtual links by considering the end-to-end traffic demand $a_{i,j}$ and the respective physical hop distance $H_{i,j}$ together. Similarly, the Minimum-delay Logical topology Design Algorithm (MLDA) [RS96] first selects the virtual links, which connect physical neighbor nodes, and then applies HLDA. In contrast, virtual topology design minimizing transit traffic should not only select virtual links with a small physical hop distance but preferably those with a very high hop distance (cf. the path-based virtual topology design in Section 6.6).

Regarding the dimensioning of virtual links, most heuristics treat traffic demands as flows and assume that the required capacity of a virtual link depends linearly on the respective offered traffic. However, this neglects the effects of statistical multiplexing exploited in packet-switched or burst-switched client layer networks—both electronic and optical. Based on mathematical models which relate a QoS objective to a link dimensioning, virtual links could be dimensioned for a specific QoS level. This is, e.g., shown for SDH/WDM networks in [KG03] using the Erlang loss formula and for IP-WDM and IP-SDH-WDM networks in [Gal05] using effective bandwidth schemes.

2.4.2 Virtual and Physical Topology Dimensioning

In order to dimension and quantitatively compare network architectures with different virtual topologies this section describes a formalization of the network dimensioning process. Based on this flow analysis formulation, the required resources of the client layer virtual topology and the server layer physical topology are derived. In order to explain the key trade-offs of sparse and dense virtual topologies, two limiting cases with a sparse and dense virtual topology are analyzed and the governing effect of economy of scales, commonly also referred to as statistical multiplexing gain, is discussed.

2.4.2.1 Formulation of the Network Dimensioning Process

The network dimensioning process applies a flow analysis as shown in Figure 2.10 to translate traffic demand values into client layer and server layer resources respecting QoS objectives. Network dimensioning can be formulated as a sequence of matrix transformations which represent routing and resource computation. Quadratic matrices of order *n*, the number of nodes, describe traffic demands, network topology, and link resources as input to and output of these transformations [Gal05, Köh05a].

The traffic matrix, $\mathbf{A} = (a_{i,j})$, contains unidirectional traffic demands in terms of data rates (e.g., in Gbps) or offered traffic in the pseudo-unit Erlang (Erl). The total traffic demand offered to the network is given by $A = ||\mathbf{A}||$, where the norm $||\cdot||$ denotes the sum of all matrix elements. As link load and resource matrices cannot only be the output of a transformation but also the input to a consecutive step, they are treated like traffic matrices in a unified way and are included in the term *demand*.

Network topologies, either virtual or physical, are represented by adjacency matrices, $\mathbf{T}$, as

$$\mathbf{T}: \qquad t_{i,j} = \begin{cases} 1 & \text{if link } i \rightarrow j \text{ exists} \\ 0 & \text{if link } i \rightarrow j \text{ does not exist.} \end{cases} \tag{2.1}$$

For scenarios with uniform traffic demands or full-mesh link topology, the matrix **E** is defined as a short-hand notation for a matrix with all elements being one except the zero values on the diagonal.

$$\mathbf{E}: \qquad e_{i,j} = \begin{cases} 1 & \text{for } i \neq j \\ 0 & \text{for } i = j. \end{cases} \tag{2.2}$$

The result of routing the demands in the network is described by the transformation, $\phi(\mathbf{A},\mathbf{T})$ according to the applied routing scheme, e.g., shortest path routing. It takes the demand matrix, **A**, and the topology matrix, **T**, as input, and returns a matrix with demand loads on the respective links of the topology as result. In case of a full-mesh topology and shortest path routing, ϕ represents the identity transformation, i.e., $\phi(\mathbf{A},\mathbf{E}) = \mathbf{A}$. Also, for fixed routing schemes, the transformation is linear with respect to the traffic matrix, i.e.,

$$\phi(\mathbf{A}_1+\mathbf{A}_2,\mathbf{T}) = \phi(\mathbf{A}_1,\mathbf{T})+\phi(\mathbf{A}_2,\mathbf{T}) \tag{2.3}$$

$$\phi(\alpha\mathbf{A},\mathbf{T}) = \alpha\phi(\mathbf{A},\mathbf{T}), \alpha \geq 0. \tag{2.4}$$

Resource computation is defined by the transformation $\gamma(\mathbf{A})$ to map a demand matrix, **A**, into a resource matrix. Again, note that demand loads on links can be interpreted as traffic demands. This mapping does not only depend on the QoS requirements to meet but also on the traffic characteristics and link model used in the dimensioning process. The resource computation is greatly simplified by assuming independence of links as well as the homogeneous case, in which the traffic statistics of all demands and the QoS constraints on all links are assumed to be identical. In this case, $\gamma(\cdot)$ applies the same dimensioning function $c(\cdot)$ to each element of the demand matrix to compute the resources $c(a_{i,j})$ necessary to carry the demand $a_{i,j}$.

Models and formulas frequently used for dimensioning are the Erlang-B formula [ID03, Küh04b] for multi-server loss systems under Poisson traffic or the effective capacity formulas according to Kelly [Kel96], Guerin [GAN91], and Norros [Nor95] for single-server delay systems under different traffic characteristics. These effective bandwidth formulas are compared in [BC00] and applied to IP-over-WDM network dimensioning and architecture comparison in [KG03, GKZM05, Gal05].

The mean hop distance of shortest paths is a useful metric in the analysis and design of physical and virtual network topologies. However, the term *mean hop distance* is ambiguous. For this discussion, following definitions for mean hops distance apply:

$$d_{\mathrm{G}}(\mathbf{T}) = \frac{||\phi(\mathbf{E},\mathbf{T})||}{||\mathbf{E}||} \quad \text{for the topology graph} \tag{2.5}$$

$$d_{\mathrm{A}}(\mathbf{A},\mathbf{T}) = \frac{||\phi(\mathbf{A},\mathbf{T})||}{||\mathbf{A}||} \quad \text{for routed traffic demands} \tag{2.6}$$

$$d_{\mathrm{D}}(\mathbf{A},\mathbf{T}) = \frac{||\phi(\gamma(\mathbf{A}),\mathbf{T})||}{||\gamma(\mathbf{A})||} \quad \text{for individually dimensioned traffic demands} \tag{2.7}$$

For the case of uniform traffic demands, all definitions for mean hop distance correspond to each. For a full-mesh topology, the mean hop distance is obviously always one. In the reference core network used in Chapter 6 [BGH+04], the physical topology graph has a mean hop distance $d_{\mathrm{G}}(\mathbf{T}) = 2.64$, while $d_{\mathrm{A}}(\mathbf{A},\mathbf{T}) = 2.42$ and $d_{\mathrm{D}}(\mathbf{A},\mathbf{T}) = 2.49$ for the traffic demands of 2006. The observation that in realistic network scenarios d_{A} is slightly less than d_{G} is also reported in

[BSG04]. The fact that $d_\mathrm{A} \approx d_\mathrm{D}$ will be exploited for the virtual topology dimensioning in Chapter 5.

The capacity of transport circuits (e.g., STM connections or lightpaths) as well as the granularity of network equipment (e.g., WDM transmission systems or fibers) commonly require discretization steps for demands and resources in the dimensioning process. Beyond rounding all values up or down, advanced algorithms could be applied which globally minimize the difference of discretized and original results [BBB$^+$03, HHN]. In the process outlined above, this can be also modeled by a transformation on the respective matrices. However, in order to avoid side-effects of discretization, the following discussion does not consider this step at all.

This formalization will be used in the following subsections to characterize and analyze principal virtual topology designs as well as to dimension burst transport networks in Chapter 5.

2.4.2.2 Economy of scales

The economy of scales describes the effect that under dynamic traffic a demand αa requires less resources than α times the resources required for the demand a, i.e., $c(\alpha a) < \alpha \cdot c(a)$. This also implies that fewer resources are required for the sum of traffic demands a_1 and a_2 compared to the case with individual dimensioning, i.e., $c(a_1 + a_2) < c(a_1) + c(a_2)$.

In terms of traffic matrices and the dimensioning transformation, this can be expressed by the inequalities

$$||\gamma(\alpha \mathbf{A})|| < \alpha ||\gamma(\mathbf{A})|| \quad \text{for } \alpha > 1 \tag{2.8}$$

$$||\gamma(\mathbf{A}_1 + \mathbf{A}_2)|| < ||\gamma(\mathbf{A}_1)|| + ||\gamma(\mathbf{A}_2)||. \tag{2.9}$$

Therefore, in network design, less resources are needed if traffic is first routed through the network, aggregated on links and then mapped to resources than if the (relatively smaller) end-to-end traffic demands were first mapped to resources individually, then routed through the network and summed up per link:

$$||\gamma(\phi(\mathbf{A}, \mathbf{T}))|| < ||\phi(\gamma(\mathbf{A}), \mathbf{T})||. \tag{2.10}$$

2.4.2.3 Capacity Analysis of Client and Server Layer Resources

Based on this formulation, the client and the server layer resources can be dimensioned according to the process in Figure 2.10. It derives principal results for given virtual and physical topologies, and traffic demands. This analysis extends the study of different IP/SDH/WDM network architectures under static client layer traffic demands in [BSG04] by incorporating the effect of traffic dynamics.

In the client layer, demands are first routed on the virtual topology, $\mathbf{T}_\mathrm{V}$, and then mapped to resource requirements of the virtual links. Therefore, the virtual link resource matrix, $\mathbf{C}$, and the total capacity of all virtual links, C, are defined as

$$\mathbf{C} = \gamma(\phi(\mathbf{A}, \mathbf{T}_\mathrm{V})) \tag{2.11}$$

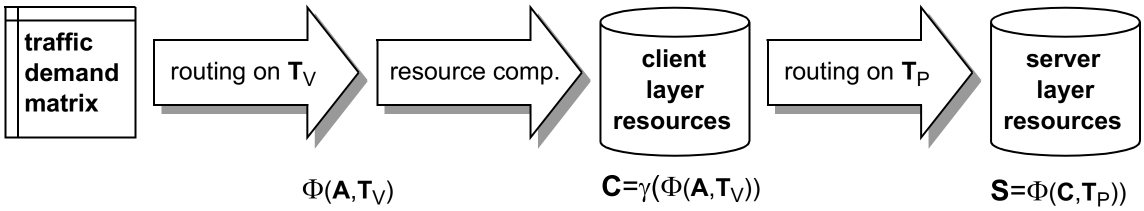


Figure 2.10: Dimensioning of client and server layer resources

$$C = ||\mathbf{C}|| = ||\gamma(\phi(\mathbf{A},\mathbf{T}_\mathrm{V}))|| \tag{2.12}$$

By expanding Eq. (2.12) and substituting d_A a more descriptive representation is obtained

$$\begin{aligned} C &= \frac{||\gamma(\phi(\mathbf{A},\mathbf{T}_\mathrm{V}))|| \cdot ||\phi(\mathbf{A},\mathbf{T}_\mathrm{V})|| \cdot ||\mathbf{A}||}{||\phi(\mathbf{A},\mathbf{T}_\mathrm{V})|| \cdot ||\mathbf{A}||} \\ &= \frac{||\gamma(\phi(\mathbf{A},\mathbf{T}_\mathrm{V}))||}{||\phi(\mathbf{A},\mathbf{T}_\mathrm{V})||} \cdot \frac{||\phi(\mathbf{A},\mathbf{T}_\mathrm{V})||}{||\mathbf{A}||} \cdot ||\mathbf{A}|| \\ &= \alpha \cdot d_\mathrm{A}(\mathbf{A},\mathbf{T}_\mathrm{V}) \cdot A \end{aligned} \tag{2.13}$$

Here, the first factor denotes the required overprovisioning of the virtual topology capacity for dynamic traffic

$$\alpha = \frac{||\gamma(\phi(\mathbf{A},\mathbf{T}_\mathrm{V}))||}{||\phi(\mathbf{A},\mathbf{T}_\mathrm{V})||}. \tag{2.14}$$

Consequently, under dynamic client layer traffic, a virtual topology with optimized resource requirements can be obtained by minimizing the product of overprovisioning factor and mean hop distance in Eq. (2.13). In contrast, under static traffic, minimizing the mean hop distance alone is sufficient as $\alpha = 1$ [BSG04].

The dimensioning of the server layer is obtained by routing the resource demands of the client layer virtual topology onto the server layer physical topology, $\mathbf{T}_\mathrm{P}$. It is assumed that the same routing scheme is applied in both layers. Thus, the server layer dimensioning and its total capacity are expressed by **S** and S respectively

$$\mathbf{S} = \phi(\mathbf{C},\mathbf{T}_\mathrm{P}) \tag{2.15}$$

$$S = ||\mathbf{S}|| = ||\phi(\mathbf{C},\mathbf{T}_\mathrm{P})||. \tag{2.16}$$

For brevity, only static virtual topologies are discussed here. However, dynamic adaptation of virtual link capacities could be considered by extending Eq. (2.15) with an appropriate dimensioning function for mapping physical layer link loads to physical layer resources [Gal05].

Again, Eq. (2.16) can be rewritten by suitable expansions and substitutions as well as by plugging in Eq. (2.13)

$$\begin{aligned} S &= \frac{||\phi(\mathbf{C},\mathbf{T}_\mathrm{P})|| \cdot ||\mathbf{C}||}{||\mathbf{C}||} \\ &= d_\mathrm{A}(\mathbf{C},\mathbf{T}_\mathrm{P}) \cdot ||\mathbf{C}|| \\ &= d_\mathrm{A}(\mathbf{C},\mathbf{T}_\mathrm{P}) \cdot \alpha \cdot d_\mathrm{A}(\mathbf{A},\mathbf{T}_\mathrm{V}) \cdot A. \end{aligned} \tag{2.17}$$

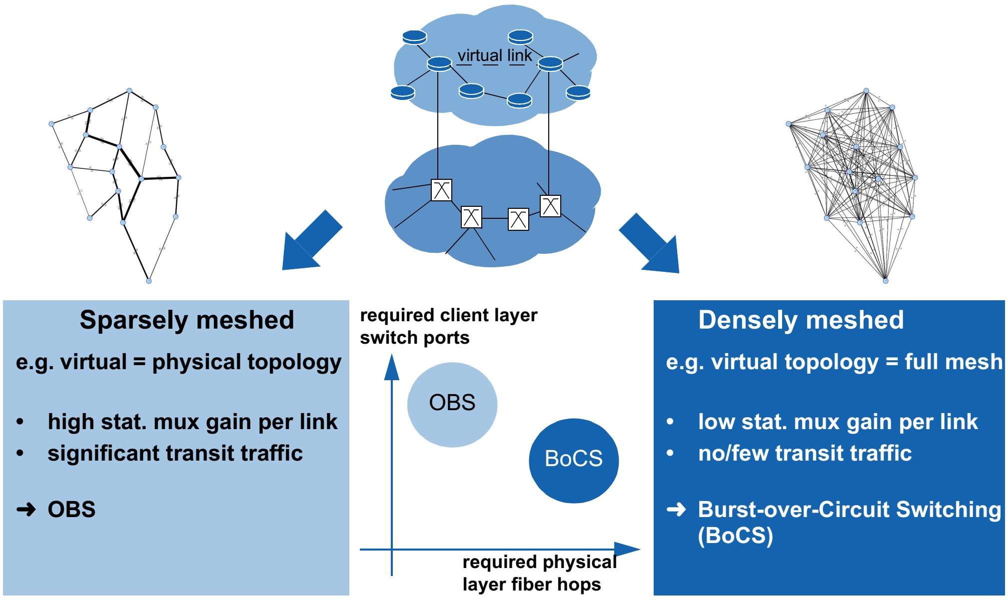


Figure 2.11: Qualitative comparison of sparsely and densely-meshed virtual topologies

Here, $d_A(\mathbf{C}, \mathbf{T_P})$ represents the fact that a virtual topology link in general spans several hops in the underlying physical topology. In realistic network scenarios, a low mean hop distance in the virtual topology leads to a higher mean hop distance in the server layer and vice versa as shown in the following subsection. In WSNs, virtual topology links are provisioned as lightpaths spanning several links in the physical fiber topology. Thus, the server layer total capacity, S, is measured in number of fiber hops for the remainder of this thesis. In Figure 2.3, a fiber hop corresponds to an OMS section in OTN [G.872].

Summarizing, under dynamic client layer traffic, an efficient architecture regarding the server layer is determined by a small product of client layer overprovisioning factor as well as of the mean hop distances of client and server layer. Again, for static traffic only the latter two factors were relevant.

2.4.3 Capacity Comparison for Two Limiting Virtual Topologies

Virtual topology design with the primary objective of minimizing transit traffic in the client layer to reduce its resource requirements is commonly termed *traffic offloading* and in the context of WDM realized by *optical bypassing* (cf. Section 2.2.4). In order to identify and quantify the conditions, in which such traffic offloading successfully reduces client layer resources, this subsection compares the two limiting virtual topologies depicted in Figure 2.11. The case of a sparse virtual topology which is identical to the physical topology does not consider any traffic offloading. In contrast, a dense/full-mesh virtual topology represents the case with complete traffic offloading [SBG04]. Regarding OBS and client-server hybrid optical networks, the sparse virtual topology corresponds to a pure OBS approach while the dense/full-mesh approach represents BoCS. From these two architectures, conclusions can be also derived for the design and expected performance of *intermediate* approaches.

In the first case (subscript p), the virtual and the physical topology are identical, i.e., $\mathbf{T}_\mathrm{V} = \mathbf{T}_\mathrm{P} = \mathbf{P}$ and all virtual links span exactly 1 physical link with $d_\mathrm{A}(\mathbf{C}, \mathbf{T}_\mathrm{P}) = 1$, leading to

$$C_\mathrm{P} = \alpha_\mathrm{P} \cdot d_\mathrm{A}(\mathbf{A}, \mathbf{P}) \cdot A \quad (2.18)$$
$$S_\mathrm{P} = \alpha_\mathrm{P} \cdot d_\mathrm{A}(\mathbf{A}, \mathbf{P}) \cdot A. \quad (2.19)$$

In the second case (subscript fm), the full-mesh virtual topology with $\mathbf{T}_\mathrm{V} = \mathbf{E}$ and thus $d_\mathrm{A}(\mathbf{A}, \mathbf{E}) = 1$ yields

$$C_\mathrm{FM} = \alpha_\mathrm{FM} \cdot A \quad (2.20)$$
$$S_\mathrm{FM} = d_\mathrm{A}(\mathbf{C}, \mathbf{P}) \cdot \alpha_\mathrm{FM} \cdot A. \quad (2.21)$$

Because of the economy of scales, the resource utilization in the first case, which aggregates more traffic on fewer virtual links compared to the full-mesh case, is always higher. Therefore, the overprovisioning factor α_P is always lower leading to:

$$\alpha_\mathrm{P} < \alpha_\mathrm{FM}. \quad (2.22)$$

Consequently, under dynamic traffic, a full-mesh virtual topology is not always attractive to reduce the client layer resources. From the comparison of client layer resources C_P and C_FM in Eq. (2.18) and Eq. (2.20) follows that the full-mesh virtual topology only requires less resources if

$$\frac{\alpha_\mathrm{FM}}{\alpha_\mathrm{P}} < d_\mathrm{A}(\mathbf{A}, \mathbf{P}), \quad (2.23)$$

i.e., the penalty in the overprovisioning factor has to be smaller than the gain due to the reduced mean hop distance.

From a network design point of view, Eq. (2.23) can be satisfied in the following situations:

- The overprovisioning factors α_FM and α_P have comparable values. This is the case with highly aggregate traffic for which statistical multiplexing gain is high and thus both factors are relatively close to one (cf. Section 6.7). Also, this can be achieved by highly efficient architectures which allow to transport even small traffic demands with a high statistical multiplexing gain.

- The mean hop distance is relatively large which is the case in typical core network topologies with many nodes and moderate connectivity. Here, the mean hop distance reaches values of 2.5 to 3.5 as can be seen from [MCL$^+$03, BGH$^+$04].

[BC00, Gal05] both provide quantitative substantiation for the first situation. Also, [Gal05] shows that in an IP core network, the impact of the reduced statistical multiplexing gain due to a full-mesh virtual topology is only minor due to large electronic buffers and the highly aggregated traffic streams. Hence, significant savings have been shown in several electronic network scenarios while they have not been systematically quantified for optically (burst) switched networks yet. However, it can be also concluded from Eq. (2.23) that the maximum gain that can be realized by such an offloading approach is $d_\mathrm{A}(\mathbf{A}, \mathbf{P})$ for the case of $\alpha_\mathrm{FM}/\alpha_\mathrm{P} \to 1$.

So far, the focus was on the client layer resources, which are commonly assumed to be more expensive and thus should be minimized. Completing this discussion, the comparison of the server layer resources S_{P} and S_{FM} in Eq. (2.19) and Eq. (2.21) exhibits a penalty for the full-mesh scenario as $\alpha_{\mathrm{P}} < \alpha_{\mathrm{FM}}$ always applies while both mean hop distances usually have comparable values and approximately cancel out. Thus, the overprovisioning factors determine the resource requirements in the server layer. This again motivates the design of efficient network architectures in which α_{P} and α_{FM} are comparably small.

These qualitative arguments on the design and dimensioning of virtual topologies founded the basis for the design of the OBTN architecture introduced in Chapter 4. Chapter 5 introduces a virtual topology dimensioning approach for OBTN based on the formulation above. Then, Chapter 6 quantifies the resource requirements of OBS, BoCS, and OBTN applying this dimensioning.

3 Optical Burst Switching Architectures

This chapter introduces the fundamental concepts and building blocks of optical burst switching (OBS) in greater detail. It reviews and classifies the scientific literature published to date, which were essential for the design of the optical burst transport network architecture (OBTN) introduced in Chapter 4.

The sections in this chapter are organized following the path an optical burst takes through the network as shown in Figure 3.1 from burst assembly in the edge nodes, via resource reservation for bursts to scheduling and contention resolution in core nodes. It presents node architectures for OBS and their integrated evaluation from a performance and a technology point of view. Finally, existing proposals for QoS conclude this chapter.

3.1 Aggregation and Assembly in OBS Edge Nodes

As introduced in Section 2.3.3, traffic aggregation and burst assembly constitute defining concepts for OBS. They are performed in the OBS edge nodes at the interface between the electronic and the optical domain. The following subsections cover burst assembly parameters and algorithms as well as their impact on traffic characteristics and QoS. They also relate and discuss the sometimes contradicting contributions in this field.

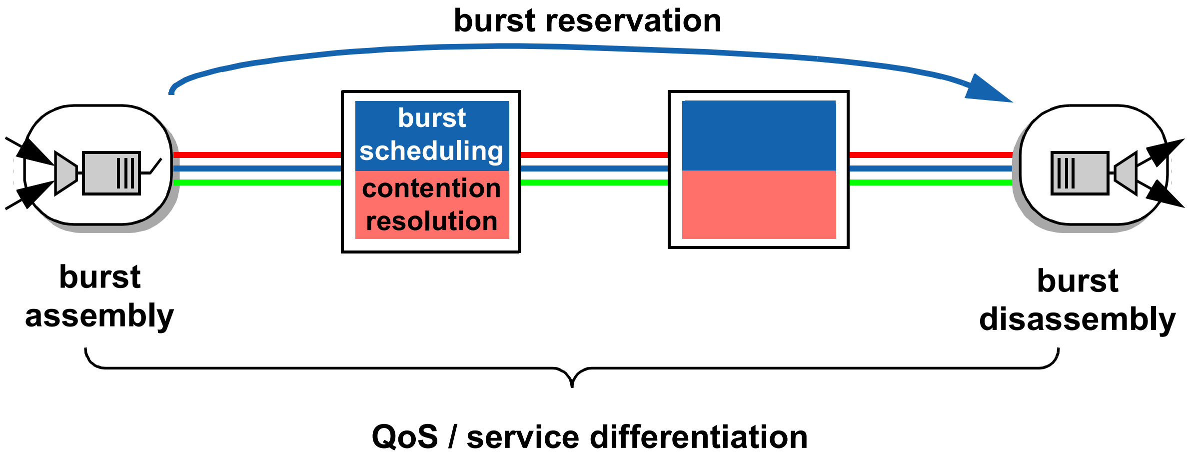


Figure 3.1: Building blocks of OBS networks along an end-to-end path

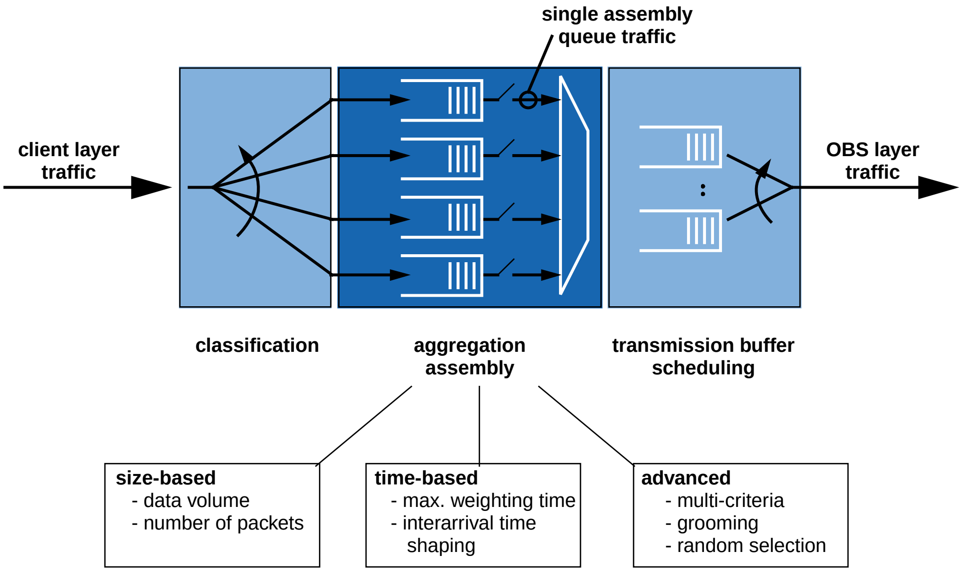


Figure 3.2: Components and strategies for burst assembly in an OBS edge node

3.1.1 OBS Edge Nodes

In an ingress edge node as depicted in Figure 3.2, incoming traffic is first classified, e.g., based on destination, QoS or MPLS forward equivalence class (FEC). Then it is aggregated in separate queues until the burst assembly algorithm decides to form a new burst. At that time, also the burst control header packet with destination address or path identifier as well as service class and burst length is generated. Both are moved to a transmission buffer and scheduling stage to be sent out into the network.

The main benefit of burst assembly is that it aggregates packets into larger containers so that they can be transported in the optical network with less guard band overhead. Aggregation also reduces the requirements towards forwarding and switching performance of core switches [Jun06] and thus decouples the switching granularity client and server layer networks. It ensures that higher bit-rates in the core do not necessarily require shorter switching times but can be compensated by aggregating more traffic into one burst instead. Finally, the burstification process encapsulates traffic for transparent switching and thus allows to transport different client network technologies over the same OBS network. In Figure 3.2, this means that the classification and aggregation process is client layer specific while the optical bursts are not.

The benefits of aggregation and assembly come at the cost of additional functionality in the ingress (assembly) and egress (disassembly) OBS edge nodes. As aggregation, assembly, and scheduling is performed in the electronic domain, their realization is less challenging regarding technology than the realization of OBS core nodes. Architectures and realizations of OBS edge nodes, particularly for traffic aggregation, burst assembly, and burst transmitter modules, are presented in [NKSO04b, NKSO04a, Bra05, Koe05].

3.1.2 Burst Assembly Schemes

Assembly algorithms limiting the waiting time of packets by a threshold on the burstification time are classified as *time-based* and lend themselves ideally to control the edge node delay of real-time traffic [CCK05, IST05]. In contrast, assembly schemes ensuring a minimum and/or maximum burst size by respective thresholds are classified as *size-based* [YCQ02b, Lae02]. They use the amount of data [VZJC02, XVC00] or the number of packets [OHK02] in the respective assembly queue. They are normally motivated by realization constraints, e.g., a minimum optical transmission unit [GCT00] or a fixed container size (e.g., G.709 frames [G.709] for Frame Switching [IST04b]).

Advanced algorithms combine both time and size criteria [GCT00, VCR00] or adapt the assembly parameters to the traffic dynamics or to the network state in order to reach multiple objectives regarding QoS, realization, and operation [Dol04, OHK02, CLCQ02]. Also, padding [YCQ02b, SEL02] or burst grooming concepts [FZJ05], i.e., joining traffic of different assembly queues, are introduced to obtain a minimum burst size in a low traffic load situation.

Finally, the assembly scheme in [dVRG04] applies a random selection scheme to decide with a constant probability upon each packet arrival whether the burst should be assembled or not. This scheme is devised to produce a Poisson burst arrival process by random sampling [Küh04b] under the assumption of a Poisson IP packet arrival process.

3.1.3 Characteristics of Assembled Burst Traffic

The parameterization of the burst assembly algorithm greatly affects the characteristics of burst traffic, which again affect network performance, e.g., as shown in [DG01] for a multi-class OBS system using offset-based QoS (Section 3.6). The input traffic of the client layer together with burst assembly algorithms completely specify the stochastic process of the burst size and the burstification time. However, queuing and scheduling of assembled bursts before burst transmission further shape the burst arrival process to the network [Car04]. For instance, in [CVR02] the burst transmission time is controlled to shape the burst arrival process and obtain a Poisson process.

Work on traffic characterization and modeling of burst assembly can be classified into work on single assembly queues and on entire assembly modules with multiple queues. Regarding the single assembly queue case, [dVRG04] presents comprehensive formulas for the burst size distribution, burstification time, number of packets per burst as well as the respective moment generating functions. It studies time-based, size-based, and random selection assembly schemes under a Poisson packet arrival process and general packet size distribution. Earlier, [Lae02] analytically derived expressions for the burst size and interarrival time for time-based, size-based and combined assembly schemes for a Bernoulli packet arrival process. [YCQ02a] and [YCQ02b] model combined time and size-based assembly in light and heavy load scenarios under short and long-range dependent traffic. Finally, [YCQ02a, KOA05] analyze the burst size distribution and Hurst parameter under self-similar traffic by simulations.

Size-based and time-based assembly algorithms directly control either the burst size or the burstification time. Thus, they approximately lead to a constant distribution with respect to the threshold criterion [Lae02]. The remaining distribution, either for the burstification time or the burst size, follows from the sum of packet interarrival times or packet sizes, respectively. Thus, for a large number of packets per burst, the variable burstification times and packet sizes are modeled by a Normal distribution following the central limit theorem [dVRG04, Lae02, IA02, YCQ02b]. In addition, [Lae02] shows that a Gamma distribution also provides a good fit. These results are shown to be rather robust with respect to the scaling behavior for long-range dependent traffic and mainly affect how fast the distributions converge to the Normal distribution [YCQ02b]. In case of combined time and size-based assembly, the distributions, which were obtained for the individual schemes, are weighted with the probabilities of using the respective assembly criteria [Lae02].

Using burst assembly to shape and actively smooth bursty and self-similar Internet traffic [CB97, WTSW97] attracted great interest. [GCT00] originally claimed that burst assembly can reduce the degree of self-similarity. However, this thesis was disproved by analysis and simulation for both synthetic traffic and traffic traces in several independent papers [IA02, HDG03, AIMM03, YLC+04]. Burst assembly can smooth traffic on short time-scales critical to the burst layer (typically milliseconds) and below and thus improve performance compared to unaggregated Internet packet traffic [IA02, YCQ02a, YXM+02]. Still, it cannot reduce the self-similarity of network traffic, which is commonly measured on higher time-scales (typically seconds and above). Only the self-similarity of the stochastic process describing the burst control packets can be reduced by time-based assembly [HDG03]. But this does not influence the data channels in OBS, which are more relevant for QoS [HMQ+05].

OBS edge nodes always employ more more than one assembly queue. Thus, the traffic characteristics for entire assembly modules with several queues are more important for network modeling. In [IA02], purely time-based assembly is analyzed under self-similar IP traffic modeled by Fractional Brownian Motion (FBM) [Nor95]. The burst size is shown to follow a Normal distribution with the variance depending on the Hurst parameter of the FBM. The burst interarrival time is approximated by a Poisson distribution with sufficiently good quality from few tens of assembly queues on. Also, simulations show that the loss probability obtained from Poisson arrivals (with a general service time distribution) serves as an upper bound for FBM traffic with pure time-based assembly. In [HMQ+05], similar results are presented for an M/Pareto self-similar IP traffic model [NZA99] using simulation.

Based on these contributions, the performance evaluations in Chapter 6 and Appendix A all assume burst arrivals following a Poisson process. Section A.1 shows for an OBS node with a fiber delay line buffer that Normal and Gamma distributions for the burst transmission time, parameterized according to the aggregation of IP packet traces, yield the same burst loss probability as a negative exponential distribution. Thus, the negative exponential distribution is used for simulation studies in Chapter 6.

3.1.4 Impact of Burst Assembly on TCP

If bursts and the transported IP packets are lost due to contention, there are no mechanisms inside the OBS network to recover or retransmit the data. Therefore, higher layers have to ensure the integrity of the end-to-end data transport. In today's Internet, the transmission control protocol (TCP) is predominantly used for this purpose [Com06]. The TCP congestion control mechanism can yield performance degradation in networks with non-negligible IP packet loss. Therefore, possible effects of OBS networks on TCP performance are widely studied. Here, traffic classification and aggregation as well as burst assembly play an important role. While early work reported partly positive, e.g., [CLCQ02, YXM$^+$02, GSSC03], partly negative [GSSC03] impact on TCP goodput [1], recent contributions isolated the key effects more clearly.

In [DL05], TCP flows are classified with respect to burst assembly algorithms. A flow is called *fast*, if its bit-rate is high enough so that all TCP segments of an entire TCP congestion window can be sent within the assembly time of one burst; it is called *slow*, if they in average contribute at most a single TCP segment to a burst. If the OBS network loses the burst of a slow TCP flow, following TCP segments of the same flow in other successfully transmitted bursts can reveal the loss and trigger the retransmission of this segment. This behavior is identical to an isolated packet loss in an IP network. In contrast, if the network loses the burst of a fast TCP flow and thus usually a complete TCP window, a retransmission timeout may be needed (with a minimum size of 1 s [RFC2988]). Such a timeout would automatically reset the window to its initial small slow-start value. However, instead of a penalty, this clustered loss of TCP segments yields a so-called *correlation benefit*. It improves TCP goodput for burst loss probabilities in the range of 10^{-4} to 10^{-2} [DL05]. In the context of alternative routing, [SPG05] also clearly shows the strong influence of the number of TCP segments of a TCP flow per burst on TCP performance. For high propagation delay variations, the aggregation of several TCP segments into a burst reduces reordering of TCP segments and improves goodput.

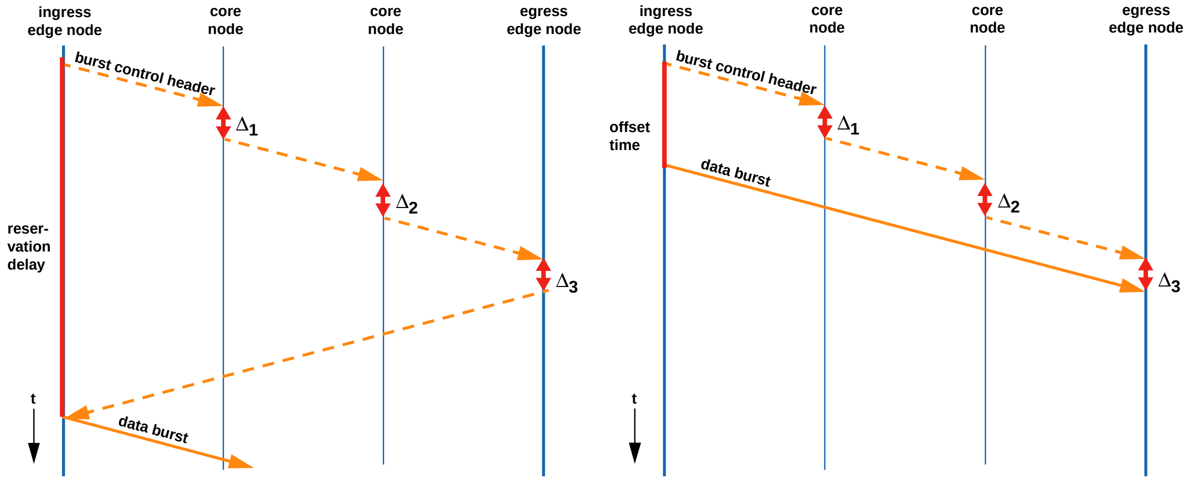
In practical OBS transport network scenarios, the bandwidth bottleneck usually exists in the access network and not in the core network. Also, many TCP flows share an assembly queue such that slow TCP flows, for which no correlation effect exists, constitute the more relevant case. Therefore, Chapter 6 uses the burst loss probability as the primary performance metric and does not study TCP performance in detail. For the low target burst loss probabilities of $P_\mathrm{L} = 10^{-4}$ and 10^{-5} it can be assumed that OBS does not interfere with TCP.

3.2 Resource Reservation in OBS

OBS reserves resources per burst by signaling control information outband in a so-called burst control packet, e.g., on a separate control wavelength. The term burst reservation has an end-to-end scope, while burst scheduling refers to the selection and management of resources inside the OBS core nodes (cf. Section 3.3). As resources are reserved per burst, signaling and reservation functionality has to be provided for both a connectionless and for a connection-oriented communication concept.

For high-speed networks, two fast resource reservation techniques are proposed [Tur92, Wid94, Wid95, KW95], namely *tell-and-wait* (TAW) and *tell-and-go* (TAG). In ATM, these reservation techniques were realized by the *Fast Reservation Protocol* (FRP) with delayed (DT) or immediate transmission (IT) [Hui88, OON88, BT92, Bri98], respectively. [HM95] then introduced these resource reservation schemes into optical networks.

[1] Goodput defines the net throughput of user data across a TCP connection.



(a) Tell-and-wait (b) Tell-and-go

Figure 3.3: Tell-and-wait and tell-and-go burst reservation schemes

TAW in Figure 3.3(a) implements classical end-to-end reservation with acknowledgments requiring an additional pre-transmission delay of at least one round-trip time. Resource management in TAW can be either performed centrally as in wavelength-routed OBS (WR-OBS) [DKKB00, DB02] or in a distributed way using a setup protocol as presented in [WPRT99, BRPS02]. As bursts are only sent out upon reception of a positive acknowledgment, burst loss inside the network can be avoided at the cost of a potentially longer pre-transmission delay. For wide-area transport networks with a diameter of hundreds or thousands of kilometers, propagation delay reaches tens of milliseconds (cf. Figure 2.5). This leads to scalability limitations due to significant amounts of reservation information stored in core nodes. Also, in case switches in intermediate nodes are already set during this reservation phase, the amount of bandwidth wasted therein is much higher than the bandwidth actually needed for burst transmission because of the large bandwidth delay product.

This observation motivated the tell-and-go schemes [Wid94, HM95], which trade-off the lossless property of the end-to-end reservation by a shorter pre-transmission time and potentially higher bandwidth efficiency. In TAG, burst transmission is not delayed until the edge node receives an acknowledgment of successful end-to-end reservation. Instead, transmission is initiated immediately when or shortly after the burst was assembled and the control information was sent out as shown in Figure 3.3(b). In the context of OBS, TAG reservation is often called *one-pass reservation*. It differs from the original TAG [Wid94] in that it does not include a retransmission functionality. Retransmission per link is infeasible as it violates transparency and retransmission from OBS ingress to egress is prohibitive due to the high bandwidth delay product.

If the transmission of the data burst in TAG is delayed with respect to the control packet, this delay is referred to as *offset time*. Here, the key idea is to pre-compensate the processing delays Δ_i, which control packets experience in intermediate nodes. This offset is reduced in all intermediate nodes along the path. In this case, information on the number of nodes on the path, i.e.,

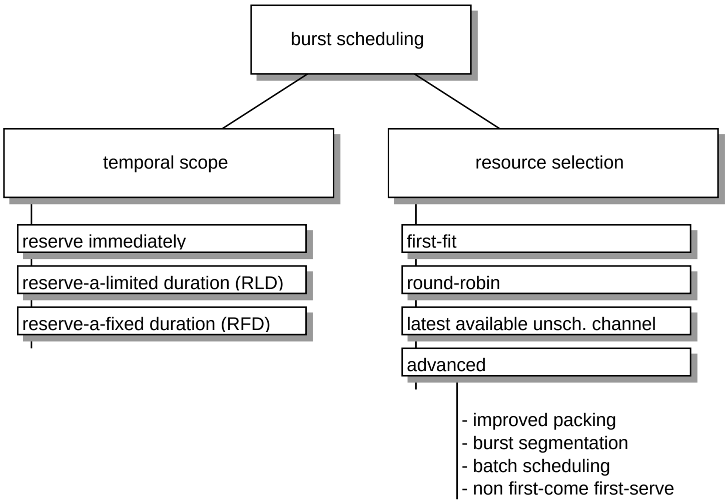


Figure 3.4: Classification scheme and key realizations for burst scheduling

the total expected processing time, is needed to determine the offset time. In contrast to this *basic* offset time for delay compensation, an additional QoS offset for service differentiation may be used (cf. Section 3.6).

Depending on round-trip time, switching time, and burst transmission time either TAW or TAG are favorable [DL01, QY99, DGSB01, BRPS02]. For WAN scenarios with large round-trip times and short burst transmission times (cf. the discussion in Section 2.3.1), TAG is favorable and thus used in the following. However, in a pure metropolitan area network (MAN) scenario with much shorter distances, TAW presents the better solution [Z^+03] though.

3.3 Burst Scheduling

During the end-to-end resource reservation process, bursts have to be scheduled in the OBS core nodes for transmission on the network links or for internal buffering. Strategies for burst scheduling can be classified based on (i) the time at which and the duration for which resources are scheduled for a burst [Gau00, DGSB01] and (ii) the algorithms for resource selection. This classification and key realizations are shown in Figure 3.4.

3.3.1 The Temporal Scope of Burst Scheduling

The most basic scheme occupies a resource immediately from the time the control packet requests the resource until the time at which it is explicitly released, either by a second control packet or by an in-band terminator. Thus, a burst also occupies otherwise idle resources for the offset time between control packet and data burst can lead to inefficient resource usage. Still, this scheme is beneficial if the burst transmission time is either large compared to the offset time or not known at the time of resource scheduling. Also, implementation is most simple as the only information that has to be kept record of in core nodes is whether a wavelength is currently available or not [San02]. The Just-in-time (JIT) [WPRT99] or JumpStart [BRPS02] burst

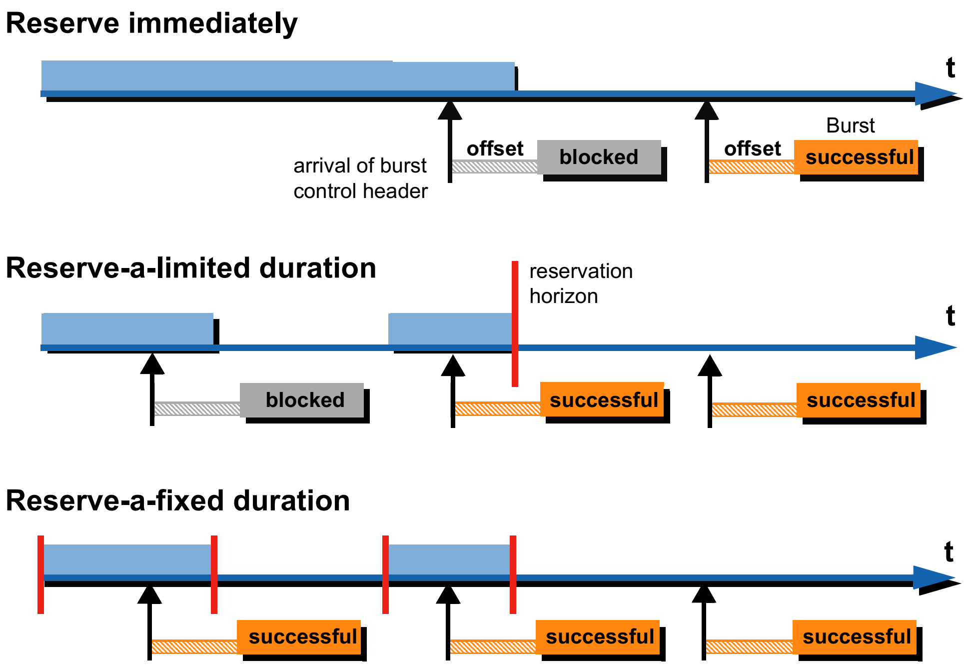


Figure 3.5: Principal burst scheduling schemes

reservation schemes employ this class of burst schedulers as the burst transmission time is large compared to the offset time and usually not known in their research network scenario.

As an improvement regarding efficiency, *Reserve-a-limited duration* (RLD) schemes only schedule resources until the expected end of burst transmission. This allows to schedule resources for bursts arriving in the future ahead of time and increase utilization by overlapping the offset time with the burst transmission time of a previous burst. Regarding realization, this scheme is more complex, as it requires the sending edge node to exactly signal the start and end time of burst transmission. Also, the core node has to record the time at which each resource will become idle again, the so-called reservation horizon. Horizon [Tur99], its descendant Assured Horizon [Dol04], and the Latest Available Unscheduled Channel (LAUC) algorithm without void filling outlined in [XVC00] are the most prominent RLD schemes.

Finally, *Reserve-a-fixed duration* (RFD) schemes consider the exact start and end time of bursts for resource scheduling. With this extension, bursts can be scheduled in voids between already reserved bursts in order to reduce resource overhead and fragmentation. Just-enough-time (JET) [YQ97] and the LAUC algorithm with void filling (VF) in [XVC00] are the original RFD schemes. The improvement in resource efficiency of RFD compared to RLD comes at the cost of an increased realization complexity. Similar to RLD, control information for RFD schemes has to comprise the exact start and end time of burst transmission. The scheduling module in the core node has to track the start and end times of all scheduled bursts. In order to leverage any performance advantage, it also has to account for and reuse voids in the scheduling process.

For the three fundamental burst scheduling schemes immediate reservation, reserve-a-limited duration, and reserve-a-fixed duration, Figure 3.5 shows successful and failed burst reservation attempts. The wavelength channels show the occupancy and reservation status at the time the control packet arrives

In literature, these principal burst scheduling schemes have been extensively analyzed regarding performance and realization complexity [Gau00, DGSB01, TR05]. The following paragraphs summarize and relate key findings on burst scheduling.

In case no offset time between control header and data burst is used at all, the three basic scheduling schemes have the same performance [DGSB01, DG01, BS05b]. RFD schemes only yield superior performance over RLD schemes in the presence of voids which are generated by scheduling requests with large offset time differences. For instance, scenarios with coarse-grain offset reduction in core nodes or with additional QoS offset [YQ00, Gau00, DG01] (cf. Section 3.6) exhibit such large variations in offset times. Also, OBS nodes, which reserve an output wavelength channel and a fiber delay line buffer at the same time (Section 3.4.3.4), produce voids and thus benefit from RFD [XVC00, CP02, COS03]. In any case, voids can only be reused if bursts with a smaller offset time exist and have short enough burst transmission time to fit into the voids [DG01].

In [San02, ZXC02, JG03a], first realizations of RFD schedulers in fast programmable logic are presented. In an integrated analysis, the latter contribution shows the QoS performance as well as design scalability of the burst scheduler. [Jun05] presents the novel burst scheduling strategy Pre-estimate burst switching (PEBS) which achieves a much lower realization complexity at only marginal QoS performance degradation. Software-based burst scheduling solutions can only be applied to burst transmission times of milliseconds or above [XQLX03, MRZ04, XQLX04].

3.3.2 Resource Selection in Burst Scheduling

Apart from the timing information used for burst scheduling, the search and selection of available resources defines the second important degree of freedom. Basic strategies search and schedule available resources according to *first-fit* (FF) or *round-robin* discipline. More complex strategies select the resource, which only becomes available just before the start of the burst to be scheduled. They are commonly referred to as *Latest available unscheduled channel* (LAUC) [XVC00] and have a higher computational complexity [LQC04].

Apart from principal burst scheduling schemes, a large number of proposals has been published which either optimize the scheduling process in time or the resource selection strategy. The most important approaches are surveyed and classified in [LQC04]. In the following, only the most important classes of approaches are discussed.

Strategies minimizing resource fragmentation by an optimal packing of bursts on resources beyond LAUC are extensively studied, e.g., in [XQLX03, ISNS02], although most of them only yield limited improvements. An alternative concept is burst segmentation, in which parts of bursts are preempted in order to achieve an overall improved QoS performance [VJ02a, DEL02, NRV$^+$03]. Here, it is assumed that the intact parts of incomplete bursts can still be used in the receiving edge node.

In order to minimize processing delay and to limit complexity in the burst scheduling module of OBS core nodes, most approaches process requests strictly on a first-come first-serve basis. However, some proposals queue control packets in core nodes for a short time either to schedule bursts in batch mode [CEBCS03, CBD$^+$03, KA05] or to serve the request queue in a different order, e.g., to account for the actual data burst arrival order [CP02, TMC03, BS05b] or class of service [YZV01, LL02]. Theoretical performance bounds and a complexity analysis of this class of approaches are presented in [LQXX04].

A comprehensive discussion of work in the field of burst scheduling, the design and comparison of different realization architectures, as well as the PEBS strategy [Jun05] and its analysis are presented in the dissertation [Jun06]. In order to benefit from the RFD capabilities without the realization complexity of the LAUC selection strategy, all performance evaluations in this thesis use RFD first-fit. In literature, this combination is also known as *just-enough-time* (JET) [YQ97].

3.4 Contention Resolution

This section first introduces principal contention resolution schemes in optical networks. Then, it presents details and the state-of-the-art in literature on wavelength conversion, fiber delay line buffering, and alternative/deflection routing. Finally, it discusses combined contention resolution schemes and their degrees of freedom.

3.4.1 Principal Contention Resolution Schemes

Contention resolution comprises all mechanisms to deal with resource conflicts among two or more bursts, e.g., if they attempt to reserve the same wavelength channel for an overlapping time-interval [YMYD00, Gau02b, Dix03, Gau04]. Because of one-pass reservation and statistical multiplexing in the optical domain efficient contention resolution in OBS core nodes is essential in order to achieve high QoS. In case a contention cannot be resolved, all contending bursts but one have to be discarded. Without loss of generality, the following discussion concentrates on the contention between two bursts but all mechanisms also apply to contentions among more than two bursts.

In principle, contention situations in OBS can be resolved in one or several of the following three physical domains:

- **Wavelength domain:** If the resource under contention is operated in WDM a wavelength converter can shift one of the contending bursts to another wavelength channel. Thus, wavelength conversion removes the *wavelength continuity constraint* which otherwise requires bursts to use the same wavelength throughout an all-optical network.
- **Time domain:** With buffering one of the contending bursts can be delayed until a resource becomes available. Buffering can be either performed in the electronic or in the optical domain. Electronic buffers offer a higher flexibility [BHS02, COS03] and can overcome contention situations which last longer than the delay achievable with optical buffers [OSHT01, LP04]. However, electronic buffers also require O/E and E/O conversions and thus disrupt the otherwise transparent burst transport. Here, optical FDL buffers are advantageous despite their smaller flexibility.

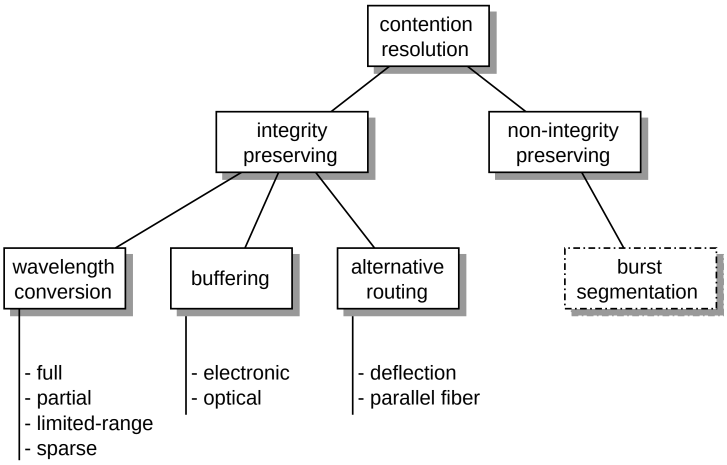


Figure 3.6: Principal contention resolution schemes and options

In contrast to electronic random access memory, fiber delay lines (FDL) only offer a fixed delay and data bursts leave the FDL in the same order in which they entered (sequential buffer). However, several bursts can share the same FDL using WDM. Recently, concepts for slowing down light have been devised, which vary dispersion characteristic of semiconductor media [CHKKC03, TKCH05]. These reports stimulated research on all-optical routers with tunable delay [YY05, MMW^{+}05]. However, [CHKKC03] reports on possible delays of up to 8.7 ns at 10 Gbps based on theoretical models. For practical OBS scenarios this is still orders of magnitude too small (cf. Section 3.4.3).

- **Space domain:** In alternative routing, often referred to as deflection routing, one of the contending bursts is sent via a different output and thus on a different route toward the destination node. The idea here is to use the entire network as a shared resource for contention resolution. In multi-fiber networks, in which neighboring nodes are connected by several parallel fibers, the space domain can also be exploited by transmitting contending bursts on a different fiber of the same output.

All former contention resolution schemes preserve the integrity of bursts, i.e., their either successfully transmit or drop bursts. In contrast, *burst segmentation* tries to resolve contention by only dropping the actually contending parts of a burst [DEL02, VJ02a, NRV^{+}03]. This approach, which is also called *composite burst switching*, has already been shortly mentioned in Section 3.3 to improve packing in burst scheduling. It can also implement a QoS differentiation scheme as outlined in Section 3.6 by deciding which of the contending bursts to segment or partly preempt. However, burst segmentation not only needs to recover segmented bursts at the disassembly node. It also requires additional functionality in the core nodes to control and signal the segmentation process. Therefore, burst segmentation will not be considered for the remainder of this thesis.

Figure 3.6 shows the classification of principal contention resolution schemes and advanced options. The following subsections discuss these three physical contention resolution domains as well as their combinations in greater architectural and operational detail. In addition, they review the state-of-the art as well as key modeling and performance evaluation work.

The architecture of optical nodes is more closely interrelated with the switching technology and component functionality than the architecture of electronic nodes. For instance, several architectures use tunable wavelength converters to perform part of the switching functionality. Consequently, not all mechanisms for contention resolution and particularly their extensions are applicable to all node architectures and trade-offs regarding performance and realization complexity cannot always be generalized. For instance, the *tune-and-select* (TAS) node architecture [Buc05] inherently relies on full wavelength conversion (cf. Section 3.5). Using a TAS node with a shared wavelength converter pool leads to substantial scalability problems [Mue03] which is not be the case in other architectures. In order to still be able to obtain fundamental results for the performance of contention resolution strategies in OBS nodes, abstract node models are used defining key dimensioning parameters and functionality. Then, based on results for abstract node models, the integrated evaluation of technology and performance provides important insight for specific node architectures [BGPS03, GBPS03, GBP05].

3.4.2 Wavelength Conversion

From a switching point of view, *full* wavelength conversion enables a node to connect any wavelength channel of any input to any wavelength channel of any output. In addition, from a resource sharing point of view, wavelength converters transform a number of dedicated wavelength channels to a group of shared channels. Under dynamic traffic, the resulting performance improvement can be explained by the statistical multiplexing gain of a multi-server system compared to a single-server system with the same offered traffic per server. As a further benefit not related to contention resolution, wavelength converters can regenerate, i.e., re-amplify, reshape or even re-time, the optical signal [WPW$^+$02] and thus enable longer transmission distances without an O/E/O network element.

Currently, wavelength converters still constitute an expensive component as they either mandate highly advanced technology, if realized all-optically, or require a receiver/transmitter pair, if realized electro-optically, [WCA$^+$00, EM00, LFB$^+$05]. Depending on the switch architecture and the wavelength conversion strategy fixed or tunable wavelength converters (TWC) have to be employed. While widely tunable wavelength converters provide the highest flexibility, wavelength converters with a fixed output wavelength trade-off the technological complexity and cost of TWCs by reduced flexibility [Yoo96, WPW$^+$02].

In order to reduce the cost of OBS nodes with tunable converters, two concepts are commonly applied, which are also known from wavelength-switched networks [RM98, Spä02]: As not all wavelength converters are simultaneously used at all times, *partial* conversion installs wavelength converters in pools, which are then shared per output fiber or node. For slotted operation with [DMJD97, DHS98, DJMS98] and without FDL buffers [LL95, EL00], analytical performance models of partial conversion are available. However, the unslotted operation can only be treated in the bufferless case [MNSW02a, MNSW02b, Küh04a, AK04, MZA05]. In contrast to partial conversion, *limited-range* conversion employs wavelength converters with a smaller spectral conversion range [YLES96, ZVZ$^+$04, ELS05, RZVZ06] and thus lower cost. Finally, *sparse* wavelength conversion installs converters only in a subset of nodes in the network [SAS96].

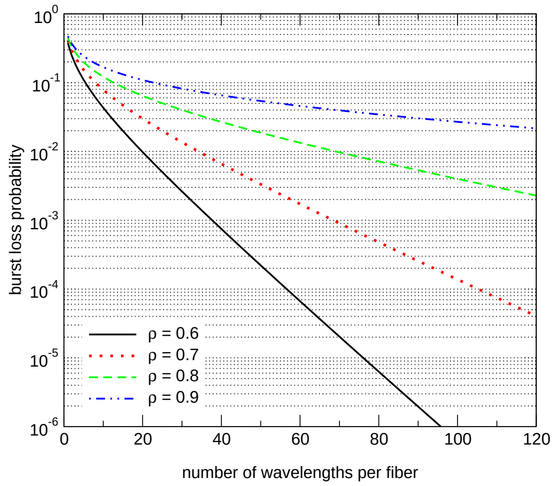


Figure 3.7: Burst loss probability vs. number of wavelengths for a bufferless OBS node with full wavelength conversion (M/G/n loss model and different offered load values ρ)

In the context of wavelength routed networks, wavelength conversion has proven its effectiveness in reducing lightpath blocking probabilities, increasing network availability, and minimizing the additional resources needed to provide network survivability. In these networks, wavelength converter usage can be effectively controlled as lightpaths are set up end-to-end. This allows for globally optimized routing and resource selection. Compared to the case of full conversion, often as little as $10-20\,\%$ of the converters are required to achieve the same QoS level [RM98, SB98, Spä00, Spä02]. In OBS, this cannot be achieved due to the one-pass reservation [Gau04]. Also, sparse wavelength conversion is not effective, as OBS lacks the required end-to-end view of the path. Despite these conceptual differences, many performance models and results for optical crossconnect nodes still apply to OBS because they were derived for isolated nodes or did not model the effects of end-to-end reservation [CCB$^+$00, MNSW02a, MNSW02b, Küh04a].

In many scenarios wavelength conversion alone cannot resolve contention. The Erlang B formula for the M/G/n model, which is valid for WDM resources with Poisson arrivals, can be used to illustrate this. Figure 3.7 shows that a large number of wavelengths is required to achieve a low burst loss probability at a high utilization. For instance, approx. 100 wavelengths per output fiber are needed for a burst loss probability of 10^{-6} at a relative offered load per wavelength channel of 0.6. To overcome this limitation, further contention resolution schemes should be applied in addition to wavelength conversion.

3.4.3 Fiber Delay Line Buffers

This section introduces the main characteristics of single FDLs as well as FDL buffers and classifies them regarding their operation and architecture. The state-of-the art on contention resolution with FDL buffers is presented by classifying the scientific literature in Table 3.1 and Table 3.2 comprehensively. Finally, the combined scheduling of FDL and output fiber resources is looked at. Closely related to this section, Appendix A presents performance results for the FDL buffers used in the evaluation of the OBTN architecture.

3.4.3.1 Fiber Delay Lines

A fiber delay line simply consists of a fiber segment. It delays a data burst according to the propagation time of the signal in the fiber medium as defined by its length and the speed of light in the fiber (200,000 km/s). Thus, data bursts can only be stored in an FDL for a fixed, pre-determined duration which is limited by physical constraints as discussed below. Like any optical fiber, FDLs can accommodate several parallel WDM channels, which increases their buffer capacity effectively. Consequently, a single FDL is characterized by its delay T_F and the number of WDM channels W_F. For a bit-rate b the effective capacity per wavelength channel follows as $B_{FDL} = T_F * b$. For instance, an FDL with delay $40\,\mu s$ as used below has a capacity per wavelength of 50 kB.

Several physical impairments like attenuation, chromatic dispersion and non-linear effects limit the practically feasible length of an FDL. As a first estimate, neglecting dispersion or higher order signal impairments, the maximum length of an FDL is limited by the power budget. Assuming that realistically only a single erbium doped fiber amplifier (EDFA) is deployed in conjunction with an FDL, its maximum length today corresponds to a typical EDFA span of approx. 80 km. From this, again, a maximum delay of approx. $400\,\mu s$ can be derived. Performance results as presented in Section A.1 show that typical FDL buffer delays should be in the order of a few mean burst transmission times. Consequently, FDL buffers can only be reasonably applied if the mean burst transmission time is less than approx. $100\,\mu s$. As the delay introduced by such FDL buffers is negligibly small compared to other critical delays, e.g., propagation delay in WANs or terminal processing latencies, they are not relevant for QoS in practical end-to-end delay budgets.

3.4.3.2 FDL Buffer Architectures

This section primarily classifies FDL buffer architectures with respect to their operation in time, internal structure, flow direction, location in the node, and resource sharing scheme. This classification scheme and key realizations are depicted in Figure 3.8. A first classification criterion is whether the node and the FDL buffer operate in a slotted or unslotted mode [HCA98]. Historically, optical packet switches with fixed size packets (initially ATM cells), for which FDL buffers were first studied, operated in slotted mode [CFK+96, HCG+98, GRG+98a]. In these scenarios, FDL buffers were mostly designed and scheduled to emulate the behavior of electronic buffers as closely as possible. However, the dominance of IP with variable-length packets and the concept of burst assembly motivated unslotted, asynchronous operation, with variable-length optical containers [Cal00].

In order to overcome the limitations of the strictly deterministic delay of a single FDL, buffers comprising multiple FDLs of different delay have been devised. Such FDL buffers can be classified in *single-stage* and *multi-stage* architectures [HCA98]. In single-stage buffers [XVC00, TYC+00, CHA+01, Gau02a], a variety of fixed delays is realized by arranging F FDLs of different length in parallel as depicted in Figure 3.9(a). Single-stage buffers are also commonly referred to as *fixed-delay* FDL buffers as their delay characteristic is determined by the set of fixed delays $T_{F,i}$ of the individual FDLs. They are further specified by the number of WDM channels per FDL $W_{F,i}$ and the total number of FDL buffer ports, i.e., the sum of all WDM chan-

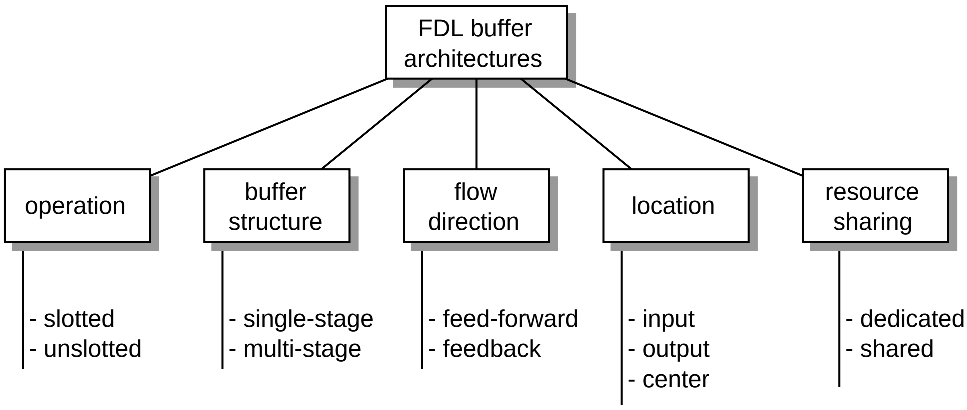


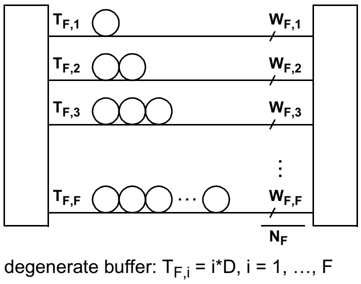
Figure 3.8: Classification of FDL buffer architectures

nels $N_\text{F} = \sum_{i=1}^{F} W_\text{F,i}$. This figure defines the maximum number of bursts that can enter/leave the buffer at the same time. Thus, it is not only meaningful regarding the required size of the switch matrix but for certain operational scenarios also regarding performance (cf. Section A.1). Usually, single-stage buffers are dimensioned as so-called *degenerate buffers* in which FDLs have linearly increasing delays $T_\text{F,i} = i \cdot D$ for $1 \leq i \leq F$ [Kar93, YQ00, Gau04], where D represents the *delay granularity*.

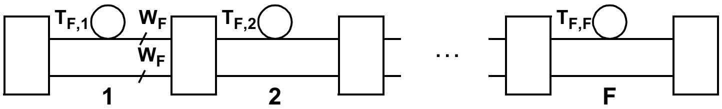
In contrast, multi-stage buffers [CFK$^+$96, HCG$^+$98, CFS00, YQ00, CCR00] cascade FDLs and optical switches as shown for an example configuration with identical delays and 2x2 switches in Figure 3.9(b). The number of FDL stages F and the number of wavelengths in the FDLs W_F define the structure of this buffer. If multi-stage buffers are operated in WDM the switches have to be wavelength selective to avoid severe performance degradation. The delay a burst experiences is determined by the sum of the FDL delays traversed in the buffer. Multi-stage buffers are frequently called *variable delay* FDL buffers as a burst can be switched out of the FDL buffer at any intermediate switch, i.e., the delay can be varied even after a burst entered the buffer.

While the basic *switched delay line* (SDL) configuration [CFK$^+$96] uses FDLs of identical length D, extensions with non-identical lengths trade-off buffer depth and hardware requirements. For instance, in the *logarithmic* delay line configuration [CFK$^+$96, HCA$^+$97, CFS00] the FDL delays grow from stage to stage according to $T_\text{F,i} = 2^i \cdot D$ for stage i with D as the basic delay element. Even more flexible multi-stage buffers can be built by increasing the number of FDLs per stage and the diversity of their respective delays as presented in [CFK$^+$96]. Also, several such multi-stage buffers with different maximum delays can be hierarchically combined into a single-stage FDL buffer. These buffers are frequently used in conceptual OBS studies to represent highly flexible and fine-grain optical buffers offering arbitrary delays [YQ00, LM04]. Finally, the *switch with large optical buffers* (SLOB) architecture in principle integrates a huge multi-stage buffer with the switching matrix of an optical packet switch [HCG$^+$98].

While the management of multi-stage switches is tractable for slotted operation with fixed size optical packets, for which these architectures have mostly been proposed for, leveraging their flexibility becomes a major challenge for unslotted operation with variable size bursts. Also, as multi-stage switches accumulate signal distortion introduced by the cascaded switch elements, they are less scalable and therefore not considered in the following.



(a) Single-stage buffer



(b) Multi-stage buffer

Figure 3.9: Principal FDL buffer architectures

In OBS switches, FDL buffers can be located at the input, at the output, as a shared output buffer in the center of the switch or as a recirculation buffer. Compared to electronic store-and-forward switches, the location is less performance-critical though. For instance, the typical head-of-line blocking problem of electronic input buffered switches is alleviated by WDM as other bursts on the same wavelength can bypass a burst while being delayed in the FDL buffer. Also, the coordinated scheduling of resources in the FDL and on the output fiber or the possibility of additional wavelength converters [YQ00] diminish performance differences regarding the location and allow to focus on realization arguments.

Furthermore, FDL buffers can be applied in *feed-forward* or *feedback* configurations shown in Figure 3.10 in a share-per-node architecture. In feed-forward buffers [Gau02a, BPV03, VTCJ03, CCRZ04, ZLJ06], bursts are delayed while flowing towards the output of the node whereas they are delayed in feedback buffers while being sent back to an earlier stage of the node. Consequently, bursts can only pass a feed-forward buffer once while they can recirculate through a feedback buffer several times. On the one hand, feedback buffers implementing *recirculation buffers* [CHA$^+$01, Gau02a, YMD03, CCC$^+$04, CZC$^+$04] can be seen as an implementation variant of a multi-stage buffer and thus as an attractive alternative towards higher delay flexibility. On the other hand, bursts accumulate signal distortion in each recirculation loop which limits the possible number of recirculations [GBPS03]. Initially, recirculation buffers in slotted nodes have also been considered to implement prioritization schemes by deciding whether a newly arriving burst or a burst leaving the buffer should be dropped in case of contention [CHA$^+$01]. However, such schemes lost importance as they are less practical with unslotted operation and do not support the coordinated scheduling described in Section 3.4.3.4.

Regarding resource sharing, FDL buffers can be either dedicated, e.g., to a single input or output or be shared per node. As a key advantage, shared FDL buffers significantly reduce the node size while preserving performance because of statistical multiplexing gains (cf. Section A.2) [GBPS03].

As already motivated, FDL buffers are frequently deployed in combination with wavelength converters in order to improve contention resolution performance and reduce FDL capacity requirements. Architectures can be classified based on whether they allow wavelength conversion in front of the FDL (thus $\lambda_{FDL} = \lambda_{out}$), behind the FDL ($\lambda_{in} = \lambda_{FDL}$), at both locations, or whether they do not allow wavelength conversion at all ($\lambda_{in} = \lambda_{FDL} = \lambda_{out}$). If all wavelengths

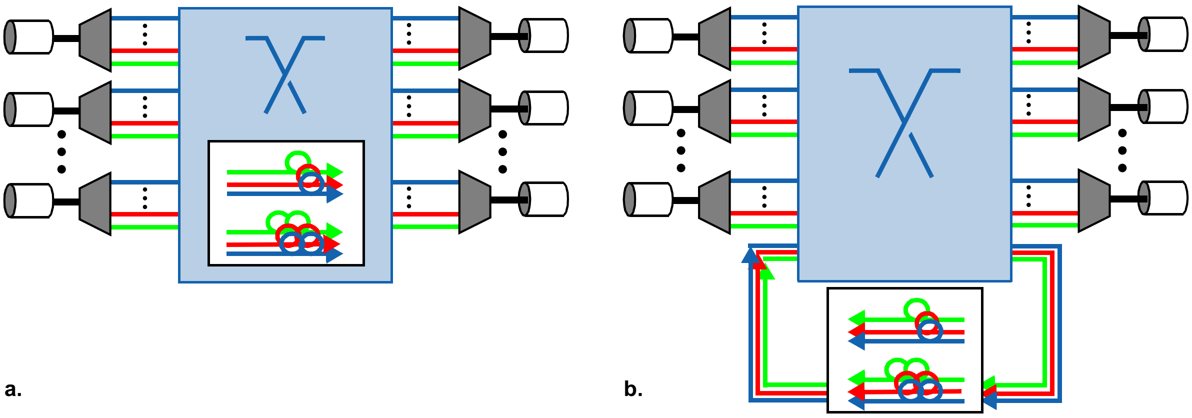


Figure 3.10: FDL buffers in a. feed-forward and b. feedback configuration

of the input or output fiber are supported in the FDL buffer, dedicated FDL buffers cannot benefit from wavelength conversion. In contrast, shared FDL buffers can already realize some statistical multiplexing gain even without wavelength conversion or with wavelength conversion only in front or only behind the buffer. However, their maximum capacity gain can only be leveraged by wavelength converters in front *and* behind the shared FDL buffer. As this usually can be more economically realized by a feedback compared to a feed-forward configuration, the former is preferable (cf. Figure 3.16).

Based on these architectural arguments, a share-per-node, single-stage feedback buffer with only one recirculation will be considered for all studies in this thesis.

3.4.3.3 Summary of the State-of-the-Art on FDL Buffers

Table 3.1 for slotted and Table 3.2 for unslotted operation summarize the key literature on the architecture and application of FDL buffers in OPS and OBS. They use the classification introduced in the previous subsection and shown in Figure 3.8. The table is structured by horizontal lines according to the main criteria. Following criteria and short notations are used, additional specific information is included in footnotes:

- **Location**: input (I), output (O) or center (C) buffer
- **Sharing**: isolated buffer (iso), buffer dedicated to input/output (D) or shared per node (S)
- **Delay variability**: fixed (F), variable by multi-stage (V) or variable by recirculation (R)
- **Structure**: single-stage (S) or multi-stage(M)
- **Flow**: feed-forward (FF) or feedback (FB)
- **Conversion**: number of possible wavelength conversions in case of shared FDL buffers
- **Evaluation**: analysis (A) or simulation (S)

Table 3.1: Classification of literature on FDL buffers with slotted operation

	Location	Sharing	Delay Variability	Structure	Flow	Con-version	Evalu-ation
[HCA98]	All	All	F, V, R	All			-
[CFK+96]	-	Iso	V	M[1]	-		A
[CFS00]	-	Iso	V	M[2]	-		A
[LB03]	-	Iso	F	S	-		A, S
[VHLL+04]	-	Iso	F	S	-		A
[FWB04]	-	Iso	F	S	-		A, S
[FLB05]	-, O	Iso, D	F	M[3]	-, FF		A
[CHA+01]		D	F, R	S	FF, FB		A
[BPV03]	I	D	F	S	FF		S
[WWB03]	O, C	D, S	F	S	FF	-, 0	A
[HCG+98]	C	S	V	M		1/stage	A
[Kar93]	C	S	R	S	FB		A
[CCC+04]	C	S	R	S	FB	2	A, S
[DHB+04]	C	S	R	S	FB	2[4]	A, S
[LHC05]	C	S	R	S	FB	2	S
[JHLC05]	C	S	R	S	FB	2	S

1. single slot per FDL
2. multiple slots per FDL
3. two-stage structure
4. due to feedback architecture

The review of the published literature on FDL buffers shows that the focus shifted over time: From exactly mimicking the functionality of electronic buffers with FDLs to exploiting FDL buffers as simple delay elements with advanced scheduling. Also, FDL buffers with unslotted operation receive considerably more attention following an overall trend in node design. Analytical performance models for FDL buffers are still in their infancy. Most of the basic models for slotted FDL buffers [LB03, VHLL+04, FWB04, FLB05] are restricted to isolated FDL buffer scenarios, i.e., do not include the effects of wavelength conversion or resource sharing.

Regarding FDL buffers with unslotted operation, several approximate analyses for dedicated feed-forward buffers [Cal00, APW04, APW05, RLFB05b] obtain excellent matching with simulation results. This is also the case for the only model for shared buffers [ZLJ05, ZLJ06] in a node without wavelength conversion. However, these models all concentrate on burst/packet scheduling schemes without void-filling (cf. Section 3.4.3.4), which leads to a major performance penalty especially for long FDL delays T_F. The model published in [RC05] analyzes different burst scheduling schemes, including void filling, but is limited to a single FDL with a single wavelength channel, and a single output wavelength. Also, it is only validated for extremely high burst loss probabilities. For the multi-stage buffer with variable, continuous delay and void-filling scheduling, [YQ00] applies the $M/M/n/n+s$ delay model for electronic buffers, for which the number of spaces in the buffer s was chosen to be the number of FDL buffer ports N_F. As expected, this model cannot comprehensively capture the particular behavior of FDLs. It

Table 3.2: Classification of literature on FDL buffers with unslotted operation

	Location	Sharing	Delay Variability	Structure	Flow	Con-version	Evalu-ation
[XVC00]	I, (C)[1]	D, (S)[1]	F	S	FF		S
[LW05]	I	D	F	S	FF		S
[YQ00]	I	D	V, (F)[2]	M+S, (S)[2]	FF		S, A
[FFWG02]	I	D	V	M	FF		A, S
[VTCJ03]	I, O	D	V	M	FF		S
[Cal00]	O	D	F	S	FF		A, S
[Gau02a]	O	D	F, R	S	FF, FB		S
[APW04]	O	D	F	S	FF		A, S
[APW05]	O	D	F	S	-		A, S
[RC05]	O	D	F [3]	S	FF		A, S
[RLFB05a, RLFB05b]	O	D	F	S	FF		A, S
[YMD03]	O	D	R	S	FB		A, S
[LM04]	O	D	V[2]	M[2]	FF		A, S
[TYC+00]	C	S	F	S	FF	2	S
[ISNS02]	C	S	F	S	FF	1	S
[CP02]	C	S	F	S	FF	2	S
[GBPS03]	C	S	F	S	FB[4]	2[5]	S
[Gau04]	C	S	F	S	FB[4]	2[5]	S
[CCRZ04]	C	S	F	S	FF	1	S
[CMR+05]	C	S	F	S	FF, FB[4]	1	S
[ZLJ05, ZLJ06]	C	S	F	S[6]	FF	-	A, S
[XQLX04]	-	-	F	S	FF	unspec.	S
[CZC+04]	C	S	R	S	FB	2[5]	S
[CCR00]	C	S	V	M	FF	1	S
[LP04]	C	S[7]	F	S	FF	2	S

1. only dedicated input buffer studied
2. only variable, continuous delay studied, realized by parallel switched delay lines in single-stage buffer
3. single FDL case
4. only single recirculations
5. due to feedback architecture
6. described as (single stage) variant of SLOB
7. both an FDL and an electronic buffer are deployed, information in table only applies to FDL buffer

only showed few correspondence with simulation results for FDL buffers with a more realistic parameterization.

Summarizing, FDL buffers are an active field of research with a large variety of architectures but still only basic performance models. Particularly, no model is available for shared, fixed-delay WDM FDL buffers as evaluated from a performance and technology point of view in [GBPS03] and proposed for OBTN. Therefore, simulation is used to evaluate the performance of FDL buffers in Appendix A and of OBTN in Chapter 6.

3.4.3.4 FDL Buffer Scheduling

Apart from the physical architecture of FDL buffers, their operation and scheduling is a major performance factor. For slotted operation, batch-scheduling of FDLs has been treated intensively. However, for unslotted operation with variable length bursts, scheduling received far less attention.

A key classification criterion for scheduling schemes is whether the node operates in a slotted or unslotted mode. Historically, optical packet switches with fixed size packets (ATM cells), for which FDL buffers were first studied, operated in slotted mode. Also, important packet switch architectures like KEOPS assembled all traffic into fixed size optical containers suitable for slotted operation [GRG$^+$98a]. In these scenarios, FDL buffers were mostly designed and scheduled to emulate the behavior of electronic buffers as closely as possible. However, the dominance of IP with variable-length packets and the objective of minimizing assembly complexity and overhead motivated unslotted, often also called asynchronous operation with variable-length optical containers [Cal00].

If the delay in the FDL buffer is known in advance, a coordinated scheduling of resources in the FDL buffer and the output fiber can be performed. This is possible for single-stage feed-forward buffers and feedback buffers with a pre-determined number of recirculations. Such strategies can improve the performance of contention resolution by avoiding situations in which bursts are first buffered and then discarded afterwards because of another contention on the output fiber. In [Sch01, Gau02a], the coordinated strategies are called *PreRes* in contrast to the uncoordinated *PostRes*.

An additional advantage of PreRes schemes is that they earlier schedule the output fiber for bursts, which have to be buffered. This is explained in Figure 3.11 for a contention resolution scenario with a single output wavelength and a single buffer FDL of delay T_F. In Figure 3.11(a), a burst control packet arrives at time t_a while the data burst arrives delayed by the offset time δ at time $t_a + \delta$. As an already scheduled burst blocks the newly arriving burst from direct transmission, the latter is first sent to the FDL and only then switched to the output wavelength (Figure 3.11()b). By already scheduling the output wavelength at time t_a, i.e., with an expanded offset $\delta + T_F$ prior to burst transmission, this scheduling request has a higher success probability than unbuffered bursts with the smaller offset time δ. In order to ensure that this prioritization due to buffering does not propagate to downstream nodes, the burst control packet can be sent out by the core node one offset time prior to burst departure.

This prioritization mechanism is the same as the one used in offset-based QoS outlined in Section 3.6. In addition, typical offset values for QoS differentiation and typical FDL delays both amount to a few mean burst transmission times. Thus, the FDL scheduling and QoS differentiation strategies can interfere as analyzed in [Sch01, Gau02a] and should not be used together.

As indicated in Section 3.3, reserve-a-fixed duration schemes allow burst scheduling with void-filling. In PreRes, the wavelength channel on the output fiber is only needed at the time the burst leaves the buffer. It is not needed during the time in the buffer and can be used by other bursts. Therefore, RFD schemes can better utilize the output fiber resources than RLD (no void filling). More specifically, RFD schemes lead to performance improvements when increasing the buffer

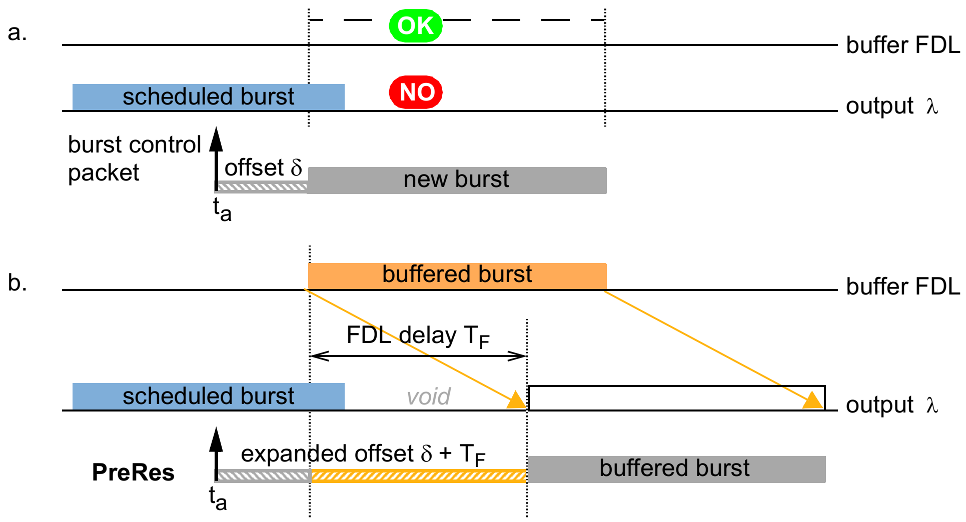


Figure 3.11: FDL scheduling: a. Burst blocking event and b. buffer scheduling with PreRes

delay *FDLDelay*. In contrast RLD schemes only benefit up to an optimal delay but beyond that suffer from the unused voids [XVC00, TYC$^+$00, Gau02a, ISNS02, CP02].

In single-stage FDL buffers with multiple delays, another task during the scheduling process is to select a proper delay. Here, most approaches select the shortest possible delay. In case of a PostRes scheme they use the shortest available FDL and in case of a PreRes scheme the shortest FDL for which also an output wavelength can be scheduled. Although this approach minimizes individual buffering delays it can lead to reordering inside a flow. However, this can only happen in the unlikely case that the FDL delay is longer than the inter-burst time in a flow. Nevertheless, advanced wavelength and delay selection strategies try to preserve a strict FIFO discipline. In [CCRZ04, CMR$^+$05], several approaches to this *Wavelength and Delay Selection* (WDS) problem are presented for RLD scheduling schemes.

Based on these arguments, a PreRes strategy is used for the coordinated scheduling of FDL buffer and output scheduling in Chapter 6 and Appendix A. Also, reserve-a-fixed duration scheduling schemes with void filling are employed and the minimal possible delay is selected.

3.4.4 Alternative/Deflection Routing

Because of the realization complexity of FDL buffers and their limited flexibility, alternative/deflection routing is frequently considered as an effective and simple to implement contention resolution scheme. However, the review of the state-of-the-art shows (i) that it only yields a limited performance improvement if applied alone, (ii) that it exhibits a strong dependence on the offered traffic, and (iii) that it requires significant efforts to control the deflection process. Nevertheless, it can be an attractive concept complementing other contention resolution schemes.

The general concept that a node sends traffic on a different route towards its destination in case of contention is referred to as alternative routing in telecommunication networks. Instead, most literature on OBS uses the term deflection routing frequently used in computer networks.

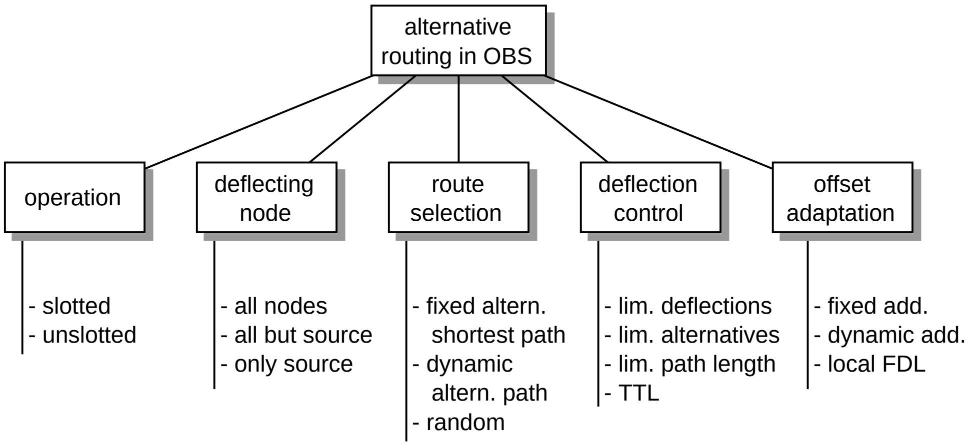


Figure 3.12: Classification of alternative routing schemes in OBS

interconnection networks. In this section, both terms are used interchangeably while following chapters will use the term alternative routing. The classification scheme used in the following and and key realizations are depicted in Figure 3.12.

Deflection routing has proven to be a simple yet effective contention resolution scheme in *slotted* networks with regular topologies, e.g., torus networks like the famous (bidirectional) Manhattan Street Network (MSN). On the one hand, slotted operation allows optimal local scheduling. On the other hand, regular topologies ensure that many equal length alternate routes exist and deflected traffic does not additionally load the network by long detours.

Early work on deflection routing in *unslotted* networks claimed severely reduced throughput and excessive deflections for some traffic [BFB93]. However, [CCF01] systematically compared the different effects of slotted and unslotted operation for a bidirectional MSN It identified following two reasons for the performance differences and proposed solutions. First, asynchronous packet arrivals prohibit optimal joint scheduling. To overcome this problem, they propose to exploit information of future arrivals derived from an FDL in the node needed for local insertion anyways. Second, unslotted operation with fixed-length packets [BFB93] can lead to resonance effects in nodes which would not occur with variable-length packets. Still, such instability has been proven for a symmetrical four node network with variable-length bursts and limited-range wavelength conversion in [ZLVR$^+$04]. However, such behavior was neither found in other published nor in own work in the case of OBS with full wavelength conversion, variable-length bursts, and irregular mesh topologies [GKS04b, GKZM05]. As can be seen from Table 3.3, most proposals proactively approach this topic and closely control deflections to avoid excessively long routes or even instability in case of high load.

A further classification criterion considers the nodes that are allowed to deflect bursts. Most literature allows deflection in all nodes, while [WMA02, HSKS03, LSKS05] exclude the source node. There, bursts can be stored in the electronic domain there and thus do not require deflection. In contrast, [TVJ03, YR05] restrict route alternatives to the source node and allow no further deflections inside the network.

In general alternate routes in irregular networks comprise more hops than the primary route, so that deflected bursts impose additional traffic on the network. This leads to a typical two-state behavior with high performance improvements for low load and even performance degradation for high load. The transitions between the two states can occur rather abruptly when changing the offered load due to the positive feedback loop as shown in [NM73]. Thus, strategies to control deflection and to avoid excessively long routes are essential.

A first class of approaches targets the selection of the alternate route or, from a local point of view, of the deflection interface is an important design criterion. Fixed alternative shortest paths with [CTT99] or without [ZLVR$^+$04] length constraints, as well as random [BBPV03, WMA02, CWXQ03] and dynamic, state-dependent optimal routes are proposed. Among the latter, the decision is based on the current load of the adjacent links [LKS$^+$05] or the loads [TVJ03, YY05], the expected distance [OT05] or even expected QoS on the alternate routes [LSKS05]. In order to account for implementation and processing constraints as well as for the inherent imprecision of detailed network state information due to the large propagation delays in core networks, simplicity and robustness of decision schemes should have priority over small performance improvements.

The second class limits the number of deflections or the resources a burst may use. Here, a fixed maximum [HL05, KKK02] or a threshold further conditioned on the position on the path or the burst size [HSKS03] are proposed. Similarly, a deflection can only be allowed with a fixed [CWXQ03] or dynamic, load-dependent probability [LSKS05]. In [ZLVR$^+$04], trunk reservation is proposed to enforce priority of undeflected bursts primarily at high load. Finally, the hop-count of routes can be limited by introducing a time-to-live field (TTL) [BBPV03, BFB93, WMA02, Sch04]. In a comparative study, [Sch04] analyzes the effect of limiting the number of deflections, the number of route alternatives as well as strategies for loop detection and avoidance. For a German and Pan-European reference network, it concludes that moderate limits, e.g., allowing only two deflections, are enough to control performance at high load values but not degrade performance at low load values.

A problem specific to OBS is an *insufficient offset time*. It occurs if the processing delay in core nodes is compensated by offset adaptation and a burst catches up with its control packet on a long alternate route. It can be either solved by dynamic offset adaptation using FDLs [HLH02, HL05] or by an additional pre-estimate routing offset [KKK02, ZLVR$^+$04]. Such schemes are not required if the processing delay of core nodes is not compensated by offset adaption but by short FDLs in each input. This is one reason why the short FDL is proposed to compensate processing in OBTN.

Typical propagation delays of links in WAN scenarios are large compared to typical mean burst transmission times (milliseconds vs. tens of microseconds). Therefore, deflection decisions in adjacent nodes can be considered to be independent. This explains the observation in [Sch04] that the propagation delays of links have no impact on the overall performance. In contrast, such impact has been observed in LAN or MAN scenarios where the propagation delays are comparable to the burst transmission time [WMA02, BFB93].

Most contributions on deflection routing concentrate on the burst loss probability or the throughput of the OBS network as performance metrics. However, they neglect possible impact of routing-induced jitter on performance. For instance, burst and thus IP packet reordering can potentially lead to duplicate acknowledgments of TCP segments if TCP is the transport protocol. Depending on the scenario, this again can degrade TCP performance. [SPG05] showed that frequent and high variations of the route length substantially reduce TCP goodput. It studies a lossless scenario with a generic topology, in which the alternate route is three times longer than the primary route. The goodput degrades starting from a deflection probability of 10^{-3} and is reduced by a factor of 3 when the deflection probability is 10^{-1}.

Table 3.3 classifies the discussed literature on alternative/deflection routing in OBS/OPS based on above criteria. Following categories and short notations are used, additional specific information is included in footnotes:

- **Operation**: slotted (S) or unslotted (U)
- **Packet length**: fixed (F) or variable (V)
- **Deflecting node**: all nodes (A), source node (S) or all but source node (A - S)
- **Route selection criterion**: fixed alternative shortest path (F), dynamic optimal routes (D) or random (R)
- **Deflection control**: limited deflections (L), alternatives (A), route length (P) or time-to-live (TTL)
- **Offset adaptation**: fixed additional routing offset (F), dynamic adaptation for deflection (D) or local FDL compensation (FDL)
- **Conversion**: yes (Y) if wavelength converters are also applied
- **FDL buffer**: yes (Y) if FDL buffer is also applied
- **Propagation delay**: scenario/assumption for typical link lengths, τ is the propagation delay normalized to the mean burst transmission time
- **Topology**: regular (R) or irregular (I) network or isolated node (Iso)
- **Evaluation**: analysis (A) or simulation (S)

Based on these contributions and above discussions, OBTN uses an alternative routing scheme to reduce the resource requirements for the FDL buffer. However, it constrains the alternate routes to specific links to avoid high or frequent route variations by design.

3.4.5 Combination of Contention Resolution Schemes

After this discussion of the individual contention resolution schemes and their architectural implications, the following subsection focuses on combinations of contention resolution schemes to improve performance or reduce implementation complexity. Section 3.4.2 already indicated that wavelength conversion alone cannot achieve the requirements for high QoS and high utilization at the same time. Thus, wavelength conversion is often complemented by FDL buffering and/or alternative routing as explained in Section 3.4.3 and in Section 3.4.4.

Table 3.3: Classification of literature on deflection/alternate routing in OPS/OBS

	Operation	Packet length	Deflect. node	Route select.	Deflect. control	Offset adapt.	Conv.	FDL	Propagation delay	Topology	Evalu-ation
[CTT99]	S	F	A	F	L[1]		Y	Y		I[2]	S
[KKK02]	S	F	A	F	L[3]	F				R[4]	S
[BBPV03]	S[5]	V	A	R	TTL		Y	Y			S
[CCF01]	U, S	V, F	A	D[6]						R[4]	S
[BFB93]	U	F	A		TTL			-, Y	$\tau \leq 1$	R[4]	S
[WMA00, WMA02]	U	V	A - S	R	TTL				$\tau < 2$	R[7]	S
[HLH02, HL05]	U	V	A	F	L[1]	D	Y	Y		I[8]	A, S
[CWXQ03]	U	V	A	R	L[9]		-, Y			Iso	A, S
[HSKS03]	U	V	A - S	-	L[10]		Y		WAN	I[11]	S
[Sch04]	U	V	A	F	L, A, TTL	FDL	Y	Y, -	$\tau < 100$	I[12]	S
[GKS04b]	U	V	A	F	-	FDL	Y	Y, -	WAN	I[12]	S
[LKS+05]	U	V	A	D[13]	-		Y[14]			I[11]	S
[LSKS05]	U	V	A - S	D[15]		D[16]	Y	Y, -		I[11]	A, S
[ZLVR+04]	U	V	A	F	A, Tr	F	Y			R,[4] I[11]	A, S
[OT05]	U	V	A	D[17]	-					R, I[11]	S
[GKZM05]	U	V	A	F	-	FDL	Y	Y, -	WAN	I[18]	S
[TVJ03]	U	V	S	F, D[19]	-		Y			I[18]	S
[YR05]	U	V	S	D[20]	-		Y			R,[4] I[11]	S

1. route length or hop-count limited
2. COST 239 network
3. limited by maximum number of deflections
4. MSN
5. mini-slots of 40 *Byte*, packets comprise at least one mini-slot
6. heuristic for optimized local matching
7. grid network with $M * M$ nodes
8. ArpaNet2 network
9. fixed deflection probability
10. four different constraints based on position on route and burst size
11. NSF network
12. Germany and COST266 CN network
13. metric of link loads
14. only for high priority traffic
15. metric of QoS performance on route to destination
16. trunk reservation
17. metric of expected distance to destination
18. US network with 24 nodes
19. load-dependent
20. several different criteria proposed

The first comparative study of basic contention resolution schemes and their combinations for unslotted, variable-length OPS networks is presented in [YMYD00] and further extended in [YMYD03, Dix03]. It clearly shows that wavelength conversion alone is significantly more effective than buffering and deflection alone as well as combined. Also, for the presented network scenarios, buffering yields better performance than deflection routing.

[YMYD03, Dix03] as well as [GKS04b, Sch04] systematically study the performance of combinations of two or three of the basic schemes. These studies are performed independently and show for different network scenarios and node configurations that the combination of wavelength conversion and buffering outperforms a combination of conversion and deflection for medium and high load values. Again, it is the strong dependence of load that renders deflection routing problematic.

In case of wavelength conversion and buffering, no performance improvements can be seen in [YMYD03] when additionally allowing deflection routing. In contrast, the results in [GKS04b] indicate quite substantial improvements. One reason for such discrepancies could be the different network dimensioning approaches as further studied in the joint paper [GKZM05]. The approach with a constant number of wavelengths per link in [YMYD03] leads to congestion around specific nodes such that additional deflection beyond wavelength conversion and buffering remains ineffective there. In contrast, the demand-based link dimensioning in [GKS04b] yields more evenly loaded links such that the additional deflection option can succeed. As both dimensioning approaches are used by operators, these results are important regarding implementation.

Combined contention resolution schemes open the additional dimension of the hunting mode, i.e., the order in which the different schemes are applied. In principle, this order can be decided based on many parameters and both statically as well as adaptively. However, considering the strict processing delay requirements [JG03b] only rather simple sequential schemes seem practical and have been proposed to date.

Figure 3.13 illustrates advance hunting modes for a node with and FDL buffer and *partial* wavelength conversion, i.e., wavelength converters are shared in a pool. The oval areas represent the output fiber (*no FDL*) and the two FDLs with several wavelength channels respectively. If a burst cannot be transmitted on the output fiber with the original wavelength, it still has three options for contention resolution and successful burst transmission: a. transmission on the original wavelength after buffering, b. transmission with wavelength conversion but without buffering, and c. transmission with wavelength conversion and buffering.

To categorize the component requirements, the dashed lines in Figure 3.13 mark four quadrants. They represent transmission on the original wavelength as well as transmission with the three combined contention resolution schemes a.–c. Buffering without wavelength conversion (option a.) only uses one FDL wavelength, while wavelength conversion without buffering (option b.) uses one wavelength converter. Option c. with wavelength conversion and buffering usually requires two resources. If an additional wavelength converter is needed after buffering three resources are used.

Figure 3.13 (center and right), illustrate two principal hunting modes for options a.–c. Note that all but the quadrant for the original wavelength contain several possibilities for successful burst

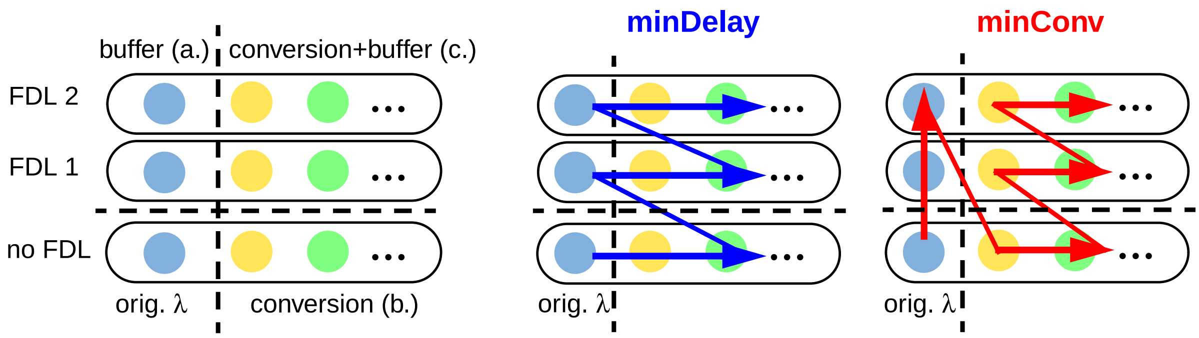


Figure 3.13: Options for combined contention resolution (left) and hunting modes for wavelength converters and FDLs (center, right)

transmission. In the hunting mode called *minDelay*, the main goal is to minimize the transfer time by preferably applying wavelength conversion. Only if there is currently no converter available or all output wavelength channels are currently busy does this mode employ the FDL buffer and start probing from the minimum delay FDL and the original wavelength. The arrow in Figure 3.13 (center), indicates the order of the contention resolution options for minDelay. In contrast, the goal of the hunting mode *minConv* is to minimize converter usage by first probing the FDL buffer on the original wavelength for contention resolution. Only if the burst cannot be reserved this way, *minConv* resorts to wavelength conversion, without and then with FDL buffering. The arrow in Figure 3.13 (right) indicates the order for contention resolution a → b → c. Apart from minDelay and minConv, other hunting modes with a different order of probing and different optimizing objectives can be defined.

[Gau04] shows that *minConv* and *minDelay* can optimize resource usage: They allow to reduce the number of converters in the pool at a given FDL buffer dimensioning or to reduce the FDL buffer capacity at a given number of converters at the same high QoS.

Based on the literature presented in this section and own previous research, the OBTN architecture introduced in the next chapter relies on a optimized combination of wavelength conversion, FDL buffering and constraint alternative routing for effective contention resolution. Section 4.4.4 introduces sequential hunting modes similar to the ones presented here for the combination of FDL buffering and constraint alternative routing in OBTN.

3.5 Switching in OBS Core Nodes

The design of fast optical switches was triggered by research on optical packet switching. Therefore, OBS nodes usually build on architectures originally devised for OPS [HA00, Nor03, Wan02, Buc05]. In the following presentation, all architectures are only discussed with respect to OBS though. This section discusses the requirements towards OBS core nodes and introduces the most promising concepts and architectures relevant to this thesis.

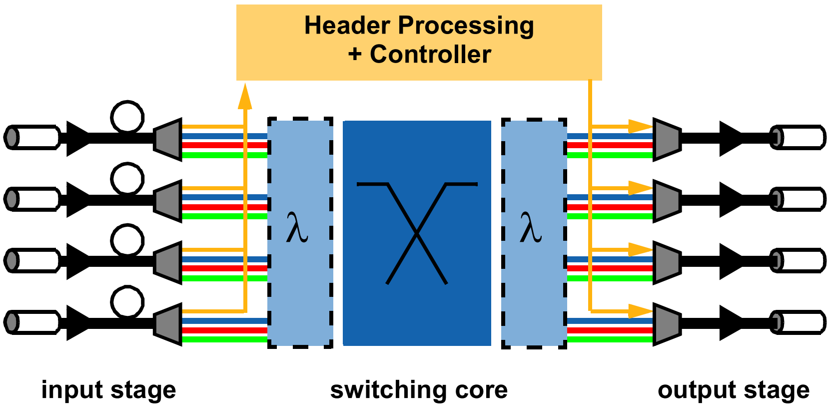


Figure 3.14: Structure of an OBS core node

3.5.1 Node Functionality, Structure and Requirements

Core nodes in OBS are responsible for switching individual bursts and for reading, processing, and writing burst control headers. The information in the burst control headers is used to make the forwarding decision and to set up the switch. The architecture and realization aspects of control and scheduling modules for OBS core nodes is not further detailed here but treated in [Jun06].

OBS nodes can be structured into input and output stages as well as a switch core as is shown in Figure 3.14 [HA00, Nor03, Buc05]. The input (output) stages amplify and demultiplex (multiplex) the WDM signal entering the node from a transmission link or a local add port. In case the processing and switch set-up delay, or at least its major part, are compensated by an FDL, this is also located in the input stage.

The switch core comprises space, wavelength, or time switching stages. Wavelength stages can be either placed at the input or at the output of the switch core or at both positions. An FDL buffer is primarily seen as a contention resolution device instead of as a time switching stage and thus not discussed here. In this thesis, each wavelength channel entering/leaving the switching core is referred to as a switch port, including add/drop and FDL buffer ports. It is used to abstract the complexity of different node architectures for dimensioning and evaluation in Chapter 5 and Chapter 6.

The key requirements for the switching core comprise functional as well as technological aspects [Buc05]. The switching core of an OBS node should be non-blocking in order to avoid any additional burst loss. More specifically, it should be strictly non-blocking as reconfiguration is impractical due to the short burst transmission time and the otherwise required additional signaling. The switch should in principle be able to support different contention resolution components such as wavelength converters or FDL buffers. In addition, it would be desirable to support multicast services at least partly, e.g., restricted to one signal copy per output on the same wavelength. High-performance and cost-efficient realizations mandate a fast switching speed, scalability towards a large number of switch ports and simple and compact implementation. Regarding technology, bit-rate transparency, low insertion loss (or even gain), low crosstalk, and polarization independence are further requirements.

3.5.2 Switching Elements

The fast switching speeds required for OBS can be realized by space and wavelength switching elements. Among the active space switching elements, semiconductor optical amplifiers (SOA) are most promising. They are used as *on*/off gates with switching times of approx. 1 ns. They amplify the signal in the *on* state and absorb the signal in the off state yielding high extinction ratios. However, they introduce noise and suffer from high power consumption. Compared to SOAs, acousto-optic or electro-optic switches are less suitable for fast and scalable OBS nodes [Buc05].

As introduced in Section 3.4.2 for contention resolution, wavelength switching elements, are realized by wavelength converters with different input and output flexibility or tunability. Wavelength converters can be realized either all-optically or electro-optically (O/E/O). All-optical wavelength converters are heavily worked in order to implement conversion at considerably lower cost than O/E/O converters. Receivers for optical burst signals inside O/E/O wavelength converters (or at local drop ports) require burst-mode capabilities.

In addition to the active switching elements, passive elements can be used for wavelength-sensitive demultiplexing/multiplexing and splitting/coupling. Arrayed waveguide gratings (AWG) provide a static interconnection pattern between input and output ports (cf. [Muk06] Chapter 2.6.5). Wavelength channels arriving on an input port are spectrally demultiplexed in the AWG and sent to different output ports. Conversely, from each input port, one specific wavelength channel appears at an output port. Thus, AWGs are used in WDM (de-)multiplexers and are attractive due to their robustness and scalability. In contrast, combiners are used in order to multiplex signals from different inputs without information on their wavelength channel. Optical splitters divide a signal into a number of signals of respectively lower power. As combiners and splitters use the same component, combination and splitting with the same number of (sub-) signals both experience the same loss penalty.

3.5.3 Switch Architectures

One option to build large optical switches is by cascading small switching elements in multi-stage configurations. Alternatively, specific optical components like splitters/combiners and AWGs can be exploited in single-stage configurations. Similar to electronics, multi-stage switching fabrics can be designed as crossbar, Benes, Spanke, and Clos architectures [RS98]. However, in contrast to digital electronic switches, optical switching elements like SOAs do not regenerate the signal. Therefore, signal degradation due to noise or crosstalk in the switches accumulates along the signal path rendering multi-stage configurations less desirable. In order to minimize the number of switching elements and the signal degradation, single-stage switching fabrics are thus more promising.

Within the class of single-stage architectures, *broadcast-and-select* (BAS) switches [GRG$^+$98a] constitute an important category. In a BAS switch with N input fibers and M wavelength channels, the entire WDM signal is split into $N*M$ signals (*broadcast*). Each of the signals is directed via an SOA *on*/*off* gate to a specific output port characterized by fiber and wavelength channel. There, N signal paths, one from each input fiber, are combined. In order to avoid conflicts, the *on*/*off* gates ensure that only one of the WDM signals reaches the combiner at any

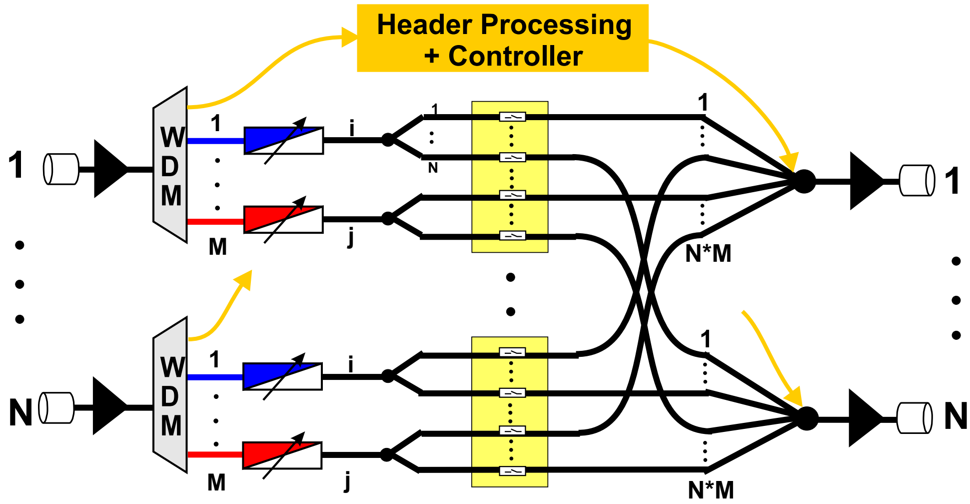


Figure 3.15: Architecture of the tune-and-select node (TAS) [Buc05]

given time. A tunable filter behind the combiner *selects* one of the wavelength channels which is then converted to the fixed output wavelength. Finally, all wavelength channels for an output fiber are multiplexed and amplified at the output. The BAS switch offers wavelength conversion and multicast/broadcast capabilities. Its switching time is defined by the SOAs and the tunable filter. The most prominent example of a BAS switch was proposed in the KEOPS project [GRG+98a, GRG+98b].

In order to avoid that the split signal is sent through the SOA with all wavelength channels, *tune-and-select* (TAS) switches reverse the order of the switching elements (cf. Figure 3.15). At the input, the WDM signal is first demultiplexed into its components. Then, it is converted to its final wavelength channel in a tunable converter. This signal is split up by N and directed via *on*/*off* gates to each output fiber. At each output, the signal paths from $N * M$ input wavelength channels are combined and amplified. Again, the *on*/*off* gates ensure that different wavelength channels enter the combiner. The TAS node also offers wavelength conversion and can multicast an input signal to all output fibers on the same wavelength.

In both architectures, MN^2 SOAs and NM wavelength switching elements are needed (tunable filters and fixed wavelength converters for BAS or tunable wavelength converters for TAS). Each active signal path only contains one SOA in *on* state avoiding the accumulation of distortion. Nevertheless, their scalability is limited by the splitting of the signal power by a factor of 1/N and 1/NM (in splitters or combiners).

AWG-based single-stage switches avoid the power splitting by using the static, wavelength-sensitive interconnection pattern as space switches. The AWG is complemented with a tunable wavelength converter at each (wavelength) input and output [HNC+99, BGP02]. The fundamental realization is in a single-plane with one large AWG as proposed in the basic WASP-NET approach [HNC+99]. In order to reduce the number of tunable wavelength converters, this design can be extended to operate in wavebands [HDH+04]. This extension applies tunable waveband converters which concurrently convert a given set of wavelengths to another set of wavelengths. Proper waveband selection and scheduling allows to achieve the cost reductions of a smaller component count at an only marginal performance degradation.

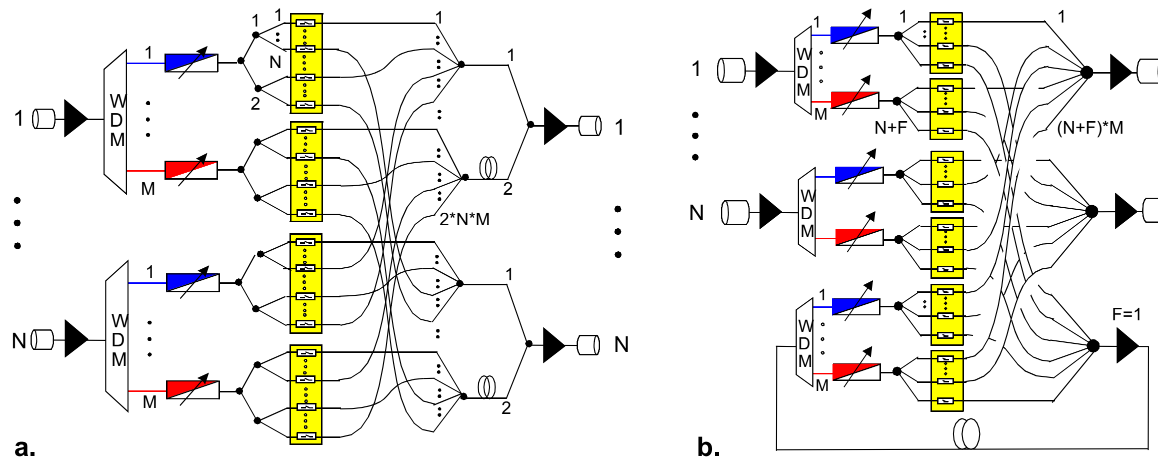


Figure 3.16: TAS nodes with a. dedicated output and b. shared FDL buffer

The size of the AWGs can be reduced by structuring the switch into several parallel planes with one AWG per plane. However, as the respective signals are combined at the output it sacrifices signal quality. The multi-plane WASPNET switch assigns the planes according to input wavelengths [HNC+99]. In contrast, extensions propose waveband planes [Nor04, Pat05] and a combination of fiber and wavelength planes [BHS02]. The complexity of the different AWG-based switches is compared in [Nor03].

These basic switching architectures can be extended or adapted with a focus on contention resolution functionality, e.g., by FDL buffers, shared wavelength converter pools or limited-range wavelength converters. As an example, Figure 3.16 depicts TAS nodes with one dedicated FDL per output and with one shared FDL. Both FDL buffers can successfully resolve contention and improve the QoS [Gau02a] (cf. Section A.1). However, the dedicated FDLs also double the number of outputs of the switch fabric and thus require to additionally split each signal by two. In contrast, the shared FDL can be treated like one additional output and thus requires to split the signal among one more output only. Consequently, there is a trade-off between signal quality and thus node scalability on the one hand and QoS on the other hand. In order to analyze this trade-off the following section outlines an integrated evaluation of performance and technology.

This discussion of switching architectures for OBS core nodes showed that several architectures are proposed for fast, non-blocking OBS switches. Analytical, simulative or prototype studies indicate that they are in principle feasible with available technology. They have been reported to scale to more than one hundred wavelength channels for 4–8 input and output fibers [BGPS03] with state-of-the art components. As technology evolves, this can be expected to further improve.

However, the complexity of the discussed architectures in terms of component count (switching elements, wavelength converters, interconnections, etc.) [Nor04] demonstrates that it is inevitable to keep the size of OBS switches to the necessary minimum. Therefore, the objective of reducing the number of switch ports is vital in node as well as in network design. In node design, effective contention resolution can help to increase the utilization and reduce the resource requirements at the same QoS. In network design, virtual topologies for OBS can effectively bypass OBS nodes to reduce the amount of transit traffic and thus switch ports. The OBTN architecture described in Chapter 4 combines both approaches to achieve the objective of fewer switch ports and thus overall reduced cost.

3.5.4 Integrated Evaluation of Performance and Technology

Commonly, performance and technology are evaluated separately. For instance, Section A.1 evaluates the QoS of an abstract OBS node model with FDL buffers under dynamic traffic. Similarly, [Buc05] analyzes the signal quality in a detailed physical TAS node model. However, neither of the separate evaluations can capture all effects together. This requires an integrated evaluation of performance and technology as proposed in [BGPS03, GBPS03, GBP05].

Figure 3.17 depicts the two research threads of performance and technology evaluation on the organizational level. Both activities evolve from the architectural stage via modeling and analysis stages to an interpretation stage. Then, the results can be integrated and the know-how can be fed back into the different stages to re-iterate on architectures, models, and analyses.

The process of such an integrated evaluation is further illustrated by quantifying the trade-off of advanced contention resolution schemes and node scalability [BGPS03, BPSG03, GBPS03, GBP05]. Starting from the node architecture and a fixed number of input/output fibers, the maximum number of wavelengths per input and output fiber is determined in a signal analysis. Here, the minimal signal quality is specified by a BER criterion. Then the maximum (static) throughput of the node is calculated as the product of the maximum number of wavelengths, the number of input fibers, and the bit-rate.

The maximum wavelength count is used to parameterize a functional model used for a teletraffic analysis or for simulation. This analysis provides the maximum achievable utilization for a given QoS under (dynamic) traffic. Finally, multiplying this utilization and the maximum throughput yields the maximum effective throughput under dynamic traffic.

Section A.2 presents the results of an integrated evaluation of performance and technology [GBPS03]. It compares the maximum and maximum effective throughput of the basic TAS node as well as TAS nodes with dedicated and shared FDL buffer.

3.6 Quality of Service Differentiation

One of the central ideas of IP-over-WDM is to allocate functionality to the layer it can be best implemented in and to minimize redundancy in functionality. As the IP layer does not inherently support QoS, optical server layer networks have been envisioned to offer QoS capabilities to the IP layer. In order to provide service guarantees or differentiation directly in OBS several approaches have been proposed which are comprehensively discussed in [Dol04]:

- *Additional QoS offset*: Offset-based schemes rely on the fact that a greater offset time translates into an earlier reservation, and thus into a higher probability of successful reservation. While the total blocking probability remains constant for RFD burst scheduling, high priority bursts which are assigned an additional QoS offset can have significantly

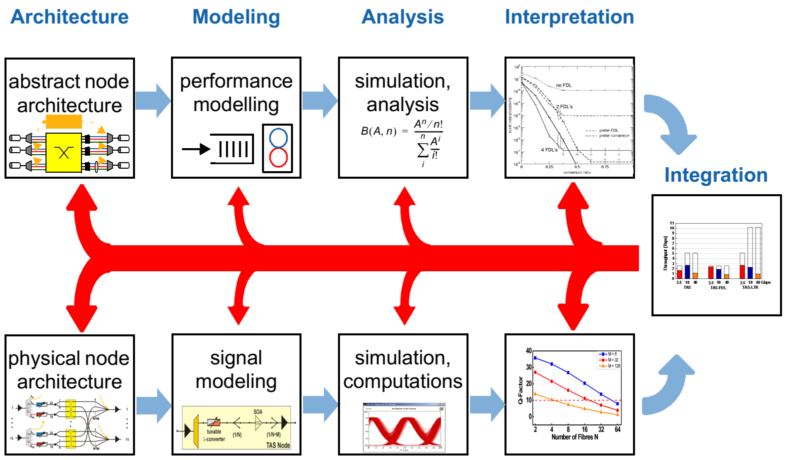


Figure 3.17: Integrated evaluation of performance (top) and technology (bottom)

reduced blocking [YQ00]. Analysis showed that the additional QoS offsets have to be in the order of a few mean burst durations [Gau00]. However, offset-based schemes have the drawback that the burst loss probability of the high priority class is very sensitive to burst size characteristics of the low priority class [DG01, BS05a]. Also, offset-based QoS can interfer with resource reservation for FDL buffers (cf. Section 3.4.3.4).

- *Preemption*: In preemption schemes, high priority bursts can preempt low priority bursts in case of a reservation conflict. An extension to this approach limits preemption to bursts violating a traffic contract [LLY02]. In general, preemption of bursts has the disadvantage that downstream nodes either waste already reserved resources or they need to be informed accordingly, which increases signaling load.

  Combining preemption and burst segmentation, the priority of a burst decides whether part of an already reserved burst is dropped or whether part of the newly arriving burst is dropped [VJ02b]. Here, an even more advanced approach allocates high priority packets at the head of a burst and low priority packets at the tail of a burst. In case of contention, preferably the tail part of a burst is preempted which then mostly concerns low priority bursts [VZJC02].

- *Intentional dropping*: Similar to schedulers in the electronic domain, some approaches provide relative service differentiation by actively dropping bursts. While applying such schemes to all traffic [CHT01] can lead to substantial unfairness for well-behaving users, combining it with trunk reservation and limiting it to out-of-profile traffic is far more attractive [DG01, Dol04].

- *(Re)scheduling of control packets*: Service differentiation can be provided by not strictly scheduling bursts on a first-come first-serve basis but according to their QoS class [YZV01, BS05b]. However, this causes additional and non-deterministic processing delays for low priority burst control packets and requires more complex reservation control modules.

- *Resource reservation*: These schemes do not grant all bursts access to all resources as in *complete sharing* but reserve some resources for high priority bursts. *Static partitioning*, *partial sharing* or *trunk reservation* are known concepts for resource reservation which can provide different degrees of isolation and performance [Dol04, PCM$^+$04, KM05].

Although no clear trend can be formulated regarding concepts for QoS provisioning in OBS, stability, robustness and ease of implementation will be key decision criterion. Thus, Chapter 4 aims at providing an overall high QoS onto which schemes for service differentiation can be flexibly built.

4 The Optical Burst Transport Network Architecture (OBTN)

This chapter introduces the Optical Burst Transport Network (OBTN) architecture [GM05a, GM05b, Gau05], which extends OBS by a dense virtual topology, constrained alternative routing, and shared overflow capacity on a selection of virtual links. It targets an optimized balance of network efficiency and node scalability at overall high QoS. Also, it combines elements of optical circuit switching and optical packet/burst switching as was also motivated by Alan Hill and Fabio Neri in their 2001 editorial for IEEE Communications Magazine [HN01]:

> … from an overall networking perspective, a hybrid solution combining the merits of fast (optical) circuit switching with those of optical packet switching may offer better cost and performance.

In this chapter, the design rationale for OBTN is outlined and the key extensions of the OBTN architecture are presented. Then, these extensions are analyzed regarding their consequences for the node and network elements. In addition, architectural variants and operational strategies are discussed. Finally, a qualitative comparison of OBTN is given with respect to the reference architectures introduced in Section 2.3.3 and Section 2.3.4 Optical Burst Swichting (OBS) and Burst-over-Circuit-Switching (BoCS).

4.1 OBTN Design Rationale and Application Scenarios

The OBTN design rationale comprises service requirements, network and node design considerations, and an important burst transport network scenario.

4.1.1 Quality of Service Requirements

A future-proof burst-switched transport network architecture has to be multi-service capable, i.e., support a wide variety of client services. It should neither be designed for a subset of applications with specific QoS requirements, nor should client layer networks and protocols depend on its characteristics. Instead, it should provide a balanced, high service quality regarding loss probability, delay and jitter satisfying standardized network performance objectives as further discussed in Section 5.3.2. Also, it should not lead to performance problems in predominantly

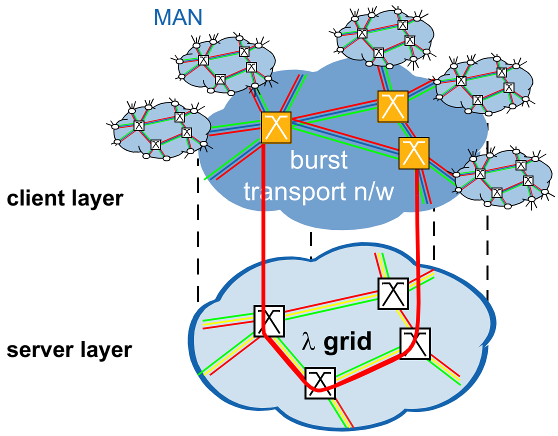


Figure 4.1: Combination of OBS and WSN in a client-server architecture

used client layer protocols. For instance, burst reordering inside the network should be low in order not to interfere with TCP congestion control (cf. Section 3.4.4). If these generic criteria are met, service providers can produce a wide range of transport services at a reasonably high network utilization [IST04a]. On top of such a high quality yet best-effort network architecture, QoS differentiation mechanisms as outlined in Section 3.6 can be deployed to produce services which require absolute or relative guarantees.

4.1.2 Cost-aware Network Design

As motivated in Chapter 2 and Chapter 3, the design and analysis of burst-switched architectures should always consider network and node complexity, i.e., the *cost of transport* and the *cost of switching* [GBPS03, Cha05b]. On the one hand, the still limited functionality and mostly *analog* nature of agile optical switching indicate that its implementation cost will remain high for some years to come (cf. Section 3.5). On the other hand, the cost of transport will further decrease due to the huge available capacity of optical fiber networks. Accounting for both trends, switching will continue to dominate the cost of optical networks in general and of OBS networks in particular. Therefore, it is crucial to primarily minimize the number of ports of OBS nodes even at the cost of additional network capacity (cf. Section 2.4.2). These trends also materialize in techno-economic studies of IP/WDM multi-layer networks which focus on the number of switch ports while mostly neglecting required transport capacity in the network [SBG04, Köh05b, Gal05].

In IP/WDM networks, WSNs are frequently used to provision a virtual topology of lightpaths which allow to bypass client layer nodes and thus to offload them from transit traffic. Similarly, optically burst-switched networks could reduce their resource requirements if combined with a WSN in a client-server relation. Thus, Figure 4.1 depicts a burst-switched and a wavelength-switched network as a client-server hybrid optical network Section 2.3.4.

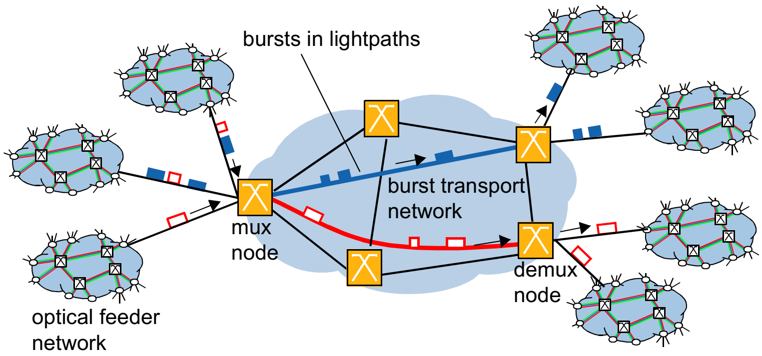


Figure 4.2: Transport network scenario for the interconnection of optical feeder networks

4.1.3 Network Scenario for Burst Transport

The interconnection of huge core routers is often considered the main application scenario driving a migration towards OBS in transport networks. However, statistical multiplexing of the highly aggregate traffic flows exchanged by such routers can only yield a limited capacity improvement (cf. Section 2.4.3). Therefore, this scenario may not justify the transition from circuit-switched to burst-switched network architectures. Also, WSNs can be expected to remain an integral part of future transport networks due to their scalability, manageability, and multi-service capabilities. Thus, any migration towards OBS will most likely be as a client to a WSN instead of in a greenfield infrastructure. Consequently, the following application case and migration scenario envisions burst-switched feeder networks, e.g., optical MANs, and particularly their efficient and seamless interconnection across an WSN core network.

To better support broadband access and high-speed metro networks, future optical networks could reach farther out towards the user than they currently do. Instead of repeated aggregation and consolidation of bursty data traffic in the electronic domain, it could be assembled and already handed over to the optical domain at the edge of the feeder network [LSDD$^+$01, HGJ05, DDC$^+$03]. As a consequence, traffic at the ingress to the optical network would be less aggregate and more bursty, and better suit the ideas of burst transport. Such a collapse of a network hierarchy level can be seen as strategic step towards CapEx reductions.

If the traffic demand between pairs of individual feeder networks across the core network do not reach full wavelength granularity, it would be inefficient to interconnect them directly with lightpaths. Instead, optical bursts from a group of attached feeder networks could be multiplexed at the edge of the core network for transport across the existing WSN as is depicted in Figure 4.2. In this scenario, the burst-switched core network nodes would mainly multiplex/demultiplex optical burst traffic at the edge of the WSN.[1] Optical bursts with the same multiplexing/demultiplexing core node would not have to be switched in intermediate core nodes but could be transported together through the WSN in a direct lightpath instead as also illustrated in Figure 4.2. Such an interconnection of burst-switched nodes with direct lightpaths is also proposed in [OSHT01, CSBO03] as an OBS/OPS introduction scenario (cf. Section 2.3.4) and referred to as *burst-over-circuit-switching* (BoCS) here.

[1]While current proposals for optical MANs mostly assume an electronic node as gateway to the core network, it is assumed here that future network solutions can also employ optical gateways.

Concluding, the outlined burst transport network scenario relies on a virtual topology of lightpaths to interconnect burst-switched multiplexing/demultiplexing nodes and to migrate burst-switching into transport networks. Regarding business models of service providers [Muk06], this interconnection would be attractive to both incumbent operators, who already own the underlying core WSN and seek to cost-efficiently interconnect burst-switched feeder networks. Also, it can be used by competitive operators who operate networks in different metro areas but have to lease the transport capacity to interconnect them.

4.2 Network Architecture

The previous two subsections both motivated that a virtual topology defined by lightpaths should be employed to design efficient burst transport networks. Based on this design rationale, the following section introduces the fundamental concepts of the OBTN network architecture and its integration with the WSN layer network.

4.2.1 Virtual Topology of OBTN

OBTN nodes multiplex/demultiplex optical bursts, which originate from the attached optical feeder networks, into lightpaths for transport across a wavelength-switched core network as illustrated in Figure 4.2. These OBTN nodes and the WSN server layer nodes are assumed to be collocated in a point-of-presence (POP) in all following discussions.

The lightpaths serve as virtual links to interconnect the OBTN nodes in a dense or even full-mesh virtual topology minimizing transit traffic in the burst-switched layer. The discussion in this chapter as well as the modeling and most of the evaluation work in the following chapters use a full-mesh virtual topology. However, OBTN is not restricted to a full-mesh virtual topology. Section 6.6 describes and evaluates sparser virtual topologies based on traffic demand and path length criteria. Still, in order to allow for the constrained alternative routing and the assignment of respective capacity introduced below, OBTN nodes should always be connected by a virtual link if they are neighbors in the physical topology. These virtual links are referred to as *single-hop virtual links* because they only span one hop in the physical topology.

The principal dimensioning trade-offs and limiting scenarios for virtual topologies under dynamic traffic were already discussed in Section 2.3.4 for OBS and quantified in Section 2.4.3. These discussions showed that there are two counteracting effects on capacity requirements when increasing the connectivity of a virtual topology: a positive effect of minimized transit traffic and a negative effect of reduced statistical multiplexing gain per link. In summary, they can either decrease or increase the capacity requirements of the client layer and thus of the OBTN nodes. Thus, in order to sustainably reduce the node size and cost, OBTN also employs contention resolution by wavelength conversion and FDL buffering inside the nodes as well as by alternative routing on the network level.

4.2.2 Constrained Alternative Routing

The concept of contention resolution by alternative/deflection routing, discussed in Section 3.4.4, is also applied in OBTN. While most findings on alternative/deflection routing are also valid here, a key difference is the fact that OBTN employs a densely meshed virtual topology compared to the sparsely meshed physical topology in OBS.

This subsection focuses on the routing schemes, i.e., on how to obtain the set of suitable routes between a pair of nodes [ITG96][2]. In particular, it discusses the computation of a constrained set of alternative routes. Below, Section 4.4.2 further details route selection, i.e., which route to pick from a set of suitable routes.

In principle, alternative routing in densely meshed (virtual) topologies can cause operational and performance problems. The large number of alternative routes, can lead to significant route and thus end-to-end delay variations. This is not only due to the high variation in the number of virtual links per route ranging from 1 to $n-1$ in a network with n nodes. It is also due to the fact that the virtual links can span several hops in the physical topology.

For illustration, in a network with n nodes and a full-mesh topology, the number of alternative routes between a node pair is given by Eq. (A.2) in [Spä02] as:

$$R_{\mathrm{FM}} = 1 + \sum_{l=2}^{n-1} \prod_{k=2}^{l} (n-k). \tag{4.1}$$

This leads to a large number of routes even for a relatively small number of nodes n, as can be easily seen from the upper bound $(n-2)(n-2)!$ for this relation.

However, a dense virtual topology also opens up interesting design opportunities. The key advantage is that for a given node pair several alternative routes can be found, which use other virtual links than the primary route but still traverse the same physical links. By constraining alternative routing with this criterion, all bursts experience the exact same end-to-end propagation delay independent of whether they are transported on their primary route or on alternate routes. Furthermore, as the propagation delay dominates in WAN scenarios, delay variations and burst reordering can be inherently avoided.

Figure 4.3 illustrates such constrained alternative routes for a network with five nodes connected by lightpaths in a full-mesh virtual topology. The left sub-figure depicts the physical fiber links and shows how the (bidirectional) virtual links are assigned to them. For traffic from node 1 to node 5, the primary route is defined by the direct lightpath traversing the three physical links A, B, and C. In case this primary route is blocked, a total of 16 alternative routes exist in this network based on Eq. (4.1) spanning two or more virtual links. The right sub-figure shows the virtual topology of the same network as well as the set of constrained alternative routes for traffic from node 1 to node 5. The alternate route 1 uses the virtual links from node 1 to node 2 (fiber A) as well as from node 2 to node 5 (fibers B and C) while the alternate route 2 comprises the three virtual links corresponding to the fiber links A, B, and C. Regarding the local forwarding of bursts, Figure 4.3 (right) also shows that nodes 1 and 2 can select from two outgoing virtual

[2] For alternative routing, this set of routes is also referred to as the set of alternative routes including the primary route and all alternate routes.

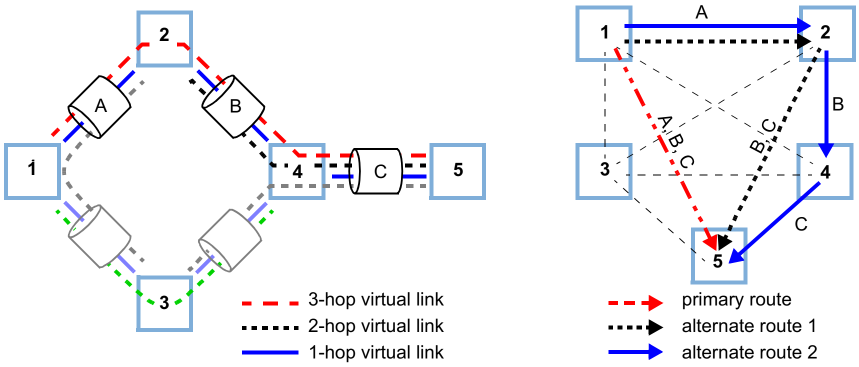


Figure 4.3: Physical fiber topology and bidirectional lightpaths as virtual links (left) and full-mesh virtual topology with primary and alternative routes from node 1 to 5 (right)

links for traffic from node 1 to node 5 respectively, while node 4 only has a single route to send traffic to node 5.

The maximum number of alternative routes, which follow the same links in the physical topology as a given primary route, can be estimated for a full-mesh virtual topology by adding up the number of possible combinations to construct such routes with virtual links. Assuming that the direct virtual link spans H fiber hops and bypasses $H-1$ intermediate nodes, the maximum number of such constrained routes R_{CR} (including the primary route on the direct virtual link) follows as

$$R_{\mathrm{CR}} = \sum_{i=0}^{H-1} \binom{H-1}{i} = 2^{H-1} \tag{4.2}$$

This number is substantially smaller than the number of all possible routes R_{FM}. Thus, constrained alternative routing also reduces the processing requirements in nodes and improves stability similar to other approaches controlling deflection (cf. Section 3.4.4). Note that the actual number of constrained alternative routes can be smaller even for a full-mesh virtual topology depending on the routing of virtual links in the physical topology. For instance, in Figure 4.3 the virtual link from node 1 to node 4 is routed via node 3 instead of via node 2. Thus, it cannot be used for traffic from node 1 to node 5. Such configurations can always occur with non-shortest path routing schemes, e.g., least-cost routing. Also, they occur with shortest path routing schemes if equal length routes exist without a unique tie breaking condition. However, such configurations can either be avoided by explicit routing of virtual links in the physical topology or by using shortest path routing with a unique route length criterion.

Regarding implementation, such a set of alternative routes can be obtained by first computing the primary path from source to destination in the physical topology. Then, a topology graph is set up with the source and destination node as well as all OBTN nodes which are traversed or by-passed by the primary route. Also, all virtual links between those nodes, which are in parallel to the primary route, are entered as edges into this graph. Then, this graph is used to compute the set of alternative routes for the given source-destination node pair with a k-shortest path algorithm [Epp98].

4.2.3 Shared Overflow Capacity

In a network dimensioning process as described in Section 2.4.2, node and network resources are allocated according to traffic demands and primary routes. Introducing alternative routing shifts some traffic from those virtual links, which were dimensioned for this traffic, to other virtual links, which were not. Such an undesired mismatch of routing schemes and dimensioning assumptions can cause a substantial performance degradation as shown in [GKZM05]. To account for traffic on alternative routes, some extra capacity should be allocated in the network. This extra capacity, which is termed *overflow capacity*, should be aggregated on a small number of links and shared among many flows in order to be most effective,

With constrained alternative routing, corresponding alternative routes follow the same fiber links in the physical topology. In addition, the virtual topology was characterized to always contain the (physical) single-hop virtual links, which connect neighboring OBTN nodes. Consequently, one of the alternative routes is always defined by the concatenation of single-hop virtual links. Conversely, a single-hop virtual link is in the set of alternative routes for all primary routes traversing it. These observations and the fact that there are only few single-hop virtual links motivate the allocation of the overflow capacity: It should be assigned to the single-hop virtual links because it can be effectively share there. Thus, this approach is taken by OBTN to combine constrained alternative routing and virtual topology dimensioning.

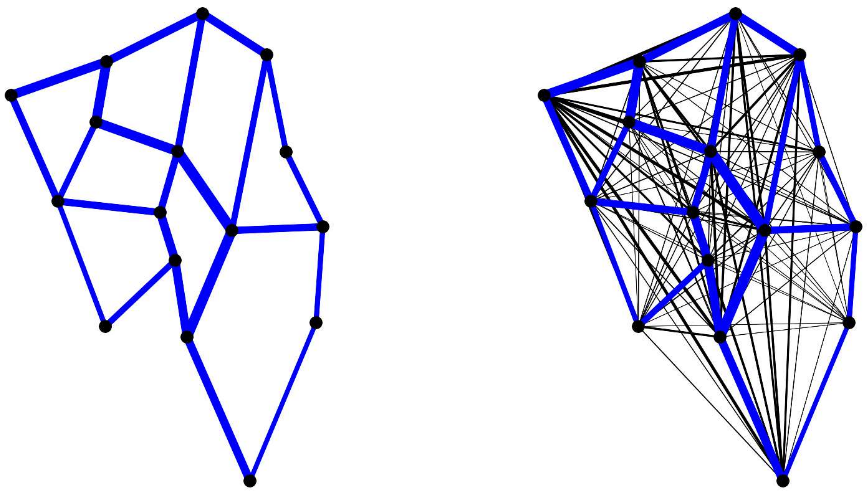
Finally, there is one more argument. The discussion of Figure 4.3 on local forwarding options in constrained alternative routing showed that traffic between neighboring nodes in the physical topology cannot benefit from alternative routing. However, the additional capacity on these virtual links counterbalances this effect.

For the physical topology of the Pan-European reference core network (CN) [MCL$^+$03], Figure 4.4 shows an example OBTN virtual topology with a full-mesh of lightpaths (thin) and with the single-hop virtual links being allocated additional shared overflow capacity (thick, line width not drawn to scale). The dimensioning of virtual links and the assignment of shared overflow capacity is further detailed in Section 5.2.

4.2.4 OBTN Example

Figure 4.5 illustrates the components of OBTN namely the dense virtual topology, the constrained alternative routing, and the shared overflow capacity. Also, it explains how OBTN relates to OBS and BoCS. For a network segment with five nodes, the three subfigures 4.5(a)–(c) depict the virtual topology of the OBTN layer (top half) as well as the interconnection of nodes by lightpaths in the WSN (bottom half), respectively. In both views, lightpaths are depicted as solid arrows. As the WSN layer is assumed to be static, the figure leaves out optical cross-connects for clarity and only shows the physical fiber infrastructure.

For a simple, quantitative comparison, this example includes traffic demands and assumes principal relations on resource sharing. Unidirectional traffic demands $A_{i,j}$ exist from OBTN nodes 1 and 2 to nodes 3 and 5 as well as from node 3 to node 4 and from node 4 to node 5 as depicted by the dashed arrows in Figure 4.5(a) (bottom). For simplicity, statistical multiplexing is modeled by assuming that each demand requires four wavelength channels whenever it shares link



(a) Physical topology (b) OBTN virtual topology

Figure 4.4: Physical and OBTN virtual topology for the Pan-European reference network

resources with other demands and 6 wavelength channels whenever it uses dedicated link resources. Under this assumption, the sub-figures also list example numbers of transmitting burst-switched ports as well as the number of fiber hops in the underlying physical infrastructure—both, per node or link and for the entire network segment. For the physical fiber infrastructure, only one of the nodes 1 or 2 is drawn but both are considered regarding the number of fiber hops as indicated by the labels "2x".

Figure 4.5(a) depicts the OBS case, in which all traffic is multiplexed hop-by-hop onto the links of the physical topology and thus all resources are shared. Figure 4.5(b) shows the BoCS case, which employs direct lightpaths for the traffic demands from nodes 1 and 2 to node 5. Although these lightpaths traverse the same fiber links in the physical topology they bypass intermediate burst-switched nodes. Comparing both scenarios, it can be seen that the total number of burst-switched ports is lower in BoCS. Here, the benefit of bypassing is greater than the penalty of reduced statistical multiplexing. However, BoCS requires more fiber hops because the reduced statistical multiplexing gain solely decides on that. This comparison exhibits the trade-offs first discussed in Section 2.3.4 and quantified in Section 2.4.2.

For the case of OBTN, Figure 4.5(c) also shows the direct lightpaths used for traffic demands from nodes 1 and 2 to node 5. In addition, it illustrates how OBTN allocates a small amount of overflow capacity to the single-hop virtual links. These links are used as alternative routes in case of contention in nodes 1 and 2, respectively. Here, Figure 4.5(c) (bottom) demonstrates again for traffic from node 1 and node 2 to node 5, that the direct lightpaths and the alternative route via nodes 3 and 4 follow the same fiber links in the physical fiber infrastructure.

In Figure 4.5(c), it is assumed that alternative routing reduces the resource requirements of the direct lightpaths connecting nodes 1 and 2 with node 5 from 6 to 5 wavelength channels, respectively. Also, the overflow capacity (dashed arrows) can be rather small as it is shared among all traffic on these single-hop virtual links. In the figure, it is equivalent to 0.5 wavelength

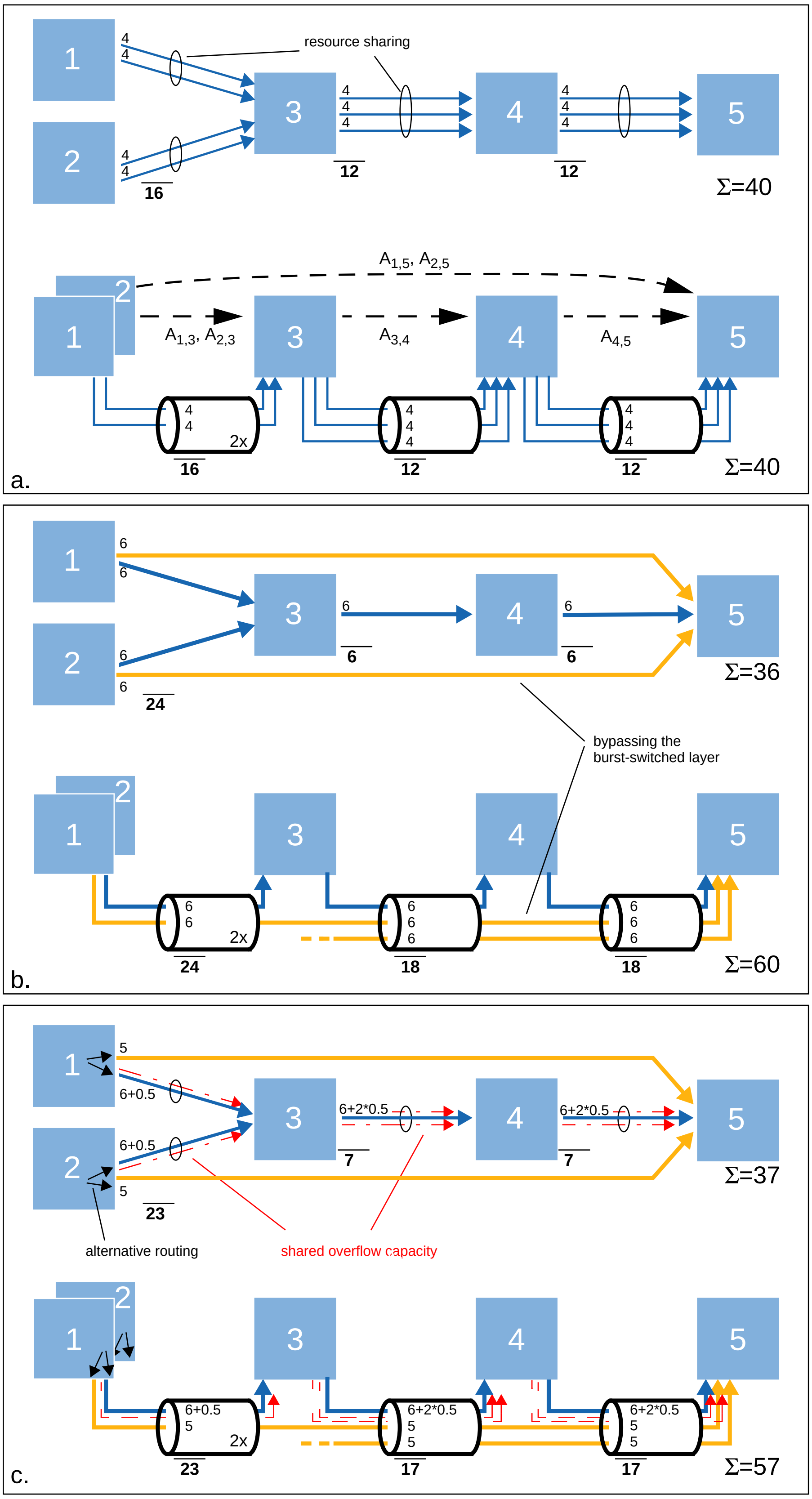


Figure 4.5: Example network segment with five nodes using a. OBS, b. BoCS, and c. OBTN

channels and not rounded up—in a network rounding will not introduce substantial errors due to aggregation. Based on these numbers, OBTN reduces the number of burst-switched ports with respect to OBS almost as effective as BoCS. However, the penalty in the number of hops with respect to OBS is smaller than for BoCS.

The numbers given in this example are solely illustrative and neither their absolute values nor the absolute or relative differences should be considered representative. In realistic networks, reductions in the number of burst ports and penalties regarding the number of fiber hops depend on the absolute size and spatial distribution of traffic demands as well as on the physical network topology (the five segment network has only a mean hop distance for routed traffic demands of 1.67). Chapter 6 presents results for a Pan-European reference network while [GM05a] studied a single node scenario modeling a variety of network characteristics. Summarizing their results, OBTN can successfully reduce the number of burst-switched ports compared to OBS without the penalty in fiber hops of BoCS.

4.2.5 Interaction of OBTN and the Underlying WSN

In the OBTN network architecture, all lightpaths for the resulting virtual topology are provisioned by the underlying WSN layer. In the following, it is always assumed that the capacity of the WSN is no bottleneck as the WSN serves many customers and carries many different client layer networks. Lightpaths can be set up by applying standard mechanisms and protocols of the ASON or GMPLS control planes (cf. Section 2.2.5). Beyond connectivity, dynamic provisioning can offer advanced functionality and flexibility to the OBTN layer. It can adapt virtual link capacities to traffic variations as described in Section 2.2.2 or to achieve survivability towards network failures using restoration schemes of the WSN layer. For these purposes, OBTN nodes should participate in the control plane of the WSN either as clients in an overlay or augmented model or as peers in a peer model (c.f. Section 2.2.5). However, the selection of constrained alternative routes requires either a peer or an augmented control plane model. In the former case, OBTN can access the full state information of the WSN server layer and identify suitable alternative routes on its own. In the latter, the information exchange between the layers has to support constraint-based routing, e.g., by rereturning a list of parallel virtual links.

A consequence of a virtual topology with a large number of links but with fewer wavelengths per link is that outband signaling of burst control headers on a dedicated wavelength per link is no longer efficient. However, alternative out-of-band signaling schemes or advanced inband schemes [CW03] should be applied instead. In the former case, the control information is multiplexed onto the data signal using an orthogonal modulation format (cf. Section 4.3.2). In this case, the transmission and switching infrastructure of the WSN also has to be transparent with respect to those signals. Similarly, the physical network infrastructure has to support burst-mode transmission, although bursts are transported inside the lightpaths.

4.3 Node Architecture

The OBTN network architecture defines requirements for OBTN nodes regarding switch and contention resolution functionality as well as signaling and node control. Edge nodes, which are

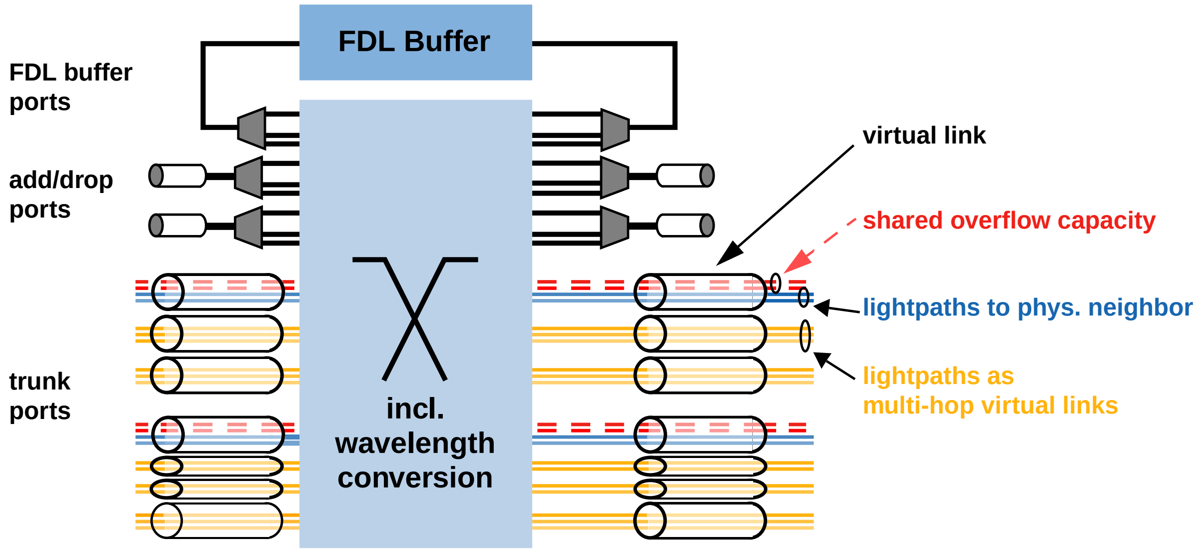


Figure 4.6: Model of an OBTN node

responsible for burst assembly and disassembly as well as the generation of the control header packet, are considered part of the optical feeder network. As they do not fundamentally differ from those in OBS as described in Section 3.1, they are not further detailed here.

4.3.1 Switch and Contention Resolution Functionality

The OBTN nodes have the task of multiplexing/demultiplexing and switching optical burst traffic as outlined in Section 4.1.3. Incoming bursts can either originate in the attached optical feeder networks, where burst assembly was performed, or arrive from other OBTN nodes on lightpaths. They are either switched to a virtual link belonging to one of the constrained alternative routes towards their destination or dropped to their destination feeder network.

As an additional option, an OBTN node can also have local add and drop ports with burst assembly and disassembly. Here, as an advanced feature, local burst transmission could be coordinated with the outgoing ports in order to prevent contention in the source node. However, as the focus lies on the transport network scenario outlined in Section 4.1.3, local add/drop ports are not considered in the following.

Figure 4.6 depicts a model of an OBTN node with trunk ports, add/drop ports, and an FDL buffer. It shows the virtual links comprising lightpaths to physical neighbors, multi-hop lightpaths to distant OBTN nodes, and lightpaths for shared overflow capacity. They are colored according to the example in Figure 4.5. Although OBTN nodes are connected to a large number of nodes in the network the principal node design does not deviate from that of OBS. This is due to the fact that switching matrix structure and complexity are determined by the physical instead of the logical interconnection pattern. Thus, existing proposals for OBS nodes can be applied to OBTN as well, including the SOA-based or AWG-based designs discussed in Section 3.5. More detailed studies on node realization or an integrated performance and technology evaluation are beyond the scope of this thesis and open for future work.

In order to achieve high network resource efficiency despite the reduced statistical multiplexing gain per network link OBTN nodes employ both wavelength conversion and a simple FDL buffer. Figure 4.6 shows an OBTN node model including the FDL buffer. Still, as FDL buffers cannot be integrated, and require additional switch ports, their FDL and wavelength count, as well as their delay should be kept to the necessary minimum [GBPS03].

As the WSN network layer is assumed to be a service-independent server layer a close integration of OBTN and WSN switching nodes is not envisioned. Thus, the WSN node architecture is not further detailed here. Nevertheless, in future work, advanced WSN node architectures with only partial burst-mode capabilities trading-off full flexibility and realization complexity could be considered.

4.3.2 Signaling and Node Control

Control packet signaling and node control functionality like burst scheduling are design aspects beyond the pure switching functionality, which also require considerable realization effort. As discussed in Section 4.2.5, control headers should not be transported on a dedicated wavelength per network link in OBTN but instead use alternative outband or inband schemes [Nor03]. A promising technique for outband signaling is to impress the control information on the same wavelength, which carries the intensity modulated data payload, but with a different, orthogonal modulation format. In [VZC+03, KSM+02], frequency-shift keying (FSK) and differential phase-shift keying (DPSK) are proposed for this purpose and, in case of FSK, also successfully demonstrated. Also, sub-carrier modulation (SCM) [BCR+99] is a feasible alternative solution that has been reported in literature and been experimentally demonstrated. With SCM, the control information modulates a radio-frequency, which again modulates the optical carrier creating two side-bands around the optical center frequency.

As only few control information is signaled on the orthogonal control channel it can be modulated with a lower bit-rate reducing physical impairments. Also, this channel can be operated in continuous mode as it is purely a point-to-point link between OBTN nodes. Consequently, such transmitters and receivers can be realized far less expensive [BCR+99, YMK05] than line-speed burst-mode equipment. This allows them to be deployed in all OBTN nodes for all wavelengths. As an additional advantage, such schemes facilitate the monitoring of control headers and data bursts in order to ensure their consistency, which is a major problem when using dedicated wavelengths for control header signaling.

The advanced inband signaling scheme for control packets proposed in [CW03] interleaves burst header packets with data bursts. By default, a switch listens to all wavelengths to hunt for and process control packets. As all data bursts are announced by a control packet with a certain offset time, the switch exactly knows when in the future data bursts will arrive and can thus change from the listening to a switching mode immediately before burst arrival. Compared to the outband schemes, this, however, requires that all wavelength channels are terminated to read control packets which would diminish the advantages of transparent optical switching.

Burst scheduling can be realized in the same way as in OBS nodes with one scheduling module per node output [JG03b, Jun04, Jun06]. Using one scheduling module per node output requires extensions to account for several virtual links per output. In contrast, a scheduling module per *virtual* link does not require such extensions but instead leads to more inter-module coordination for alternative routing and mandates that the scheduling logic be reconfigured/reprogrammed upon virtual topology changes. Similarly, in the case of FDL buffering, the burst scheduling modules for the FDL buffer and the respective output link have to be coordinated—also as in regular OBS.

4.4 Architectural Variants and Operational Strategies

The OBTN architecture introduces several degrees of freedom regarding the dense virtual topology, constrained alternative routing, and shared overflow capacity. Beyond the above description of the principal concepts, this section extends the discussion towards more advanced architectural and operational options. First, this section discusses basic strategies for virtual topology design, which decide on the interconnection of OBTN nodes by lightpaths. Then, it further details routing and forwarding and looks at resource sharing concepts for virtual link capacity. Finally, hunting modes for resource selection inside the OBTN nodes are discussed.

4.4.1 Virtual topology design

So far, the OBTN virtual topology has only been characterized to be dense and to always contain the single-hop virtual links connecting neighbor nodes in the physical topology. Although densely-meshed virtual topologies are common in core networks, scalability, management, and resource utilization arguments motivate a lower connectivity. Therefore, two alternative approaches are introduced and evaluated in Section 6.6 to show that the virtual topology can be less densely-meshed without fundamental disadvantages.

In the context of this thesis, the OBTN virtual topology is assumed to be static or to be pseudo-static, i.e., only adapted on time-scales much higher than end-to-end delays of bursts. Nevertheless, virtual topology reconfiguration/reoptimization as proposed for WSN [RR00, NTM00, GM03] could be considered in future work.

4.4.2 Routing and Forwarding

Section 4.2.2 already discussed the principal concept of the constrained alternative routing scheme used in OBTN. Here, the route selection process with respect to the responsible node(s) and the maximum number of alternative routes is discussed.

Similar to burst reservation and scheduling, the route is selected from the constrained alternative set of routes on a per burst basis in OBTN, i.e., in case a burst cannot take its primary route, it selects an alternate route. A major classification criterion for route selection is which nodes decide on the route. In source routing, only the source node is allowed to switch to an alternate route, which leaves more control with the originating OBTN node and can be more simply integrated into MPLS or GMPLS control planes. In contrast, link-by-link routing, where a link refers to a virtual link, also allows intermediate nodes to decide on the further route to the destination node. As the latter approach works greedily in intermediate nodes it offers more opportunities for successful transmission, can be expected to yield a lower burst loss probability, and is thus used in OBTN. For the purpose of link-by-link routing, routing information specified in the set of constrained alternative routes can be extracted into local forwarding tables to speed-up forwarding look-up. In general, such forwarding tables list a set of next-hop virtual links for bursts of a given source and destination node pair.

The route selection process is further characterized by the maximum number of alternate routes to be used and the order of route selection. Except for overload situations, increasing the number of alternate routes can yield a reduced burst loss probability [Sch04]. Still, as it also imposes additional load on the node control module, a suitable balance has to be found here. Prominent examples for the order of route selection use fixed sequential, random or load-dependent schemes. In OBTN, selecting alternate routes sequentially according to their length in virtual links can control the transit traffic induced by alternate routes and thus support the overall design objectives. However, selecting the routes according to link dimensioning, e.g., with shared overflow capacity, may also be beneficial. Allowing only the alternate route along the single-hop virtual links leads to an exact match of routing and allocation of shared overflow capacity. In all other cases, the alternate route either matches this allocation only partly or the alternate route uses virtual links from which other contending bursts are in turn shifted to an alternate route. There, these bursts may again benefit from the shared overflow capacity.

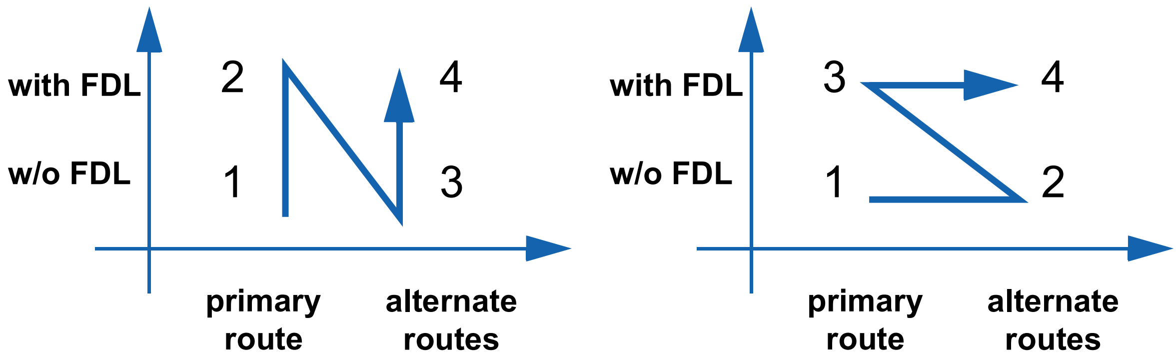
As the constrained set of routes is assumed to be fixed over longer time intervals and as the routes are assumed to be hunted sequentially, routing in OBTN can be summarized as fixed alternative routing.

4.4.3 Resource sharing

The *shared overflow resources* in the network could either be completely shared among all traffic traversing a virtual link or be restricted to bursts on alternate routes. In addition, the partitioning on the link could be managed by admission control strategies like partial sharing, trunk reservation, virtual or full partitioning [Ros95]. While sharing would improve the statistical multiplexing gain and yield a lower burst loss probability, partitioned resources could be, e.g., used for differentiated provisioning or recovery. Following the design objectives of OBTN, only the complete sharing approach is considered in the following in order to achieve a high network efficiency. Nevertheless, the other approaches can be subject of future work.

4.4.4 Hunting modes

The OBTN node control assigns resources inside the node and on the virtual links on a per burst basis. In OBTN, a burst initially attempts a transmission on the primary route without wavelength conversion and without buffering as was already motivated for OBS in Section 3.4. In case this fails because no idle resources are available or because the output link supports different wavelength channels, the contention situation has to be resolved by wavelength conversion, buffering in an FDL or transmission on the alternate route. The process and order in which these contention resolution options are used is again referred to as *hunting* (cf. Section 3.4.5).



(a) FDL-Alt (b) Alt-FDL

Figure 4.7: Hunting modes for OBTN considering FDL buffering and alternate routes

In the following, it is always assumed that the node uses wavelength conversion as the primary contention resolution strategy. Not only has it been shown to be most effective (cf Section 3.4), also do several burst-switched node architectures inherently provide this capability (cf. Section 3.5 and [Buc05]). In case the nodes do not offer full wavelength conversion, shared converter pools can be introduced and dimensioned to yield similar performance at reduced converter count (cf. Section 3.4 and [Gau04]). Consequently, the remaining dimensions of hunting are FDL buffering and transmission on alternate routes.

In a two-dimensional representation, Figure 4.7 depicts two sequential hunting modes. The first mode is called *FDL-Alt* (Figure 4.7(a)) and selects the primary route first without, and then with, the buffer; only if this fails, it uses the alternate routes first without, and then with, the buffer. The second mode is referred to as *Alt-FDL* (Figure 4.7(b)). It first tries the primary route and then the alternate routes, first without the buffer and only if this fails, with the buffer. Obviously, both hunting modes correspond to each other for OBTN nodes in which no alternate route is available.

In addition to sequential hunting, integrated hunting could improve the performance by selecting the optimal contention resolution strategy per burst according to some fitness function. Also, hunting modes could adapt to traffic and network load situations by exploiting state information of the primary route, the first hop of different alternate routes or the FDL buffer. Finally, hunting strategies can depend on the QoS class, e.g., a low priority class may only use a subset of the options or a different order. However, such approaches increase the coordination effort and the realization complexity of the node control and are thus not considered in the following.

4.5 Qualitative Discussion

After the previous sections presented the fundamental concepts of OBTN, this section summarizes the chapter and compares OBTN qualitatively with related architectures.

Based on the classification in Section 2.3.4 and [GvBK+06], OBTN belongs to the client-server type of hybrid optical network architectures as it employs the burst-switched client layer on top of a wavelength-switched server layer. OBTN targets the burst transport network scenario outlined in Section 4.1.3 for a scalable integration of optical feeder and core networks as well as for migration of burst-switching into WSN. Also, it follows a cost-aware design for burst-switched networks respecting the decreasing cost for network capacity and dominating cost for switch ports. Finally, it aims at an overall high QoS by applying effective contention resolution schemes.

4.5.1 Comparison to OBS and Hybrid Optical Network Architectures

OBTN is most closely related to OBS and BoCS and thus first compared to them regarding performance, resource requirements, and flexibility. After the comparison with OBS and BoCS the OBTN architecture is also shortly contrasted with parallel and integrated hybrid optical networks (cf. Section 2.3.4 and [GvBK+06]).

OBTN inherited several basic architectural components like burst reservation, burst scheduling, and contention resolution from OBS but incorporates additional concepts on a network level. While almost all work on OBS deals with isolated core network scenarios in green-field deployments, OBTN builds on the existing WSN by setting up a virtual topology, by using constrained alternative routing, and by allocating shared overflow capacity. In principle, OBS and OBTN can both achieve a very low burst loss probability if respective contention resolution schemes are applied. However, the constrained alternative routing in OBTN avoids the route and thus propagation delay variability inherent to most deflection routing schemes in OBS. Regarding dimensioning, OBS can, on the one hand, be expected to require less fiber hops for a given offered traffic and burst loss probability as it benefits from a higher statistical multiplexing gain. On the other hand, the large amount of transit traffic in OBS nodes also leads to a higher number of switch ports, on which OBTN economizes. Both effects will be quantified in Section 6.1.

OBTN is also closely related to the two network architectures outlined in [OSHT01] which deploy OPS networks in combination with WSN—from an architectural point of view OPS and OBS are interchangeable here. The OPS layer either functions as a complete network layer or as an interconnection network between aggregation edge nodes. From the paper it can be derived that the former architecture uses lightpaths as virtual links to form an only sparsely meshed virtual topology. In contrast, the latter BoCS architecture sets up a dense or full-mesh virtual topology completely eliminating switching in intermediate nodes.

OBTN extends the BoCS approach by allowing constrained alternative routing and assigning the shared overflow capacity. These advanced concepts allow to better utilize resources than BoCS and will be shown to reduce both, the required network capacity and the number of switch ports. Also, they offer a higher adaptivity due to the flexibility of the routing scheme. This performance advantage, however, comes at the cost of a higher complexity in node control and routing.

OBTN specifically supports a *grow-as-you-go* evolution of OBS networks by first upgrading shared single-hop virtual links for smaller demand increases and later off-loading qualifying traffic demands to end-to-end virtual links. Similarly, an initial BoCS network can be gradually expanded and improved regarding QoS by adding resources to the shared single-hop virtual links and adopting the constrained alternate routing.

The major difference between OBTN and the virtual optical networks (VON) [QY99], polymorphic multi-service optical networks (PMON) [dM$^+$04] as well as the hybrid OBS proposals [LWZ$^+$03, XQYD03] is that the latter deploy the burst-switched and the wavelength-switched network in parallel instead of in a client-server fashion. Edge nodes select a network technology based on explicit user request, traffic characteristics or QoS requirements, there.

The hybrid optical network architectures ORION [vBCC$^+$03] and OpMigua [Bjo04] closely integrate an electronically packet-switched network layer with a wavelength-switched network layer. For instance, ORION inserts packets into transmission gaps of bypassing lightpaths, which are used by different services and are not terminated in the respective node. Thus, this concept requires a fully integrated control plane for routing and faultless delivery. Also, it relies on specialized devices to physically mark the packets and extract them in downstream nodes. Similarly, the OpMigua approach transports packet-switched and circuit-switched traffic together and uses a different polarization state to classify the traffic in nodes and to direct it to the respective switching matrix. Summarizing, the key difference here is that OBTN does not apply this extreme level of data plane and of control plane integration.

4.5.2 Scalability and Typical Dimensionings

The dense or even full-mesh virtual topology of OBTN and BoCS raises the so-called n^2 problem, i.e., the problem that a full-mesh interconnection of n network nodes requires in the order of n^2 network links. This mostly limits the scalability for growing n. However, current WDM networks actually motivate densely meshed virtual topologies because of their large number of wavelengths per fiber (80–160 are available in products) and an only limited number of nodes (usually 10–20). In addition, the concept of traffic off-loading and optical bypassing is supported by the fact that network capacity is relatively inexpensive compared to node equipment.

Even more, the multiplexed traffic demands exchanged between all MANs of a pair of multiplexing/demultiplexing nodes (cf. Figure 4.2) can very well amount to several wavelength channels per virtual link. This can be seen from a short estimation based on the average number of MANs per core node, n_{MAN}. Typical values of n_{MAN} range between 2 and 5 which can be derived from statistics on the number of MANs and the number of points-of-presence [MNVW04]. Hence, even traffic demands between pairs of MANs only amounting to a fraction of a lightpath capacity, e.g., 20 %, can result in an aggregate traffic demand exchanged between a pair of multiplexing/demultiplexing nodes equivalent to several lightpaths assuming n_{MAN}^2 traffic relations among them. The situation that several parallel wavelengths instead of only one at higher bitrate interconnect core nodes is also backed by technology and transmission arguments. Some experts expect 10 Gbps equipment to stay in carriers' networks for several years, maybe even decades, similar to the *Jumbo jet* in airline fleets [Lan03].

The OBTN network architecture introduced in this section is modeled in Chapter 5 and quantitatively evaluated in Chapter 6. The focus of these evaluations is to demonstrate that OBTN meets the expectations towards a future optical network architecture in general and the above design rationale in particular.

5 Modeling and Dimensioning of Burst Transport Networks

The previous chapters described existing architectures for burst transport networks, namely Optical Burst Switching (OBS) and Burst-over-Circuit Switching (BoCS), and introduced the new Optical Burst Transport Network architecture (OBTN). Section 4.5 already discussed that OBS and BoCS can be considered as extreme cases of an OBTN architecture regarding virtual topology dimensioning. Based on this interpretation, this chapter presents a unified resource modeling and dimensioning approach. This allows to compare the three architectures comprehensively in Chapter 6.

First, this chapter describes and justifies an abstract resource model for burst-switched nodes and the unified modeling approach for OBS, BoCS, and OBTN. Then, it defines a unified process for dimensioning client and server layer resources and discusses implementation constraints. Finally, the last section looks at the performance evaluation methodology including key performance and resource metrics, and the evaluation scenario.

5.1 Unified Resource Modeling for OBS, BoCS, and OBTN

The design rationale for OBTN in Section 4.1 stressed that resource and cost arguments be considered in network design. Particularly, the focus is on the complexity of the burst-switched nodes in terms of the number of switch ports. Therefore, this section outlines a unified model and metrics for client layer and server layer resources of all three architectures.

As OBS, BoCS, and OBTN nodes require identical switch functionality (cf. Section 4.3) they can be realized based on the same node architecture. They only differ with respect to node control functionalities, such as routing or hunting. Thus, the resource requirements of the three burst-switched architectures can be compared based on a unified abstract node and network model.

In order to be independent of specific burst switch architectures and technologies (cf. Section 3.5), node complexity is measured in terms of the total number of switch ports[1]. From this complexity metric an architecture-specific analysis regarding component count, signal regeneration requirements or cost can be derived in a second independent step. Here, an integrated evaluation as discussed in Section 3.5.4 can be applied.

[1] In this thesis, only the sending switch ports are counted assuming symmetrical configurations.

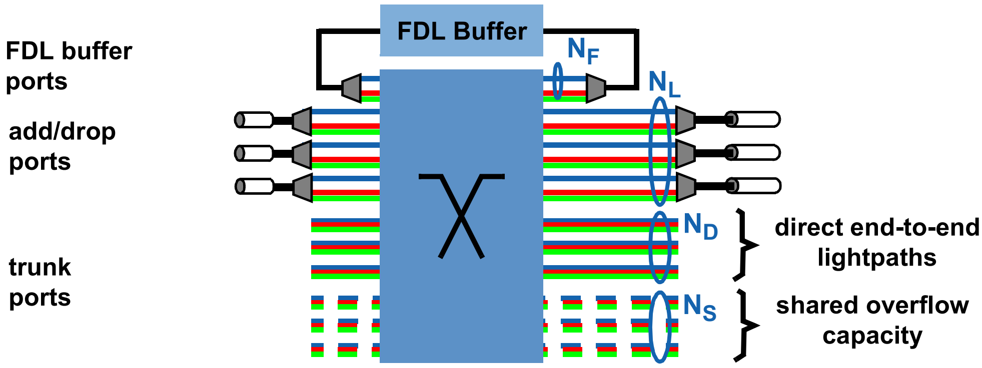


Figure 5.1: Unified dimensioning view of an OBS, BoCS, and OBTN node

As can be seen from Figure 5.1, the switch ports of OBS, BoCS, and OBTN nodes comprise FDL buffer ports, add/drop ports, and trunk ports, which connect the node to other nodes via virtual links. The number of FDL buffer ports, N_{F}, and the number of local add/drop ports, N_{L}, can be derived from performance objectives for contention resolution and from local add/drop traffic demands, respectively. In contrast, the number of trunk ports is obtained from the virtual topology design and the dimensioning process outlined in Section 2.4.2. For OBTN, there are two types of virtual links (cf. Figure 4.6) according to their functionality. The total number of switch ports terminating direct end-to-end lightpaths is denoted as N_{D} while the total number of switch ports used for shared overflow capacity is denoted as N_{S}.

In addition to the resources of the burst-switched client layer nodes, the server layer resources required to provision the virtual links are also modeled. These virtual links are provisioned by the underlying WSN as lightpaths. In order to stay independent of the implementation of the server layer nodes, the resources required for a lightpath are measured in terms of the total number of fiber hops it spans in the network. This number abstracts both server layer node and link resources. The dimensioning and all evaluations assume that nodes in the client and server layer are collocated in a point-of-presence (POP). Furthermore, the computation of the number of fiber hops does not model any additional server layer nodes between POPs. Similar to the number of trunk ports above, the number of fiber hops comprises a number S_{D} used for direct end-to-end lightpaths and a number S_{S} used for shared overflow capacity.

The WSN server layer is typically reconfigured on time-scales ranging from seconds to minutes or even far above. Thus, it can be realized by equipment with different reconfiguration speed including optical cross-connects, managed patch-panels or static interconnections. Typical time-scales of the burst-switched client layer, which range from microseconds to milliseconds, are much smaller than those of the WSN server layer. Therefore, the WSN layer can be modeled to be static from the point of view of the burst-switched layer. Consequently, the server layer dimensioning outlined in Section 2.4.2 is applicable.

As mentioned above, the distinct difference in terms of virtual topology dimensioning between OBS, BoCS, and OBTNis the amount of shared overflow capacity: BoCS applies a dense or full-mesh virtual topology with direct end-to-end lightpaths and is thus equivalent to OBTN *without* any shared overflow capacity. In contrast, OBS only interconnects neighbor nodes in the physical topology and is thus equivalent to OBTN *with only* shared overflow capacity. Thus, the share of (physical) network capacity assigned as shared overflow capacity, β, can be used to characterize the three burst transport architectures by a single unified figure:

$$\beta = \frac{S_S}{S_S + S_D} = \begin{cases} 0 & \text{BoCS} \\ 0 < \beta < 1 & \text{OBTN} \\ 1 & \text{OBS} \end{cases} . \tag{5.1}$$

Note that this unified modeling relates to the virtual topology dimensioning. OBS, BoCS, and OBTN can differ regarding operational strategies, e.g., in the their routing flexibility.

Although the WSN in general is assumed to act as a server layer to many client layer networks, the following discussion assumes that its capacity is tightly dimensioned for the considered client layer network. Consequently, effects due to the discrete granularity of network equipment are not modeled.

5.2 Unified Dimensioning for OBS, BoCS, and OBTN

This section explains the dimensioning of client layer and server layer resources as used in the following performance evaluation. Building on the above classification of ports this section first treats the virtual and physical topology dimensioning and then looks at the FDL buffer and local add/drop resources.

5.2.1 Client Layer and Server Layer Resources

The dimensioning process in Section 2.4.2 is used to determine the trunk ports in the client layer and the fiber hops in the server layer. However, to date, there are no analytical models for burst-switched nodes with asynchronous, WDM multi-FDL buffers using PreRes buffer scheduling (cf. Section 3.4.3). Therefore, the virtual and physical network dimensioning cannot exactly compute the link capacities required to transport traffic demands at a given QoS. In the used formalization, this means that there is no suitable dimensioning function $\gamma(\cdot)$ available to analytically relate offered traffic, QoS, and link resources. Therefore, the client layer and server layer resources are instead dimensioned for a given overprovisioning factor by using an alternate dimensioning function $\tilde{\gamma}(\cdot)$ such that the QoS is (approximately) the same on all network links. Then, a systematic performance evaluation with differently scaled node and network capacities identifies the dimensioning delivering the required target QoS.

The proposed dimensioning approach works for β values covering OBTN ($0 < \beta < 1$), OBS ($\beta = 1$), and BoCS ($\beta = 0$) based on the unified resource model introduced in Section 5.1. Apart from β, the traffic demand matrix $\mathbf{A}$ and the target overprovisioning factor α (cf. Eq. (2.14)) serve as input parameters. In addition, a densely meshed topology graph $\mathbf{T}_V$ is used for the direct end-to-end lightpaths and the physical topology graph $\mathbf{T}_P$ for the shared overflow capacity. Finally, the routing of traffic onto the topology is given as the routing function ϕ.

The unified dimensioning process for OBTN is depicted in Figure 5.2. It derives the client layer capacities for the basic dense virtual topology (direct end-to-end lightpaths), $\mathbf{C}_D$, and for the shared overflow capacity, $\mathbf{C}_S$, individually from separate resource pools. Then, they are added

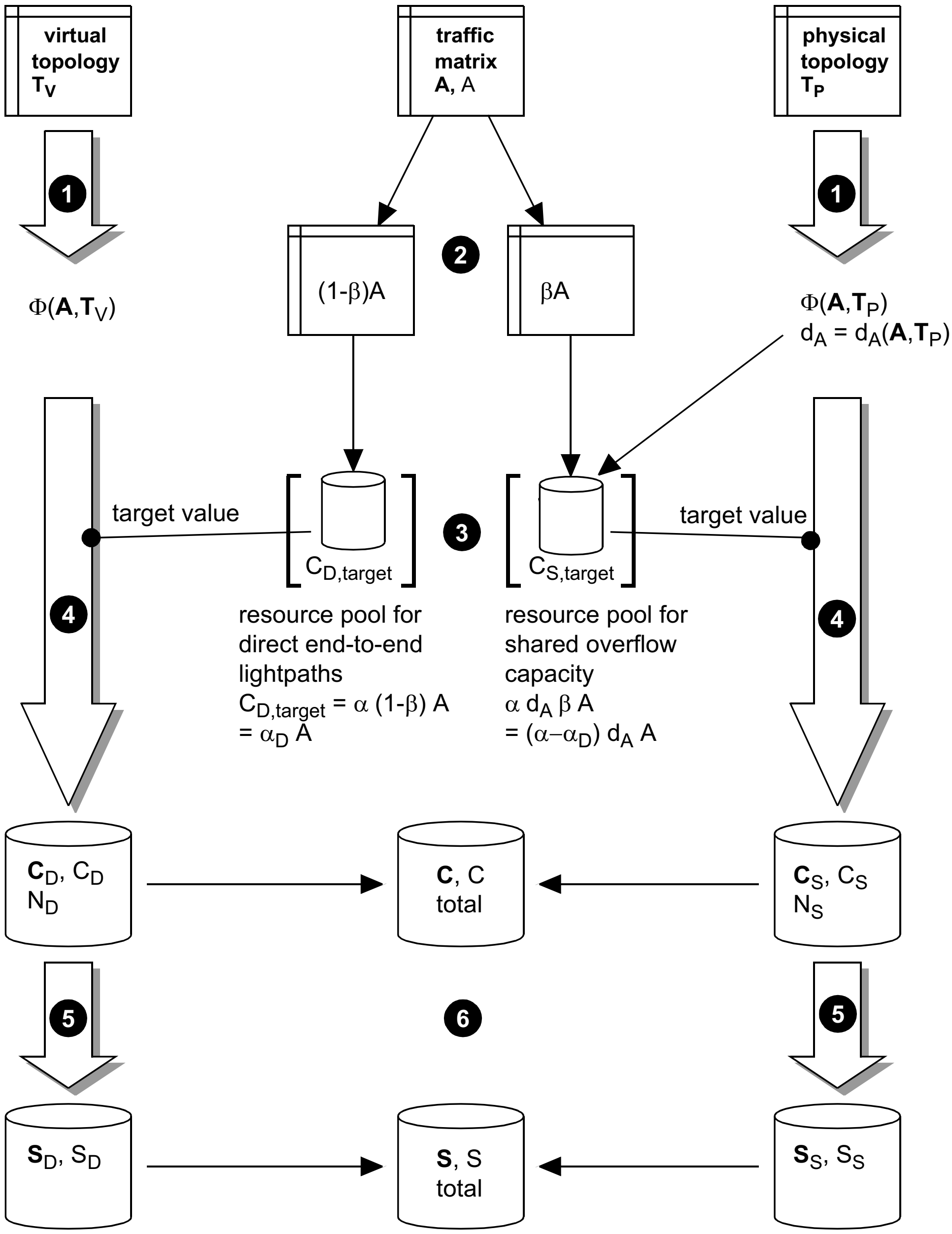


Figure 5.2: Dimensioning process for OBTN client and server layer resources

to obtain the final virtual topology dimensioning. These client layer resources are individually mapped to the server layer and added up, respectively. For the extreme case of BoCS this means that only direct end-to-end lightpaths are dimensioned while for OBS only shared overflow capacity is considered.

The following steps further detail the dimensioning process for the two resource pools of direct end-to-end lightpaths and shared overflow capacity:

1. The traffic demand matrix $\mathbf{A}$ is routed on the respective topology graphs $\mathbf{T}_\mathrm{V}$ and $\mathbf{T}_\mathrm{P}$ to obtain offered load values $\phi(\mathbf{A},\mathbf{T}_\mathrm{V})$ and $\phi(\mathbf{A},\mathbf{T}_\mathrm{P})$ for all virtual topology links[2]. At the same time, the mean hop distance in the physical topology $d_\mathrm{A}(\mathbf{A},\mathbf{T}_\mathrm{P})$ is calculated.

2. The total traffic demand A is split into two parts according to β. The first part $(1-\beta)\cdot A$ is used to dimension the direct end-to-end lightpaths. The second part $\beta \cdot A$ is used to dimension the shared overflow capacity.

3. The total target capacity of the virtual topology links are calculated for both resource pools based on Eq. (2.13) as:

$$C_\mathrm{D,target} = \alpha \cdot (1-\beta) \cdot A \tag{5.2}$$
$$C_\mathrm{S,target} = \alpha \cdot \beta \cdot d_\mathrm{A}(\mathbf{A},\mathbf{T}_\mathrm{P}) \cdot A. \tag{5.3}$$

   Instead of varying both $C_\mathrm{S,target}$ and $C_\mathrm{D,target}$, the amount of direct end-to-end lightpath capacity $C_\mathrm{D,target}$ is fixed in the evaluations. This allows to only vary the amount of shared overflow capacity $C_\mathrm{S,target}$. It is achieved by varying β together with α in Eq. (5.2) and Eq. (5.3) such that

$$\alpha_\mathrm{D} = (1-\beta)\alpha = \text{const.} \tag{5.4}$$

   Consequently, Eq. (5.2) and Eq. (5.3) can be rewritten as

$$C_\mathrm{D,target} = \alpha_\mathrm{D} \cdot A \tag{5.5}$$
$$C_\mathrm{S,target} = (\alpha - \alpha_\mathrm{D}) \cdot d_\mathrm{A}(\mathbf{A},\mathbf{T}_\mathrm{P}) \cdot A. \tag{5.6}$$

4. The virtual topology links are dimensioned individually such that all links offer the same QoS and that the total capacities match the target values derived from Eq. (5.2) and Eq. (5.3), respectively. The algorithm to meet the QoS and the capacity constraints is further detailed below step 6.

$$\mathbf{C}_\mathrm{D} = \tilde{\gamma}(\phi(\mathbf{A},\mathbf{T}_\mathrm{V})) \quad \text{s.t.} ||\mathbf{C}_\mathrm{D}|| = \mathrm{C}_\mathrm{D,target} \tag{5.7}$$
$$\mathbf{C}_\mathrm{S} = \tilde{\gamma}(\phi(\mathbf{A},\mathbf{T}_\mathrm{P})) \quad \text{s.t.} ||\mathbf{C}_\mathrm{S}|| = \mathrm{C}_\mathrm{S,target} \tag{5.8}$$

   As each virtual link is terminated by exactly one sending port, the number of trunk ports for direct end-to-end lightpaths N_D and for shared overflow capacity N_S can be calculated for the individual nodes as the row sums of $\mathbf{C}_\mathrm{D}$ and $\mathbf{C}_\mathrm{S}$, respectively.

5. The virtual links are routed onto the physical topology to obtain the dimensioning matrices $\mathbf{S}_\mathrm{D}$ and $\mathbf{S}_\mathrm{S}$ of the WSN

$$\mathbf{S}_\mathrm{D} = \phi(\mathbf{C}_\mathrm{D},\mathbf{T}_\mathrm{P}) \tag{5.9}$$
$$\mathbf{S}_\mathrm{S} = \phi(\mathbf{C}_\mathrm{S},\mathbf{T}_\mathrm{P}) = \mathbf{C}_\mathrm{S} \tag{5.10}$$

   as well as the total number of fiber hops S_D and S_S, respectively.

6. Finally, the dimensioning matrices for the direct end-to-end lightpaths and the shared overflow capacity are added up to obtain $\mathbf{C} = \mathbf{C}_\mathrm{D} + \mathbf{C}_\mathrm{S}$ for the client layer and $\mathbf{S} = \mathbf{S}_\mathrm{D} + \mathbf{S}_\mathrm{S}$ for the server layer resources.

[2] Here, a demand matrix scaled to the target capacity could also be used leading to slightly different results. For instance, Section 6.6 takes this approach to simplify implementation there. However, as the differences are minimal this degree of freedom is not further detailed.

For all dimensionings in Chapter 6, the inverse of the Erlang-B formula is used as the alternative dimensioning function $\tilde{\gamma}$ in step 4. The required number of wavelengths (servers) $c(a_{i,j})$ for a virtual link carrying the traffic $a_{i,j}$ is computed for the target burst loss probability P_L assuming Poisson arrivals and general burst transmission times as

$$c(a_{i,j}) = B^{-1}(a_{i,j}, P_L). \tag{5.11}$$

The implemented dimensioning algorithm iteratively adapts the link capacities until the total target capacity is met and all links have (approximately) the same QoS. Note that the resulting equilibrium QoS value has no further relevance for the network performance analysis as the Erlang-B formula only serves as an alternative dimensioning function but does not exactly model the burst-switched nodes with FDL buffers. When suitable performance models become available for such nodes, they can be easily integrated into this process instead.

As virtual links can only comprise an integer number of lightpaths, it is in general impossible to exactly match the constraints regarding QoS and total target capacity at the same time. Thus, optimized rounding functions for minimizing the effects of residual capacities [BBB+03, HHN] are used to obtain sufficiently good matching. Regarding Eq. (5.7) and Eq. (5.8), these functions are assumed to be included in the dimensioning matrix transformation $\tilde{\gamma}(\cdot)$.

5.2.2 FDL Buffer and Add/Drop Ports

So far, this dimensioning approach only determines the number of trunk ports but not the number of FDL and add/drop buffer ports. The structure and dimensioning of the FDL buffers is modeled to be homogeneous across all n nodes. Thus, the total number of FDL buffer ports in the network equals $N_{\text{F,total}} = \sum_{i=1}^{n} N_{\text{F,i}} = n \cdot N_{\text{F}}$.

As the focus of the following performance evaluation is on the core network, the attached optical feeder networks are not modeled in detail. Instead, they are abstracted by the respective traffic demands exchanged with the respective core node and the required number of add/drop ports. In order to avoid side-effects because of rounding, the total number of add ports in the network is calculated. The optical feeder networks are modeled to all carry their traffic with the same utilization ρ_{MAN} on the add/drop ports. Thus, for a total traffic demand of A injected into the core network, the total number of add ports in all n core nodes can be modeled as $N_{\text{L,total}} = \sum_{i=1}^{n} N_{\text{L,i}} = \lceil A/\rho_{\text{MAN}} \rceil$.

5.3 Performance Evaluation Methodology

While the previous subsections concentrate on the modeling and dimensioning of client layer and server layer resources, this subsection introduces the evaluation methodology used in Chapter 6. This comprises the simulation model and environment, the evaluation metrics, and the evaluation scenario.

5.3.1 Simulation Model and Environment

In order to analyze the performance of the three burst transport architectures OBS, BoCS, and OBTN, event-driven simulation is used. Abstract simulation model which reflects exactly the required level of detail for the different model components allows for a thorough, flexible, and comprehensive evaluation. In addition, intelligent component behavior described by state-dependent or adaptive algorithms, e.g., the scheduling algorithms for FDL buffers, can be directly accounted for. Finally, simulation offers access to a broad range of statistics fostering detailed system understanding and explanation. Simulation is also the methodology of choice as there are still no analytical models for the used FDL buffers with integrated scheduling as discussed in Section 3.4.

The functionality, the time-scales, and QoS metrics, which are relevant for the following performance evaluation, all focus on the burst-level. Therefore, the software tool simulates the transport of individual optical bursts in the core network. Particularly, components and functionality for contention resolution by wavelength conversion, FDL buffering, and alternative routing are modeled in detail and can be parameterized flexibly. In contrast, the assembly of bursts from packets in the edge nodes of the optical feeder networks and the feeder networks themselves are abstracted by burst traffic generators attached to the core nodes.

The simulation model is implemented using the IKRSimLib [IKR] which provides the simulation control, the basic model and communication components, as well as I/O support. Also, the simulation tool applies the IKRSimLib capabilities regarding random variate generation and statistical analysis. For instance, the latter functionality provides the 95 % confidence intervals shown in the diagrams of Chapter 6 based on the batch-means method. The dimensioning approach described in Section 5.2 for client layer and server layer resources is integrated into the simulation tool using a class library for WDM network dimensioning [Köh05a].

5.3.2 Evaluation Metrics

This subsection introduces and classifies the metrics that quantify the QoS performance and realization complexity in Chapter 6. Also, it explains how resource metrics can be obtained for a given target QoS value.

5.3.2.1 Performance and Complexity Metrics

While *direct* metrics like burst loss probability, transfer delay or jitter quantify the achieved QoS, *indirect* metrics like the distribution of the number of virtual links traversed support explanations.

As current transport networks are mostly static and circuit switched, QoS requirements directly applicable to burst/packet-switched optical networks do not exist yet. However, they can be derived from ITU-T's recommendation on provisional IP QoS classes and network performance objectives [Y.1541]. This standard defines performance objectives for international end-to-end IP network paths and thus for entire IP client layer networks. As all specified classes require an

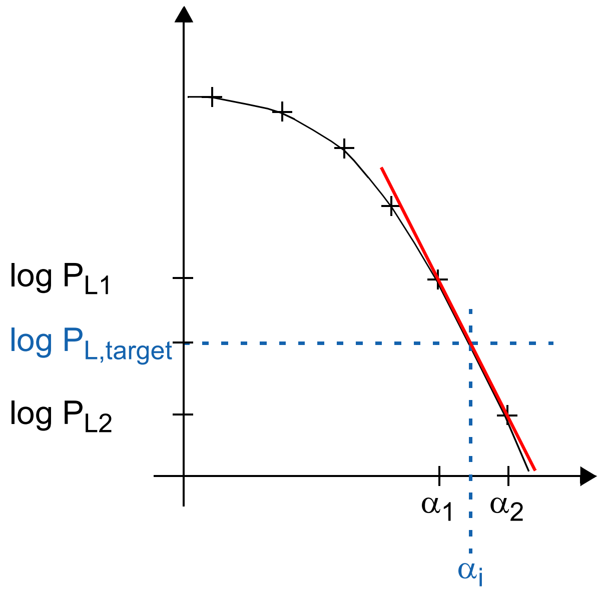


Figure 5.3: Linear interpolation to obtain the required overprovisioning factor for a given burst loss probability

IP packet loss ratio (IPLR) less than 10^{-3} end-to-end, the burst loss probability of the underlying burst-switched transport network has to be even lower. In order to have enough safety margin, mean loss probabilities of 10^{-4} to 10^{-5} for a burst-switched network or of 10^{-5} to 10^{-6} for a single burst-switched node are reasonable (cf. the path model in [DG01]).

Upper bounds on mean IP packet transfer delay (IPTD) are between 100 ms for the most demanding and 1 s for the least demanding class. However, the 100 ms IPTD criterion can only be met for path lengths less than 20,000 km, for which the propagation delay alone is approx. 100 ms. As discussed in Chapter 3 and Appendix A, characteristic delays in OBS, e.g., FDL delays are dominated by the propagation delay. Thus, burst-switched nodes are not critical here and the propagation delays and the delays introduced by IP routers decide whether the IPTD values can be met. The IP delay variation (IPDV) is only specified for the two real-time classes with values of 50 ms. Focusing again on the dominating propagation delay in an OBS network, a delay variation of 50 ms corresponds to 10,000 km of differential path length. As such high values can be easily avoided by suppressing extensive burst deflections, the jitter introduced by a burst-switched network is not critical for IPDV. Consequently, the mean burst loss probability within the burst-switched network is used as direct QoS metric.

Commonly, simulation studies on network performance concentrate on the analysis of direct QoS metrics and only consider a very small set of resource constraints. In contrast, Chapter 6 also quantifies the client layer and server layer resources measured by number of burst-switched ports as well as number of fiber hops for a given QoS level. This is particularly important for the comparison of various architectures like OBS, BoCS, and OBTN, which trade off different resource types. The evaluations consider the number of trunk ports as well as the total number of switch ports (Figure 5.1). It is interesting to also separately study trunk ports, which terminate virtual links to other core nodes, as they may require additional equipment such as amplifiers.

However, by including local add/drop and FDL buffer ports allows to even the effects of different MAN utilizations and FDL buffer architectures are captured.

5.3.2.2 Resource Metrics for Target QoS Values

As the simulations use the number of switch ports and fiber hops as input parameters and yield QoS metrics as results, the resource requirements for a given target loss probability cannot be directly obtained. To overcome this deficiency, a post-processing step interpolates between the resource requirements of QoS results in the vicinity of the target QoS value.

Figure 5.3 illustrates this in a diagram showing the logarithm of the burst loss probability for several values of the overprovisioning factor α. As the logarithmic curves of the burst loss probability only exhibit a minimal bend for the targeted low burst loss probabilities, an interpolation line yields sufficiently good results. Thus, the figure also shows the linear interpolation of the value pairs $(\alpha_1, \log P_{L1})$ and $(\alpha_2, \log P_{L2})$ which are closest to the given target loss probability $P_{\mathrm{L,target}}$. The intersection of the interpolation line and the curve $y = \log P_{\mathrm{L,target}}$ yields the approximation of the overprovisioning factor α_i required for the target loss probability $P_{\mathrm{L,target}}$.

In OBTN, simulations are performed for a given overprovisioning factor α_{D} of the direct end-to-end capacity. Thus, the ratio of shared overflow capacity β_i required to meet a QoS objective follows from the intersected overprovisioning factor α_i and α_{D} as $\beta_i = 1 - \alpha_{\mathrm{D}}/\alpha_i$ (cf. Eq. (5.4)). The total number of trunk ports C_i and fiber hops S_i can be calculated from Eq. (2.13) and Eq. (2.17) in combination with Eq. (5.5) and Eq. (5.6) as

$$C_i = \alpha_{\mathrm{D}} \cdot A + (\alpha_i - \alpha_{\mathrm{D}}) \cdot d_{\mathrm{A}} \cdot A \quad (5.12)$$

$$S_i = \alpha_{\mathrm{D}} \cdot d_{\mathrm{A}} \cdot A + (\alpha_i - \alpha_{\mathrm{D}}) \cdot d_{\mathrm{A}} \cdot A \quad (5.13)$$

where d_{A} is the mean hop distance for the traffic demands of the physical topology (cf. Eq. (2.6)). The total number of switch ports is finally obtained by adding up the total number of FDL buffer ports $N_{\mathrm{F,total}}$, the total number of local add/drop ports $N_{\mathrm{L,total}}$, and the total number of trunk ports C_i to improve readability, the indices i (intersected) are omitted in Chapter 6.

5.3.3 Evaluation Scenario

The evaluation scenario is designed to be realistic while at the same time avoiding parameter space explosion by introducing appropriate model abstractions based on literature and the author's previous work [Gau00, DGSB01, GDSB01, DG01, Gau02a, Gau02b, BGPS03, GKS04a, Gau04, GKZM05]. It comprises a physical network topology, traffic demands and traffic characteristics as well as the node functionality. The parameter decisions for the FDL buffers are additionally based on the performance evaluation in Section A.1.

- **Network topology:** All simulations are performed for the Pan-European reference core network topology (CT) specified by the LION and COST266 projects [COS03, MCL$^+$03]. The CT network topology depicted in Figure 5.4 comprises 16 nodes in major European metro areas and 23 bidirectional fiber links. With a mean node degree of 2.88 and a

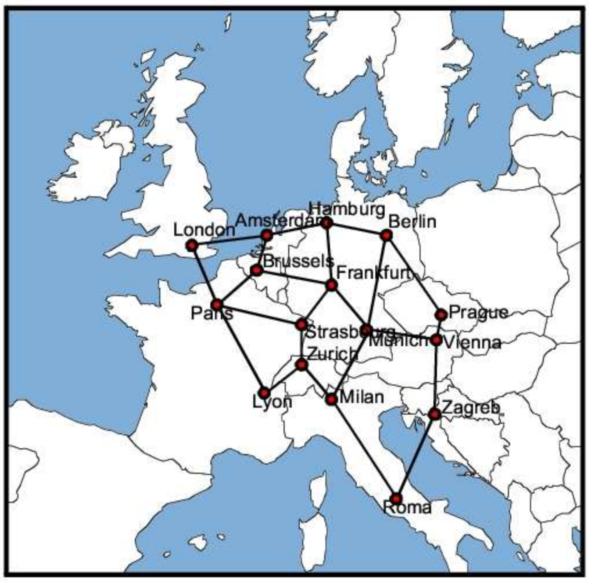


Figure 5.4: Pan-European reference network topology

mean hop distance of 2.64 it can be considered representative for optical core networks [BGH+04].

It is assumed that a burst-switched core node is deployed in each of the core network points-of-presence (POP) in order to dynamically switch traffic from and to the attached optical feeder networks and to switch core network transit traffic (where applicable). In case the wavelength-switched server layer uses optical cross-connects to switch light-paths, it is assumed that OXCs are collocated in the core POPs. Consequently, there is always exactly one fiber hop between burst-switched nodes, which are neighbors in the physical topology.

- **Traffic demands:** The traffic demand matrix used for network dimensioning and simulation is derived from the population-based demand model published with the reference network scenario [MCL+03]. The demand model computes the amount of voice, transaction and Internet traffic exchanged between core network nodes based on the number of citizens, companies, and Internet hosts estimated for the regions surrounding the nodes. For each type of traffic, the model specifies relations capturing the dependence of the demands on the distance between their source and destination nodes. For the evaluations, a total traffic demand for the year 2006 of approximately 9.9 Tbps is used. This is equivalent to a total of $A_0 = 990$ Erlang (wavelength equivalents) for 10 Gbps line-rate and leads to an average of 4.124 Erlang per core node pair.

- **Traffic generation and traffic characteristics:** The optical feeder networks are assumed to be independent of each other and are abstracted by their traffic demands. Traffic flows arriving from those networks to the core node are thus implemented as independent traffic generators. They generate bursts according to a Poisson process following the arguments in [IA02] for time-based assembly with a wide range of IP traffic characteristics. In this core network context, this assumption is further supported by the superposition of traffic from many independent burst assembly queues and the focus on small, burst-level time-scales. Due to the Poisson process, all bursts, which are generated in the feeder networks of one core node and destined to an arbitrary feeder network of a second core node, can be aggregated into a single generator.

The burst transmission time follows a negative exponential distribution[3]. The mean value is chosen to be $h = 10\,\mu$s, i.e., a mean burst length of 12,500 Byte for the bit-rate of 10 Gbps.

- **Node functionality:** The evaluations use the optical burst switch model depicted in Figure 4.6 with full wavelength conversion capability. Therefore, wavelength conversion can always be used as the primary contention resolution strategy.

  Unless stated differently, each node is equipped with an FDL buffer with one FDL ($F = 1$) operated in WDM with $W_\mathrm{F} = 32$ wavelength channels, i.e., with $N_\mathrm{F} = 32$ ports. Increasing the number of wavelengths for a single FDL beyond this values did not yield further improvements (cf. Figure A.3(b)). The FDL delay for the single FDL buffers and the delay granularity for degenerate multi-FDL buffers is chosen to be 4 mean burst transmission times. Bursts on the output wavelength channels and on the FDLs are scheduled based on a first-fit reserve-a-fixed duration algorithm with void filling (cf. Section 3.3). Such a node could, e.g., be implemented by a tune-and-select architecture (TAS) as shown in Figure 3.15 and Figure 3.16 for $T_\mathrm{F} = D = 4h = 40\,\mu$s.

Additional information on special evaluation scenarios is introduced in the context of the respective evaluations in Chapter 6.

[3] This is motivated by the preparatory studies in Section A.1. They showed that the results are rather insensitivity with respect to the burst length distribution of assembled burst traffic.

6 Evaluation of Burst Transport Architectures

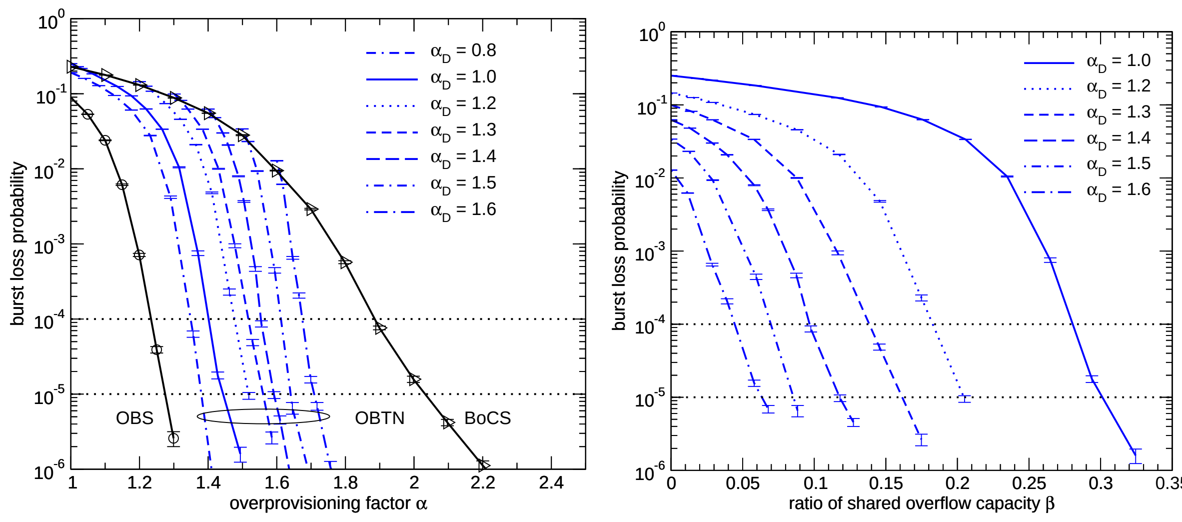
This chapter evaluates OBTN and its architectural components regarding the key criteria of QoS and resource efficiency set forth in the design rationale in Section 4.1. In addition, it compares OBTN with the existing solutions OBS and BoCS to quantify the QoS performance as well as the requirements for client layer and server layer resources.

The first section of this chapter discusses fundamental results on the OBTN, OBS, and BoCS QoS performance and the client layer and server layer resources required to achieve a target QoS level. Then, Section 6.2 broadens the comparison by also considering advanced OBS architectures including alternative routing. Section 6.3 takes a closer look at the hunting modes inside the OBTN nodes. This is then extended in Section 6.4 to compare OBTN and advanced BoCS architectures regarding the impact of routing flexibility and shared overflow capacity. Advanced studies regarding the design of OBTN nodes in Section 6.5 and of the virtual topology in Section 6.6 further complement the evaluation. Finally, Section 6.7 quantifies the impact of the total traffic demand in the network.

6.1 Principal Performance of OBTN, OBS, and BoCS

This section analyzes OBTN, OBS, and BoCS regarding the QoS and resource metrics defined in Section 5.3.2 in order to provide a fundamental understanding of the advantages and disadvantages of the respective architectures.

In all three architectures, the nodes employ the same FDL buffers with a single FDL and the parameters outlined in Section 5.3.3. In order to ensure that all bursts follow the same physical path through the network, the reference OBS and BoCS architectures use fixed shortest path routing and OBTN uses the constrained alternative routing. For now, routing in OBTN can only use one alternative route in each node, which traverses single-hop virtual link to the neighbor in the physical topology. As outlined in Section 4.4.2, the alternative routes then exactly match the assignment of shared overflow capacity. These assumptions will be generalized in succeeding sections to comprehensively compare the three architectures.



(a) Impact of the overprovisioning factor (b) Impact of the ratio of shared overflow capacity

Figure 6.1: Comparison of the mean burst loss probability for OBS, BoCS, and OBTN

6.1.1 Evaluation of the QoS

Figure 6.1(a) depicts the mean burst loss probability vs. the overprovisioning factor α for OBS, BoCS, and OBTN. The overprovisioning factor relates the total network capacity to the capacity required for static traffic. It can be seen that the burst loss probability drops much faster for OBS than for BoCS. It reaches the targeted low burst loss probabilities of $P_L = 10^{-4}$ or 10^{-5} for overprovisioning factors of only 1.2–1.3 compared to 1.9–2.1 with BoCS. This performance advantage can be attributed to the higher statistical multiplexing gain per virtual link in OBS and directly supports the qualitative discussion in Section 2.4.3.

The figure also shows the results for OBTN with seven different dimensionings of the direct end-to-end lightpath capacity. For a systematic comparison, the capacity of the direct end-to-end lightpaths is fixed by the respective overprovisioning factor $\alpha_D = (1-\beta) \cdot \alpha$ in Eq. (5.4) while the ratio of shared overflow and total network capacity, β, increases along the curves of constant α_D with increasing α.

As targeted, the curves for OBTN lie in-between OBS and BoCS. Also, the curves for OBTN have a much steeper slope than BoCS comparable to OBS. The curves are sorted according to α_D such that architectures with few direct end-to-end lightpath capacity (small α_D) and more shared overflow capacity lie closer to OBS. The architectures with more direct end-to-end lightpath capacity (large α_D) but fewer shared overflow capacity are closer to BoCS. For the plotted α_D values, they require overprovisioning factors of 1.33–1.75 to meet the target loss probabilities of 10^{-4} or 10^{-5}.

It can be seen that for $\alpha = \alpha_D$ the results for OBTN closely match the respective results for BoCS as the architectures do not use any shared overflow capacity for this dimensioning (β). Thus, they only differ in the (slightly) higher routing flexibility of OBTN. In OBTN, the steep decrease of the burst loss probability for increasing α is due to the higher routing flexibility and the high effectiveness of the shared overflow capacity as detailed in Section 6.4.

While Figure 6.1(a) did not explicitly contain the ratio of shared overflow capacity β, Figure 6.1(b) depicts the burst loss probability vs. β for different values of α_D in OBTN. OBS ($\beta = 1$) and BoCS ($\beta = 0$) are not included though. It can be seen that the smaller the amount of direct end-to-end capacity is, the more shared overflow capacity is needed to yield the same burst loss probability. For the depicted dimensioning, OBTN only requires a relatively small share of overflow capacity with β values in the range of 0.05–0.3 to meet the target QoS.

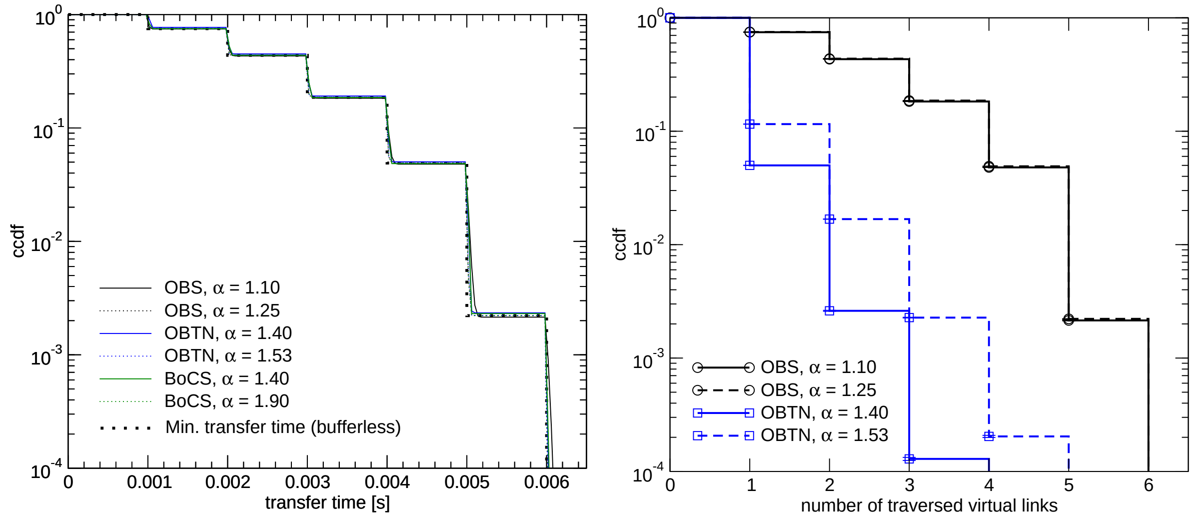
The two diagrams in Figure 6.1 together show that OBTN can reach very low burst loss probabilities with different combinations of direct end-to-end lightpath capacity and shared overflow capacity. However, the two capacity types have different impact on the amount of client layer and server layer resources. Therefore, the optimal dimensioning regarding client layer and server layer resources requires a more detailed evaluation for given QoS objectives which is presented in Section 6.1.2.

In order to extend the QoS comparison to the latency of the core network, Figure 6.2(a) shows the complementary cumulative distribution function (CCDF) of the transfer time in the core network for OBS, BoCS and OBTN ($\alpha_D = 1.4$). Here, the propagation delays of all links in the physical fiber topology (cf. Figure 5.4) are chosen to be 1 ms to isolate delay components of signal propagation and FDL buffering. This value corresponds to a link length of 200 km according to a speed-of-light in the fiber of 200,000 km/s. The relatively small link delay is selected as a practical upper bound on the delay jitter introduced by FDLs. All scenarios with link lengths > 200 km experience less impact of FDL buffer delays ($T_F = 40\,\mu$s) on the total transfer delay by large propagation delays.

In Figure 6.2(a), the transfer time CCDF is given for the three architectures at two overprovisioning factors each. One rather low α value is chosen to represent a scenario with a high burst loss probability of 2–6 % ($\alpha = 1.1$ for OBS and $\alpha = 1.4$ for BoCS and OBTN). The other value leads to a burst loss probability of about $P_L = 10^{-4}$. It can be seen that the CCDF of the transfer time does not vary significantly. This behavior is expected as OBS and BoCS do not use any alternative routing and OBTN only uses constrained alternative routing, i.e., all routes are fixed with respect to the fiber links traversed.

Figure 6.2 also depicts the CCDF of a minimal delay reference scenario, which is bufferless and lossless and thus not biased by higher burst loss probabilities for longer routes. In the figure, the CCDF of the different architectures can hardly be distinguished from the CCDF of the minimum transfer time. Consequently, the delays introduced by the FDL delays do not substantially contribute to the transfer time. They are clearly dominated by the propagation delays in the network, even for the relatively small link delays assumed here. Based on this observation and due to the fact that all physical fiber links have the same delay, the CCDF of the transfer time closely matches the CCDF of the number of traversed physical links which is thus not shown here.

In contrast to the transfer time, the CCDF of the number of traversed virtual links differs between OBS, BoCS, and OBTN. While in OBS virtual and physical links correspond to each other, BoCS and OBTN both employ virtual links spanning several physical links. Figure 6.2(b)



(a) CCDF of the transfer time in the core network

(b) CCDF of the number of traversed virtual links

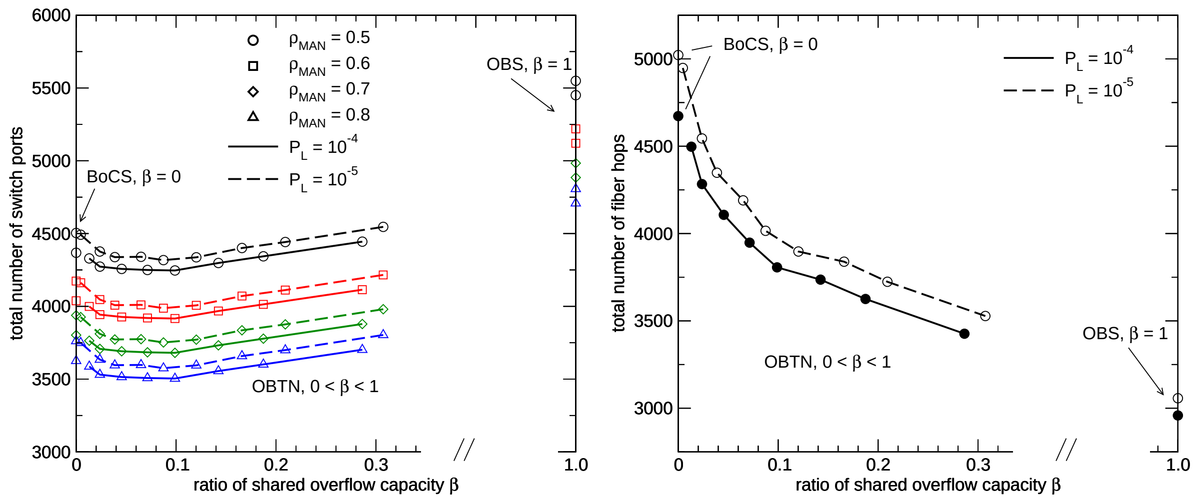
Figure 6.2: CCDFs of transfer time and virtual link count for OBS, BoCS, and OBTN

shows this difference for OBS and OBTN using the same overprovisioning factors as in Figure 6.2(a). As bursts in the BoCS scenario without alternative routing can always only traverse a single virtual link, the results for BoCS are not discussed here. For OBS, the number of virtual links does not change when increasing α. As additional resources do not lead to a higher number of successfully transmitted bursts on long routes, the loss probability is not very sensitive to the route length. In contrast, the additional shared overflow capacity in the case of OBTN with $\alpha = 1.53$ compared to $\alpha = 1.4$ enables more bursts to take constrained alternative routes and thus shifts the CCDF towards a higher number of virtual links. Still, due to the constrained alternative routing, Figure 6.2(a) does not show any negative impact on the transfer time distribution.

6.1.2 Evaluation of the Required Client Layer and Server Layer Resources

In order to assess the architectures in terms of their client layer and server layer resource requirements, the number of switch ports and fiber hops are calculated from the simulation results for the two burst loss probabilities 10^{-4} and 10^{-5} as described in Section 5.3.2. Figure 6.3(a) presents the total number of burst switch ports versus the ratio of shared overflow resources β using the unified resource model introduced in Section 5.1. The number of add/drop ports is calculated from the total traffic injected into the network $A = 990$ Erlang and different values of the utilization ρ_{MAN} in the attached metro networks. For instance, $\rho_{MAN} = 0.6$ defines that $990/0.6 = 1650$ switch ports in the network are required for local add/drop ports (cf. Section 5.3.3). The total number of FDL buffer ports amounts to 512. As the number of switch ports and fiber hops are derived from the interpolated QoS results (cf. Section 5.3.2), the respective figures do not contain confidence intervals.

Although OBS ($\beta = 1$) above showed the best QoS performance for a given overprovisioning factor, it clearly requires the highest number of burst-switched ports independent of the target



(a) Client layer resources

(b) Server layer resources

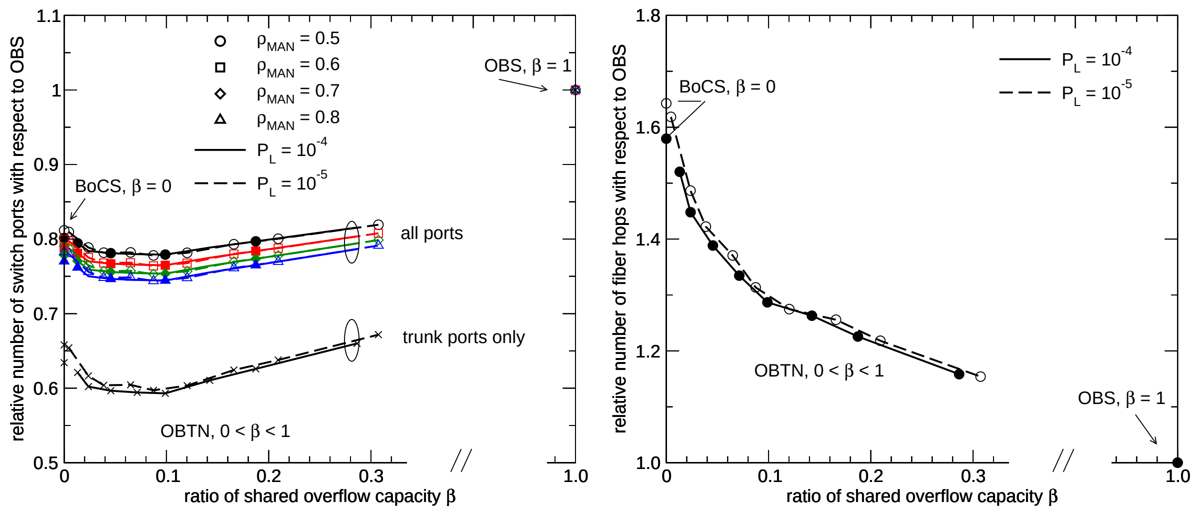
Figure 6.3: Comparison of the absolute number of resources for OBS, BoCS, and OBTN

burst loss probability and the MAN utilization. In contrast, BoCS ($\beta = 0$) and OBTN ($0 < \beta < 1$) both require fewer ports. In OBTN, β values in the range of 0.2–0.3 lead to results comparable to BoCS while smaller values of β even yield improvements. For higher values of β, more capacity is assigned to the single-hop virtual links which increases the number of switch ports towards the values for OBS. This principal behavior follows from the superposition of the effects of statistical multiplexing gain per virtual link and transit traffic in nodes (cf. Section 2.4.3). The minimum in the number of switch ports for OBTN shows that it successfully reduces the transit traffic with a limited penalty in statistical multiplexing gain.

While the absolute numbers increase slightly, when demanding a lower target burst loss probability, the principal behavior is mostly independent of it for the depicted target burst loss probabilities. As a higher utilization ρ_{MAN} in the attached metro networks only reduces the number of add/drop ports, the impact is identical for all architectures as can be seen from the vertical parallel shift of the respective curves.

In order to quantify the relative improvement with respect to OBS, Figure 6.4(a) shows the number of switch ports for all architectures normalized to the number of switch ports for the respective OBS scenario, i.e., the scenario with the same MAN utilization and target loss probability. The results show that OBTN reduces the total number of switch ports by 22–26 % with respect to OBS depending on the MAN utilization and the ratio of shared overflow capacity. The improvement is higher for small values of β and maximal around $\beta = 0.1$. Again, it can be seen that the relative reductions in the number of switch ports are independent of the target loss probabilities $P_{\mathrm{L}} = 10^{-4}$ and 10^{-5}.

In addition to the total number of switch ports, Figure 6.4(a) also depicts the number of trunk ports, i.e., only those switch ports which terminate virtual links to other core nodes. The number of trunk ports is an interesting metric as those ports commonly require additional equipment, e.g., amplifiers. As the number of trunk ports does not comprise any add/drop ports, it is independent of the MAN utilization ρ_{MAN}. Here, the improvements of OBTN compared to OBS



(a) Client layer resources (b) Server layer resources

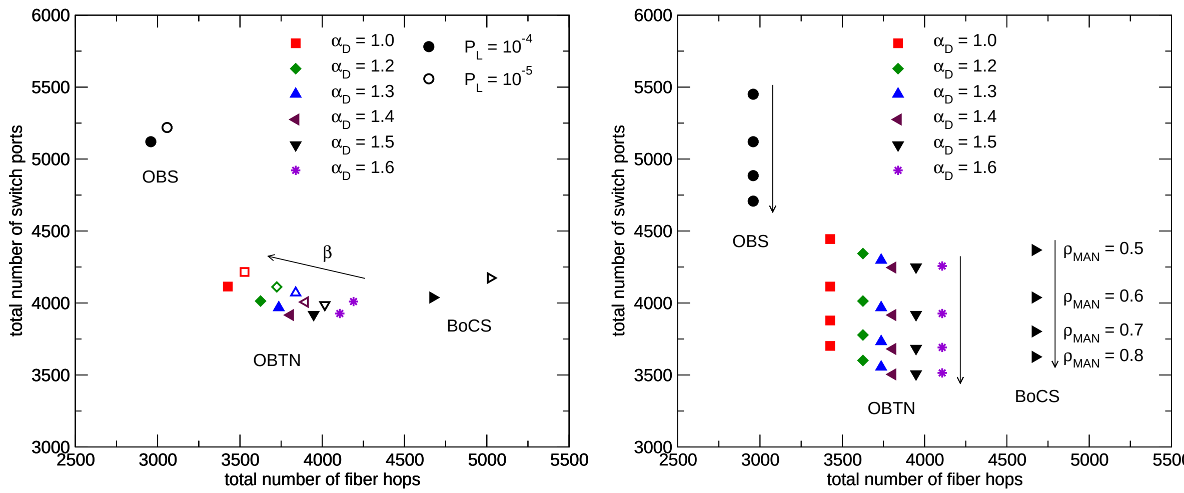
Figure 6.4: Comparison of the required resources relative to OBS

are greater and reach values of up to 40 % around $\beta = 0.1$. This is due to the fact that the optimized virtual topology design directly impact the number of trunk ports while the total number of switch ports has the significant constant offset of the FDL buffer and add/drop ports. The results are again mostly independent of the target loss probability with minor differences for BoCS and OBTN for very low β values.

Section 2.4.3 derived that the maximum gain of a full-mesh virtual topology regarding client layer capacity equals the mean hop distance d_A of the physical topology. Thus, a relative number of trunk ports with respect to OBS of $1/2.42 \approx 0.41$ would serve as a lower bound in Figure 6.4. It can be seen that the best OBTN parameterizations lie approx. 50 % above this value.

So far, the discussion of the resource requirements focused on the number of switch ports showing that OBTN and BoCS can achieve a considerable reduction compared to OBS. However, as already quantified in Section 2.4.3, dense virtual topologies yield a penalty regarding the required number of fiber hops. Therefore, in order to construct the complete picture, Figure 6.3(b) plots the total number of fiber hops vs. the ratio of shared overflow capacity β while Figure 6.4(b) is again normalized to the respective OBS values. This figure can directly be obtained by dividing the intersecting overprovisioning factors α_i of the respective architecture and of OBS for the target loss probability.

In both figures, the dependence of the number of fiber hops on β can be described in two regimes, one for β values smaller than approx. 0.1 with a steep decrease and one for higher β values with a less steep decrease. In the first regime, with small β and large α_D values or for BoCS the penalty regarding the statistical multiplexing gain per virtual link dominates. It leads to a 30–60 % higher number of fiber hops. In contrast, OBTN with a higher β but a smaller α_D value can achieve a better performance due to the shared overflow capacity. This leads to a penalty of only 17–30 % compared to OBS. The figures show that the absolute numbers



(a) Impact of target loss probability ($\rho_{\text{MAN}} = 0.6$) (b) Impact of the MAN utilization ($P_{\text{L}} = 10^{-4}$)

Figure 6.5: Integrated comparison of client layer and server layer resources

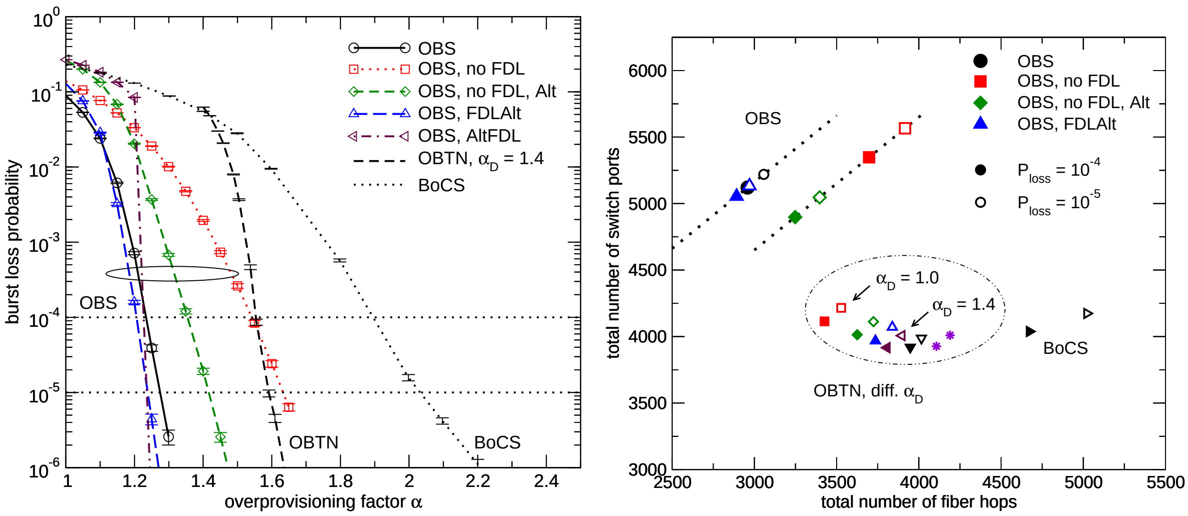
increase when demanding the lower target burst loss probability while the principal behavior and the relative values remain almost unchanged.

Combining the conclusions of Figure 6.4(a) and (b) already shows that OBTN offers an attractive trade-off between OBS and BoCS. Even for rather small amounts of shared overflow capacity it successfully reduces the required number of switch ports compared to OBS. Compared to BoCS, it can realize a much smaller penalty in the server layer resources. For $\beta \approx 0.1$ (this corresponds to $\alpha_{\text{D}} = 1.4$), the improvement in switch ports is maximum while the penalty in fiber hops is still limited.

Finally, Figure 6.5 integrates both views by plotting the total number of switch ports vs. the total number of fiber hops required to meet a target burst loss probability for OBS, BoCS, and OBTN with different α_{D} values. In Figure 6.5, each architecture and dimensioning is represented by one symbol, solid for $P_{\text{L}} = 10^{-4}$ and empty for $P_{\text{L}} = 10^{-5}$. Figure 6.5(a) concentrates on the impact of the target burst loss probability for a fixed MAN utilization of $\rho_{\text{MAN}} = 0.6$. In contrast, Figure 6.5(b) depicts the impact of different ρ_{MAN} for a fixed target loss probability of $P_{\text{L}} = 10^{-4}$.

Figure 6.5(a) clearly shows the extreme resource requirements of OBS and BoCS. OBS has the smallest number of fiber hops but also the greatest number of switch ports and BoCS with the greatest number of fiber hops and a small number of switch ports. In-between OBS and BoCS lie the different OBTN realizations with a low number of switch ports and a medium number of fiber hops. For both target loss probabilities, the figure shows that OBTN dimensionings with $\alpha_{\text{D}} = 1.4$ ($\beta \approx 0.1$) and $\alpha_{\text{D}} = 1.5$ ($\beta \approx 0.087$) yield the smallest number of switch ports with still substantially fewer fiber hops than BoCS. Thus, OBTN with $\alpha_{\text{D}} = 1.4$ is used in most of the following detail studies.

For OBS and OBTN, demanding the lower burst loss probability requires a rather small number of additional ports and fiber hops, e.g., 2 % more fiber hops for OBTN with $\alpha_{\text{D}} = 1.4$. In con-



(a) Impact on burst loss probability

(b) Impact on client layer and server layer resources

Figure 6.6: Comparison with advanced OBS architectures

trast, the lower target loss probability has a more substantial impact on the number fiber hops for BoCS, here approx. 7.5 %.

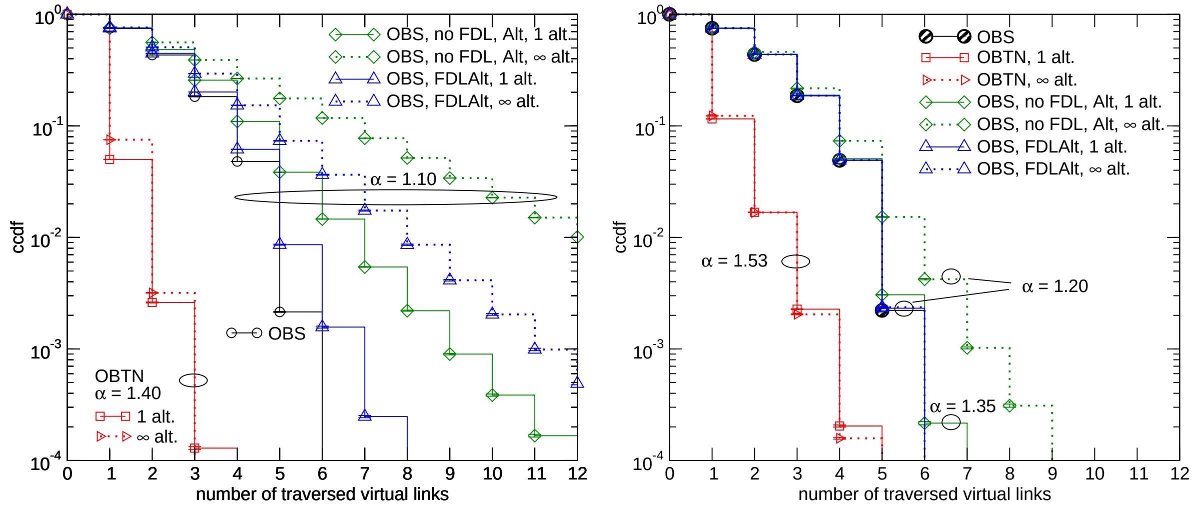
Figure 6.5(b) concludes these studies by evaluating the impact of the MAN utilization ρ_{MAN}. Here, a higher utilization directly reduces the number of switch ports without any changes to the number of fiber hops.

Summarizing, compared to OBS, suitable OBTN parameterizations can reduce the number of trunk ports by up to 40 % and the total number of switch ports by 22-26 % depending on ρ_{MAN}. At the same time, the experience a penalty in the number of fiber hops of less than 28 %. Applying the cost relations for WSN scenarios outlined in Section 4.1.2, in which bandwidth and thus server layer resources are considered a commodity and switch ports the major cost driver, OBTN constitutes an effective solution to reduce the overall cost.

6.2 Comparison of OBTN and Advanced OBS Architectures

The principal comparison presented so far focused on an OBS architecture applying an FDL buffer but no alternative routing referred to as *OBS* here. This section extends the comparison with BoCS and OBTN ($\alpha_D = 1.4$) to bufferless as well as advanced OBS architectures with alternative routing to complement the picture.

This study includes bufferless OBS without (*OBS, no FDL*) and with alternative routing (*OBS, no FDL, Alt*) as well as OBS with FDL buffers and alternative routing. In the latter scenario, two hunting modes, i.e., the order of buffering and alternative routing for contention resolution, are studied (cf. Section 3.4.5). *OBS, FDLAlt* probes the buffer first and only if this fails considers alternative routing, while *OBS, AltFDL* uses the inverse order. Alternative routing is configured



(a) α values with higher burst loss probabilities (b) α values with burst loss probabilities $\approx P_L$

Figure 6.7: CCDF of the number of traversed virtual links for advanced OBS architectures

to offer one alternate route in addition to the primary route in every node. The alternate route is computed as the shortest path to the destination which does not include the blocked output link of the primary route. These configurations have been shown to yield optimal or close to optimal results while avoiding excessive deflections in [Sch04] and have also been applied in [GKS04b].

6.2.1 Evaluation of QoS

Figure 6.6(a) depicts the burst loss probability vs. the overprovisioning factor for the different OBS realizations as well as for BoCS, and OBTN. It first shows that bufferless OBS performs significantly worse that OBS with FDL buffers and requires approximately the same overprovisioning factors as OBTN to reach the target loss probabilities. Introducing alternative routing for bufferless OBS first leads to the well-known performance degradation for small overprovisioning factors (high network load) [NM73, Spä02, GKS04a] and only then improves the performance for higher overprovisioning factors (lower network load). Nevertheless, bufferless OBS with alternative routing cannot achieve the same performance as the reference OBS architecture with an FDL buffer and shortest path routing.

Regarding the two OBS architectures with FDL buffer and alternative routing, the figure exhibits very different behavior. *OBS, FDLAlt* yields only limited performance degradation for low and only limited improvement for high values of α compared to *OBS*. In contrast, *OBS, AltFDL* shows the typical two-state behavior already discussed in Section 3.4.4 with severe degradation for α values below a certain threshold and a rapidly dropping burst loss probability above that threshold. Summarizing, OBS architectures with FDL buffer, both with or without alternative routing, perform comparably well around the target loss probabilities and significantly better than bufferless OBS architectures.

The discussion of Figure 6.2(a) showed that the reference OBS architecture without alternative routing and OBTN have the same transfer time CCDF. However, OBS with alternative routing can obviously lead to a higher variability regarding the routes, thus regarding the number of traversed (virtual and physical) links, and finally regarding the transfer time. For OBS, the CCDF of the transfer time closely matches the CCDF of the number of traversed physical and virtual links, such that only the latter is discussed here. Figure 6.7 shows the CCDF of the number of traversed virtual links for the OBS and OBTN reference architectures. It also includes OBS with alternative routing, both with FDL buffer and bufferless. While Figure 6.7(a) depicts the results for lower overprovisioning factors α leading to relatively high burst loss probabilities, Figure 6.7(b) uses α values leading to burst loss probabilities close to the target loss probability $P_L = 10^{-4}$.

In Figure 6.7(a), the smaller α values and thus the smaller capacity triggers more contention situations in nodes. Thus, alternative routing leads to significantly more traversed virtual links compared to the reference OBS architecture. This effect is more pronounced when allowing an unlimited number of route alternatives instead of only one. Also, it is always stronger for the bufferless OBS architecture. These route length variations are no longer negligible and directly lead to high transfer time variations. The higher capacity used in the cases depicted in Figure 6.7(b) yields less contention situations to be resolved by alternative routing and thus less traversed virtual links. Except for the bufferless OBS case with an infinite number of route alternatives, the results are even comparable to the reference OBS architecture. However, route length and thus transfer time are not controlled by design as in OBTN but also depend on the network capacity introducing an additional constraint to be considered in the dimensioning process.

6.2.2 Evaluation of the Required Client Layer and Server Layer Resources

While Figure 6.6(a) and Figure 6.7 focus on the QoS performance, Figure 6.6(b) adds the results bufferless and advanced OBS architectures to the results already presented in Figure 6.5(a). In OBS, each virtual link requires one switch port and spans exactly one fiber hop. Thus, OBS architectures with FDL buffer lie on the line $C_i = S_i + N_{F,total} + N_{L,total}$, while those without FDL buffer lie on the line $C_i = S_i + N_{L,total}$.

Compared to the reference OBS architecture, it can be seen that bufferless OBS without alternative routing requires a greater number of switch ports as well as a greater number of fiber hops. Compared to OBTN ($\alpha_D = 1.4$), the number of switch ports at $P_L = 10^{-5}$ is approx. 40 % higher while the number of fiber hops is the same. Bufferless OBS with alternative routing reduces the number of switch ports only slightly with respect to OBS but still requires additional fiber hops. With respect to OBTN with $\alpha_D = 1.0$ it uses 20 % more switch ports at only approx. 3 % fewer fiber hops. Finally, introducing alternative routing to OBS with an FDL buffer (*OBS, FDLAlt* vs. *OBS*) yields hardly any resource advantage, neither regarding the number of switch ports nor the number of fiber hops.

Concluding, OBS with FDL buffer but without alternative routing is the preferable solution among the different variants of OBS. This follows based on the QoS metrics burst losses and transfer time as well as based on the resource arguments.

6.3 Impact of Hunting Modes in OBTN

The OBTN hunting modes *AltFDL* and *FDLAlt* introduced in Section 4.4.4 define in which order alternative routing and FDL buffering are applied in order to resolve a contention situation[1]. To be most effective, the different hunting modes should comply with and support all other architectural decisions. Therefore, this section analyzes which hunting mode should be preferably applied. Also this section founds the basis for the systematic analysis of routing and overflow capacity in Section 6.4.

Figure 6.8(a) depicts the burst loss probabilities of the AltFDL and FDLAlt hunting modes for the overprovisioning factors $\alpha_D = 1.0$, 1.4, and 1.6. It shows that the burst loss probability of FDLAlt decreases slower than that of AltFDL when increasing α and thus β. Obviously, FDLAlt cannot exploit the additional shared overflow capacity as effectively. Also, the slope of the FDLAlt curves is steeper for higher than for lower α_D values, which causes the curves to intersect. In contrast the AltFDL curves are basically in parallel.

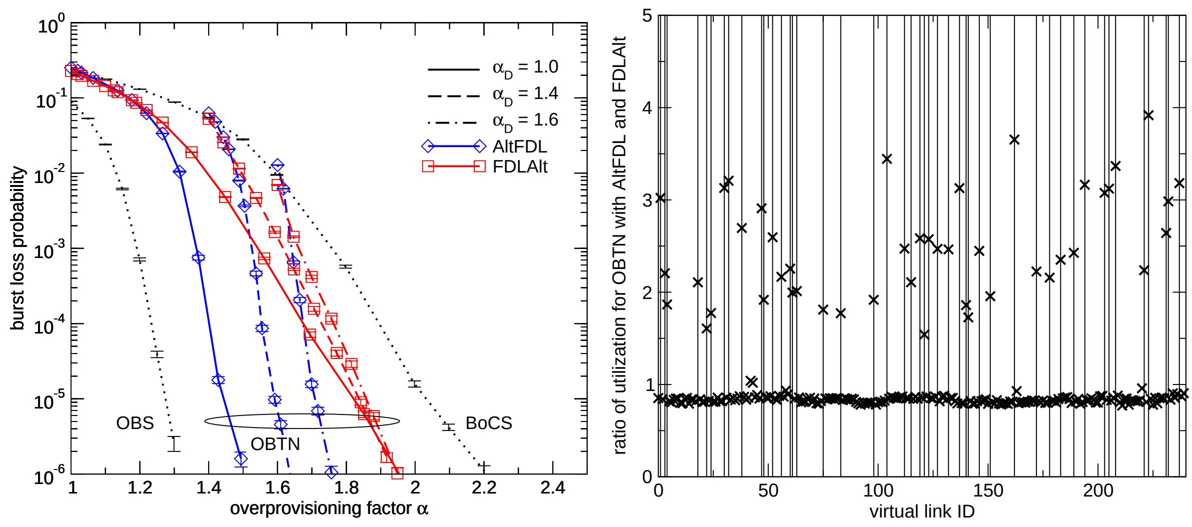
Both effects can be explained by analyzing the relationship of the hunting mode and the shared overflow capacity. As AltFDL preferably uses alternative routing, it directly benefits from the overflow capacity, which is assigned to the virtual links traversed by the alternative routes. In contrast, FDLAlt preferably uses the FDL buffer and thus only benefits from the added overflow capacity if buffering fails.

The fact that AltFDL utilizes the shared overflow capacity much better than FDLAlt can be seen by comparing the mean link utilization of all virtual links for two scenarios with comparable QoS. Namely, Figure 6.8(b) depicts the ratio of the mean link utilizations for $\alpha_D = 1.4$ and the overprovisioning factors $\alpha \approx 1.61$ for AltFDL and at $\alpha = 1.91$ for FDLAlt. The virtual links are listed along the x-axis numbered from virtual link $(1,2)$ with ID 0 to $(16,15)$ with ID 239. Those virtual links, which connect neighbor nodes in the physical topology and thus have shared overflow capacity are marked by the vertical lines. It can be seen that all virtual links except the ones with overflow capacity show similar ratios. However, on the links with shared overflow capacity AltFDL can achieve a much higher utilization than FDLAlt because it more frequently uses this capacity for alternative routing.

Finally, the different slopes of the curves for FDLAlt and their intersection are explained. As the FDL buffer is used first in FDLAlt, the buffer fills up and remains clogged for small α_D values. Thus, contention resolution combining the FDL buffer and an alternative route fails. This renders the alternative routing and the shared overflow capacity rather ineffective leading to the small slope. However, for higher α_D values, the statistical multiplexing on the direct end-to-end links is more effective. Thus, the buffer is less clogged and contention resolution combining the FDL buffer and alternative routing succeeds leading to the steeper slope.

Summarizing, the hunting mode AltFDL better fits the contention resolution and capacity assignment approach of OBTN and should thus be used. As the dimensioning of node and network resources is independent of the hunting mode, it is not further analyzed here.

[1] Note that in both schemes wavelength conversion is still always probed first.



(a) Burst loss probability

(b) Comparison of virtual link utilization

Figure 6.8: Analysis of the OBTN hunting modes AltFDL and FDLAlt

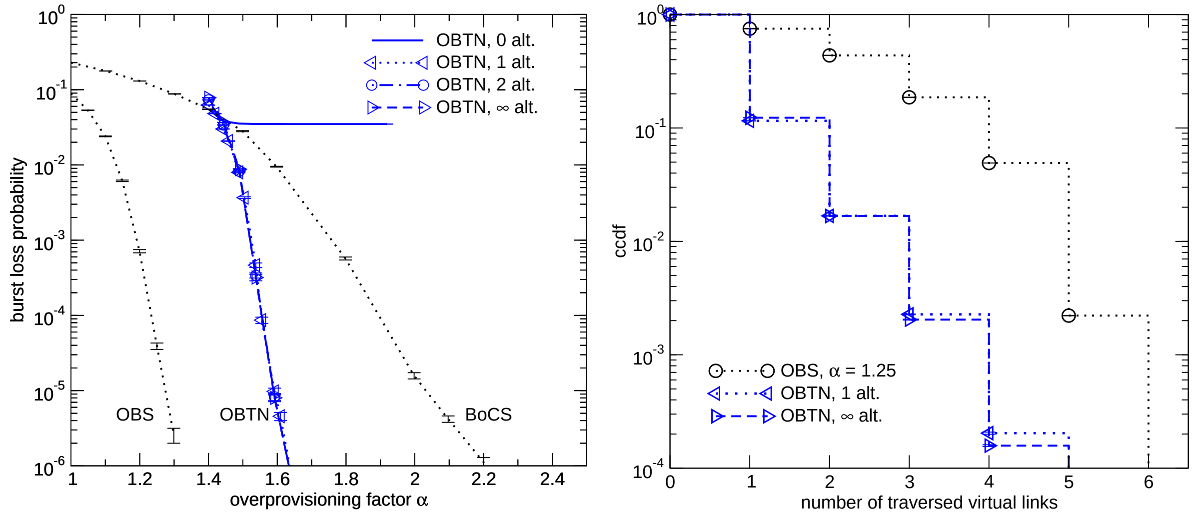
6.4 Constrained Alternative Routing and Shared Overfbw Capacity

This section concentrates on the effectiveness of the defining OBTN concepts of constrained alternative routing and shared overflow capacity. For a systematic analysis, the effects of routing flexibility and shared overflow capacity in OBTN are isolated and compared to the BoCS architectures.

6.4.1 Evaluation of the Routing Flexibility in OBTN

All previous evaluations of OBTN used constrained alternative routing with the restriction of only one alternative route, which traverses the virtual link to the neighbor in the physical topology. This restriction is lifted in the following study to analyze the impact of higher routing flexibility. The comparison of OBTN with one, two and an unrestricted number of alternatives per node is shown in Figure 6.9(a) for $\alpha_D = 1.40$. It can be seen that allowing only one alternative practically achieves the same performance as the case with 2 and an unrestricted number. This is explained by Figure 6.9(b) for the case of $\alpha_D = 1.40$ with $\alpha = 1.53$ already studied in Figure 6.2(b). Here, the CCDF of the number of traversed virtual links is approximately the same for 1 and an unrestricted number of alternatives. In the unrestricted case, only few bursts take multi-hop virtual links which otherwise would reduce the number of traversed virtual links. Thus, intermediate OBTN core nodes are hardly offloaded. Consequently, as a smaller number of alternatives facilitates the local forwarding decision, it is preferable when combining performance and realization arguments.

Figure 6.9(a) also shows the case of OBTN without constrained alternative routing, i.e., with only shortest path routing. In this case, the shared overflow capacity cannot perform its original task of providing capacity for bursts on alternate routes. Therefore, the additional shared over-



(a) Burst loss probability (b) CCDF of the number of traversed virtual links

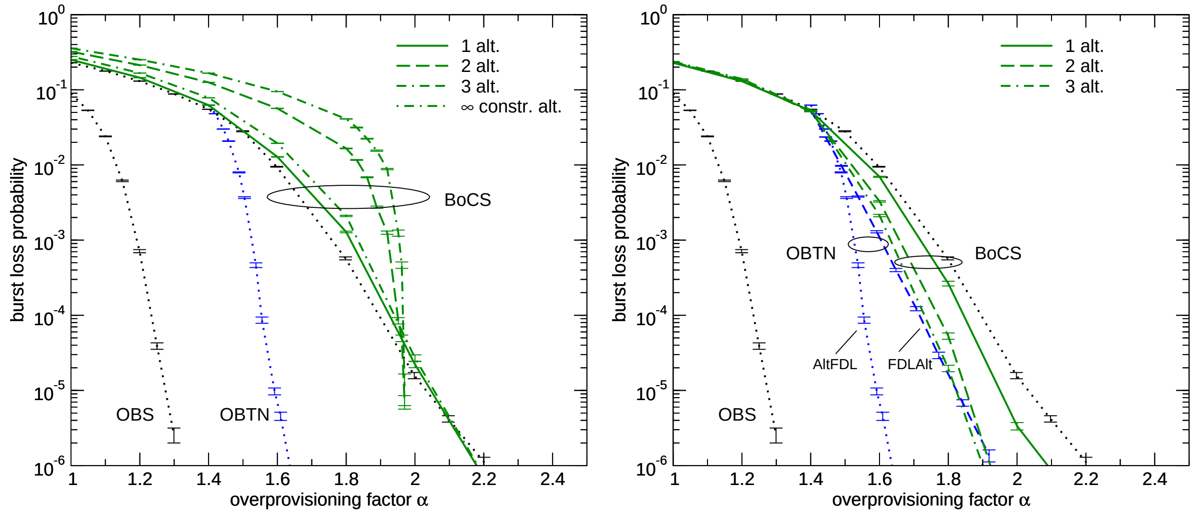
Figure 6.9: Impact of the maximum number of constrained alternate routes in OBTN

flow capacity, which is assigned with growing α, reduces the burst loss probability only over a very small range before it floors out onto a rather high value. This can be explained by the fact, that only the virtual links connecting physical neighbor nodes benefit from this capacity and achieve a lower burst loss probability. However, all other virtual links do not benefit at all and their performance eventually dominates the mean burst loss probability.

6.4.2 Comparison of BoCS and OBTN

This section analyzes the effects of adding constrained and unconstrained alternative routing to BoCS and compares the performance with OBTN in order to study the role of the shared overflow capacity. As the hunting modes play an important role here, the two cases of primarily using FDL buffers (*FDLAlt*) and primarily using alternative routing (*AltFDL*) are treated separately in Figure 6.10. For reference, the OBS, BoCS, and OBTN architectures are shown as dotted lines.

Figure 6.10(a) depicts the case of OBTN FDLAlt ($\alpha_D = 1.4$) and BoCS FDLAlt with a different number of alternate routes. The route selection is implemented such that the case of only one alternate route corresponds to the constrained alternative routing approach. It can be seen that introducing alternative routing to BoCS as the primary contention resolution strategy (AltFDL) leads to increased congestion and a higher loss probability up to high overprovisioning factors. This performance degradation is small for only one alternate route and the case of constrained alternative routing with an unrestricted number of alternatives. However, it is substantial for the cases with two and three alternate routes. Only for rather high values of the overprovisioning factor, do the virtual links have enough capacity so that an overall improvement in the burst loss probability can be achieved. In contrast, OBTN assigns all capacity beyond $\alpha = \alpha_D = 1.4$ as shared overflow capacity to virtual links which carry traffic on alternate routes. Therefore, these



(a) *AltFDL* hunting mode (b) *FDLAlt* hunting mode

Figure 6.10: Impact of the routing scheme and the capacity assignment in BoCS and OBTN

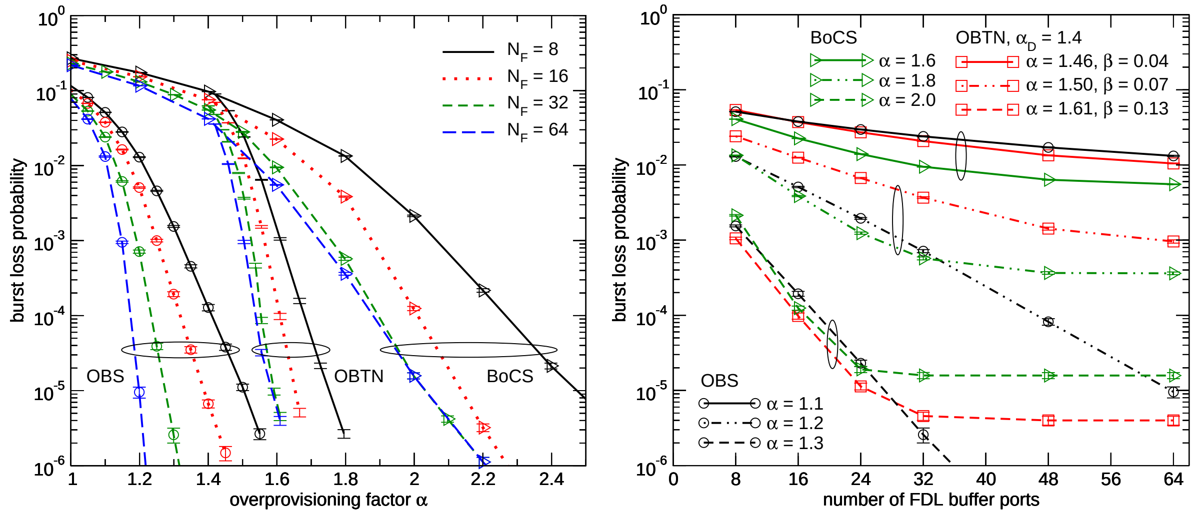
resources optimally support the alternative routing scheme and immediately reduce the burst loss probability.

In contrast to AltFDL, FDLAlt can improve the performance compared to the reference BoCS architecture Figure 6.10(b) shows that increasing the number of alternate routes reduces the burst loss probability compared to BoCS. For three alternate routes, the performance is comparable to OBTN with FDLAlt but is still clearly worse than OBTN with AltFDL. The figure also demonstrates that constrained alternative routing in BoCS is outperformed by unconstrained alternative routing. As there is no shared overflow capacity the number of alternative routes is the deciding factor.

As a summary, alternative routing with several alternate routes as the primary contention resolution scheme in dense virtual topologies is unstable for AltFDL. For FDLAlt, alternative routing yields improvements which are comparable to OBTN FDLAlt. However, it still cannot reach the performance of OBTN AltFDL. Consequently, alternative routing, the hunting mode, and the assignment of shared overflow capacity have to fit together to leverage their full potential.

6.5 Impact of the FDL Buffer Design

The structure and dimensioning of the FDL buffer is an important part of the node design and has direct impact on the number of switch ports. Therefore, this section looks at the impact of the number of buffer ports for a single FDL as well as the number of FDLs for a fixed number of buffer ports. All previous studies use an FDL buffer with a single FDL operated in WDM with 32 wavelengths. Although this parameterization is motivated by the single node evaluations presented in Section A.1, the following results confirm that this decision is also advantageous in the studied network context.



(a) Varying network dimensioning

(b) Fixed network dimensioning

Figure 6.11: Impact of the number of FDL buffer ports ($F = 1$)

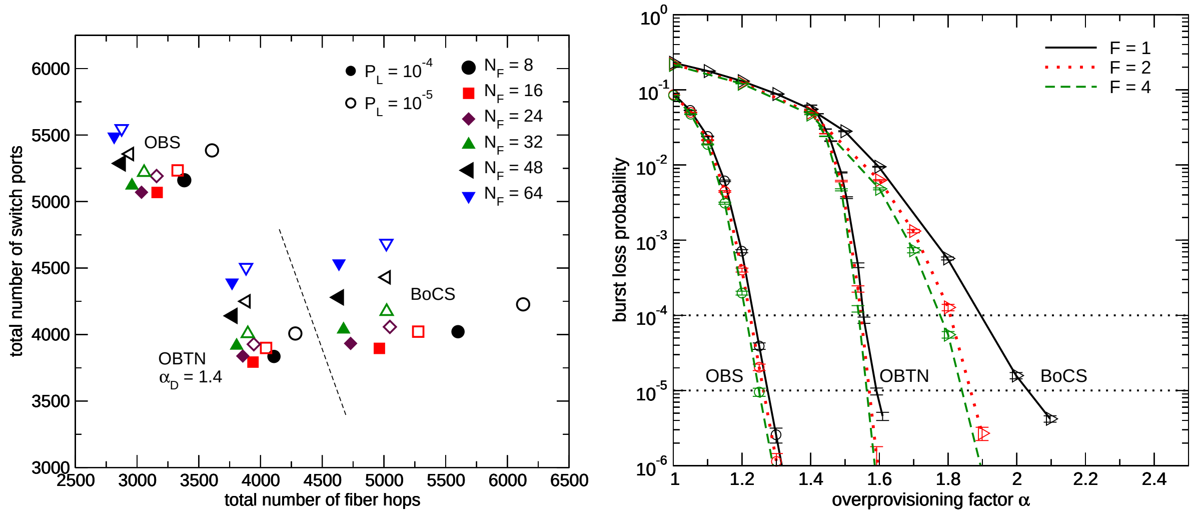
Figure 6.11 depicts the burst loss probability vs. the overprovisioning factor for a single FDL ($F = 1$) with a varying number of buffer ports N_F, i.e., a varying number of wavelengths per FDL. These network performance results for OBS, BoCS, and OBTN ($\alpha_D = 1.4$), support the findings for a single OBS node in Figure A.3[2]. Increasing the number of buffer ports improves the QoS for a given α or reduces the overprovisioning factor required for a given QoS. Especially BoCS benefits from additional buffer ports, because of its otherwise low contention resolution performance (cf. the improvement from $N_F = 8$ to 16). While increasing the number of FDL buffer ports beyond 32 leads to further improvements in the OBS network scenario, it has only minor effects for BoCS and OBTN.

In order to more precisely demonstrate the impact of a higher number of FDL buffer ports for a given network dimensioning, Figure 6.11(b) depicts the burst loss probability vs. the number of buffer ports for different fixed values of α. For low loss probabilities (lower two groups of curves), the loss probability floors out for a higher number of buffer ports in BoCS and OBTN but not yet in OBS. In contrast, for a high loss probability (upper group) all three architectures benefit only slightly and exhibit saturation. It can be further concluded that for BoCS and for OBTN with a low loss probability, values around $N_F = 32$ can be called optimal regarding QoS.

Regarding dimensioning, the FDL buffer *directly* contributes to the number of switch ports. Also, it *indirectly* affects both the number of switch ports and the number of fiber hops through overprovisioning factor α required to meet a QoS criterion. Therefore, the optimal configuration regarding the overall number of switch ports can very well be different from the optimal number considering QoS only.

Figure 6.12(a) presents the total number of switch ports vs. the total number of fiber hops for different values of N_F and the target burst loss probabilities of 10^{-4} and 10^{-5}. Starting

[2] As an overprovisioning factor of $\alpha > 1$ corresponds to an offered load value less than 1 the results match despite the different visual impression.



(a) Client layer and server layer resources ($F = 1$)

(b) Impact of number of buffer FDLs ($N_F = 32$)

Figure 6.12: Impact of different FDL buffer configurations

from $N_F = 8$, increasing the number of buffer ports first reduces both the number of switch ports and fiber hops. Then, with higher values of N_F the marginal gain of an FDL buffer in terms of a lower overprovisioning factor diminishes. This leads to a slower reduction in fiber hops and an increase in the total number of switch ports. Finally, BoCS and OBTN exhibit no further improvements in fiber hops but a steep increase in switch ports, which originates in the saturation effect discussed in conjunction with Figure 6.11(b). While values of $N_F = 16$ and 24 yield minimal or close to minimal total number of switch ports in all architectures, 32 buffer ports still reduce the number of fiber hops with a limited penalty in switch ports.

Regarding the impact of the target loss probability, Figure 6.12(a) shows only very marginal impact except for $N_F = 8$ and for the BoCS results. This behavior for BoCS can be explained by the slow improvement of the burst loss probability with an increasing overprovisioning factor in Figure 6.11(a) for these cases.

Apart from the number of FDL buffer ports, the number of FDLs in the buffer F is an important parameter regarding realization. Figure 6.12(b) varies F while keeping the total number of buffer ports, i.e., the product of buffer FDLs and wavelengths per FDL, constant at $N_F = 32$. Except for BoCS with $F = 1$, increasing the number of buffer FDLs yields no significant performance benefit. As the number of buffer ports remains constant, there is also no significant impact on the number of switch ports or fiber hops. Despite the same number of buffer ports, the different FDL buffer architectures are still not equivalent regarding node realization and scalability (cf. Section 3.5.4 and Section A.2). For instance, broadcast-and-select node architectures (cf. Section 3.5) scale substantially worse in terms of the number of output fibers than the number of wavelengths per fiber. As an additional buffer FDL leads to an increase in complexity comparable to that of an additional output fiber, fewer buffer FDLs with a higher number of wavelength channels yield realization and scalability advantages. Thus, it can be concluded that $F = 1$ is the best choice here.

Table 6.1: Total number of virtual links for different virtual topologies

	$A_{thr} = 0$	$A_{thr} = 1$	$A_{thr} = 2$	$A_{thr} = 3$
OBS	46	-	-	-
OBTN $H_{thr} = 1$	240[a]	218	188	150
OBTN $H_{thr} = 2$[b]	240[a]	218	188	150
OBTN $H_{thr} = 3$	170	150	128	106
BoCS	240[a]	-	-	-

a. full-mesh virtual topology
b. differs from $H_{thr} = 1$ only in the dimensioning process

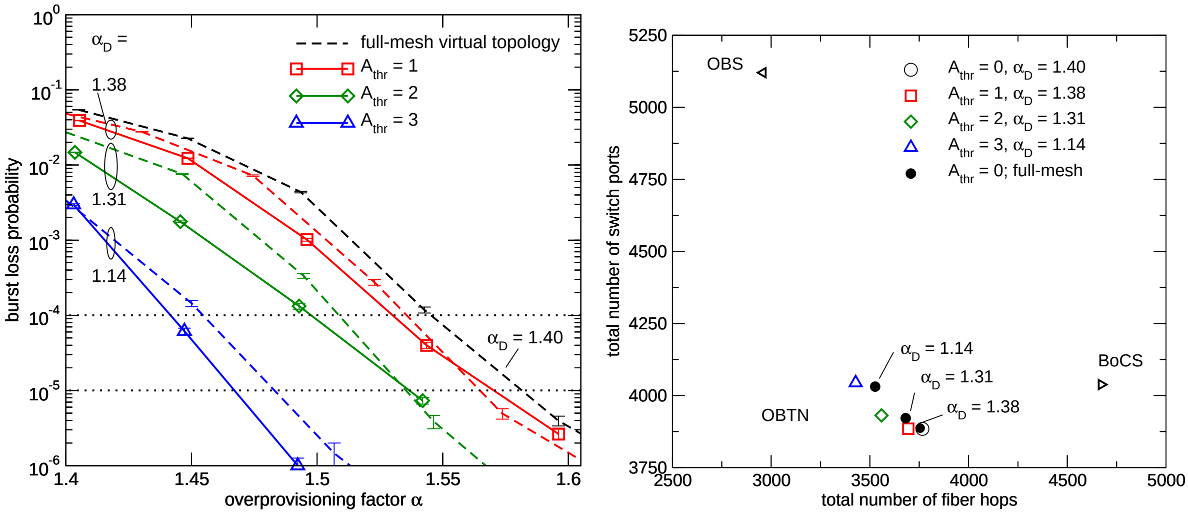
The evaluation of the impact of FDL buffer design on client layer and server layer resources demonstrates that a QoS-centric view alone is not sufficient. Thus, the resource model used here helps to quantify important trade-offs to derive proper design decisions. Beyond this level of abstraction, models for specific physical node architectures are required to also additionally study the impact of technological effects (cf. Section 3.5.4).

6.6 Impact of the Virtual Topology Design

In all previous sections of this chapter a full-mesh virtual topology was used for OBTN. While this approach is commonly applied in WDM core networks to minimize transit traffic, a sparser virtual topology can also be attractive as discussed in Section 4.4.1. Reasons are not only better scalability and simplified management but can also be resource efficiency. For instance, virtual links carrying only little traffic achieve a low statistical multiplexing gain and thus require a high overprovisioning factor. This effect can lead to a situation in which this penalty exhausts the gain of reducing the transit traffic. Therefore, this section studies virtual topology design approaches, in which virtual links are selected depending on traffic demands and/or depending on the number of physical hops of a virtual link. As discussed in Section 2.4, these criteria are also used in virtual topology design for WSN.

6.6.1 Demand-based Virtual Topology Design

As virtual links with only little traffic can reduce or diminish the advantages of a dense virtual topology, this first alternative approach only selects virtual links for those OBTN core node pairs which exchange more traffic than a given traffic demand threshold A_{thr}. The traffic demand matrix used for all previous evaluations contains traffic demands from 0 to 12.9 Erlang (wavelength channel equivalents). In order to only avoid the most inefficient virtual links, small traffic demand thresholds A_{thr} = 1, 2, and 3 Erlang are used in the following (they are written without the pseudo-unit Erlang to improve the clarity of the graphs). As can be seen from Table 6.1 (row with $H_{thr} = 1$), the number of virtual links decreases from 240 in the full-mesh case to 218, 188, and 150 for A_{thr} = 1, 2, and 3 respectively.



(a) Burst loss probability

(b) Client layer and server layer resources ($P_L = 10^{-4}$)

Figure 6.13: Impact of traffic demand thresholds on virtual topology design

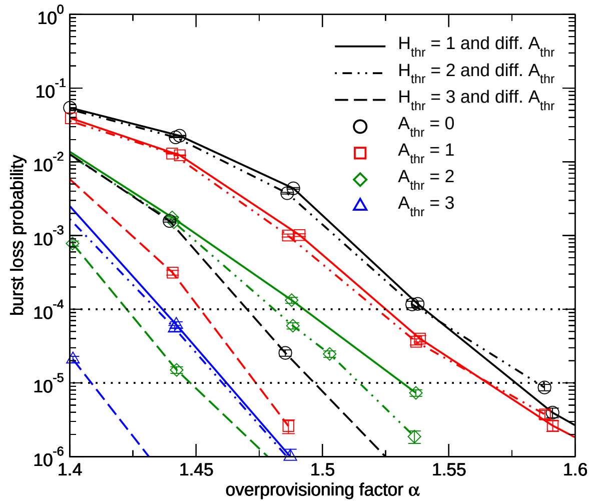
The dimensioning process outlined in Section 5.2 and depicted in Figure 5.2 is also applied here[3]. However, the resource pool for direct end-to-end lightpaths is reduced by the share of traffic demands that do not qualify for a virtual link. These resources are instead shifted to the resource pool for shared overflow capacity. This transfer of capacity corresponds to a reduced overprovisioning factor of the direct end-to-end lightpaths α_D and an increased ratio of shared overflow and total network capacity β.

Figure 6.13(a) shows the burst loss probability vs. the overprovisioning factor α for the different values of A_{thr}. Here, the dimensioning process for the demand-based virtual topology design starts with $\alpha_D = 1.4$ which is reduced to 1.38, 1.31, and 1.14 for A_{thr} = 1, 2, and 3 respectively. Thus, for a systematic analysis, the figure also depicts the results for the respective values of α_D with a full-mesh virtual topology.

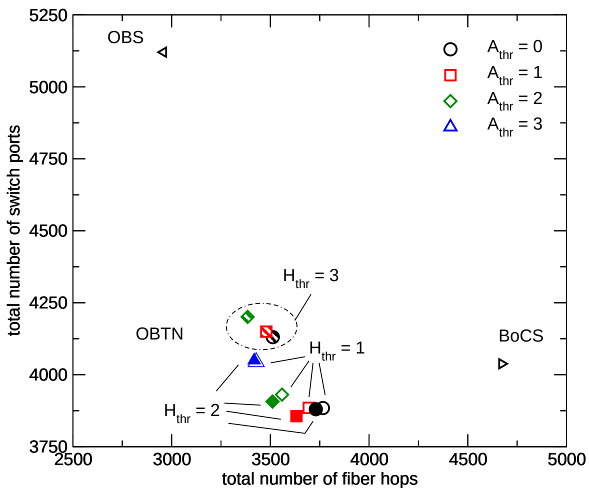
Compared to the full-mesh virtual topology with $\alpha_D = 1.4$ the demand-based designs significantly improve the QoS and thus only require a smaller overprovisioning factor α for a target QoS level. Regarding the full-mesh virtual topology with the same α_D value, demand-based virtual topology designs also yield a slight improvement for most overprovisioning factors.

In order to complete this study, Figure 6.13(b) depicts the total number of switch ports vs. the total number of fiber hops for OBTN with a target burst loss probability of $P_L = 10^{-4}$. For reference, the results for OBS and BoCS are also included. Compared to a full-mesh OBTN with $\alpha = 1.4$, the demand-based design reduces the total number of fiber hops (up to 9 % for $A_{thr} = 3$) with some penalty in the number of switch ports (up to 4 % for $A_{thr} = 3$). Compared

[3] In order to implement this threshold function in the dimensioning process most efficiently, a slightly different order in the capacity computation is used. This led to minimal differences in the results depicted in Figure 6.12 compared to Figure 6.1(a) and Figure 6.5.



(a) Burst loss probability



(b) Client layer and server layer resources ($P_L = 10^{-4}$)

Figure 6.14: Impact of path length thresholds on virtual topology design

to the full-mesh topologies with reduced α_D values, only marginal reductions in the number of fiber hops can be observed though.

Summarizing, the studied demand-based virtual topology design allows to reduce the number of virtual links: For $A_{thr} = 1$, the number of virtual links decreases by 10 % without substantial changes in client layer and server layer resources. In contrast, $A_{thr} = 2$ and 3 reduce the number of virtual links by 22–37 % at a smaller number of hops but (slightly) more switch ports.

6.6.2 Path-based Virtual Topology Design

While the demand-based virtual topology design avoids virtual links carrying only little traffic, the path-based design selects virtual links based on their number of physical hops. In order to primarily set-up virtual links for traffic demands generating a large amount of transit traffic, the studied approach only selects virtual links which span at least H_{thr} links in the physical topology. By definition, $H_{thr} = 1$ selects all possible virtual links and thus corresponds to the full-mesh case. Note that a threshold of $H_{thr} = 2$ leads to a full-mesh virtual topology as the single hop virtual links are always deployed for shared overflow capacity. However, for $H_{thr} = 3$, the number of virtual links is reduced to 170 as listed in Table 6.1.

The dimensioning of the path-based virtual topology follows the same adaptations to the dimensioning process in Section 5.2 discussed above. Although $H_{thr} = 1$ and $H_{thr} = 2$ both lead to a full-mesh virtual topology, they still differ in the dimensioning. In the former case, part of the resources for single hop virtual links are computed from the resource pool of end-to-end resources. In the latter case, they are all computed from the pool of shared overflow capacity yielding a higher statistical multiplexing gain.

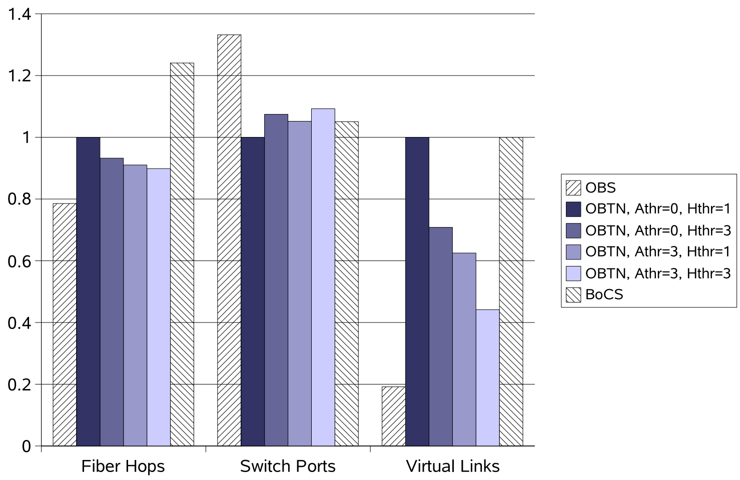


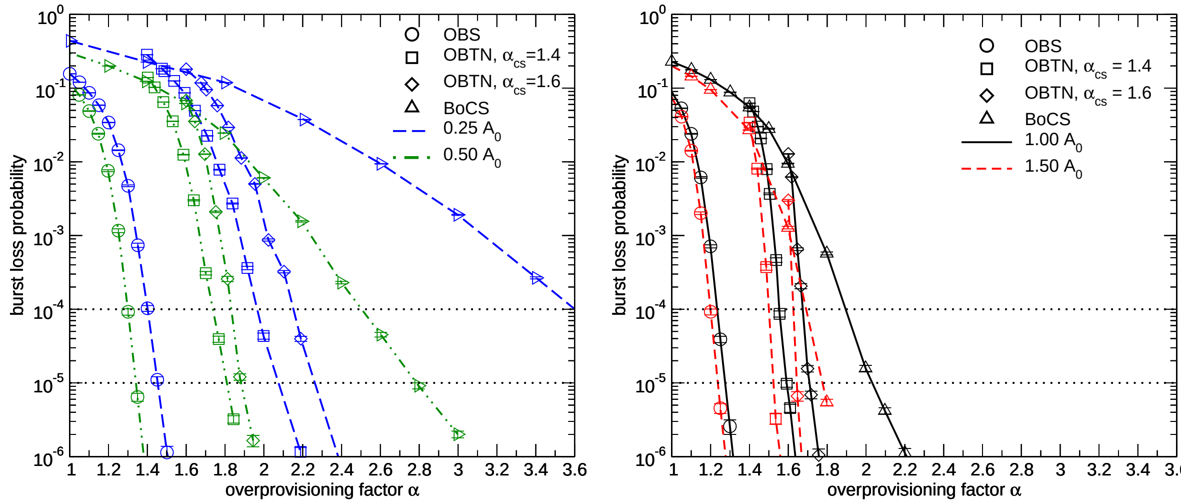
Figure 6.15: Overall comparison of the number of fiber hops, switch ports, and virtual links

As an extension to the pure path-based approach, the combination of path-based and demand-based virtual topology design can further leverage the problem of virtual links carrying too little traffic. Table 6.1 shows that a threshold of $H_{\text{thr}} = 3$ reduces the number of virtual links even more significantly from 218, 188, and 150 to 150, 128, and 106 for $A_{\text{thr}} = 1, 2, 3$ respectively.

Figure 6.14 analyzes the burst loss probability and the client layer and server layer resources for purely path-based ($A_{\text{thr}} = 0$) and combined schemes ($A_{\text{thr}} = 1, 2, 3$). Figure 6.14(a) shows independent of A_{thr} that the difference in the dimensioning process for $H_{\text{thr}} = 1$ and $H_{\text{thr}} = 2$ has only marginal impact on the QoS. In contrast, selecting virtual links with $H_{\text{thr}} = 3$ leads to considerable performance improvements for all demand thresholds. Regarding the client layer and server layer resources in Figure 6.14(b), the differences between $H_{\text{thr}} = 1$ and $H_{\text{thr}} = 2$ are again marginal. Consequently, the effect of the separate vs. integrated dimensioning for single-hop links is negligible. However, the improvements are more pronounced for $H_{\text{thr}} = 3$. In the latter case, the number of fiber hops decreases while the number of switch ports increases, both by about 7 %.

In order to summarize the section on the impact of virtual topology design, Figure 6.15 compares the number of fiber hops, switch ports, and virtual links for OBS, different OBTN virtual topologies, and BoCS. All numbers are normalized to the reference case of OBTN with a full-mesh virtual topology.

The figure shows that the demand-based and path-based virtual topology design approaches can reduce the number of virtual links in OBTNand thus simplify management. At the same time, they achieve small improvements in fiber hops at approximately the same penalties in switch ports. Although detailed relations of capital and operational expenditures finally have to decide on the optimal configuration, it can be concluded that OBTN is not limited to a full-mesh virtual topology and also performs well with a lower connectivity.



(a) $A < A_0$

(b) $A \geq A_0$

Figure 6.16: Impact of the total traffic demand on the burst loss probability

6.7 Impact of Traffic Demands

The trade-off regarding client layer and server layer resources for OBS, BoCS, and OBTN is controlled by the statistical multiplexing gain per virtual link and the amount of transit traffic in nodes (cf. Section 2.4.3). Because of the economy of scale, a network, which is dimensioned for an absolutely higher total traffic demand, achieves a better link utilization and thus requires relatively fewer client layer and server layer resources (cf., e.g., the figure in Chapter 4.1-1.4.4 in [Küh04b]). Also, the higher the traffic demand for a virtual link, the slower grows the utilization when scaling up the traffic demand. Consequently, a higher total traffic demand favors architectures with fewer traffic per virtual link like BoCS or OBTN and penalizes OBS. In order to quantify this effect, this section analyzes the sensitivity of the results obtained so far with respect to higher and lower traffic demands.

Figure 6.16 depicts the burst loss probability vs. the overprovisioning factor α for four different traffic demand values. Figure 6.16(a) compares OBS, BoCS, and OBTN for 0.25 and 0.5 times the total traffic demand A_0 used so far (cf. Section 5.3.3), while Figure 6.16(b) shows the results for the reference traffic demand A_0 and a 1.5 times higher demand. The figures confirm that smaller traffic demands lead to higher overprovisioning factors because of the reduced statistical multiplexing gain, particularly for BoCS and partly for OBTN. In contrast, a higher traffic demand reduces the required overprovisioning factor because of the higher statistical multiplexing gain which can again be best seen for BoCS.

The consequences of changing the total traffic demand on client layer and server layer resources are shown in Figure 6.17. Relative to OBS, lower traffic demands penalize BoCS and OBTN regarding the number of switch ports while a difference in the number of fiber hops remains. A higher traffic demand, however, increases the advantage in switch ports of OBTN and BoCS. Here, the OBTN realization with $\alpha_D = 1.4$ still needs less fiber hops than BoCS—although the margin becomes smaller.

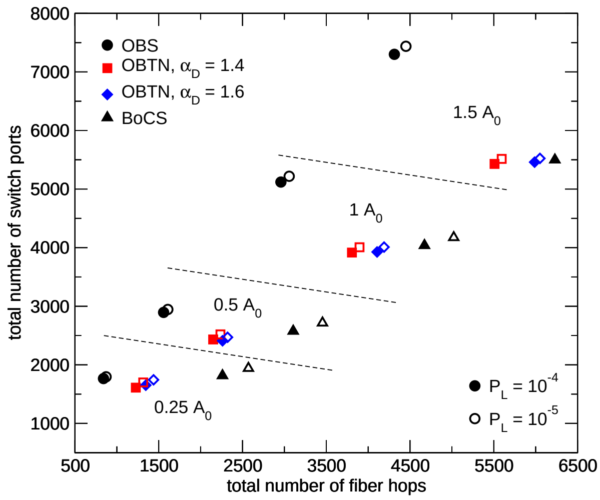


Figure 6.17: Impact of the total traffic demand on the client layer and server layer resources

Summarizing, OBS can be considered the natural choice for small traffic demands while BoCS ultimately becomes the natural choice for high traffic demands. OBTN provides an attractive solution for intermediate traffic demands where neither OBS nor BoCS can offer optimal solutions. Thus, it can be the architecture of choice to support network evolution for increasing traffic demands. In addition, the traffic demands in wavelength equivalents (or Erlang) do not continuously increase despite sustained traffic growth (in Gbps). Instead, increasing traffic demands trigger a migration towards higher bit-rates per wavelength channel, in transport networks typically by a factors of four. These transitions again reduce the traffic demand in wavelength equivalents and lead to such *intermediate* traffic demands rendering a solution between OBS and BoCS attractive.

7 Conclusions and Outlook

This thesis presents the design, modeling, and evaluation of the optical burst transport network architecture (OBTN). The architecture is motivated by the need for flexible, scalable, and cost-efficient transport in next generation networks. In addition, it is stimulated by the research activities towards highly dynamic optical network infrastructures.

OBTN defines a network architecture to transport and switch optical burst data in a core network. The design objectives for the OBTN architecture are (i) an overall high quality of service, (ii) a network design allowing for cost-efficiency and scalability, and (iii) a network evolution perspective based on the current wavelength-switched networks. These objectives are achieved by combining selected concepts, architectures, and strategies of optical burst and optical packet switching as well as of multi-layer traffic engineering.

In order to provide the background information for the design of OBTN, Chapter 2 introduced the general characteristics, requirements, and trends for next generation transport networks. Also, it discussed the concept of layering in next generation networks and its application in layer networks for the virtualization of transport resources. Consequently, virtual topology design and dimensioning were analyzed to quantify the trade-offs regarding connectivity and resource requirements. Chapter 2 also reviewed the fundamental technologies as well as currently emerging data and control plane architectures for optical transport networks. This presentation was then extended towards a long-term perspective. It described architectural constraints and classification criteria for highly dynamic optical network architectures. These criteria were used to characterize the *fast optical circuit switching*, *optical burst switching*, and *optical packet switching* architectures. Then, hybrid optical network architectures were discussed as a framework to combine wavelength-switched and optical burst/packet-switched networks.

Chapter 3 discussed the state of research and technology for optical burst switching to structure the design space and identify promising approaches. Thus, it presented the requirements for the different functions in an OBS network and classified the proposed architectures and mechanisms. Particularly, it addressed contention resolution which is necessary to achieve a high QoS in burst-switched networks. Here, wavelength conversion, fiber delay line buffering, alternative/deflection routing, and their combinations were looked at. It was concluded that wavelength conversion is a promising primary contention resolution strategy but should be complemented by FDL buffering and/or alternative routing. Thus, architectures, parameters, and operational strategies for FDL buffers were discussed in detail. This was supported by Appendix A which analyzed the performance of shared FDL buffers for different configurations and traffic characteristics. The review of alternative/deflection routing showed that it can only support other contention resolution schemes if it was closely controlled, i.e., if extensive deflections and route variations are avoided. Finally, architectures and realization aspects for burst-switched core nodes were presented to understand their resource and scalability constraints.

Chapter 4 presented the design rationale for OBTN and explained how OBTN combines a burst-switched client layer network with a wavelength-switched server layer network. Then, it introduced its fundamental concepts, namely the dense virtual topology, constrained alternative routing, and shared overflow capacity. These components were analyzed regarding their consequences for the overall node and network architecture. Further architectural details and variants as well as operational strategies of OBTN were discussed to complete the presentation. Finally, a qualitative discussion of OBTN with respect to optical burst switching and hybrid optical networks concluded this chapter.

Chapter 5 described a unified resource model which allows to dimension and evaluate burst-switched architectures with different virtual topologies. Also, it detailed the dimensioning process for OBTN. Then, it addressed the simulation methodology and the reference evaluation scenario used in Chapter 6. It discussed metrics for node and network resources as well as for QoS performance. Finally, it derived QoS objectives for burst-switched core networks.

Chapter 6 evaluated OBTN and compared it with the two burst-switched reference architectures OBS and *Burst-over-Circuit-Switching*. OBS uses a sparse virtual topology while BoCS employs a full-mesh virtual topology. The evaluations in Section 6.1 showed that for the same high target QoS, suitable OBTN dimensionings require substantially less resources in burst-switched nodes than OBS and slightly less than BoCS. This improvement comes at the cost of higher resource requirements compared to OBS in the underlying wavelength-switched server layer. However, applying the cost relations for wavelength-switched networks, in which bandwidth is considered a commodity and client layer resources the major cost driver, OBTN yields an overall cost reduction. The best results for OBTN were obtained when approximately 10 % of the network capacity was assigned as shared overflow capacity.

In Section 6.2, the comparison of OBTN and OBS was extended towards OBS architectures without an FDL buffer and OBS architectures with alternative routing. It was demonstrated that bufferless OBS with and without alternative routing requires approximately the same amount of server layer resources as OBTN. However, it consumes more client layer resources. For OBS with an FDL buffer, alternative routing did neither impact the client layer nor the server layer resources substantially.

In Section 6.3, the analysis of different hunting modes for contention resolution in OBTN identified that alternative routing should be probed before the FDL buffer. Furthermore, the effectiveness of constrained alternative routing and of the shared overflow capacity in OBTN was assessed by isolating them in Section 6.4. This was achieved by comparing OBTN and BoCS which use the same virtual topology but differ in the routing flexibility and network dimensioning. The evaluations showed that applying any of the concepts alone to BoCS did not yield any or only limited improvements or even produced severe penalties. However, if applied together as OBTN, they harmonize and effectively improve performance.

Then, Section 6.5 compared different FDL buffer architectures and dimensionings. It quantified the trade-off of improved contention resolution, thus reduced node and network resources and the additional node resources required for the FDL buffer. The results showed that increasing the number of FDL buffer ports up to approx. 16–24 led to fewer node and network resources. Beyond this dimensioning, node resources increased with diminishing reductions regarding resources in the underlying server layer.

Section 6.6 demonstrated that OBTN is not restricted to the full-mesh virtual topology used for all previous evaluations. A traffic demand-based and a path-length-based virtual topology design approach were introduced and investigated. For one example parameterization, the number of virtual links could be more than halved with a small penalty in client layer resources and a small gain in server layer resources.

Finally, Section 6.7 studied the impact of changing total traffic demands on the performance and on resource requirements of all three architectures. It showed that when increasing the total traffic demand, BoCS becomes increasingly efficient. For small traffic demands OBS becomes more attractive. However, OBTN provides an attractive solution for intermediate traffic demands where neither OBS nor BoCS can offer optimal solutions.

Concluding, OBTN was shown to offer an overall high QoS, to effectively reduce the node resources of the burst-switched client layer, and to perform well in a wavelength-switched network context.

Further work could extend the OBTN architecture by mechanisms for QoS differentiation and survivability. Dynamic and reconfigurable virtual topologies could be incorporated to adapt to traffic demand variations. As applicable analytical performance models for FDL buffers become available, analytical performance models could be defined for OBTN. Then, virtual topology designs could be based on exact performance models which could further optimize the resource requirements. Finally, the resource model which currently uses the number of switch ports and fiber hops as metrics could be extended towards a larger set of client layer and server layer resources. Such more detailed, yet implementation-specific resource model could then also allow for cost comparisons.

A Evaluation of OBS Nodes with Shared FDL Buffers

The discussion in Section 3.4.3 introduced several FDL buffer architectures and their design parameters. Based on QoS and complexity arguments, it motivated the application of a shared single-stage feedback FDL buffer with coordinated output and buffer scheduling. In order to obtain a deeper understanding of those architectures and to support the parameter selection for Chapter 6, this section looks closer at this type of buffer. It evaluates the design parameters regarding their effects on performance, first by qualitative arguments, then quantitatively by simulation. The results presented in Section A.1 summarize several previous studies on the performance of FDL buffers [Sch01, Gau02a, GBPS03, Gau04, GKS04b]. Section A.2 discusses results with an integrated evaluation of performance and technology outlined in Section 3.5.4 and first reported in [BGPS03, GBPS03, GBP05].

A.1 Performance of OBS Nodes with Shared FDL Buffers

This section focuses on the performance of shared FDL buffers with a single fixed delay or a set of fixed delays. An example for this type of buffer is the single-stage feedback FDL buffer allowing only one recirculation (Figure 3.16(b)). Because of the feedback architecture and full wavelength conversion capability assumed in the OBS node, e.g., the TAS node in Figure 3.15, wavelength conversion is available before and after buffering. Thus, the buffer capacity can be most flexibly shared. Resources on the output fiber and in the buffer are scheduled in a coordinated way. In the *PreRes* strategy, bursts are only buffered if they can be scheduled on the output fiber for the time when they leave the buffer.

A.1.1 Qualitative Discussion

Figure A.1 illustrates a simple performance model for the coordinated scheduling of a single FDL and a WDM output fiber. Although not depicted, the FDL buffer is assumed to be shared among all output fibers so that buffering requests arrive from all output fibers. A newly arriving burst cannot be directly scheduled to the output fiber, which happens with probability P_{O}, the node attempts to schedule an FDL and a wavelength channel on the output fiber together. The scheduling of the output fiber wavelength channel is needed after buffering with delay T_{F} and fails with probability $P_{\mathrm{O}}^{*}(T_{\mathrm{F}})$. The scheduling of a wavelength channel in the respective FDL is

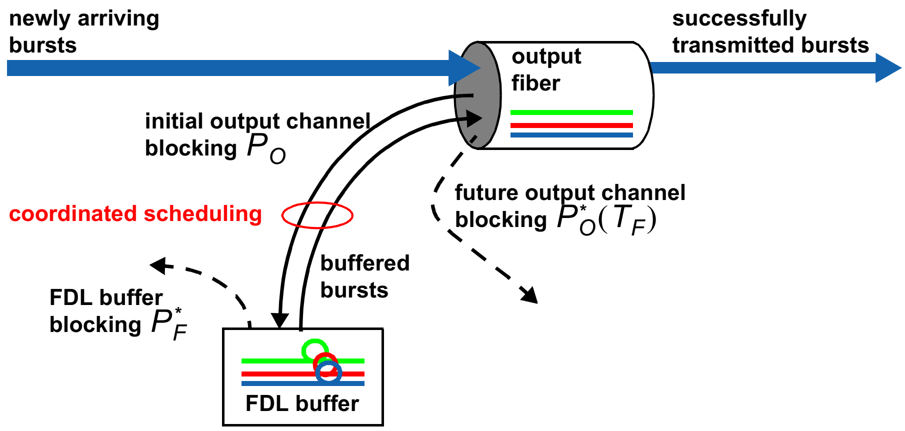


Figure A.1: Blocking on the output wavelength channels and in the FDL buffer

blocked with probability P_{F}^*. In both cases, the super-script * indicates that the probabilities are conditioned on the blocking event at burst arrival.

Based on this model, a burst is finally lost with probability P_{L} if both its direct and delayed scheduling is blocked. In general, this can be expressed as:

$$P_{\mathrm{L}} = P_{\mathrm{O}} \cdot [1-(1-P_{\mathrm{O}}^*(T_{\mathrm{F}}))(1-P_{\mathrm{F}}^*)] \qquad \text{(A.1)}$$

and for the practically interesting cases with $P_{\mathrm{F}}^* \ll 1$ and $P_{\mathrm{O}}^*(T_{\mathrm{F}}) \ll 1$ by the approximation:

$$P_{\mathrm{L}} \simeq P_{\mathrm{O}} \cdot (P_{\mathrm{O}}^*(T_{\mathrm{F}}) + P_{\mathrm{F}}^*) \qquad \text{(A.2)}$$

The multi-FDL case can be treated accordingly by considering the joint probability for the output fiber and for the buffer FDL.

The relation Eq. (A.2) allows to isolate the effects of traffic intensity and output fiber dimensioning from those of the FDL buffer dimensioning. In extreme regimes, either the output channel blocking probabilities P_{O} and $P_{\mathrm{O}}^*(T_{\mathrm{F}})$ or the buffer blocking probability P_{F}^* determine the system behavior. In-between these two regimes, a superposition of the effects can be observed in the simulations. For a single wavelength channel and a dedicated buffer, this relation has been extended to a quantitative analysis in [RC05].

The initial and future output channel blocking probabilities P_{O} and $P_{\mathrm{O}}^*(T_{\mathrm{F}})$ in Eq. (A.2) are closely correlated. First, because they describe the state of the same output fiber at two close time instances. Second, because $P_{\mathrm{O}}^*(T_{\mathrm{F}})$ is conditioned on an initial blocking event and thus focuses on the time of or immediately following a busy period. Therefore, a higher buffering delay, T_{F}, increases the probability that the busy period ended. This, again, reduces the blocking probability $P_{\mathrm{O}}^*(T_{\mathrm{F}})$ until it floors out due to saturation. Here, a higher delay flexibility can offer additional opportunities for successful scheduling.

P_{O} and $P_{\mathrm{O}}^*(T_{\mathrm{F}})$ are both strongly influenced by the traffic intensity and the number of wavelength channels on the output fiber. Therefore, a different behavior can be observed in low and high load regimes. The probability P_{F}^* of not being able to reserve a wavelength channel in the FDL depends on the traffic offered to the FDL buffer and its capacity. The following evaluations show

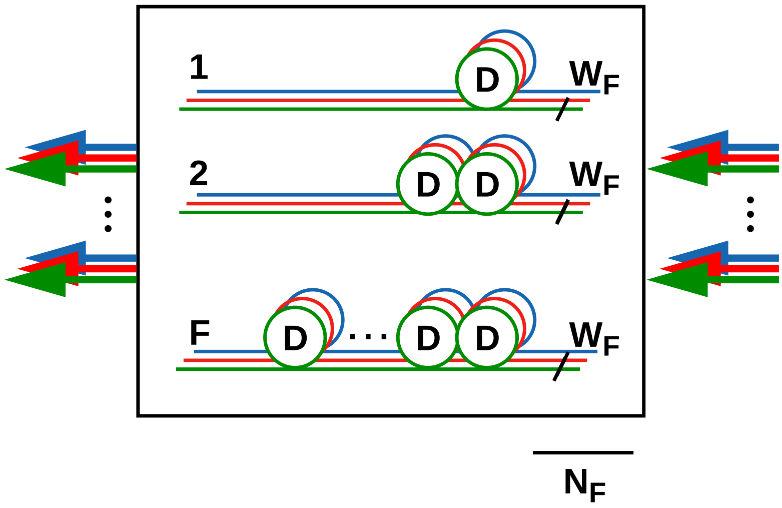


Figure A.2: Architecture and parameters of the degenerate multi-FDL buffer

that the number of buffer ports N_F is a decisive parameter. Therefore, FDL buffers with a small and large number of ports are distinguished below. The following performance evaluation is structured according to those principal regimes. Its quantitative results support the conclusions of the former qualitative discussion.

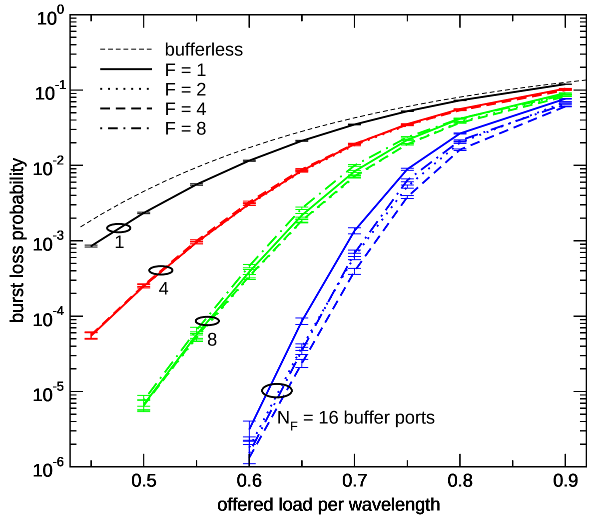
A.1.2 Simulation Scenario

All simulations are performed for a single OBS node with $N = 4$ input and output fibers each carrying $M = 16$ wavelength channels, i.e., a total of 64 input and output wavelength channels. Unless stated differently, the studies apply FDL buffers as depicted in Figure A.2 with up to $N_F = 64$ ports. The degenerate FDL buffers are structured in $F = 1, 2, 4$, and 8 FDLs with delays $T_{F,i} = i \cdot D$ for $1 \leq i \leq F$.

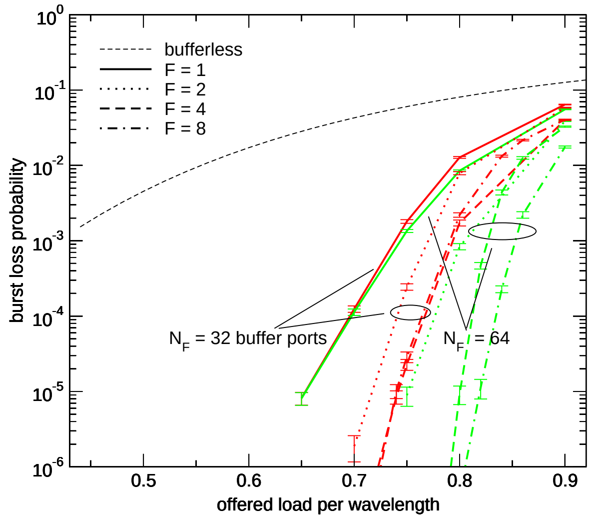
Bursts arrive to the node following a Poisson process. The destination of bursts is uniformly distributed over all output fibers. Also, selection of the wavelength on which a burst arrives at the node follows a uniform distribution. For the burst transmission time, the simulations use a negative exponential distribution with a mean value of $h = 10\,\mu$s, i.e., a mean burst length of $12,500$ Byte for a bit-rate of $b = 10$ Gbps. The wavelength channels on the output fiber and in the FDLs are scheduled together according to *PreRes*. Finally, all resources use a first-fit reserve-a-fixed duration algorithm with void filling (cf. Section 3.3).

A.1.3 Performance Evaluation Results

The evaluations first explain the principal operational regimes of OBS nodes with shared FDL buffer. Then they focus on the effects of the number of FDL buffer ports, the delay granularity, and the structure of the FDL buffer. Finally, the influence of the distribution of the burst transmission time is studied with parameterizations obtained from IP traffic traces.



(a) $N_F \leq 16$



(b) $N_F = 32, 64$

Figure A.3: Burst loss probability for different numbers of FDL buffer ports

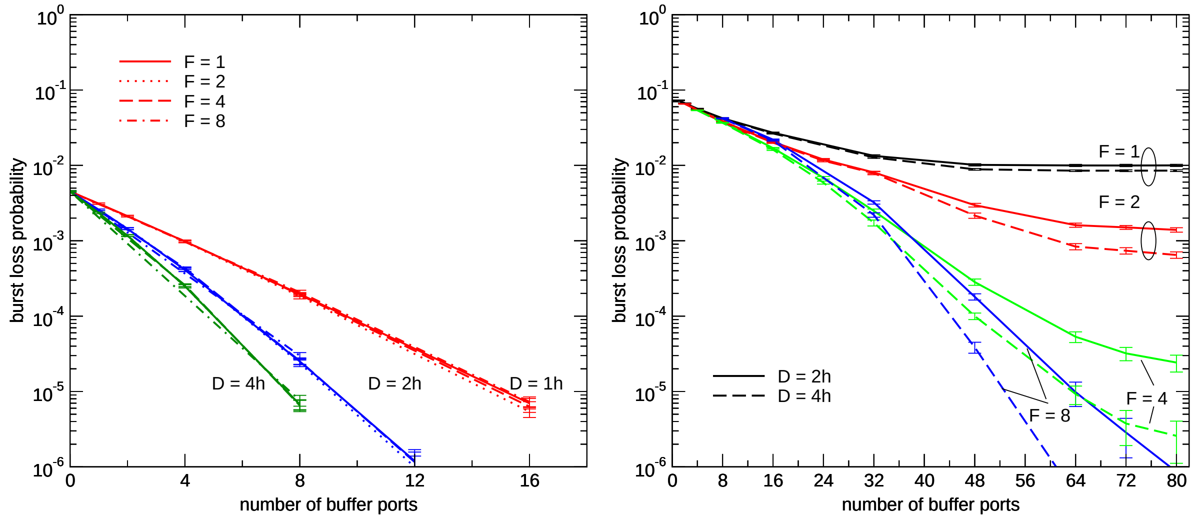
A.1.3.1 Impact of the Offered Load

Figure A.3 depicts the mean burst loss probability vs. the offered load per wavelength in the case without and with FDL buffer. The burst loss probability for the bufferless case is obtained from the M/G/n loss model. The delay granularity, i.e., the delay of the shortest buffer FDL, is chosen to be 4 mean burst transmission times ($D = 4h = 40\,\mu$s). In the left subfigure, results are presented for up to 16 FDL ports, while the right subfigure shows the cases of 32 and 64 ports. The parameter sets correspond to the operational regimes of a small and a large number of FDL buffer ports as discussed in the qualitative discussion above.

The figures show that the performance improves over the entire range of load compared to the bufferless case. For small load values, the burst loss probability is reduced by several orders of magnitude. In contrast, the improvement is only limited for high load situations, in which the initial and future output blocking probabilities are so high that the buffer is flooded and even a high capacity cannot resolve contention effectively.

A.1.3.2 Impact of the FDL Buffer Dimensioning

Figure A.3 shows that the burst loss probability decreases with an increasing number of FDL buffer ports. For only few buffer ports, the FDL buffer is the sole bottleneck. Thus, the structure of the FDL buffer, i.e., the number of FDLs F, has no or in the case of 16 ports only minor impact. In contrast, for many buffer ports, the FDL buffer does not constitute the bottleneck any more. Consequently, the structure of the FDL buffer and the delay flexibility become significant. On the one hand, increasing the buffer capacity for a single FDL beyond 32 does not yield any improvements because no additional buffered bursts can be scheduled with this delay. On the other hand, increasing the delay flexibility by a higher number of FDLs in the buffer reduces the burst loss probability. However, as the blocking probability decreases, the buffer capacity



(a) $\rho = 0.5$

(b) $\rho = 0.8$

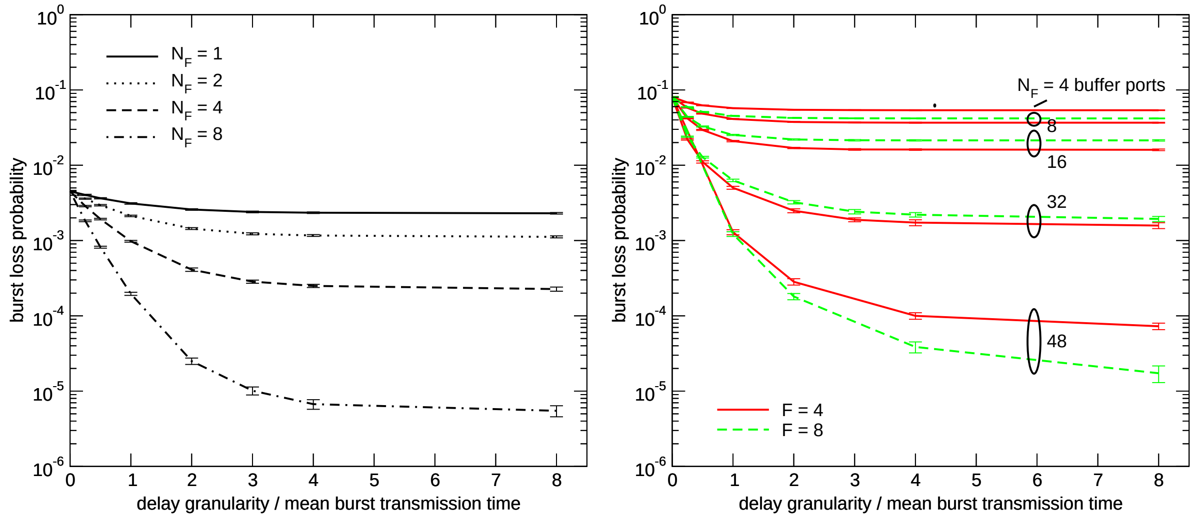
Figure A.4: Effect of the buffer structure for different delay granularities

becomes an issue again: While doubling the number of FDLs from 4 to 8 is ineffective for 32 buffer ports, it is highly effective for 64 buffer ports.

For low and high load values ($\rho = 0.5$ and $\rho = 0.8$), Figure A.4(a) and Figure A.4(b) analyze the impact of the number of buffer ports and the delay granularity ($D = h, 2h, 4h$). Both figures extend the above finding, that the structure of the FDL buffer has negligible impact for few buffer ports, to smaller delay granularities. Again, this can be seen from the overlapping curves.

For $\rho = 0.5$, increasing the number of ports N_F from 1 to 16 reduces the burst loss probability by orders of magnitude. In contrast, this has only little effect for $\rho = 0.8$ as future scheduling is blocked on the highly utilized output fibers. Still, when increasing the number to 80, the buffer becomes effective and both the burst loss probability and the number of buffer ports at which the curves floor significantly depend on the buffer structure. For a large number of buffer ports, the influence of the delay granularity increases. In Figure A.4(b), a buffer with 4 FDLs and a delay granularity of $D = 4h$ achieves a burst loss probability comparable to a buffer with 8 FDLs and $D = 2h$. Here, arguments concerning the realization complexity favor the implementation with fewer FDLs but higher delay granularity.

Regarding the realization complexity, the number of FDL in the buffer should be minimized as is shown in Section A.2. Thus, if the QoS and network utilization requirements allow to use an FDL buffer with only few ports, the solution with a single FDL is clearly beneficial. However, if FDL buffers with more ports have to be employed, the best solution from a realization point of view has to be obtained from an integrated evaluation of performance and technology (cf. Section 3.5.4 and Section A.2).



(a) $\rho = 0.5$ (b) $\rho = 0.8$

Figure A.5: Effect of the delay granularity for different numbers of buffer ports

A.1.3.3 Impact of the Delay Granularity

Figure A.5 depicts the dependence of the delay granularity in a more detailed way, again for offered loads per wavelength of $\rho = 0.5$ and 0.8. For the lower load and a small number of buffer ports, the case of a single FDL is selected based on realization arguments. For a load of 0.8 FDL buffers with 4 and 8 FDLs are used. In both cases, increasing the delay granularity D reduces the burst loss probability until saturation at approx. 4 mean burst transmission times is reached. Obviously, this delay value is optimal to not only overcome the busy period of the output fiber but also to distribute the buffered bursts on the output fiber uniformly over time.

As FDL buffers were originally used in slotted nodes, for which a delay granularity of one packet transmission time (slot) is the natural choice, this parameterization was also used in many studies on unslotted nodes. However, previous results in [Gau02a, GBPS03] as well as the results presented here motivate a higher delay granularity. Thus, a delay granularity of $D = 4h$ is used for all simulation studies in Chapter 6.

A.1.3.4 Impact of the Distribution of the Burst Transmission Time

In order to validate the previous assumption of a negative exponential distribution for the burst transmission time, the sensitivity of the results with respect to different distributions is studied.

The performance is studied for 4 general distributions with the same mean burst transmission time but different coefficients of variation c_T. The negative exponential distribution ($c_T = 1$) is compared with a shifted negative exponential ($c_T = 0.5$) and two second order hyperexponential ($c_T = 1.5$ and $c_T = 2$) distributions[1]. For these distributions, Figure A.6(a) shows that in

[1]The hyperexponential distribution satisfies the symmetry condition $ph_1 = (1-p)h_2$ where p is the branch probability and h_i are the mean values of the respective phases.

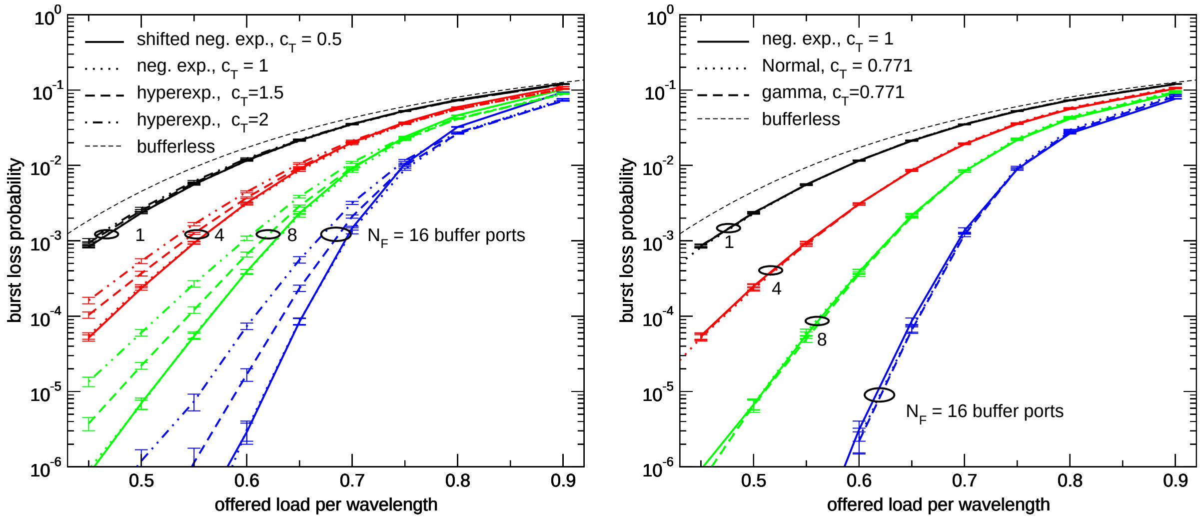


Figure A.6: Effect of the distribution of the burst transmission time

principle a coefficient of variation greater than 1 yields a higher burst loss probability compared to the negative exponential and shifted negative exponential case. Also, it can be seen that the differences become more prominent as the number of buffer ports increases. However, the burst assembly process can closely control the burst statistics.

As discussed in Section 3.1, several publications show that with time-based assembly the distribution of the burst transmission time converges to a Normal distribution [Lae02, IA02, YCQ02b, dVRG04]. In addition, [Lae02] shows that a Gamma distribution also provides a good fit[2]. For the following study, both Normal and Gamma distributions are used and parameterized as if resulting from time-based assembly of the Internet packet trace of the uplink of a student dormitory network at the University of Stuttgart [Sas04]. The packet trace has a mean bit-rate during the considered period of 25 Mbps. It is aggregated in intervals of 4 ms in order to model time-based assembly. At the same time, this assembly period generates a mean burst length of $12,500$ Byte used above. In the assembled burst trace, the coefficient of variation of the burst length was measured to be $c_{\mathrm{T}} = 0.771$, i.e., smaller than for the negative exponential distribution.

Figure A.6(b) depicts the burst loss probability vs. the offered load per wavelength for the negative exponential (cf. Figure A.3), the Normal and the Gamma distribution. In addiion, the bufferless case according to the M/G/n loss model is plotted, for which the distribution of the burst transmission time has no effect. It can be seen that the burst loss probability is insensitive with respect to these distributions of the burst transmission time. These results were again found to be insensitive with respect to the FDL buffer structure. Therefore, they are only presented for an FDL buffer with a single FDL. Based on this observed approx. insensitivity, a negative exponential burst transmission time is a valid modeling assumption and is used in all simulations studies in Chapter 6.

[2]The advantage of the Gamma distribution is its strictly positive support. For the Normal distribution this has to be enforced by resampling and adapting the mean value to match the offered load generated with the negative exponential and Gamma distributions, respectively.

A.2 Integrated Evaluation of Performance and Technology

This section presents the results of an integrated evaluation of performance and technology as described in Section 3.5.4. These results were initially published in [BGPS03, GBPS03, GBP05]. It quantifies the impact of different architectural variants of tune-and-select (TAS) OBS nodes on maximum and maximum effective throughput.

A.2.1 Evaluation Scenario

The integrated evaluation comprises the following three TAS node configurations which are depicted in Figure 3.15 and Figure 3.16:

- **TAS**: basic TAS architecture
- **TAS-dFDL**: TAS architecture with one dedicated FDL per output fiber
- **TAS-shFDL**: TAS architecture with $F = 1,2,3$ or 4 FDLs in a shared FDL buffer

The TAS nodes are studied in an isolated scenario with $N = 4$ input and output fibers. They are modeled with state-of-the-art components as characterized in [GBPS03]. The signal analysis models a reference signal path from a 3R wavelength converter (cf. Section 2.2.2) in one node to the 3R wavelength converter in the adjacent node. Along this path, it considers noise and crosstalk as physical impairments for bit-rates of 10 Gbps and 40 Gbps. It uses a threshold for the maximum acceptable BER of 10^{-22} in order to have enough safety margin for impairments not considered yet, e.g., SOA dynamics.

While the dedicated FDL buffers have a delay of $T_F = 2h = 20\,\mu s$, the shared FDL buffer is degenerate with a delay granularity of $D = 2h = 20\,\mu s$. Both FDL buffers coordinate the FDL and output fiber scheduling according to the PreRes strategy (cf. Section 3.4.3.4). For the performance and technology evaluation, the number of wavelengths in the FDL buffer and on the input/output fibers is the same. The TAS nodes with FDL buffers primarily use wavelength conversion to resolve contentions and only resort to the FDL buffer if this fails. For this isolated node scenario, the maximum acceptable burst loss probability is chosen to be 10^{-6}, again to have enough margin for a network scenario.

A.2.2 Results of the Integrated Evaluation

Figure A.7 depicts the maximum number of wavelengths per input or output fiber, the maximum throughput (wire bars) and the maximum effective throughput (solid bars). First, the evaluation shows that the maximum throughput and the maximum effective throughput values are in the range of 2.5–6 Tbps for both 10 Gbps and 40 Gbps. Therefore, the nodes with 40 Gbps obviously cannot benefit from the higher bit-rate regarding the maximum throughput values. This is due to the fact that at 40 Gbps noise degrades the signal more severly.

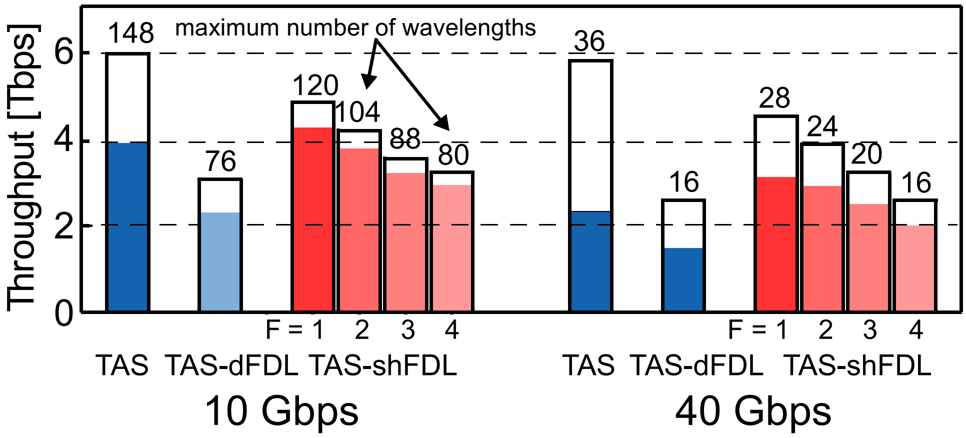


Figure A.7: Results of the integrated evaluation of performance and technology [GBPS03]

From a pure performance perspective, Figure A.3 shows that FDL buffers reduce the burst loss probability. Also, it demonstrates, that for a high number of buffer ports multi-FDL buffers are advantageous. Figure A.7 shows that introducing a dedicated or shared FDL buffer and increasing the number of FDLs F leads to higher signal degradation and thus a smaller maximum throughput compared to TAS. Still, the improved contention resolution performance and the higher achievable utilization partly pay off for the maximum effective throughput. Here, TAS-shFDL with 1 FDL for 10 Gbps and TAS-shFDL with up to 3 FDLs for 40 Gbps outperform the basic TAS architecture. TAS-dFDL yields the lowest maximum and maximum effective throughput for both bit-rates.

Bibliography

[ABMR06] D. Allan, N. Bragg, A. McGuire, and A. Reid. Ethernet as carrier transport infrastructure. *IEEE Communications Magazine*, 44(2):134–140, 2006.

[AIMM03] J. Aracil, M. Izal, D. Morato, and E. Magana. Multiresolution analysis of optical burst switching traffic. In *Proceedings of the IEEE International Conference on Networks (ICON)*, pages 409–412, 2003.

[AK04] N. Akar and E. Karasan. Exact calculation of blocking probabilities for bufferless optical burst switched links with partial wavelength conversion. In *Proceedings of BroadNets 2004*, pages 110–117, 2004.

[Ams83] S. Amstutz. Burst switching–an introduction. *IEEE Communications Magazine*, 21(8):36–42, 1983.

[Ams89] S. Amstutz. Burst switching–an update. *IEEE Communications Magazine*, 27(9):50–57, 1989.

[APW04] R. Almeida, J. Pelegrini, and H. Waldman. Optical buffer modelling for performance evaluation considering any packet inter-arrival time distribution. In *Proceedings of the IEEE International Conference on Communications (ICC)*, volume 3, pages 1771–1775 Vol.3, 2004.

[APW05] R. Almeida, J. Pelegrini, and H. Waldman. A generic-traffic optical buffer modeling for asynchronous optical switching networks. *IEEE Communications Letters*, 9(2):175–177, 2005.

[ATM06] ATM-Forum. Public specifications of the ATM-Forum on the Internet, 2006. `http://www.mfaforum.org/ftp/pub/approved-specs/`.

[BBB+03] M. Barry, S. Bodamer, S. Brink, L. Levrau, and J. Späth. Dynamic multi-layer networks - architecture, modelling, tool support. MultiTeraNet Dynamo Deliverable, Marconi GmbH, Backang, November 2003.

[BBE+05] I. Baldine, A. Bragg, G. Evans, M. Pratt, M. Singhal, D. Stevenson, and R. Uppalli. Jumpstart deployments in ultra-high-performance optical networking testbeds. *IEEE Communications Magazine*, 43(11):S18–S25, 2005.

[BBPV03] M. Baresi, S. Bregni, A. Pattavina, and G. Vegetti. Deflection routing effectiveness in full-optical IP packet switching networks. In *Proceedings of the IEEE International Conference on Communications (ICC)*, volume 2, pages 1360–1364 vol.2, 2003.

[BBR+03] D. Blumenthal, J. Bowers, L. Rau, H.-F. Chou, S. Rangarajan, W. Wang, and K. Poulsen. Optical signal processing for optical packet switching networks. *IEEE Communications Magazine*, 41(2):S23–S29, 2003.

[BC00] S. Bodamer and J. Charzinski. Evaluation of effective bandwidth schemes for self-similar traffic. In *Proceedings of the 13th ITC Specialist Seminar on IP Traffic Measurement, Modeling and Management*, pages 21.1–21.10, Monterey, CA, September 2000.

[BCR+99] D. Blumenthal, A. Carena, L. Rau, V. Curri, and S. Humphries. All-optical label swapping with wavelength conversion for WDM-IP networks with subcarrier multiplexed addressing. *IEEE Photonics Technology Letters*, 11(11):1497–1499, 1999.

[BDL+01a] A. Banerjee, J. Drake, J. Lang, B. Turner, K. Kompella, and Y. Rekhter. Generalized multiprotocol label switching: an overview of routing and management enhancements. *IEEE Communications Magazine*, 39(1):144–150, 2001.

[BDL+01b] A. Banerjee, L. Drake, L. Lang, B. Turner, D. Awduche, L. Berger, K. Kompella, and Y. Rekhter. Generalized multiprotocol label switching: an overview of signaling enhancements and recovery techniques. *IEEE Communications Magazine*, 39(7):144–151, 2001.

[BFB93] F. Borgonovo, L. Fratta, and J. Bannister. Unslotted deflection routing in all-optical networks. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, pages 119–125, Houston, November 1993.

[BGH+04] A. Betker, C. Gerlach, R. Huelsermann, M. Jaeger, M. Barry, S. Bodamer, J. Späth, C. M. Gauger, and M. Köhn. A set of typical transport network scenarios for network modelling. In *Proceedings of the 5. ITG Symposium on Photonic Networks*, pages 65–72, Leipzig, May 2004.

[BGP02] S. Bregni, G. Guerra, and A. Pattavina. Optical packet switching of IP traffic. In *Proceedings of the Optical Network Design and Modelling Conference (ONDM)*, Torino, February 2002.

[BGPS03] H. Buchta, C. M. Gauger, E. Patzak, and J. Saniter. Limits of effective throughput of optical burst switches based on semiconductor optical amplifiers. In *Proceedings of the Optical Fiber Communication Conference (OFC)*, Atlanta, March 2003.

[BHJD96] M. Bischoff, M. Huber, O. Jahreis, and F. Derr. Operation and maintenance for an all-optical transport network. *IEEE Communications Magazine*, 34(11):136–142, 1996.

[BHS02] S. Bjornstad, D. Hjelme, and N. Stol. An optical packet switch design with shared electronic buffering and low bit rate add/drop inputs. In *Proceedings of the International Conference on Transparent Optical Networks (ICTON)*, volume 1, pages 69–72, 2002.

[Bjo04] S. Bjornstad. Packet switched hybrid optical networks. In *Proceedings of the 6th International Conference on Transparent Optical Networks (ICTON)*, Wroclaw, June 2004.

[BM96] D. Banerjee and B. Mukherjee. A practical approach for routing and wavelength assignment in large wavelength-routed optical networks. *IEEE Journal on Selected Areas in Communications*, 14(5):903–908, 1996.

[BM00] D. Banerjee and B. Mukherjee. Wavelength-routed optical networks: linear formulation, resource budgeting tradeoffs, and a reconfiguration study. *IEEE/ACM Transactions on Networking*, 8(5):598–607, 2000.

[BPSG03] H. Buchta, E. Patzak, J. Saniter, and C. M. Gauger. Maximal and effective throughput of optical switching nodes for optical burst switching. In *Proceedings of the 4. ITG Symposium on Photonic Networks*, Leipzig, May 2003.

[BPV03] S. Bregni, A. Pattavina, and G. Vegetti. Architectures and performance of AWG-based optical switching nodes for IP networks. *IEEE Journal on Selected Areas in Communications*, 21(7):1113–1121, 2003.

[Bra05] F. Braun. Entwurf einer Basisarchitektur fuer einen OBS-Randknoten und Realisierung mit programmierbarer Logik. Diploma thesis, University of Stuttgart, 2005.

[Bri98] U. Briem. *Verbindungslose Datenkommunikation ueber ATM-Weitverkehrsnetze: Architekturen, Protokolle und Verkehrsleistung*. Dissertation, University of Stuttgart, Stuttgart, 1998.

[BRPS02] I. Baldine, G. N. Rouskas, H. G. Perros, and D. Stevenson. Jumpstart: a just-in-time signaling architecture for WDM burst-switched networks. *IEEE Communications Magazine*, 40(2):82–89, February 2002.

[BS05a] N. Barakat and E. Sargent. Analytical modeling of offset-induced priority in multiclass OBS networks. *IEEE Transactions on Communications*, 53(8):1343–1352, 2005.

[BS05b] N. Barakat and E. Sargent. Dual-header optical burst switching: a new architecture for WDM burst-switched networks. In *Proceedings of IEEE INFOCOM*, volume 1, pages 685–693, 2005.

[BSG04] S. Bodamer, J. Späth, and C. Glingener. An efficient method to estimate transponder count in multi-layer transport networks. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, pages 1780–1785, Dallas/TX, December 2004.

[BT92] P. E. Boyer and D. P. Tranchier. A reservation principle with applications to the ATM traffic control. *Computer Networks and ISDN Systems*, 24(4):321–334, 1992.

[Buc05] H. Buchta. *Analysis of Physical Constraints in an Optical Burst Switching Network.* Dissertation, Technical University of Berlin, Berlin, April 2005.

[Cal00] F. Callegati. Optical buffers for variable length packets. *IEEE Communications Letters*, 4(9):292–294, 2000.

[Car04] J. Cartagena. Scheduling of assembled traffic in IP-over-Photonics edge nodes. Student thesis, University of Stuttgart, September 2004.

[CB97] M. E. Crovella and A. Best. Self-similarity in World Wide Web traffic: evidence and possible causes. *IEEE/ACM Transactions on Networking*, 5(6):835–846, December 1997.

[CBD+03] J. Cheyns, E. V. Breusegem, C. Develder, A. Ackaert, M. Pickavet, and P. Demeester. Performance improvement of an internally-blocking optical packet/burst switch. In *Proceedings of the IEEE International Conference on Communications (ICC)*, Anchorage, May 2003.

[CCB+00] T. Y. Chai, T. H. Cheng, S. Bose, C. Lu, and G. Shen. Analytical model for a WDM optical cross-connect with limited conversion capability. *IEEE Communications Letters*, 4(11):369–371, 2000.

[CCC+04] F. Callegati, W. Cerroni, G. Corazza, C. Develder, M. Pickavet, and P. Demeester. Scheduling algorithms for a slotted packet switch with either fixed or variable length packets. *Photonic Network Communications*, 8(2):163–176, September 2004.

[CCF01] T. Chich, J. Cohen, and P. Fraigniaud. Unslotted deflection routing: a practical and efficient protocol for multihop optical networks. *IEEE/ACM Transactions on Networking*, 9(1):47–59, February 2001.

[CCK05] J. Choi, J. Choi, and M. Kang. Dimensioning burst assembly process in optical burst switching networks. *IEICE Transactions on Communications*, E88-B(10):3855–3863, 2005.

[CCR00] F. Callegati, G. Corazza, and C. Raffaelli. Design of a WDM optical packet switch for IP traffic. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, volume 2, pages 1283–1287 vol.2, 2000.

[CCRZ04] F. Callegati, W. Cerroni, C. Raffaelli, and P. Zaffoni. Wavelength and time domain exploitation for QoS management in optical packet switches. *Computer Networks*, 44(4):569–582, 2004.

[CEBCS03] S. Charcranoon, T. El-Bawab, H. Cankaya, and J.-D. Shin. Group-scheduling for optical burst switched (OBS) networks. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, volume 5, pages 2745–2749, 2003.

[CFK+96] I. Chlamtac, A. Fumagalli, L. Kazovsky, P. Melman, W. Nelson, P. Poggiolini, M. Cerisola, A. Choudhury, T. Fong, R. Hofmeister, C.-L. Lu, A. Mekkittikul, I. Sabido, D.J.M., C.-J. Suh, and E. Wong. CORD: contention resolution by delay lines. *IEEE Journal on Selected Areas in Communications*, 14(5):1014–1029, 1996.

[CFS00] I. Chlamtac, A. Fumagalli, and C.-J. Suh. Multibuffer delay line architectures for efficient contention resolution in optical switching nodes. *IEEE Transactions on Communications*, 48(12):2089–2098, 2000.

[CHA+01] M. Chia, D. Hunter, I. Andonovic, P. Ball, I. Wright, S. Ferguson, K. Guild, and M. O'Mahony. Packet loss and delay performance of feedback and feed-forward arrayed-waveguide gratings-based optical packet switches with WDM inputs-outputs. *IEEE/OSA Journal of Lightwave Technology*, 19(9):1241–1254, 2001.

[Cha03a] V. W. S. Chan, editor. *IEEE Journal on Selected Areas in Communications: Optical communications and networking series*, volume 21, October 2003.

[Cha03b] V. W. S. Chan, editor. *IEEE Journal on Selected Areas in Communications: Optical communications and networking series*, volume 21, November 2003.

[Cha04] V. W. S. Chan, editor. *IEEE Journal on Selected Areas in Communications: Optical communications and networking series*, volume 22, November 2004.

[Cha05a] V. W. S. Chan, editor. *IEEE Journal on Selected Areas in Communications: Optical communications and networking series*, volume 23, August 2005.

[Cha05b] V. W. S. Chan. Editorial: Optical communications and networking series. *IEEE Journal of Selected Areas in Communications*, 23(8):1441–1443, August 2005.

[CHH94] P. Cochrane, R. Heckingbottom, and D. Heatley. The hidden benefits of optical transparency. *IEEE Communications Magazine*, 32(9):90–97, 1994.

[CHKKC03] C. Chang-Hasnain, P.-c. Ku, J. Kim, and S.-l. Chuang. Variable optical buffer using slow light in semiconductor nanostructures. *Proceedings of the IEEE*, 91(11):1884–1897, 2003.

[CHT01] Y. Chen, M. Hamdi, and D. Tsang. Proportional QoS over OBS networks. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, San Antonio, November 2001.

[Cin03] T. Cinkler. Traffic and λ Grooming. *IEEE Network*, March 2003.

[CLCQ02] X. Cao, J. Li, Y. Chen, and C. Qiao. Assembling TCP/IP packets in optical burst switched networks. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, Taipei, November 2002.

[CMR+05] F. Callegati, G. Muretto, C. Raffaelli, P. Zaffoni, and W. Cerroni. A framework for performance evaluation of OPS congestion resolution. In *Proceedings of Optical Network Design and Modelling Conference (ONDM)*, pages 242–249, 2005.

[Com06] D. E. Comer. *Internetworking with TCP/IP*, volume 1. Prentice Hall International, 5th edition, 2006.

[COS03] COST 266. Final report: Advanced infrastructures for photonic networks, 2003. http://www2.cordis.lu/cost/src/266_indivpage.htm.

[CP02] J.-B. Chang and C.-S. Park. Efficient channel-scheduling algorithm in optical burst switching architecture. In *Proceedings of the Workshop on High Performance Switching and Routing*, pages 194–198, 2002.

[CSBO03] C. T. Chou, F. Safaei, P. Boustead, and I. Ouveysi. A hybrid optical network architecture consisting of optical cross connects and optical burst switches. In *Proceedings of the IEEE International Conference on Computer Communications and Networks*, pages 53–58, Dallas/TX, October 2003.

[CTT99] G. Castanon, L. Tancevski, and L. Tamil. Routing in all-optical packet switched irregular mesh networks. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, volume 1B, pages 1017–1022, 1999.

[CVR02] H. Chaskar, S. Verma, and R. Ravikanth. Robust transport of IP traffic over WDM using optical burst switching. *Optical Networks Magazine*, 3(4), July/August 2002.

[CW03] B. Chen and J. Wang. Hybrid switching and p-routing for optical burst switching networks. *IEEE Journal on Selected Areas in Communications*, 21(7):1071–1080, September 2003.

[CWXQ03] Y. Chen, H. Wu, D. Xu, and C. Qiao. Performance analysis of optical burst switched node with deflection routing. In *Proceedings of the IEEE International Conference on Communication*, Anchorage, May 2003.

[CZC$^+$04] K. Cheng, Y. Zhu, X. Cao, D. H. K. Tsang, and D. T. K. Tong. Architecture for an optical burst switch with WDM shared output buffer. *Journal of Optical Networking*, 3(6):417–432, 2004.

[DB02] M. Dueser and P. Bayvel. Analysis of a dynamically wavelength-routed optical burst switched network architecture. *IEEE/OSA Journal of Lightwave Technology*, 20(4):574–585, April 2002.

[DDC$^+$03] L. Dittmann, C. Develder, D. Chiaroni, F. Neri, F. Callegati, W. Koerber, A. Stavdas, M. Renaud, A. Rafel, J. Sole-Pareta, W. Cerroni, N. Leligou, L. Dembeck, B. Mortensen, M. Pickavet, N. Le Sauze, M. Mahony, B. Berde, and G. Eilenberger. The European IST project DAVID: a viable approach toward optical packet switching. *IEEE Journal on Selected Areas in Communications*, 21(7):1026–1040, 2003.

[DEL02] A. Detti, V. Eramo, and M. Listanti. Performance evaluation of a new technique for IP support in a WDM optical network: Optical composite burst switching (OCBS). *IEEE/OSA Journal of Lightwave Technology*, 20(2):154–165, February 2002.

[DG01] K. Dolzer and C. M. Gauger. On burst assembly in optical burst switching networks—a performance evaluation of just-enough-time. In *Proceedings of the 17th International Teletraffic Congress (ITC 17)*, pages 149–160, Salvador, Brazil, December 2001.

[DGSB01] K. Dolzer, C. M. Gauger, J. Späth, and S. Bodamer. Evaluation of reservation mechanisms for optical burst switching. *AEÜ International Journal of Electronics and Communications*, 55(1):18–26, January 2001.

[DHB$^+$04] C. Develder, B. V. Houdt, C. Blondia, M. Pickavet, and P. Demeester. Analytical MMAP-based bounds for packet loss in optical packet switching with recirculating fdl buffers. *Photonic Network Communications*, 8(2):149–161, September 2004.

[DHS98] S. L. Danielsen, P. B. Hansen, and K. E. Stubkjaer. Wavelength conversion in optical packet switching. *IEEE/OSA Journal of Lightwave Technology*, 16(12):2095–2108, December 1998.

[Dix03] S. Dixit. *IP over WDM: Building the Next-Generation Optical Internet*. Wiley, April 2003.

[DJMS98] S. L. Danielsen, C. Joergensen, B. Mikkelsen, and K. E. Stubkjaer. Analysis of a WDM packet switch with improved performance under bursty traffic conditions due to tuneable wavelength converters. *IEEE/OSA Journal of Lightwave Technology*, 16(5):729–735, May 1998.

[DKKB00] M. Dueser, R. Kozlovski, R. Killey, and P. Bayvel. Design trade-offs in optical burst switched networks with dynamic wavelength allocation. In *Proceedings of the 26th European Conference on Optical Communication (ECOC 2000)*, volume 2, pages 23 – 24, 2000.

[DL01] A. Detti and M. Listanti. Application of tell and go and tell and wait reservation strategies in a optical burst switching network: a performance comparison. In *Proceedings of the IEEE International Conference on Telecommunications*, Bucharest, June 2001.

[DL05] A. Detti and M. Listanti. Amplification effects of the send rate of TCP connection through an optical burst switching network. *Optical Switching and Networking*, 2(1):49–69, May 2005.

[dM$^+$04] I. de Miguel et al. Polymorphic architectures for optical networks and their seamless evolution towards next generation networks. *Photonic Network Communications*, 8(2):177–189, September 2004.

[DMJD97] S. L. Danielsen, B. Mikkelsen, C. Joergensen, and T. Durhuus. WDM packet switch architectures and analysis of the influence of tuneable wavelength converters on the performance. *IEEE/OSA Journal of Lightwave Technology*, 15(2):219–227, February 1997.

[Dol04] K. Dolzer. *Mechanisms for quality of service differentiation in optical burst switched networks*. Dissertation, University of Stuttgart, Stuttgart, 2004.

[DR00] R. Dutta and G. N. Rouskas. A survey of virtual topology design algorithms for wavelength routed optical networks. *Optical Networks*, 1(1):73–89, January 2000.

[dVRG04] M. de Vega Rodrigo and J. Goetz. An analytical study of optical burst switching aggregation strategies. In *Proceedings of the Third International Workshop on Optical Burst Switching (WOBS)*, San Jose, October 2004.

[EDLW04] G. Eilenberger, L. Dembeck, W. Lautenschläger, and J. Wolde. Burst-Techniken für das optische Internet. In *Proceedings of the 5. ITG Symposium on Photonic Networks*, pages 137–142, Leipzig, May 2004.

[EL00] V. Eramo and M. Listanti. Packet loss in a bufferless optical WDM switch employing shared tunable wavelength converters. *IEEE/OSA Journal of Lightwave Technology*, 18(12):1818–1833, December 2000.

[ELS05] V. Eramo, M. Listanti, and M. Spaziani. Resources sharing in optical packet switches with limited-range wavelength converters. *IEEE/OSA Journal of Lightwave Technology*, 23(2):671–687, 2005.

[EM00] J. Elmirghani and H. Mouftah. All-optical wavelength conversion: technologies and applications in DWDM networks. *IEEE Communications Magazine*, 38(3):86–92, 2000.

[Epp98] D. Eppstein. Finding the k shortest paths. *SIAM Journal on Computing*, 28(2):652–674, 1998.

[FFWG02] P. Fan, C. Feng, Y. Wang, and N. Ge. Investigation of the time-offset-based QoS support with optical burst switching in WDM networks. In *Proceedings of the IEEE International Conference on Communications (ICC)*, volume 5, pages 2682–2686, 2002.

[FLB05] D. Fiems, K. Laevens, and H. Bruneel. Performance analysis of an all-optical packet buffer. In *Proceedings of the Optical Network Design and Modelling Conference (ONDM)*, pages 221–226, 2005.

[FML$^+$89] M. Farries, P. Morkel, R. Laming, T. Birks, D. Payne, and E. Tarbox. Operation of erbium-doped fiber amplifiers and lasers pumped with frequency-doubled Nd:YAG lasers. *IEEE/OSA Journal of Ligthwave Technology*, 7(10):1473–1477, 1989.

[FPS02] H. Feng, E. Patzak, and J. Saniter. Methods for stabilizing the gain of EDFAs in burst switching optical networks. *Photonic Network Communications*, 4(2):151–166, May 2002.

[FWB04] D. Fiems, J. Walraevens, and H. Bruneel. Discrete time queueing analysis of a fibre delay line structure with granularity. In *Proceedings of the Optical Network Design and Modelling Conference (ONDM)*, Gent, Belgium, February 2004.

[FZJ05] F. Farahmand, Q. Zhang, and J. Jue. Dynamic traffic grooming in optical burst-switched networks. *IEEE/OSA Journal of Ligthwave Technology*, 23(10):3167–3177, 2005.

[G.694.1] ITU. Rec. G.694.1: Spectral grids for WDM applications: DWDM frequency grid, International Telecommunication Union, ITU-T, June 2002.

[G.694.2] ITU. Rec. G.694.2: Spectral grids for WDM applications: CWDM wavelength grid, International Telecommunication Union, ITU-T, December 2003.

[G.702] ITU. Rec. G.702: Digital hierarchy bit rates, International Telecommunication Union, ITU-T, November 1988.

[G.7041] ITU. Rec. G.7041/Y.1303: Generic Framing Procedure (GFP), International Telecommunication Union, ITU-T, December 2001.

[G.7042] ITU. Rec. G.7042/Y.1305: Link capacity adjustment scheme (LCAS) for virtual concatenated signals, International Telecommunication Union, ITU-T, November 2001.

[G.707] ITU. Rec. G.707/Y.1322: Network node interface for the synchronous digital hierarchy (SDH), International Telecommunication Union, ITU-T, October 2000.

[G.709] ITU. Rec. G.709/Y.1331: Interfaces for the Optical Transport Network (OTN)breu, International Telecommunication Union, ITU-T, March 2003.

[G.7713.2] ITU. Rec. G.7713.2/Y.1704.2: Distributed Call and Connection Management: Signalling mechanism using GMPLS RSVP-TE, International Telecommunication Union, ITU-T, 2003.

[G.805] ITU. Rec. G.805: Generic functional architecture of transport networks, International Telecommunication Union, ITU-T, March 2000.

[G.807] ITU. Rec. G.807/Y.1302: Requirements for automatic switched transport networks (ASTN), International Telecommunication Union, ITU-T, 2001.

[G.8080] ITU. Rec. G.8080/Y.1304: Architecture for the automatically switched optical network (ASON), International Telecommunication Union, ITU-T, 2001.

[G.809] ITU. Rec. G.809: Functional architecture of connectionless layer networks, International Telecommunication Union, ITU-T, March 2003.

[G.872] ITU. Rec. G.872: Architecture of optical transport networks, International Telecommunication Union, ITU-T, 2001.

[Gal05] A. Galicher. Optical network dimensioning with effective bandwidth theory. Diploma thesis, University of Stuttgart, Stuttgart, Month 2005.

[GAN91] R. Guerin, H. Ahmadi, and M. Naghshineh. Equivalent capacity and its application to bandwidth allocation in high-speed networks. *IEEE Journal on Selected Areas in Communications*, 9(7):968–981, September 1991.

[Gau00] C. M. Gauger. Untersuchung von Reservierungsverfahren für Optical Burst Switching. Diploma thesis, University of Stuttgart, Stuttgart, 2000.

[Gau02a] C. M. Gauger. Dimensioning of FDL buffers for optical burst switching nodes. In *Proceedings of the Optical Network Design and Modelling Conference (ONDM)*, Torino, February 2002.

[Gau02b] C. M. Gauger. Performance of converter pools for contention resolution in optical burst switching. In *Proceedings of the Optical Networking and Communications Conference (OptiComm)*, Boston, October 2002.

[Gau04] C. M. Gauger. Optimized combination of converter pools and FDL buffers for contention resolution in optical burst switching. *Photonic Network Communications*, 8(2):139–148, September 2004.

[Gau05] C. M. Gauger. Optical burst switching—the network perspective. In *Proceedings of SPIE Asia-Pacific Optical Communications (APOC)*, Shanghai, November 2005.

[GBP05] C. M. Gauger, H. Buchta, and E. Patzak. Integrated evaluation of performance and technology - throughput of optical burst switching nodes under dynamic traffic. *paper in preparation*, 2005.

[GBPS03] C. M. Gauger, H. Buchta, E. Patzak, and J. Saniter. Performance meets technology - an integrated evaluation of OBS nodes with FDL buffers. In *Proceedings of the 1st International Workshop on Optical Burst Switching*, Dallas/TX, October 2003.

[GCT00] A. Ge, F. Callegati, and L. S. Tamil. On optical burst switching and self-similar traffic. *IEEE Communications Letter*, 4(3), March 2000.

[GDSB01] C. M. Gauger, K. Dolzer, J. Späth, and S. Bodamer. Service differentiation in optical burst switching networks. In *Proceedings of the 2. ITG Symposium on Photonic Networks*, pages 124–132, Dresden, March 2001.

[GG05] S. Gunreben and C. M. Gauger. Dynamic bandwidth adaptation in NG SDH/WDM transport networks using LCAS. In *Proceedings of the Networks and Optical Conference (NOC)*, July 2005.

[GKS04a] C. M. Gauger, M. Köhn, and J. Scharf. Comparison of contention resolution strategies in OBS network scenarios. In *Proceedings of the 6th International Conference on Transparent Optical Networks (ICTON)*, volume 1, pages 18–21 vol.1, 2004.

[GKS04b] C. M. Gauger, M. Köhn, and J. Scharf. Performance of contention resolution strategies in OBS network scenarios. In *Proceedings of the 9th Optoelectronics and Communications Conference/3rd International Conference on the Optical Internet (OECC/COIN)*, Yokohama/Japan, July 2004.

[GKZM05] C. M. Gauger, M. Köhn, J. Zhang, and B. Mukherjee. Network performance of optical burst/packet switching: The impact of dimensioning, routing and contention resolution. In *Proceedings of the 6. ITG Symposium on Photonic Networks*, Leipzig, May 2005.

[GLM+00] O. Gerstel, B. Li, A. McGuire, G. Rouskas, K. M. Sivalingam, and Z. Zhang, editors. *IEEE Journal on Selected Areas in Communications: Protocols and architectures for next generation optical WDM networks*, volume 18, October 2000.

[GM03] A. Gencata and B. Mukherjee. Virtual-topology adaptation for WDM mesh networks under dynamic traffic. *IEEE/ACM Transactions on Networking*, 11(2):236–247, 2003.

[GM05a] C. M. Gauger and B. Mukherjee. Optical burst transport network (OBTN) - a novel architecture for efficient transport of optical burst data over lambda grids. In *Proceedings of IEEE High Performance Switching and Routing (HPSR)*, Hong Kong, May 2005.

[GM05b] C. M. Gauger and B. Mukherjee. Why we should consider virtual topologies for OBS. In *Proceedings of the 5th International Workshop on Optical Burst Switching (WOBS)*, Boston/MA, October 2005.

[GPTB05] S. Gorshe, G. Parsons, M. Truskowski, and I. Busi. Editorial for special issue on Ethernet WAN transport. *IEEE Communications Magazine*, 43(11):134–135, 2005.

[GPTV04] S. Gorshe, G. Parsons, M. Truskowski, and M. Vissers. Ethernet WAN transport. *IEEE, Communications Magazine*, 42(3):62–63, 2004.

[Gre04] P. Green. Fiber to the home: the next big broadband thing. *IEEE Communications Magazine*, 42(9):100–106, 2004.

[GRG+98a] P. Gambini, M. Renaud, C. Guillemot, F. Callegati, I. Andonovic, B. Bostica, D. Chiaroni, G. Corazza, S. Danielsen, P. Gravey, P. Hansen, M. Henry, C. Janz, A. Kloch, R. Krahenbuhl, C. Raffaelli, M. Schilling, A. Talneau, and L. Zucchelli. Transparent optical packet switching: network architecture and demonstrators in the KEOPS project. *IEEE Journal on Selected Areas in Communications*, 16(7):1245–1259, 1998.

[GRG+98b] C. Guillemot, M. Renaud, P. Gambini, C. Janz, I. Andonovic, R. Bauknecht, B. Bostica, M. Burzio, F. Callegati, M. Casoni, D. Chiaroni, F. Clerot, S. Danielsen, F. Dorgeuille, A. Dupas, A. Franzen, P. Hansen, D. Hunter, A. Kloch, R. Krahenbuhl, B. Lavigne, A. Le Corre, C. Raffaelli, M. Schilling, J.-C. Simon, and L. Zucchelli. Transparent optical packet switching: the European ACTS KEOPS project approach. *Journal of Lightwave Technology*, 16(12):2117–2134, 1998.

[GSSC03] S. Gowda, R. Shenai, K. Sivalingam, and H. Cankaya. Performance evaluation of tcp over optical burst-switched (OBS) WDM networks. In *Proceedings of the IEEE International Conference on Communications (ICC)*, volume 2, pages 1433–1437 vol.2, 2003.

[GvBK+06] C. M. Gauger, E. van Breusegem, P. Kühn, M. Pickavet, and P. Demeester. Hybrid optical network architectures: Bringing together packets and circuits. *IEEE Communications Magazine*, 44(8), August 2006.

[HA00] D. Hunter and I. Andronovic. Approaches to optical internet packet switching. *IEEE Communications Magazine*, 38(9):116–122, 2000.

[Haa93] Z. Haas. The 'staggering switch': an electronically controlled optical packet switch. *IEEE/OSA Journal of Ligthwave Technology*, 11(5):925–936, 1993.

[HCA+97] D. Hunter, D. Cotter, R. Ahmad, W. Cornwell, T. Gilfedder, P. Legg, and I. Andonovic. 2x2 buffered switch fabrics for traffic routing, merging, andshaping in photonic cell networks. *IEEE/OSA Journal of Ligthwave Technology*, 15(1):86–101, 1997.

[HCA98] D. Hunter, M. Chia, and I. Andonovic. Buffering in optical packet switches. *IEEE/OSA Journal of Ligthwave Technology*, 16(12):2081–2094, 1998.

[HCG+98] D. Hunter, W. Cornwell, T. Gilfedder, A. Franzen, and I. Andonovic. SLOB: a switch with large optical buffers for packet switching. *IEEE/OSA Journal of Ligthwave Technology*, 16(10):1725–1736, 1998.

[HDG03] G. Hu, K. Dolzer, and C. M. Gauger. Does burst assembly really reduce self-similarity? In *Proceedings of the Optical Fiber Communication Conference (OFC)*, Atlanta, March 2003.

[HDH+04] Y. Huang, D. Datta, J. Heritage, Y.-C. Kim, and B. Mukherjee. A novel OBS node architecture using waveband-selective switching for reduced component cost and improved performance. In *The 17th Annual Meeting of the IEEELasers and Electro-Optics Society (LEOS)*, volume 1, pages 426–427 Vol.1, 2004.

[HGJ05] G. Hu, C. M. Gauger, and S. Junghans. Performance analysis of the CSMA/CA MAC protocol in the DBORN otpical MAN network architecture. In *Proceedings of the 19th International Teletraffic Congress (ITC 19)*, Beijing/CN, 2005.

[HHN] A. Hamilton, T. Hare, and H. Niemeyer. Quotenverfahren mit Restausgleich nach groessten Bruchteilen.

[HL05] C.-F. Hsu and T.-L. Liu. On deflection routing in optical burst-switched networks. *Journal of High Speed Networks*, 14(4):341–362, 2005.

[HLH02] C.-F. Hsu, T.-L. Li, and N.-F. Huang. Performance analysis of deflection routing in optical burst-switched networks. In *Proceedings of IEEE INFOCOM*, New York, June 2002.

[HM95] G. Hudek and D. Muder. Signaling analysis for a multi-switch all-optical network. In *Proceedings of the IEEE International Conference on Communication*, pages 1206 – 1210, June 1995.

[HMQ+05] G. Hu, G. Muretto, F. Querzola, C. Gauger, and C. Raffaelli. Traffic and performance analysis of optical packet/burst assembly with self similar traffic. In *Proceedings of the 7th International Conference on Transparent Optical Networks (ICTON)*, volume 1, pages 189–192, 2005.

[HN01] A. Hill and F. Neri. Optical switching networks: from circuits to packets. *IEEE Communications Magazine*, 39(3):107–108, 2001.

[HNC+99] D. Hunter, M. Nizam, M. Chia, I. Andonovic, K. Guild, A. Tzanakaki, M. O'Mahony, L. Bainbridge, M. Stephens, R. Penty, and I. White. WASP-NET: a wavelength switched packet network. *IEEE Communications Magazine*, 37(3):120–129, 1999.

[HSKS03] S. L. Hyun Sook Kim and J. S. Song. Optical burst switching with limited deflection routing. *IEICE - Transactions on Communications*, E86B(5):1550–1554, May 2003.

[Hui88] J. Hui. Resource allocation for broadband networks. *IEEE Journal on Selected Areas in Communications*, 6(9):1598–1608, 1988.

[I.121] ITU. Rec. I.121: Broadband aspects of ISDN, International Telecommunication Union, ITU-T, April 1991.

[IA02] M. Izal and J. Aracil. On the influence of self-similarity on optical burst switching traffic. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, pages 2308–2312, Taipei, November 2002.

[ID03] ITU-D. *ITU-D Handbook Teletraffic Engineering*. ITU-D, 2003.

[IET06] IETF Internet Draft. A Path Computation Element (PCE) Based Architecture. draft-ietf-pce-architecture-05.txt, work in progress, April 2006.

[IGW+02] R. Izmailov, S. Ganguly, T. Wang, Y. Suemura, Y. Maeno, and S. Araki. Hybrid hierarchical optical networks. *IEEE Communications Magazine*, 40(11):88–94, 2002.

[IKR] IKR. The IKR simulation library. Institute of Communication Networks and Computer Engineering, University of Stuttgart. `www.ikr.uni-stuttgart.de/IKRSimLib`.

[ISNS02] M. Iizuka, M. Sakuta, Y. Nishino, and I. Sasase. A scheduling algorithm minimizing voids generated by arriving bursts in optical burst switched WDM network. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, volume 3, pages 2736–2740, Taipei, November 2002.

[ISO94] ISO. International Standard 7498–1: Information technology – Open Systems Interconnection – Basic Reference Model: The basic model, ISO/IEC, 1994.

[IST04a] IST-NOBEL, Workpackage 1. Preliminiary definition of drivers and requirements for core and metro networks supporting end-to-end broadband services for all. Project Deliverable D6, EU, 6th Framework Program, Information Society Technologies, September 2004.

[IST04b] IST-NOBEL, Workpackage 3. Requirements for burst/packet networks in core and metro supporting high quality broadband services over IP. Project Deliverable D4, EU, 6th Framework Program, Information Society Technologies, June 2004.

[IST05] IST-NOBEL, Workpackage 3. Preliminary definition of burst/packet network and node architectures and solutions, section 5.2.6.2. Project Deliverable D3, EU, 6th Framework Program, Information Society Technologies, June 2005.

[ITG96] ITG-Fachgruppe 5.2.2 Systemtechnik. ITG - Empfehlung 5.2-01: Architekturen und Verfahren der Vermittlungstechnik, December 1996.

[ITG06] ITG-Fachgruppe 5.3.3 Photonische Netze. Optical Transport Networks (OTN) – Technical Trends and Assessment. ITG-Positionspapier, VDE ITG Frankfurt, March 2006.

[ITU03] ITU. Optical Transport Network (OTN) Tutorial, 2003.

[JG03a] S. Junghans and C. M. Gauger. Architectures for resource reservation modules for optical burst switching core nodes. In *Proceedings of the 4. ITG Symposium on Photonic Networks*, Leipzig, May 2003.

[JG03b] S. Junghans and C. M. Gauger. Resource reservation in optical burst switching: Architectures and realizations for reservation modules. In *Proceedings of the Optical Networking and Communications Conference (OptiComm)*, Dallas, October 2003.

[JHLC05] S. Jiang, G. Hu, S. Liew, and H. Chao. Scheduling algorithms for shared fiber-delay-line optical packet switches-part II: the three-stage clos-network case. *IEEE/OSA Journal of Ligthwave Technology*, 23(4):1601–1609, 2005.

[Jun04] S. Junghans. A testbed for control systems of optical burst switching core nodes. In *Proceedings of the Third International Workshop on Optical Burst Switching (WOBS)*, San Jose/CA, October 2004.

[Jun05] S. Junghans. Pre-estimate burst scheduling (PEBS): An efficient architecture with low realization complexity for burst scheduling disciplines. In *Proceedings of the Fifth International Workshop on Optical Burst/Packet Switching (WOBS)*, 2005.

[Jun06] S. Junghans. *Realisierbarkeit von Scheduling-Modulen in Optical Burst Switching-Kernknoten*. Monography, University of Stuttgart, 2006.

[KA05] A. Kaheel and H. Alnuweiri. Batch scheduling algorithms: a class of wavelength schedulers in optical burst switching networks. In *Proceedings of the IEEE International Conference on Communications (ICC)*, volume 3, pages 1713–1719, 2005.

[Kar93] M. Karol. A shared-memory optical packet (ATM) switch. In *Proceedings of the 6th IEEE Workshop on Local and Metropolitan Area Networks*, pages 0.205–0.211, 1993.

[Kel96] F. Kelly. *Stochastic Networks: Theory and Applications*, chapter Notes on effective bandwidths, pages 141–168. Number 4. Oxford University Press, 1996.

[KG03] M. Köhn and C. M. Gauger. Dimensioning of SDH/WDM multilayer networks. In *Proceedings of the 4. ITG Symposium on Photonic Networks*, pages 29–33, Leipzig, May 2003.

[KHFK05] S. Kota, E. Hossain, R. Fantacci, and A. Karmouch. Cross-layer protocol engineering for wireless mobile networks. part 1. *IEEE Communications Magazine*, 43(12):110–111, 2005.

[KHFK06] S. Kota, E. Hossain, R. Fantacci, and A. Karmouch. Cross-layer protocol engineering for wireless mobile networks: part 2. *IEEE Communications Magazine*, 44(1):83–84, 2006.

[KIA+05] A. Kirstädter, A. Iselt, A. Autenrieth, D. A. Schupke, R. Prinz, and B. Edmaier. Business models for next generation transport networks. *Photonic Network Communications*, 10(3):283–296, November 2005.

[KKK02] S. Kim, N. Kim, and M. Kang. Contention resolution for optical burst switching networks using alternative routing. In *Proceedings of the IEEE International Conference on Communications (ICC)*, volume 5, pages 2678–2681laevens, 2002.

[KL04] P. Knight and C. Lewis. Layer 2 and 3 virtual private networks: taxonomy, technology, and standardization efforts. *IEEE Communications Magazine*, 42(6):124–131, 2004.

[KM05] S. Kim and B. Mukherjee. Integrated congestion-control mechanism in optical burst switching networks. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, St. Louis/MO, November 2005.

[KOA05] B. Kantarci, S. Oktug, and T. Atmaca. Analyzing the effects of burst assembly in optical burst switching under self-similar traffic. In *Proceedings of the Advanced Industrial Conference on Telecommunications (ACIT)*, pages 109–114, 2005.

[Koe05] J. Koegel. Design and implementation of a burst assembly unit based on a network processor. Master thesis, University of Stuttgart, 2005.

[KP02] G. Kramer and G. Pesavento. Ethernet passive optical network (EPON): Building a next-generation optical access network. *IEEE Communications Magazine*, 40(2):66–73, February 2002.

[KSM+02] T. Koonen, S. Sulur, I. T. Monroy, J. Jennen, and H. de Waardt. Orthogonal optical labeling of packets in IP over WDM networks. In *Proceedings of the 7th European Conference on Networks and Optical Communications (NOC)*, pages 82–89, Darmstadt, May 2002.

[KW95] A. Kompella and I. Widjaja. Burst-level admission control protocols with multirate traffic andarbitrary network topology. In *Proceedings of the 4th International Conference Computer Communications and Networks*, pages 396–403, 1995.

[Köh05a] M. Köhn. Software tool for the dimensioning of transport networks, 2004–2005.

[Köh05b] M. Köhn. Comparison of SDH/SONET-WDM multi layer networks with static and dynamic optical plane. In *Proceedings of the Conference on Optical Network Design and Modelling (ONDM)*, pages 403–412, February 2005.

[Küh04a] P. J. Kühn. Exact and approximate analysis of WDM switching networks with limited number of wavelength converters. Presentation at ITG FG5.2.1. TU Hamburg-Harburg, February 2004.

[Küh04b] P. J. Kühn. Lecture notes: Teletraffic theory and engineering, University of Stuttgart, 2004.

[Lae02] K. Laevens. Traffic characteristics inside optical burst switched networks. In *Proceedings of the Optical Networking and Communications Conference (OptiComm)*, 2002.

[Lan03] D. Lanza. Jumbo optics. Light Reading - Networking the Telecom Industry, May 2003.

[LB03] K. Laevens and H. Bruneel. Analysis of a single-wavelength optical buffer. In *Proceedings of IEEE INFOCOM*, volume 3, pages 2262–2267, 2003.

[LFB+05] J. Leuthold, W. Freude, G. Boettger, P. Vorreau, A. Marculsescu, and J. Wang. Transparent networks and the role of SOA based all-optical wavelength converters and signal regenerators. In *Proceedings of SPIE Asia-Pacific Optical Communications (APOC'2005)*, Shanghai, November 2005.

[LHC05] S. Liew, G. Hu, and H. Chao. Scheduling algorithms for shared fiber-delay-line optical packet switches-part 1: the single-stage case. *IEEE/OSA Journal of Ligthwave Technology*, 23(4):1586–1600, 2005.

[LKS+05] S. Lee, L. Kim, J. Song, D. Griffith, and K. Sriram. Dynamic deflection routing with virtual wavelength assignment in optical burst-switched networks. *Photonic Network Communications*, 9(3):347–356, May 2005.

[LL95] K.-C. Lee and V. Li. Optimization of a WDM optical packet switch with wavelength converters. In *Proceedings of IEEE INFOCOM*, pages 423–430, 1995.

[LL02] D. Q. Liu and M. T. Liu. Differentiated services and scheduling scheme in optical burst-switched WDM networks. In *Proceedings of the International Conference on Networks (ICON)*, Singapore, June 2002.

[LLY02] C.-H. Loi, W. Liao, and D. Yang. Service differentiation in optical burst switched networks. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, pages 2313–2317, Taipei, November 2002.

[LM04] X. Lu and B. Mark. Performance modeling of optical-burst switching with fiber delay lines. *IEEE Transactions on Communications*, 52(12):2175–2183, 2004.

[LP04] H. Lim and C.-S. Park. An optical packet switch with hybrid buffer structure for contention resolution of asynchronous variable length packets. In *Proceedings of the Workshop on High Performance Switching and Routing*, pages 162–166, 2004.

[LQC04] J. Li, C. Qiao, and Y. Chen. Recent progress in the scheduling algorithms in optical-burst-switched networks. *Journal of Optical Networking*, 3:229–241, April 2004.

[LQXX04] J. Li, C. Qiao, J. Xu, and D. Xu. Maximizing throughput for optical burst switching networks. In *Proceedings of IEEE INFOCOM 2004*, volume 3, pages 1853–1863, 2004.

[LSDD$^+$01] N. Le Sauze, A. Dupas, E. Dotaro, L. Ciavaglia, M. Nizam, A. Ge, and L. Dembeck. A novel, low cost optical packet metropolitan ring architecture. In *Proceedings of the 27th European Conference on Optical Communication (ECOC 2001)*, volume 6, pages 66–67, Amsterdam, 2001.

[LSKS05] S. Lee, K. Sriram, H. Kim, and J. Song. Contention-based limited deflection routing protocol in optical burst-switched networks. *IEEE Journal on Selected Areas in Communications*, 23(8):1596–1611, 2005.

[LW05] Y. Luo and S. Wang. An FDL-based QoS scheduling algorithm in OBS networks. In *Proceedings of the International Conference on Communications, Circuits and Systems*, volume 1, pages 639–642, 2005.

[LWZ$^+$03] G. M. Lee, B. Wydrowski, M. Zukerman, J. K. Choi, and C. H. Foh. Performance evaluation of an optical hybrid switching system. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, pages 2508–2512, San Francisco/CA, December 2003.

[MCL$^+$03] S. D. Maesschalck, D. Colle, I. Lievens, M. Pickavet, P. Demeester, C. Mauz, M. Jaeger, R. Inkret, B. Mikac, and J. Derkacz. Pan-european optical transport networks: An availability-based comparison. *Photonic Network Communications*, 5(3):203–225, May 2003.

[Med05] A. Meddeb. Why ethernet WAN transport? *IEEE Communications Magazine*, 43(11):136–141, 2005.

[MEF] MEF, The Metro Ethernet Forum. `www.metroethernetforum.org`.

[MEF6] MEF. MEF Technical Specification 6: Ethernet Services Definition – Phase 1, The Metro Ethernet Forum, June 2004.

[MMW$^+$05] K. Merchant, J. McGeehan, A. Willner, S. Ovadia, P. Kamath, J. Touch, and J. Bannister. Analysis of an optical burst switching router with tunable multi-wavelength recirculating buffers. *IEEE/OSA Journal of Ligthwave Technology*, 23(10):3302–3312, 2005.

[MNSW02a] D. Mitra, C. Nuzman, I. Saniee, and P. Whiting. Optical crossconnect with shared wavelength conversion under dynamic loading. In *Proceedings of the Optical Fiber Communication Conference (OFC)*, pages 160–162, 2002.

[MNSW02b] D. Mitra, C. Nuzman, I. Saniee, and P. Whiting. Fixed point approximations for optical cross-connects using wavelength converters. Unpublished manuscript, 2002.

[MNVW04] S. D. Maesschalck, L. Nederlof, M. Vaughn, and R. Wagner. Traffic studies for fast optical switching in an intelligent optical network. *Photonic Network Communications*, 8(3):285–307, November 2004.

[Mod01] E. Modiano. Traffic grooming in WDM networks. *IEEE Communications Magazine*, 39(7):124–129, 2001.

[MRZ04] G. Muretto, C. Raffaelli, and P. Zaffoni. Effective implementation of void filling in OBS networks with service differentiation. In *Proceedings of the Third International Workshop on Optical Burst Switching (WOBS)*, October 2004.

[MTF05] M. Morrow, M. Tatipamula, and A. Farrel. GMPLS: the promise of the next-generation optical control plane. *IEEE Communications Magazine*, 43(7):26–27, 2005.

[Mue03] C. Mueller. Untersuchung der Effektivitaet der Wellenlängenkonversion in Optical Burst Switching Netzen. Student thesis, University of Stuttgart, Stuttgart, June 2003.

[Muk06] B. Mukherjee. *Optical WDM Networks (Optical Networks).* Springer, 2006.

[MZA05] Y. Mingwu, L. Zengji, and W. Aijun. Accurate and approximate evaluations of asynchronous tunable-wavelength-converter sharing schemes in optical burst-switched networks. *IEEE/OSA Journal of Ligthwave Technology*, 23(10):2807–2815, 2005.

[NDT02] M. Nava and C. Del-Toso. A short overview of the VDSL system requirements. *IEEE Communications Magazine*, 40(12):82–90, 2002.

[Nec02] M. Necker. Untersuchung von integrierten Verkehrslenkungsverfahren für SDH/WDM Multilayer-Netze. Diploma thesis, University of Stuttgart, 2002.

[NGB03] M. Necker, C. Gauger, and S. Bodamer. A new efficient integrated routing scheme for SDH/SONET-WDM multilayer networks. In *Proceedings of the Optical Fiber Communication Conference (OFC)*, volume 2, pages 487–488, 2003.

[NKSO04a] R. Nejabati, D. Klonidis, D. Simeonidou, and M. O'Mahony. Demonstration of ingress edge IP-optical packet router in wavelength routed optical packet switched networks. In *Proceedings of the 30th European Conference on Optical Communication (ECOC 2004)*, Stockholm, September 2004.

[NKSO04b] R. Nejabati, D. Klonidis, D. Simeonidou, and M. O'Mahony. Hybrid edge IP/optical packet generator in wavelength routed networks. In *Proceedings of the Optical Fiber and Communication Conference (OFC)*, Anaheim/CA, February 2004.

[NM73] Y. Nakagome and H. Mori. Flexible routing in the global communication network. In *Proceedings of the 7th International Teletraffic Congress (ITC 7)*, Stockholm, 1973.

[Nor95] I. Norros. On the use of fractional Brownian motion in the theory of connectionless networks. *IEEE Journal on Selected Areas in Communications*, 13(6):953–962, 1995.

[Nor03] M. Nord. Node design in optical packet and optical burst switching. In *Proceedings of the 5th International Conference on Transparent Optical Networks (ICTON)*, volume 1, pages 136–143, June 2003.

[Nor04] M. Nord. Waveband based multi-plane optical packet switch with partially shared wavelength converters. In *Proceedings of the Optical Network Design and Modelling Conference (ONDM)*, Gent, Belgium, February 2004.

[NRV+03] M. Neuts, Z. Rosberg, H. L. Vu, J. White, and M. Zukerman. Performance enhancement of optical burst switching using burst segmentation. In *Proceedings of the IEEE International Conference on Communications (ICC)*, volume 3, pages 1828–1832, 2003.

[NTM00] A. Narula-Tam and E. Modiano. Dynamic load balancing in WDM packet networks with and without wavelength constraints. *IEEE Journal on Selected Areas in Communications*, 18(10):1972–1979, 2000.

[NZA99] T. D. Neame, M. Zukerman, and R. G. Addie. A practical approach for multimedia traffic modeling. In *Proceedings of the 5th IFIP TC6 International Conference on Broadband Communications (BC 1999)*, Hong Kong, November 1999.

[OHK02] S.-Y. Oh, H. H. Hong, and M. Kang. A data burst assembly algorithm in optical burst switching networks. *ETRI Journal*, 24(4):311–322, August 2002.

[OIF04] OIF, The Optical Internet Forum. *User Network Interface (UNI) 1.0 Signaling Specification, Release 2*, February 2004.

[OM05] C. S. Ou and B. Mukherjee. *Survivable Optical WDM Networks*. Springer, 2005.

[OON88] H. Ohnishi, T. Okada, and K. Noguchi. Flow control schemes and delay/loss tradeoff in ATM networks. *IEEE Journal on Selected Areas in Communications*, 6(9):1609–1616, 1988.

[OSHT01] M. J. O'Mahony, D. Simeonidou, D. K. Hunter, and A. Tzanakaki. The application of optical packet switching in future communication networks. *IEEE Communications Magazine*, 39(3):128–135, March 2001.

[OT05] N. Ogino and H. Tanaka. Deflection Routing for Optical Bursts Considering Possibility of Contention at Downstream Nodes. *IEICE Transactions on Communication*, E88-B(9):3660–3667, 2005.

[Pat05] A. Pattavina. Multi-wavelength switching in IP optical nodes adopting different buffering strategies. *Optical Switching and Networking*, 1(1):65–75, January 2005.

[PCM+04] M. H. Phung, K. C. Chua, G. Mohan, M. Motani, and T. C. Wong. Absolute QoS signalling and reservation in optical burst-switched networks. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, Dallas/TX, November 2004.

[PKI+05] S. Pasqualini, A. Kirstadter, A. Iselt, R. Chahine, S. Verbrugge, D. Colle, M. Pickavet, and P. Demeester. Influence of GMPLS on network providers' operational expenditures: a quantitative study. *IEEE Communications Magazine*, 43(7):28–38, 2005.

[PV05] D. Papadimitriou and D. Verchere. GMPLS user-network interface in support of end-to-end rerouting. *IEEE Communications Magazine*, 43(7):35–43, 2005.

[QY99] C. Qiao and M. Yoo. Optical burst switching (OBS)—a new paradigm for an optical Internet. *Journal of High Speed Networks*, 8(1):69–84, January 1999.

[RC05] A. Rostami and S. Chakraborty. On performance of optical buffers with specific number of circulations. *IEEE Photonics Technology Letters*, 17(7):1570–1572, 2005.

[RFC2460] S. Deering and R. Hinden. Internet Protocol, Version 6 (IPv6) Specification. RFC 2460 (Draft Standard), December 1998.

[RFC2988] V. Paxson and M. Allman. Computing TCP's Retransmission Timer. RFC 2988 (Proposed Standard), November 2000.

[RFC3031] E. Rosen, A. Viswanathan, and R. Callon. Multiprotocol Label Switching Architecture. RFC 3031 (Proposed Standard), January 2001.

[RFC3036] L. Andersson, P. Doolan, N. Feldman, A. Fredette, and B. Thomas. LDP Specification. RFC 3036 (Proposed Standard), January 2001.

[RFC3209] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow. RSVP-TE: Extensions to RSVP for LSP Tunnels. RFC 3209 (Proposed Standard), December 2001. Updated by RFCs 3936, 4420.

[RFC3212] B. Jamoussi, L. Andersson, R. Callon, R. Dantu, L. Wu, P. Doolan, T. Worster, N. Feldman, A. Fredette, M. Girish, E. Gray, J. Heinanen, T. Kilty, and A. Malis. Constraint-Based LSP Setup using LDP. RFC 3212 (Proposed Standard), January 2002. Updated by RFC 3468.

[RFC3471] L. Berger. Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description. RFC 3471 (Proposed Standard), January 2003. Updated by RFCs 4201, 4328.

[RFC3472] P. Ashwood-Smith and L. Berger. Generalized Multi-Protocol Label Switching (GMPLS) Signaling Constraint-based Routed Label Distribution Protocol (CR-LDP) Extensions. RFC 3472 (Proposed Standard), January 2003. Updated by RFCs 3468, 4201.

[RFC3477] K. Kompella and Y. Rekhter. Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE). RFC 3477 (Proposed Standard), January 2003.

[RFC3630] D. Katz, K. Kompella, and D. Yeung. Traffic Engineering (TE) Extensions to OSPF Version 2. RFC 3630 (Proposed Standard), September 2003. Updated by RFC 4203.

[RFC3717] B. Rajagopalan, J. Luciani, and D. Awduche. IP over Optical Networks: A Framework. RFC 3717 (Informational), March 2004.

[RFC3945] E. Mannie. Generalized Multi-Protocol Label Switching (GMPLS) Architecture. RFC 3945 (Proposed Standard), October 2004.

[RFC4054] J. Strand and A. Chiu. Impairments and Other Constraints on Optical Layer Routing. RFC 4054 (Informational), May 2005.

[RFC4203] K. Kompella and Y. Rekhter. OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS). RFC 4203 (Proposed Standard), October 2005.

[RFC4204] J. Lang. Link Management Protocol (LMP). RFC 4204 (Proposed Standard), October 2005.

[RFC4208] G. Swallow, J. Drake, H. Ishimatsu, and Y. Rekhter. Generalized Multiprotocol Label Switching (GMPLS) User-Network Interface (UNI): Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Support for the Overlay Model. RFC 4208 (Proposed Standard), October 2005.

[RFC4209] A. Fredette and J. Lang. Link Management Protocol (LMP) for Dense Wavelength Division Multiplexing (DWDM) Optical Line Systems. RFC 4209 (Proposed Standard), October 2005.

[RFC791] J. Postel. Internet Protocol. RFC 791 (Standard), September 1981. Updated by RFC 1349.

[RLFB05a] W. Rogiest, K. Laevens, D. Fiems, and H. Bruneel. A performance model for an asynchronous optical buffer. In *Proceedings of the 19th International Teletraffic Congress (ITC 19)*, Beijing, September 2005.

[RLFB05b] W. Rogiest, K. Laevens, D. Fiems, and H. Bruneel. A performance model for an asynchronous optical buffer. *Performance Evaluation*, 62(1-4):313–330, October 2005.

[RM98] B. Ramamurthy and B. Mukherjee. Wavelength conversion in WDM networking. *IEEE Journal on Selected Areas in Communications*, 16(7):1061–1073, 1998.

[Ros95] K. W. Ross. *Multiservice loss models for broadband telecommunication networks*. Springer, 1995.

[RR00] B. Ramamurthy and A. Ramakrishnan. Virtual topology reconfiguration of wavelength-routed optical WDM networks. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, volume 2, pages 1269–1275, 2000.

[RS96] R. Ramaswami and K. Sivarajan. Design of logical topologies for wavelength-routed optical networks. *IEEE Journal on Selected Areas in Communications*, 14(5):840–851, 1996.

[RS98] R. Ramaswami and K. N. Sivarajan. *Optical Networks: a Practical Perspective*. Morgan Kaufmann Publishers, 1998.

[RZVZ06] Z. Rosberg, A. Zalesky, H. L. Vu, and M. Zukerman. Analysis of OBS networks with limited wavelength conversion. *IEEE/ACM Transactions on Networking*, 2006.

[San01] J. Saniter. Selected contributions, photonic networks. Heinrich-Hertz-Institut Report, p. 49, 2001.

[San02] M. Sanchez. Untersuchung moeglicher Realisierungen von Reservierungsverfahren fuer Optical Burst Switching-Netzknoten. Diploma thesis, University of Stuttgart, 2002.

[SAS96] S. Subramaniam, M. Azizoglu, and A. Somani. All-optical networks with sparse wavelength conversion. *IEEE/ACM Transactions on Networking*, 4(4):544–557, 1996.

[Sas04] D. Sass. Traffic trace of a the uplink of a student dormitory network, (UST-IKR 1 trace, 3h busy time starting at 320,000 sec.). NOBEL internal, further information at http://www.ikr.uni-stuttgart.de/Content/I2MP, 2004.

[SB98] J. Späth and S. Bodamer. Routing of dynamic poisson and non-poisson traffic in wdm networks with limited wavelength conversion. In *Proceedings of the 24th European Conference on Optical Communication (ECOC '98): Regular and Invited Papers*, volume 1, pages 359–360. Telefonica de Espana, Lerko Prin S.A., 1998.

[SBG04] J. Späth, S. Bodamer, and C. Glingener. Next-generation transport solutions for IP backbone networks - benefits of an ASTN-based multi-layer OTN network (invited paper). In *Proceedings of Asia-Pacific Optical Communications (APOC)*, Beijing, China, November 2004. SPIE.

[Sch01] M. Scharf. Entwurf und Bewertung von Puffermechanismen für Optical Burst Switching. Student thesis, University of Stuttgart, July 2001.

[Sch04] J. Scharf. Untersuchung von Optical Burst Switching in Netz-Szenarien. Student thesis, University of Stuttgart, February 2004.

[SEL02] L. B. Sofman, T. S. El-Bawab, and K. Laevens. Segmentation overhead in optical burst switching. In *Proceedings of the Optical Networking and Communications Conference (OptiComm)*, pages 101–108, Boston, July 2002.

[Shi05] H. Shinohara. Broadband access in Japan: rapidly growing FTTH market. *IEEE Communications Magazine*, 43(9):72–78, 2005.

[SPG05] M. Schlosser, E. Patzak, and P. Gelpke. Impact of deflection routing on TCP performance in optical burst switching networks. In *Proceedings of the 7th International Conference on Transparent Optical Networks (ICTON)*, volume 1, pages 220–223, Barcelona, June 2005.

[Spä00] J. Späth. Dynamic routing and resource allocation in WDM transport networks. *Computer Networks*, 32(5):519–538, May 2000.

[Spä02] J. Späth. *Entwurf und Bewertung von Verfahren zur Verkehrslenkung in WDM-Netzen*. PhD thesis, University of Stuttgart, Stuttgart, 2002.

[SSJ97] B. Skjoldstrup, J. Späth, and A. Jukan. Transparency in high capacity networks based on an optical layer. In *Proceedings of the COST 240 Workshop on SOA-based Components for Optical Networks*, pages 22–1–22–5, Prague, 1997.

[TBPOB05] T. Takeda, D. Brungard, D. Papadimitriou, and H. Ould-Brahim. Layer 1 virtual private networks: driving forces and realization by gmpls. *IEEE Communications Magazine*, 43(7):60–67, 2005.

[TDY05] R. Telikepally, T. Drwiega, and J. Yan. Modeling and analysis of storage area network extension solutions. In *Proceedings of the International Teletraffic Congress (ITC 19)*, pages 1019–1028, September 2005.

[TKCH05] R. Tucker, P.-C. Ku, and C. Chang-Hasnain. Fundamental limitations of slow-light optical buffers. In *Proceedings of the Optical Fiber Communication Conference (OFC/NFOEC)*, volume 3, 2005.

[TMC03] S. Tan, G. Mohan, and K. Chua. Burst rescheduling with wavelength and last-hop fdl reassignment in WDM optical burst switching networks. In *Proceedings of the IEEE International Conference on Communications (ICC)*, volume 2, pages 1448–1452, 2003.

[TR05] J. Teng and G. N. Rouskas. A detailed analysis and performance comparison of wavelength reservation schemes for optical burst switched networks. *Photonic Network Communications*, 9(3):311–335, May 2005.

[Tur92] J. Turner. Managing bandwidth in ATM networks with bursty traffic. *IEEE Network*, 6(5):50–58, 1992.

[Tur99] J. S. Turner. Terabit burst switching. *Journal of High Speed Networks*, 8(1):3–16, January 1999.

[TVJ03] G. Thodime, V. Vokkarane, and J. Jue. Dynamic congestion-based load balanced routing in optical burst-switched networks. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, volume 5, pages 2628–2632, 2003.

[TYC$^+$00] L. Tancevski, S. Yegnanarayanan, G. Castanon, L. Tamil, F. Masetti, and T. McDermott. Optical routing of asynchronous, variable length packets. *IEEE Journal of Selected Areas in Communications*, 18(10):2084–2093, October 2000.

[UOF+01] H. Ueda, K. Okada, B. Ford, G. Mahony, S. Hornung, D. Faulkner, J. Abiven, S. Durel, R. Ballart, and J. Erickson. Deployment status and common technical specifications for a B-PON system. *IEEE Communications Magazine*, 39(12):134–141, 2001.

[vBCC+03] E. van Breusegem, J. Cheyns, D. Colle, M. Pickavet, and P. Demeester. Overspill routing in optical networks: a new architecture for future-proof IP over WDM networks. In *Proceedings of the Optical Networking and Communications Conference (OptiComm)*, Dallas/TX, October 2003. SPIE.

[VCR00] S. Verma, H. Chaskar, and R. Ravikanth. Optical burst switching: a viable solution for terabit IP backbone. *IEEE Network*, 14(6):48–53, 2000.

[VHLL+04] B. Van Houdt, K. Laevens, J. Lambert, C. Blondia, and H. Bruneel. Channel utilization and loss rate in a single-wavelength fibre delay line (FDL) buffer. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, volume 3, pages 1900–1906, 2004.

[VJ02a] V. Vokkarane and J. Jue. Burst segmentation: an approach for reducing packet loss in optical burst switched networks. In *Proceedings of the IEEE International Conference on Communication*, New York City, April/May 2002.

[VJ02b] V. Vokkarane and J. Jue. Prioritized routing and burst segmentation for QoS in optical burst-switched networks. In *Proceedings of the Optical Fiber Communication Conference (OFC)*, Anaheim, March 2002.

[VPD04] J.-P. Vasseur, M. Pickavet, and P. Demeester. *Network Recovery: Protection and Restoration for Optical, SONET-SDH, IP, and MPLS*. Morgan Kaufmann Publishers, 2004.

[VTCJ03] V. Vokkarane, G. Thodime, V. Challagulla, and J. Jue. Channel scheduling algorithms using burst segmentation and fdls for optical burst-switched networks. In *Proceedings of the IEEE International Conference on Communications (ICC)*, volume 2, pages 1443–1447, 2003.

[VZC+03] K. Vlachos, J. Zhang, J. Cheyns, Sulur, N. Chi, E. Van Breusegem, I. Monroy, J. Jennen, P. Holm-Nielsen, C. Peucheret, R. O'Dowd, P. Demeester, and A. Koonen. An optical IM/FSK coding technique for the implementation of a label-controlled arrayed waveguide packet router. *IEEE/OSA Journal of Ligthwave Technology*, 21(11):2617–2628, 2003.

[VZJC02] V. M. Vokkarane, Q. Zhang, J. P. Jue, and B. Chen. Generalized burst assembly and scheduling techniques for QoS support in optical burst-switched networks. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, pages 2747–2751, Taipei, November 2002.

[VZL+03] M. Veeraraghavan, X. Zheng, H. Lee, M. K. Gardner, and W.-C. Feng. Cheetah: Circuit-switched high-speed end-to-end transport architecture. In *Proceedings of the Optical Networking and Computer Communications Conference (OptiComm)*, Dallas/TX, October 2003.

[Wan96] S. Wannenmacher. *Untersuchung und Aufbau praseodymdotierter optischer Faserverstärker*. PhD thesis, University of Stuttgart, Stuttgart, October 1996.

[Wan02] T.-S. Wang. Architectural evolution and principles of optical terabit packet switches (OTPS). *Computer Communications*, 25(6):557–576, 2002.

[WCA+00] A. E. Willner, M. C. Cardakli, O. H. Adamczyk, Y.-W. Song, and D. Gurkan. Key building blocks for all-optical networks. *IEICE - Transactions on Communications*, E83B(10):2166–2177, October 2000.

[Wid94] I. Widjaja. Random access for ATM LANs and WANs. In *Proceedings of the IEEE International Conference on Communications (ICC)*, pages 39–43, 1994.

[Wid95] I. Widjaja. Performance analysis of burst admission-control protocols. *Communications, IEE Proceedings-*, 142(1):7–14, 1995.

[WMA00] X. Wang, H. Morikawa, and T. Aoyama. Deflection routing protocol for burst switching WDM mesh networks. In *Proceedings of SPIE Terabit Optical Networking: Architectures, Control and Management Issues*, Boston, November 2000.

[WMA02] X. Wang, H. Morikawa, and T. Aoyama. Burst optical deflection routing protocol for wavelength routing WDM networks. *Optical Networks Magazine*, pages 12–19, November/December 2002.

[WPRT99] J. Y. Wei, J. L. Pastor, R. S. Ramamurthy, and Y. Tsai. Just-in-time optical burst switching for multi-wavelength networks. In *Proceedings of the 5th International Conference on Broadband Communication (BC'99)*, pages 339–352, 1999.

[WPW+02] I. White, R. Penty, M. Webster, Y. J. Chai, A. Wonfor, and S. Shahkooh. Wavelength switching components for future photonic networks. *IEEE Communications Magazine*, 40(9):74–81, 2002.

[WTSW97] W. Willinger, M. S. Taqqu, R. Sherman, and D. V. Wilson. Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level. *IEEE/ACM Transactions on Networking*, 5(1):71–86, February 1997.

[WWB03] J. Walraevens, S. Wittevrongel, and H. Bruneel. Calculation of the packet loss in optical packet switches: An analytic technique. *International Journal of Electronics and Communications*, 57(4):270–276, 2003.

[X.200] ITU. Rec. X.200: Information technology – Open Systems Interconnection – Basic Reference Model: The basic model, International Telecommunication Union, ITU-T, 1994.

[XQLX03] J. Xu, C. Qiao, J. Li, and G. Xu. Efficient channel scheduling algorithms in optical burst switching networks. In *Proceedings of IEEE INFOCOM*, San Francisco, April 2003.

[XQLX04] J. Xu, C. Qiao, J. Li, and G. Xu. Efficient burst scheduling algorithms in optical burst-switched networks using geometric techniques. *IEEE Journal on Selected Areas in Communications*, 22(9):1796–1811, 2004.

[XQYD03] C. Xin, C. Qiao, Y. Ye, and S. Dixit. A hybrid optical switching approach. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, pages 3808–3812, San Francisco/CA, December 2003.

[XVC00] Y. Xiong, M. Vanderhoute, and H. C. Cankaya. Control architecture in optical burst-switched WDM networks. *IEEE Journal of Selected Areas in Communications*, 18(10):1838–1851, October 2000.

[Y.100] ITU. Rec. Y.100: General overview of the Global Information Infrastructure standards development, International Telecommunication Union, ITU-T, 1998.

[Y.1541] ITU. Rec. Y.1541: Network Performance Objectives for IP-Based Services, International Telecommunication Union, ITU-T, 2003.

[Y.2001] ITU. Rec. Y.2001: General Overview of NGN, International Telecommunication Union, ITU-T, December 2004.

[Y.2011] ITU. Rec. Y.2011: General principles and general reference model for Next Generation Networks, International Telecommunication Union, ITU-T, October 2004.

[YCQ02a] X. Yu, Y. Chen, and C. Qiao. Performance evaluation of optical burst switching with assembled burst traffic input. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, pages 2318–2322, Taipei, November 2002.

[YCQ02b] X. Yu, Y. Chen, and C. Qiao. A study of traffic statistics of assembled burst traffic in optical burst switched networks. In *Proceedings of the Proceeding of the Optical Networking and Communications Conference (OptiComm)*, pages 149 – 159, Boston, July 2002.

[YLC+04] X. Yu, J. Li, X. Cao, Y. Chen, and C. Qiao. Traffic statistics and performance evaluation in optical burst switched networks. *IEEE/OSA Journal of Ligthwave Technology*, 22(12):2722–2738, 2004.

[YLES96] J. Yates, J. Lacey, D. Everitt, and M. Summerfield. Limited-range wavelength translation in all-optical networks. pages 954–961, San Francisco, March 1996.

[YMC+03] Q. Yan, S. D. Maesschalck, D. Colle, B. Puype, I. Lievens, M. Pickavet, and P. Demeester. Influence of the observation window size on the performance of multi-layer traffic engineering. In *Proceedings of ITCOM 2003*, 2003.

[YMD00] S. Yao, B. Mukherjee, and S. Dixit. Advances in photonic packet switching: an overview. *IEEE Communications Magazine*, 38(2):84–.94, February 2000.

[YMD03] S. Yao, B. Mukherjee, and S. Dixit. Plato: A generic modeling technique for otpical packet-switched networks. *International Journal on Wireless & Optical Communciations*, 1(1):91–101, 2003.

[YMK05] N. Yan, I. T. Monroy, and T. Koonen. All-optical label swapping node architectures and contention resolution. In *Proceedings of the Optical Network Design and Modelling Conference (ONDM)*, pages 115–123, Milano, February 2005.

[YMYD00] S. Yao, B. Mukherjee, S. J. B. Yoo, and S. Dixit. All-optical packet-switched networks: a study of contention-resolution schemes in an irregular mesh network with variable-sized packets. In *Proceedings of the Optical Networking and Communications Conference (OptiComm)*, Dallas, October 2000.

[YMYD03] S. Yao, B. Mukherjee, S. Yoo, and S. Dixit. A unified study of contention-resolution schemes in optical packet-switched networks. *IEEE/OSA Journal of Ligthwave Technology*, 21(3):672–683, 2003.

[Yoo96] S. Yoo. Wavelength conversion technologies for WDM network applications. *IEEE/OSA Journal of Ligthwave Technology*, 14(6):955–966, 1996.

[YQ97] M. Yoo and C. Qiao. Just-enough-time (JET): A high speed protocol for bursty traffic in optical networks. In *Proceedings of the IEEE/LEOS Summer Topical Meetings*, pages 26–27, Montreal, Que. , Canada, August 1997.

[YQ00] M. Yoo and C. Qiao. QoS performance in IP over WDM networks. *IEEE Journal of Selected Areas in Communications*, 18(10):2062–2071, October 2000.

[YR05] L. Yang and G. N. Rouskas. Path switching in optical burst switched networks. In *Proceedings of Networking 2005*, pages 406–418, Waterloo/CND, May 2005.

[YXM$^+$02] S. Yao, F. Xue, B. Mukherjee, S. J. B. Yoo, and S. Dixit. Electrical ingress buffering and traffic aggregation for optical packet switching and their effect on TCP-level performance in optical mesh networks. *IEEE Communications Magazine*, 40(9):55–72, September 2002.

[YY05] H. Yang and S. Yoo. All-optical variable buffering strategies and switch fabric architectures for future all-optical data routers. *IEEE/OSA Journal of Ligthwave Technology*, 23(10):3321–3330, 2005.

[YZV01] M. Yang, S. Q. Zheng, and D. Verchere. A QoS supporting scheduling algorithm for optical burst switching DWDM networks. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, San Antonio, November 2001.

[Z$^+$03] A. Zapata et al. Investigation of future optical metro ring networks based on 100-gigabit ethernet. In *Proceedings of the 4. ITG Symposium on Photonic Networks*, Leipzig, May 2003.

[ZLJ05] T. Zhang, K. Lu, and J. Jue. An analytical model for shared fiber-delay line buffers in asynchronous optical packet and burst switches. In *Proceedings of the IEEE International Conference Communications*, volume 3, pages 1636–1640, 2005.

[ZLJ06] T. Zhang, K. Lu, and J. P. Jue. Shared buffering in optical packet-switched networks. *IEEE Journal on Selected Areas in Communications*, 2006.

[ZLVR$^+$04] A. Zalesky, H. Le Vu, Z. Rosberg, E. Wong, and M. Zukerman. Modelling and performance evaluation of optical burst switched networks with deflection routing and wavelength reservation. In *Proceedings of INFOCOM 2004*, volume 3, pages 1864–1871, 2004.

[ZVZ+04] A. Zalesky, H. Vu, M. Zukerman, Z. Rosberg, and E. Wong. Evaluation of limited wavelength conversion and deflection routing as methods to reduce blocking probability in optical burst switched networks. In *Proceedings of the IEEE International Conference on Communications (ICC)*, volume 3, pages 1543–1547, 2004.

[ZXC02] S. Q. Zheng, Y. Xiong, and H. C. Cankaya. Hardware design of a channel scheduling algorithm for optical burst switching routers. In *Proceedings of SPIE ITCOM*, 2002.

[ZZM05] K. Zhu, H. Zhu, and B. Mukherjee. *Traffic Grooming in Optical WDM Mesh Networks*. Springer, 2005.

[ZZZM03] H. Zhu, K. Zhu, H. Zang, and B. Mukherjee. Cost-effective WDM backbone network design with OXCs of different bandwidth granularities. *IEEE Journal on Selected Areas in Communications*, 21(9):1452–1466, 2003.

Acknowledgments

The past almost six years at the University of Stuttgart including my research leave to the University of California at Davis were not only scientifically challenging and exciting but also extremely formative. In Stuttgart and Davis, in national and international projects as well as in the wider research community, I had the privilege to learn from and work with many distinguished colleagues. I very much enjoyed this joint work, the free exchange of ideas, and certainly all the social activities. I cherish all these unforgettable experiences and would like to express my deepest gratitude here.

Foremost, I would like to thank Prof. Paul Kühn for the opportunity to work at the Institute of Communication Networks and Computer Engineering (IKR), his confidence in all my initiatives, his scientific advice, and the supervision of this thesis. Likewise, I would like to say thank you to Prof. Biswanath Mukherjee for the great time at UC Davis, the inspiring discussions, and for co-advising this thesis—even taking part in the examination in Stuttgart.

I am very grateful to all active and former colleagues for the open, friendly, and constructive working environment at IKR. Particular thanks go to my friends Martin Köhn, Sascha Junghans, Sebastian Gunreben, Simon Hauger, Guoqiang Hu, Arthur Mutter, Detlef Sass and Joachim Scharf of the optics group, my friend and office mate Christian Hauser as well as my mentoring colleagues Stefan Bodamer, Klaus Dolzer, and Jan Späth. Also, I greatly appreciated Martin Köhn's, Marc Necker's, and Michael Scharf's help in reviewing a draft of this thesis. Last but not least, many thanks to Ulrich Gemkow for the freedom, feedback, and strong support for all the activities.

I would also like to acknowledge the great help of *my* Master students, the work of whom contributed a great deal to the tool suite and broadened my understanding of the field.

In writing my thesis, I greatly benefitted from a wide variety of perspectives which I gained from discussions or joint work with project colleagues. Therefore, I would like to thank all my colleagues in the national and international projects TransiNet, MultiTeraNet, COST 266, NOBEL, and e-Photon/ONe as well as in the research community. As it is impossible to mention all the particular reasons for saying thanks here, I restrict myself to alphabetical lists grouped by projects: Hao Buchta, Lars Dembeck, Michael Düser, Gert Eilenberger, Thomas Fischer, Andreas Gladisch, Ralph Hülsermann, Monika Jäger, Erwin Patzak, Jürgen Saniter, Michael Schlosser, Dominic Schupke, Joachim Westphal, Hagen Woesner; Steinar Bjornstad, Walter Cerroni, Tibor Cinkler, Robert Inkret, Anton Kuchar, Christian Mauz, Branko Mikac, Martin Nord, Ioannis Tomkos; Achim Autenrieth, Erik van Breusegem, Didier Colle, Jürgen Götz, Andreas Iselt, Jesus Lobo, Antonio Manzalini, Josep Pareta, Marco Schiano, Alex Stavdas; Nail Akar, Javier Aracil, Franco Callegati, Piet Demeester, Ezran Karasan, Fabio Neri, Mike O'Mahony, Carla Raffaelli; Neil Barakat, Jason Jue, Chunming Qiao, Zvi Rosberg, George Rouskas, Dan Stevenson, Vinod Vokkarane, and Moshe Zukerman.

For their generous offer to live in their house in Davis, I am very grateful to Amiya and Annmari Mukherjee. It was in their beautiful garden, where the basic idea for OBTN was born. Also, thanks to Massimo Tornatore for the great time and fun we had and for introducing me to the Italian community of Davis.

Finally, I want to say thank you to my loving wife Susanne and my dear parents and sisters. Susanne, your continuing, motivating support helped me through all the ups and downs of writing this thesis. Mum and Dad, I am very thankful for your advise and all the opportunities in my life you made possible.