Universiteit Gent
Faculteit Ingenieurswetenschappen

Vakgroep Telecommunicatie en
Informatieverwerking

Stochastische Modellering van Optische Buffers

Stochastic Modeling of Optical Buffers

Wouter Rogiest

Proefschrift ingediend
tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen

Academiejaar 2008-2009

Universiteit Gent
Faculteit Ingenieurswetenschappen

Vakgroep Telecommunicatie en
Informatieverwerking

Promotor: Prof. Dr. Ir. Herwig Bruneel

Begeleiders: Dr. Ir. Koenraad Laevens
Dr. Ing. Dieter Fiems
Prof. Dr. Evsey Morozov

Universiteit Gent (UGent)
Faculteit Ingenieurswetenschappen (FIrW)

Vakgroep Telecommunicatie en Informatieverwerking (TELIN) (IR07)
Sint-Pietersnieuwstraat 41, B-9000 Gent, België

Tel.: +32 9 264 34 12
Fax: +32 9 264 42 95
Web: http://telin.ugent.be

Dit werk is de vrucht van vier jaar onderzoek aan onderzoeksgroep SMACS. SMACS is een klinkend acroniem voor "Stochastic Modeling and Analysis of Communication Systems."

Dit document is vormgegeven met de opmaaktaal LaTeX. De template is van Bruno Volckaert, en is beschikbaar op http://latex.ugent.be.

De laatste update van dit document is gebeurd op 20 november 2008.

Proefschrift ingediend
tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen

Academiejaar 2008-2009

—Dedicated to the memory of Philippe—

Dankwoord
—Acknowledgment in Dutch—

¶ Door het schrijven van dit dankwoord sluit ik me aan bij de langlopende traditie in de wetenschappelijke wereld, om een jarenlange en vruchtbare samenwerking op mutueel exclusieve wijze neer te slaan in enerzijds een lijvige presentatie van de resultaten—de in formele stijl en in een vreemde taal gestelde turf genaamd doctoraatsproefschrift!—en anderzijds een heel summier verslag van wie me steunde onderweg—het in familiaire stijl gestelde stukje tekst genaamd dankwoord. Dat het laatste als een (in stilistische zin) verdacht aanhangsel van het eerste moet figureren, strookt niet dan gedeeltelijk met de intuïtie van deze in het Nederlands opgegroeide Europeaan. Zo graag zou ik u in informeel gesteld Nederlands aanschrijven: het ware een minste blijk van mijn overtuiging, dat er met wetenschap ook gevoelens gemoeid zijn. Daar ik evenwel doordrongen ben geraakt van de idee dat het gebruik van wetenschappelijk Engels "the free dissemination of ideas" ten goede komt, bied ik u mijn werk aan middels laatstgenoemde, en bedien me enkel in dit dankwoord van het Nederlands. (…en hoop met de bevlogen stijl van voorgaande zinnen mijn liefde voor het Nederlands te onderstrepen!)

Voor mensen die niet inhoudelijk betrokken waren bij de inhoud van mijn doctoraat, is het overigens zeer gepast om hen enkel hier te vermelden. Waarom zouden zij hun naam temidden van sommaties en integralen willen zien staan? Maar in het geval van Koenraad, Dieter, Herwig, Evsey, Koen en Benny ligt dat anders. Hoe zou dit doctoraat eruit zien, indien ik bij de door hen aangereikte ideeën, methodes en formules hun naam zou vermelden: "as suggested first by Mr. $\sim$"? Zou het dan geen plezier doen aan deze gelijkgestemden? En hoeveel vermeldingen zou dit opleveren voor elk?

Om toch een gedeeltelijk antwoord te geven op deze laatste vraag: in de top drie treffen we Herwig, Dieter en Koenraad, die me met hun grondige nalezing van vorige versies van dit proefschrift ontzettend vooruit hebben geholpen. Bovenaan in de hitlijst staat Koenraad, die ik wil bedanken voor vier jaar onvervangbare begeleiding. Koenraad, jij verstrekte geen toezicht maar inzicht, doorzicht; overzicht. Bij wijze van parafrase noteer ik, met in het achterhoofd een toost op onze samenwerking: "Queueing drink je niet alleen met verstand, maar ook met het hart." Dergelijke positieve ervaringen deelde ik ook met Dieter, die ik wil danken voor de vruchtbare en joviale samenwerking, die ongetwijfeld zeer gebaat was bij Dieters brede kennis van stochastische modellen en processen. Dankzij Herwig

werd ik niet alleen onthaald, “gekweekt” en bijgestaan door een op wetenschappelijk vlak vermaard onderzoeker, maar mocht ik bovendien rekenen op de steun van een persoonlijk geëngageerd promotor. Herwig, bedankt om deze vier memorabele jaren onderzoek mogelijk te maken. Switching language for now, I thank you, Evsey, for this extraordinary collaboration that proved to me that e-mail is an interesting medium after all—especially if your correspondent is a genuine and sympathetic expert in your field. Koen, jou wil ik danken voor de vele inzichtsvolle interventies op ons bureau. Benny, jou bedank ik met in het achterhoofd, wederom, nieuwe inzichten.

Enkele collega’s wil ik, om diverse redenen, nog eens bijzonder bedanken: Joris, Stijn, Bart, Tom, Danny, Dieter C en Sabine van SMACS; Patrick, Philippe, Davy en Annette van TELIN; Thomas, Luigi, Carla, Francesco en Frederick, mijn bureaugenoten.

Verder richt ik mijn dankbaarheid tot mijn vader en tot de nagedachtenis van mijn moeder, voor alles wat jullie gedaan hebben tot op vandaag, en daar voorbij. Ook een speciale dank aan Lut, Sofie, Mark, Marianne, Carl, Tine, Peter DB, Fabrizio, Ruben, Simon H, Benoit, Tom D, Frederik, mijn familie, vrienden van de quiz en van OD. Dit werk is opgedragen aan Philippe “VIP” De Boom. De periode gedurende dewelke dit werk geschreven werd is ook een periode van afscheid nemen en vooruit zien gebleken, en is onlosmakelijk verbonden met jouw nagedachtenis.

Onmeetbaar en onmisbaar is de steun die ik van Liesbeth heb mogen ontvangen. Ja het was werken geblazen om dit document af te krijgen, en zelfs indien het niet gelukt zou zijn was ik je nog een onbegrensde dankbetuiging verschuldigd—maar in dat geval zou het niet in een dankwoord hebben kunnen staan. Dat het ook hier kan, is een heel plezierige vaststelling.

Gent, augustus 2008
Wouter Rogiest

Table of Contents

List of Figures

List of Tables

List of Acronyms

C

cdf	cumulative distribution function, applicable to both discrete and continuous rv's
CT	Continuous Time (setting), as opposed to DT

D

DT	Discrete Time (setting), as opposed to CT

F

FCFS	First Come First Served
FIFO	First In First Out delay-line assignment algorithm
FDL	Fiber Delay Line

I

iid	independent and identically-distributed

J

JSQ	Join-the-Shortest-Queue wavelength assignment algorithm

L

LP	Loss Probability
lst	Laplace-Stieltjes transform, applicable for continuous rv's

O

OBS	Optical Burst Switching
OPS	Optical Packet Switching

P

pdf	probability density function, applicable for continuous rv's
pgf	probability generating function, applicable for discrete rv's
pmf	probability mass function, applicable for discrete rv's

R

RAM	Random-Access Memory
RR	Round-Robin wavelength assignment algorithm
RND	Random wavelength assignment algorithm
rv	random variable

T

TWC	Tuneable Wavelength Converter

W

w.p.1	With Probability 1 [→Chapter 5]

Main Variables, Parameters & Terms

Main Variables & Parameters

$\mathcal{A}$ is a set of parameters denoting a general non-degenerate FDL set.

a_i is a parameter denoting the length of FDL i, $i = 0, \ldots N$, with $a_i \in \mathcal{A}$.

A_l is a rv denoting the number of arrivals during an arbitrary slot l, in DT [→Chapter 2].

B_k is a rv denoting the burst size of burst k. This is synonymous to the transmission time, or, in queueing terms, the service time of a burst.

D is a parameter denoting the granularity of a degenerate FDL set. It only applies to such sets, and is not defined in the case of a non-degenerate FDL set.

H_k is a rv denoting the scheduling horizon of burst k, that is, the earliest time until all previous bursts will have left the system, as seen by arriving burst k.

N is a parameter denoting the optical buffer size. The number of lines in an FDL set, including the line with length zero, is $N + 1$.

R_k is a rv denoting the reactivation time following burst k [→Chapter 3].

T_k is a rv denoting the inter-arrival time of burst k, defined as the time between the arrival of burst k and the arrival of the next burst.

U_k is an auxiliary rv associated with burst k, defined as $U_k = B_k - T_k$.

V_k is a rv denoting the void size that burst k creates upon arrival, defined as $V_k = W_k - H_k$.

W_k is a rv denoting the waiting time of burst k, that is, the time that burst k spends in the buffer between the instant of arrival and the instant that transmission commences. In this work, it is used interchangeably with "delay".

X_k is a rv denoting the batch size of batch k [→Chapter 2].

Y_k is a rv denoting the unavailable period following the arrival of burst k [→Chapter 3].

Z_k is a rv denoting the number of losses during the unavailable period associated with burst k [→Chapter 3].

Terms

Burst is a term borrowed from OBS context. In an OPS context, it can simply be replaced by “packet”.

Classic buffer is used interchangeably with “RAM buffer” or “non-optical buffer”. It refers to a buffer with continuous waiting room, that doesn’t suffer performance loss from voids.

Degenerate buffer settings have FDL lengths that are multiples of the granularity D. Such settings are also called equidistant, and are opposed to non-degenerate settings.

Equidistant is a synonym for “degenerate”, and is opposed to “non-equidistant”.

FDL buffer and “optical buffer” are used interchangeably in Chapters 2 to 6 of this work.

Memoryless distribution is used as an umbrella term for the geometric distribution (in DT) and the negative-exponential distribution (in CT).

Transform function is used as an umbrella term for pgf (in DT) and lst (in CT) in this work.

Samenvatting
—Summary in Dutch—

¶ Glasvezel is het standaardmedium voor lange-afstandscommunicatie. Grote steden zijn verbonden met optische DWDM (Dense Wavelength Division Multiplexing) verbindingen, en hebben op die manier datadebieten van meer dan 10 Tbit/s per vezel ter beschikking. Pakketschakelen over deze optische verbindingen vereist evenwel dat de transmissiecapaciteit van de verbinding ook gehandhaafd kan worden bij het schakelen in de knooppunten. Aangezien huidige pakketschakelaars de data in het tragere elektronische domein verwerken, wordt het schakelen stilaan het knelpunt als het op netwerksnelheid aankomt. Nieuwe, pakketgeörienteerde technologieën zoals OPS (Optical Packet Switching) en OBS (Optical Burst Switching) kunnen dit knelpunt omzeilen, door te schakelen in het optische domein. In de resulterende optische schakelaars zijn buffers een afdoende oplossing voor conflicten, die zich voordoen wanneer twee of meer datapakketten op hetzelfde moment over dezelfde uitgang willen beschikken. Aangezien licht niet tot stilstand kan gebracht worden, buffert men de data door ze met behulp van een schakelmatrix doorheen een vertragingslijn (Fiber Delay Line, FDL) van voldoende lengte te sturen, gekozen uit een set van FDL's. Zodoende introduceert optisch schakelen een nieuw type buffer, dat het onderwerp van dit doctoraatsproefschrift uitmaakt.

Optisch schakelen (zowel OPS als OBS) bevindt zich nog steeds in de onderzoeksfase, en het tijdstip voor werkelijke (commerciële) verbreiding is alsnog een vraagteken. Zoals El-Bawab[1] het stelt: "zodra alle technologische ingrediënten volgroeid en kosteneffectief zijn, en wanneer de telecomindustrie klaar is voor de stap." Wanneer de omslag van elektronisch schakelen naar optisch schakelen zich voltrekt is moeilijk te voorspellen, maar de technische en operationele drijfveren voor de visie van de optische laag zijn authentiek: de netwerkarchitectuur heeft nood aan vereenvoudiging, flexibiliteit, intelligentie en schaalbaarheid. Optisch schakelen komt aan deze nood tegemoet, en optische buffers vormen in dit verhaal een klein maar zonder twijfel vitaal onderdeel.

In de huidige situatie heeft optisch schakelen nog steeds af te rekenen met enkele (technologische) kinderziekten, en heeft het bovendien nood aan nauwkeurige hulpmiddelen voor de prestatie-analyse van het beoogde netwerk. Dit laatste

[1] T.S. El-Bawab. *Optical Switching*. Springer, 2006.

is cruciaal, aangezien het netwerkarchitecten toestaat om de prestatiewinst af te schatten die men van de overgang van elektronisch schakelen naar optisch schakelen mag verwachten, en bovendien toestaat om het geassocieerde netto kostenplaatje te maken. Dergelijke middelen voor prestatie-analyse kunnen bekomen worden door delen van het beoogde netwerk, zoals de FDL buffers, in een stochastisch model te gieten. Stochastische modellering van FDL buffers heeft de laatste jaren veel aandacht genoten [1–37], en vormt ook de invalshoek van dit boek, dat grotendeels gebaseerd is op [1–9].

Het geheel van FDLs en schakelmatrix vormt een optische buffer, met lijnlengtes die typisch (maar niet noodzakelijk) een veelvoud zijn van een basiswaarde D, genaamd de granulariteit. Een kenmerkende eigenschap is dat bursts (of pakketten) in een optische buffer pas beschikbaar zijn wanneer ze de toegewezen vertragingslijn verlaten, in plaats van op willekeurige tijdstippen, zoals in het geval van een klassieke RAM buffer. Dit heeft tot gevolg dat er zich periodes voordoen waarin het uitgangskanaal beschikbaar is maar ongebruikt blijft, ondanks het feit dat er nog bursts in de buffer wachten om verzonden te worden. Deze periodes heten leemtes (Engels: voids), en treden niet op in het geval van klassieke buffers. Wanneer we de vergelijking maken met klassieke RAM buffers, kunnen we van deze leemtes enkel een prestatievermindering verwachten. De leemtes zijn inherent verbonden aan de structuur van de FDL buffer, en kunnen wel gereduceerd, maar niet uitgeschakeld worden.

Overigens impliceert het eindig aantal vertragingslijnen niet alleen dat er leemtes optreden, maar ook dat er een maximaal realizeerbare vertraging is, gelijk aan de lengte van de langste lijn. Dit creëert een effect dat in de wachtlijntheorie meestal ongeduld wordt genoemd, en dat resulteert in het verlies van bursts, telkens wanneer een burst een grotere vertraging moet ondergaan dan realizeerbaar is. Bemerk dat dit verschilt van de situatie bij klassieke buffers met eindige grootte, waar een bovengrens wordt opgelegd aan het aantal bursts dat gelijktijdig in de buffer kan aanwezig zijn. Wanneer we de vergelijking maken met klassieke RAM buffers, leidt ongeduld evenwel niet noodzakelijk tot een prestatieverlies. Veeleer vormt het een karakterisatie van de eindigheid van de buffer die complementair is aan de typische karakterisatie, door middel van een bovengrens op het aantal beschikbare wachtplaatsen.

In dit opzicht hebben alle prestatiemodellen in dit werk gemeen dat ze het gecombineerde effect van leemtes en ongeduld in rekening brengen. Deze modellen dienen hoofdzakelijk als werktuig voor prestatie-analyse. Meer bepaald staan ze toe om het prestatieverlies ten gevolge van leemtes te ondervangen, en dit door de ontwerpsparameters (in de eerste plaats de granulariteit) met dit doel af te regelen. De voornaamste prestatiematen zijn de verlieskans en de gemiddelde wachttijd van bursts. Deze worden bekomen voor algemene aannames voor de tussenaankomsttijden en de burst- (of pakket-)groottes, en dat voor een algemeen tijdskader (zowel discrete tijd als continue tijd). De analyse is generiek gesteld in termen van de betrokken parameters, en staat een netwerkarchitect toe om haar in

te zetten voor een waaier van toepassingen. Het meest voor de hand liggende doel is de optimalisatie van de vezellengten met het oog op minimaal verlies. Een ander typisch gebruik is de bepaling van de minimale buffergrootte, nodig om te voldoen aan gegeven prestatievereisten.

Wat de opbouw van dit werk betreft kunnen we spreken van complementaire hoofdstukken, aangezien elk een andere methodologie hanteert: met tranformaties (Hoofdstuk 2), met Markovketens (Hoofdstuk 3), met ongeduld (Engels: impatience) (Hoofdstuk 4), en met stochastische regeneraties (Hoofdstuk 5). Deze methodes worden achtereenvolgens behandeld, en staan toe om het FDL-bufferprobleem op grotendeels complementaire wijze te belichten. Naar de mening van de auteur levert de aanpak met transformaties de vanuit mathematisch oogpunt elegantste prestatie- en stabiliteitsanalyse op, die, voor een lezer met enige vertrouwdheid met wachtlijntheorie, een maximaal inzicht in de werking van optische buffersystemen toestaat. De aanpak met Markovketens staat dan weer toe om de meest accurate resultaten te bekomen, en levert in sommige gevallen zelfs een oplossing in gesloten gedaante op. De aanpak met ongeduld biedt dan weer een andere kijk, en laat de in mathematisch opzicht eenvoudigste manier van oplossen toe. Tenslotte reikt de aanpak met regeneraties ons een elegante methode aan om stabiliteitsvoorwaarden op te stellen voor een zeer brede klasse van FDL-buffersystemen.

Zodoende voegt het geheel van deze vier benaderingen een consistent en mathematisch onderbouwd raamwerk toe aan het debat omtrent optische buffering. Het kan van waarde zijn voor lezers die op zoek zijn naar elementaire inzichten over, prestatie-analyses voor, of intuïtie omtrent de werking van optische buffers. Gezien de essentiële rol van prestatie-analyse in de verbreiding van nieuwe technologieën, zien wij voor dit doctoraatsproefschrift ook een rol weggelegd als pleitbezorger voor optische buffers, en, in het algemeen, optische netwerken.

Summary

¶ Optical fiber is the standard carrier for data transport over long distances. Nowadays, major cities are connected by dense wavelength division multiplexing links (DWDM), enabling transmission capacities of over 10 Tbit/s per fiber. Packet switching over these optical links, however, requires that the transmission speeds over the links are matched by equivalent switching capacities in the nodes. As current packet switches perform data processing in the (slower) electronic domain, the switching is becoming the bottleneck in terms of network speed. Novel packet-oriented technologies like optical packet switching (OPS) and optical burst switching (OBS) promise to alleviate this bottleneck, by performing the switching in the optical domain. In the resulting optical switches, buffers provide an appropriate solution to external blocking, that occurs whenever two or more data packets contend for the same output at the same time. Since light cannot be stored "in place", data is buffered by sending it through a Fiber Delay Line (FDL) of sufficient length, chosen from a set of FDLs by means of a switching matrix. As such, optical switching gives rise to an essentially new type of buffering, that is the subject of this doctoral dissertation.

Optical switching (OPS as well as OBS) is still in its research phase, and the timing for wide-spread adoption is an as yet unanswered question. Quoting El-Bawab[2]: "as soon as all technological ingredients are fully mature and cost-effective, and when the telecommunications market is ready for this step." When exactly the shift from electrical switching to optical switching will take place is hard to predict, but the technical and operational drivers for the vision of the optical layer are genuine: the network architecture needs simplification, agility, flexibility, intelligence, and scalability. Optical networking comes up to these needs, and optical buffering, although but a small piece of this optically-switched network, surely provides a vital piece.

At this point, optical switching still needs to overcome some (technological) teething troubles, but also requires accurate tools for the performance evaluation of the envisaged network. The latter is key, since it allows network architects to predict the performance gain that the shift from electronic switching to optical switching will bring about, and further allows to draw up the outline of associated net costs. Such performance evaluation tools can be obtained by casting parts of the envisaged network, such as the FDL buffers, in a stochastic model. Stochastic

[2] T.S. El-Bawab. *Optical Switching*. Springer, 2006.

modeling of FDL buffers has received much attention in recent years [1–37], and also constitutes the line of approach of this book, that is largely based on [1–9].

The FDLs, together with the switching matrix, constitute an optical buffer system, with line lengths that are typically (but not necessarily) a multiple of a basic value D, called the granularity. An essential characteristic is that bursts (or packets) present in the optical buffer are only available when they leave the assigned delay line, rather than at arbitrary ("random") instants, like in the case of a classic RAM buffer. This results in periods during which the outgoing channel remains unused, even though bursts are still present in the buffer and awaiting transmission. These periods are called voids, and do not come about in the case of a classic buffer. When compared to a classic RAM buffer, the effect of voids is strictly a performance degradation. The voids are implied by the FDL buffer structure, and can be mitigated, but not annihilated.

On the other hand, the limited number of delay lines implies that there is a maximum achievable delay, namely the length of the longest line. This creates an effect that is commonly referred to as impatience in queueing literature, and results in burst loss, whenever a burst finds that its required waiting time is larger than the maximum achievable delay. Note that this is different from the classic case of finite buffer size, where the number of bursts that can be present in the buffer is bounded. When compared to a classic RAM buffer, the effect of impatience is not necessarily a performance degradation. Rather, it provides a characterization of buffer finiteness that is complementary to the typical limit on buffer size, namely an upper bound on the number of waiting places.

As such, taking into account the combined effect of voids and impatience is key in all the performance models presented in this work. The main scope of these models is to provide a performance evaluation tool to mitigate the performance loss due to voids, by fine-tuning the involved design parameters (in the first place, the granularity) to this end. Main performance measures of interest are the loss probability (LP) and the waiting time of bursts. These are derived for general assumptions on the inter-arrival times and burst (or packet) sizes, and also for general time setting (both discrete-time (DT) and continuous-time (CT)). The analysis is generic in terms of the involved parameters, so allowing a network architect to deploy it for a wide variety of applications. Most obvious purpose is the optimization of the fiber lengths for minimal loss. Another typical use is to determine the buffer size needed, so as to meet given performance requirements.

As for the structure of this work, the chapters are complementary, in that each represents a different methodological approach: with transform functions (Chapter 2), with Markov Chains (Chapter 3), with impatience (Chapter 4) and with stochastic regenerations (Chapter 5). These methods will be treated subsequently, and will allow to shed light on the FDL buffering problem mostly in a complementary way. In the author's opinion, the approach with transform functions yields the most elegant mathematical performance and stability results, that, for the reader

acquainted with queueing theory, allow for maximal insight into the functioning of FDL buffers. The approach with Markov Chains allows for the most accurate loss performance results, and in some cases even yields closed-form solution. The approach with impatience provides an alternate view to the system, that allows to obtain results of minimal mathematical complexity. Finally, the approach with regenerations offers an elegant way to establish stability results for a very broad class of FDL buffer systems.

As such, the whole of these four approaches adds to the debate on optical buffering a consolidated framework for performance evaluation with firm mathematical basis. It can prove an asset to readers looking for essential insight in, performance tools for, and intuition on an optical buffer's capabilities. Further, given the essential role of performance evaluation in the propagation of new technologies, this dissertation both directly and indirectly stimulates the deployment of optical buffers in actual networks, and the deployment of optical networking in general.

1

Introduction

¶ This doctoral dissertation is devoted to the performance modeling of optical buffers. As a vital part of the next-generation backbone network, optical buffers provide a hot topic in present networking research, challenging engineers to think beyond the limitations of present-generation network technology. Sometimes, this implies giving up old ideas about what a technology is, was or has been. Sometimes, this implies introducing into communications networks elements and techniques that never were employed for that purpose before. Always, this includes the evaluation of the cost and the performance of the newly proposed technology.

In this context, the present work adds to the debate on optical buffering a consolidated framework for performance evaluation with firm mathematical basis. It can prove an asset to readers looking for essential insight in, performance tools for, and intuition on an optical buffer's capabilities. Further, given the essential role of performance evaluation in the propagation of new technologies, this dissertation both directly and indirectly stimulates the deployment of optical buffers in actual networks, and the deployment of optical networking in general.

Before moving to the stochastic modeling, which forms the heart of this work, we first provide the reader with a concise introduction to optical networking and optical buffering, essential to grasp the implications of the performance results we obtained. This is the material of the current chapter.

1.1 Access Networks vs. Backbone Networks

Given the rampant ubiquity of the Internet, and the apparent dependence of economics upon it, the backbone is bound to process an immensely growing amount of traffic. In the whole of a computer network, the backbone is the central part of the network, as opposed to the peripheral parts, that constitute the access network. The access network is that part of the network where individual users connect with the network. In present Belgium, we typically think of wired technologies like for example twisted pair (ADSL, ADSL2(+), VDSL(2)) and coax (DOCSIS) and of wireless technologies, such as for example GSM, WiFi. Typically, the access network delivers service locally, to a limited number of users. As such, the connection speed (or, bandwidth) of individual users needs to be sufficient to provide good service, but the aggregated traffic of different users on the same connection does not result in very large bandwidth, since the number of users is limited. Further, note that, for instance in Belgium, the maximum amount of bandwidth that a single user can consume is limited not only by their subscription type, but also by the physical properties of the access medium itself, namely copper wire. State-of-the-art technology allows for (theoretical) speeds of up to 250 Mbit/s over twisted pair (VDSL2, as of 2008) and 400 Mbit/s over coax (DOCSIS 3.0, as of 2008), which is actually no match for the theoretical capacity of a single optical fiber, mentioned below. Therefore, private users in Belgium typically have access at rates of 1-100 Mbit/s, whereas massive deployment of Fiber to the Home (FTTH) currently enables access speeds of up to 1 Gbit/s for private users in Japan.

Opposed to this, the backbone network (or, for short, backbone) typically delivers service internationally, even globally, to a huge amount of users. The backbone thus conveys the aggregated data of thousands of users per link, and carries data world-wide, interconnecting cities by long-haul links of hundreds of km, connecting different backbone nodes. The connections between backbone nodes consist of optical fibers, that apply Dense Wavelength Division Multiplexing (DWDM, a more recent version of WDM) to carry multiple lightwaves at different frequencies on one single-mode optical fiber, realizing total data rates well beyond the Tbit/s with one fiber. As of 2008, state-of-the-art optical fiber technology enables up to 160 channels at an equal number of wavelengths, with a data rate of typically 40 Gbit/s per channel, ranging up to 160 Gbit/s per channel. To name a recent example: a total data rate of 12.8 Tbit/s was obtained over a 2550 km link, as demonstrated at ECOC 2007 in September 2007 [38]. As a reference: the first transatlantic telephone cable to use optical fiber (TAT-8) went into operation in 1988, operating at first with one wavelength, and allowing for a bandwidth of 280 Mbit/s.

1.2 Alleviate the Bottleneck

In early computer networks (1980s and earlier), the physical properties of the link itself constituted the limiting factor for connection speed. As noted by El-Bawab [39], the extensive deployment of Wavelength Division Multiplexing in the 1990s annihilated this bottleneck, so creating a new one, namely the switching. In current backbone networks, packets travel hop-by-hop in the form of light, but are converted into electricity in order to extract header data, buffer them, convert them back to light and transmit them to the next hop. As such, the switching is done by electronics, and requires intermediary conversions: from optical to electrical (O/E), and then back from electrical to optical (E/O). The combination of the latter two conversions is referred to as O/E/O conversion. Given that the latter is done on a channel per channel basis, it does not only create a tremendous multiplication of switching elements (one electronic circuit per channel), but also introduces conversion delay, while limiting flexibility and scalability [40]. Further, even with complicated electronic circuits able to handle the switching at sufficient speed, the current trend is that optical transmission capacity grows faster than electronic processing capacity [41], making electronics the most likely bottleneck of the next-generation network.

1.3 Wavelength-Routing Optical Networks

Although the solution that wavelength routing provides is not the focus of the current work, it is an optical technology that is actually deployed in present networks, and therefore worth mentioning [42–44]. The central concept in this approach is to set up wavelength paths (or lightpaths) between a source-destination pair. Once set up, such wavelength path allows data to flow in optical form, without O/E/O conversion in-between. In this context, a connection without conversion is called *transparent*, since it allows for light beams to travel through it purely in the form of light. Opposed to this, a connection with conversion is called *opaque*, since it does not allow for light beams to pass through it without intermediary conversion to electrical form. The backbone nodes in-between allowing for such transparent connection are Optical Cross Connects (OXCs).

Although this approach allows to avoid O/E/O conversion, an important difference with the optical switching considered in this work is that a wavelength-routing network is essentially circuit-switched. This means that typically, a packet travels from source to destination trough a connection that spans multiple nodes, and that is set up beforehand. This is opposed to (pure) packet-switched networks, where packets are routed locally hop-by-hop, and no (higher-level) connection is set up beforehand. To see this, consider the following example, that also occurs in [44]. A wavelength-routing network carries IP traffic, and a lightpath is created

between source-destination pair A and C, by passing from router A, over B, to C. With this given, at the IP layer, router A sees a direct connection to router C, with large capacity, associated with the wavelength of the lightpath that was set up. The IP protocol being essentially connectionless, the IP routers are unaware of the intermediate routers on the path, and just see a link between A and C at their disposition for packet forwarding. Since this link gets assigned an entire wavelength in a fixed manner, this capacity is no longer available for packet forwarding along other links. For example, if router B wants to route packets over the link between B and C, it cannot re-use the wavelength associated with the lightpath between A and C, even if no data is transmitted over it at the given time. The same goes for the situation where for example another router D would want to route packets to router C over router B. As such, an increased number of wavelengths is needed to fulfill the bandwidth demands of a given traffic matrix. This essential drawback can be partially overcome with Automatically Switched Optical Networks (ASONs), where wavelength paths are set up by the control plane, without explicit intervention of the network operator, allowing to adapt to traffic variations on medium to large timescale (several minutes to hours or days). Still, flexibility of the connection is low, and the network is operated mainly as a set of source-destination pairs, rather than as an actual topology-aware network. This results in poor usage of resources, because bottleneck links between source and destination cannot be avoided by routing over alternate links, as it can be done in topology-aware networks.

Finally, note that OXCs are available on the market today, and are typically relatively cheap in terms of cost per Tbit/s. Since they do not require O/E/O conversion, and also are not to allow fast reconfiguration times, it is possible to use low-cost technologies like Micro-Electromechanical Systems (MEMs) for the implementation [45].

1.4 Optical Packet and Burst Switching

The drawback of wavelength-routing optical networks is that it does not allow to re-use wavelengths of a lightpath for other traffic. This is actually a specific case of a more general consideration, as proposed in (the introduction of) [43]. There, Perros distinguishes between circuit-switched and packet-switched networks. Typical examples of circuit-switched networks include the above-mentioned wavelength-routing optical networks, and telephone networks. Typical examples of packet-switched networks include approaches with virtual circuits (connection-oriented networks, like ATM and frame relay) on the one hand, and approaches with datagrams (connectionless networks, like IP) on the other hand. Main advantage of circuit-switched networks is that they are easy to control and provide robust operation. However, it is well-known that circuit-switching is not very efficient from

a bandwidth perspective. Opposed to this, packet-switching allows for point-to-point connections, without making the allocation of intermediary nodes strictly necessary. This allows for statistical multiplexing, since numerous traffic flows can be treated at once from hop to hop, implying that the entire fiber can be shared by all data arriving at the node. Combining different flows results in more efficient resource usage and better overall performance. This is applied in the current electronically-switched networks (with O/E/O conversion), and provides a considerable performance advantage over circuit-switching. In this regard, although wavelength-routing optical networks do allow to circumvent the electronic bottleneck, they are compromised in that they do not allow for the indispensable statistical multiplexing, needed to maximize resource utilization.

1.4.1 Optical Packet Switching

Optical packet switching (OPS) is presented as key solution to these problems [46–51], by performing packet switching in the optical domain. On the one hand, optical switching avoids the costly O/E/O conversions that occur on the path in-between a source-destination pair, providing thus a transparent connection. On the other hand, OPS allows for flexible routing, and benefits from the advantages of packet switching in general, amongst others statistical multiplexing.

For the most flexible implementation, all operations are to be moved to the optical domain: optical packets arrive at network nodes in the form of light, the packet header is extracted optically, the header is interpreted optically, the packet is switched to the desired destination optically, and the header information is rewritten optically, all on a packet-by-packet basis. The latter description concerns OPS in its purest form, and is often referred to as all-optical packet switching. Given that especially the control is very hard to implement optically, even an implementation with electronic control is commonly referred to as OPS [48], but never as all-optical packet switching. In this regard, slightly different types of OPS coexist, but all have in common that the payload is switched optically, in packets of limited length. The continuous interest in OPS is also reflected in numerous research projects [40, 49, 51–57].

1.4.2 Optical Burst Switching

In recent years, Optical Burst Switching (OBS) [58, 59] received much attention from the research community, as it provides a practical compromise between wavelength routing and packet switching. The main motivation is to relax the stringent technological requirements that OPS imposes, mainly by (i) assembling packets into larger bursts, which reduces the burden on control, and by (ii) transmitting data and control separately and performing control electronically, which avoids the reading and writing of packet headers in the optical domain.

While the simplest instance of OPS can be understood as IP packet switching in the optical domain, OBS is somewhat more particular in its approach. Also, note that the name OBS does not identify a single architecture, in that many OBS papers consider implementations that deviate considerably from [58], that nevertheless remains the main reference for OBS. In an OBS network, packets are assembled into larger data bursts (DBs) (or, for short, bursts[1]), that are provided each with a Burst Header Packet (BHP). The assembly of bursts and creation of BHPs is done in ingress OBS edge nodes. This burstification process is typically either threshold-based or age-based (for variable packet length) [60], and (rather seldom) strictly threshold-based (for fixed-length bursts) [61]. After burstification, each burst/BHP pair is sent to their destination (an egress OBS edge node) through OBS core nodes. The burst is transmitted over a data channel, while the BHP is sent over a dedicated control channel. Depending on the control protocol, BHPs are transmitted ahead of the burst with a certain offset time, calculated beforehand so as to compensate for the processing delays of the BHPs in subsequent nodes. The burst is transmitted transparently (without O/E/O conversion), whereas the BHP goes through O/E/O conversions at every node, to allow for a (more) feasible control. The control itself consists in configuring each node so as to allow forwarding the burst to the correct output when it arrives [51, 62], so permitting the burst to reach the egress node, where the burst is disassembled. This configuration requires a reservation mechanism; especially in case multiple wavelengths are available to realize the correct forwarding, a variety of reservation mechanisms such as Horizon Scheduling [51], Latest-Available Unscheduled Channel (LAUC) [62] and Latest-Available Unused Channel with Void-Filling (LAUC-VF) [62] exist.

In this work, we will consider optical buffers in both an OBS and OPS context. For notational convenience, we will consistently use the word “burst” rather than “packet” in the following, unless we want to make explicit reference to an OPS or electronic packet switching context. We note however that the word “burst” is mostly used interchangeably with “packet” in this work, and results equally hold for OPS, and, in principle, for any optical switching design that involves optical buffering.

Finally, note that optical switching (OPS as well as OBS) is still in its research phase, and the timing for wide-spread adoption is an often-recurring, as yet unanswered question. El-Bawab provides his answer in [39] (Sect. 1.4.1): “as soon as all technological ingredients are fully mature and cost-effective, and when the telecommunications market is ready for this step.” Further on, he mentions three main reasons for the eventual adoption of optical networks. First, the true drivers for optical switching are genuine: simplifying the network architecture, while providing flexibility, intelligence and scalability. Second, the progress made in several

[1] In literature, these bursts are often referred to as data bursts (DBs). Since the shorter term “burst” does not lead to confusion, we prefer to use the latter name throughout this work.

optical technologies, especially in optical switching, is real and impressive. Third, the annual growth of the Internet is a fact. The yearly incremental increase in traffic volume might not be matched by revenue growth for carriers (due to the bussiness model), but nevertheless urges to provide additional bandwidth in the future. For further detail, we refer to [39].

1.5 Contention Resolution

In every network, burst (or packet) contention occurs whenever two or more bursts head for the same destination simultaneously. In the following, we look at three possibilities to deal with this in an optical network, one of them being buffering.

Given that OBS and OPS process individual bursts (or packets), contention naturally arises whenever two or more bursts want to use the same output on the same wavelength at the same time. Of these contending bursts, one can be transmitted on the desired wavelength immediately, while the others have to be dealt with by means of a contention resolution scheme. In classic (electrical) packet-switched networks, contention resolution is provided very simply, by a store-and-forward technique, which stores contending packets in a memory bank, and sends them out as soon as the desired output port turns available again. The memory bank is implemented by means of electronic random-access memory (RAM). Since light cannot easily be stored in a RAM buffer, a somewhat less straight-forward way of contention resolution is to be adopted. In the context of OPS, Yao and his co-authors [63, 64] identify three dimensions for contention resolution in OBS/OPS: space, wavelength and time. This corresponds to following three strategies:

1. **Deflection routing** (space), also called space deflection in [63, 64], is a multi-path routing technique, that reroutes bursts that lose the contention to other than the preferred next-hope node. The main idea is that the bursts will eventually be routed to their proper destinations. Clearly, the latter can only be the case if the network is not too severely congested. Especially in the case of severe traffic congestion, deflection routing is known to cause a significant increase in (effective) load.

2. **Wavelength conversion** (wavelength) offers a solution by converting a contending burst to another wavelength, but on the same output port. It is a very straight-forward technique, that does not cause extra burst latency, jitter, and re-sequencing problems. However, it comes at the cost of additional channel capacity, and necessitates a wavelength converter at the given point.

3. **Fiber delay line buffering** (time) allows bursts that lose the contention to wait, by sending them through a piece of fiber of appropriate length. This solution is the optical counterpart of electronic RAM, and shares some of

its benefits, in that it allows to deal with contention locally, without requiring the use of additional link capacity. However, FDL buffers have a large physical size, which implies considerable implementation cost.

Apart from these three possibilities, note that a network does not necessarily require contention resolution. In principle, a network that drops all bursts that cause contention might also function, on the condition that protocols in higher network layers are able to compensate for such losses. However, even in that case, it can be expected that a network without contention resolution will exhibit poor performance due to poor resource usage and high burst/packet loss probability.

Typically, the three presented forms of contention resolution are considered in combination. A typical contention resolution then begins with seeking an available wavelength on the preferred output port. If none is available, optical buffering is applied. If the buffer is full, an available wavelength is sought on the deflection output port. If the latter does not succeed, the burst is dropped. In general, it is accepted that a combination of wavelength conversion and FDL buffering provides the best performance in terms of resource usage [64], with deflection routing only as a possible third recourse. Therefore, the combination of wavelength conversion and optical buffering is also the most typical setting in testbed situations. The superior performance of this combination over single-wavelength optical buffering will also be confirmed by the results presented in this work, see Sect. 2.5.

The contention resolution scheme outlined here was originally proposed for OPS, but equally goes for OBS. In principle, OBS does not necessitate the use of FDL buffers [58], since the involved resource reservation mechanism should be able to operate also without buffers. However, performance results for OBS networks point out that also for burst switching, a combination of wavelength conversion and FDLs provides the best results [65], possibly with a converter pool and shared buffering [66].

In this doctoral dissertation, the focus is on FDL buffering as means of contention resolution. In most of this work, a single output wavelength is considered, except in Sect. 2.5 and 5.2, where also the combined impact of wavelength conversion and FDL buffering is considered. In queueing lingo, this amounts to a system with a single server, and to a system with multiple servers, respectively.

1.6 Buffering for Optical Networks

Similar to RAM buffering, optical buffering can be deployed at the switch output or input. Such strategies are treated first in this section. Next, given that our work focuses on an output buffer within an optical node of an optical network, it is instructive to look into a recent example of a research project that proposes exactly this setting. Then, we add a note on the implementation of FDL buffers.

1.6.1 Buffering Strategy

As a means of contention resolution, buffering can be deployed according to different strategies, that are known already for classic RAM buffers in non-optical networks. Taking the listing of [67] as a reference, we identify as typical alternatives output buffering, input buffering, shared buffering and buffering with recirculations.

While output buffering deals with contention at the output, input buffering resolves contention by buffering contending bursts before they enter the switch. It has similar implementation complexity, compared to output buffering, but lowered performance, due to (the well-known effects of) head-of-line blocking [68, 69]. Shared buffering can be understood as a variant of output buffering, but is much more complex to implement optically [67]. It also provides buffering at the output, but then for multiple fibers at once, instead of being dedicated to just one output fiber (like in the case of output buffering). It was studied in more detail, also from a performance evaluation perspective, in [33–35, 66]. The fourth option is to let bursts recirculate from the output of a space switch back to the input. When implemented optically, typically many recirculations are required, implying high loss and accumulation of amplifier noise with each recirculation. A stochastic approach to this problem can be found in [36, 37, 70].

1.6.2 A Recent Example

Many research projects propose optical switches that include fiber delay lines for the buffering [52–56, 67, 71–74]. As mentioned, we assume output buffering, and associate the buffer with one output fiber. In literature, this is a common assumption, that is applied for example recently in the IST-LASAGNE project [49, 75]. Rather than proposing a proper switch design, we prefer to take the generic OPS switch architecture proposed for the latter project as a reference, as displayed in Fig. 1.1 (Figure source: [49]). It includes contention resolution at the output, in the general form of FDL buffering and wavelength conversion. Also, since each wavelength realizes merely a link between two switches, rather than an entire connection between source and destination, it is possible that different wavelengths are realizing different routes if the protocol requires this. (This occurs in for example OBS, where the entire connection is set up beforehand.) As a result, two situations can occur:

1. The c channels on the outgoing fiber all have different destinations. Then, given that no deflection routing is applied, all different c wavelengths in the FDL buffer are used for independent destinations, and, as a result, the buffer implements c different queues that are independent. Each of the c queues can be studied separately, as a single-wavelength FDL buffer, or, in terms of queueing systems, as a single-server system.

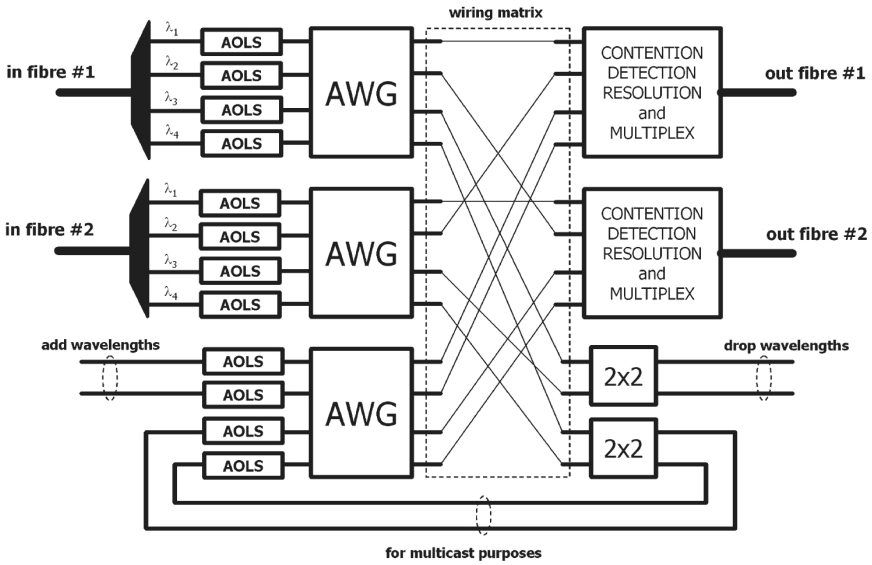


Figure 1.1: LASAGNE node architecture as proposed in the IST-LASAGNE project.

2. Of the c channels on the outgoing fiber, c_0 ($\leq c$) have a common destination. Then, again given that no deflection routing is applied, the group of c_0 wavelengths is associated with one destination, and buffer control has access to multiple wavelengths for contention resolution. This case is considered in Sect. 2.5, and results in a variety of possible scheduling algorithms. Depending on the scheduling algorithm, the buffer will benefit to a certain extent of multiplexing gain, since the use of multiple wavelengths allows to spread contending packets (or bursts) more evenly over different output wavelengths. In terms of queueing systems, the buffer system of c_0 wavelengths corresponds to a multi-server queue. As for the remaining $c - c_0$ wavelengths, similar considerations go.

Further, while a study of the entire node architecture is beyond the scope of the current work, it is instructive to mention the components involved in the LASAGNE architecture. As shown in Fig. 1.1, the fibers at the input each carry a number of wavelengths (in this case, four), that are first demultiplexed, usually by means of a simple prism. Second, each wavelength enters an AOLS block (All-Optical Label Swapping). Each AOLS block allows for correct all-optical packet forwarding, with separation of the packet payload (at 40 Gb/s) and label (at 10 Gb/s), header extraction and interpretation, new label generation and insertion. The wavelength on which the packet leaves the AOLS is further involved in the scheduling, since it is determining in the followed path through the arrayed-waveguide grating (AWG).

Traversing a wiring matrix, the last step is the contention resolution and multiplexing block, especially relevant for the current work.

Finally, note that in the IST-LASAGNE project, the aim was to implement also the contention resolution all-optically. As such, also the FDL buffers had to be provided with optical control, which is apparently more feasible if the buffer is placed at the input. For this reason, an all-optical variant of lower control complexity was placed at the input. Note however that in most cases, the fact that output buffering demands electronic control is hardly to be considered a drawback, as treated in the next note.

Note on the Control of Output FDL Buffers

Just as illustrated in Fig. 1.1, we assume output buffering. As mentioned in the previous section, the fact that output buffering is best implemented with electronic control might seem to compromise the applicability of our results. Indeed, input buffering is typically the approach in research projects that aim for all-optical implementation, with the control performed in the optical domain [40]. However, most importantly, whenever it is not a prerequisite to do the switching entirely in the optical domain, there is a general consensus [39,76] that control should be done in the electronic domain, for reasons of feasibility, speed and robustness. Further, note that there are also other considerations that make output buffering preferable over input buffering, such as the better performance, mentioned in Sect. 1.6.1. At any rate, independently of whether the control is done optically or electronically, the results of this work are valid for any optical buffer that carries the payload in optical form. As such, this includes buffer settings of OPS (not necessarily all-optical), and (potentially) all variants of OBS that contain buffers.

1.7 FDL Buffers

In the above, we mostly treated the broader context of FDL buffering; below, we shift focus to the anatomy of the buffer itself: the structure, the related performance problem, the control algorithms, and also to some alternatives to FDL buffering.

1.7.1 FDL Buffer Structure

A conceptual diagram of the buffer studied in this work is given in Fig. 1.2. Recall that this buffer is associated with one output port (or output line), and thus, bursts coming in from the input channels (3) are all destined for this output port. The fiber delay lines (1) are provided with an input and output switching matrix (2) to allow to switch bursts coming in from various input channels to any delay line. In general, the switching can be done to several output fibers (4), however, given the huge capacity of one single fiber, it is realistic to assume only one output fiber, that

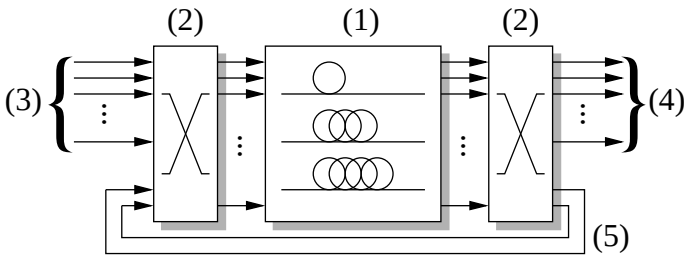


Figure 1.2: Conceptual diagram of a single-stage FDL buffer, with fiber delay lines (1), input and output switching matrix (2), input (3) and output lines (4) and with the possibility of feedback (5).

can carry one or multiple output wavelengths (see further). Taking the terminology of Hunter as a starting point [67] (also adopted in [63, 64]) FDL buffer types may be further categorized according to following properties.

- **Degenerate (equidistant) or non-degenerate (non-equidistant).** While the lengths of the FDLs can be chosen freely in principle, a common assumption is to let the lengths of the fibers be multiples of some basic value D, called the granularity (a term coined in [24]). According to the definition of [77], the former is called non-degenerate, and the latter degenerate. Completely equivalent are the terms coined in [17], where a degenerate setting is referred to as equidistant, and the non-degenerate setting as non-equidistant. The latter terms originate from the fact that a set of degenerate FDL lengths constitutes an equidistant grid when set out on a time axis, while the opposite is true when non-degenerate optical buffers are considered. In this work, we will use the somewhat more often encountered terms "degenerate" and "non-degenerate". In case of a degenerate buffer setting, the FDL set will be referred to by means of its most characteristic parameter, namely the granularity, whereas for the non-degenerate buffer setting, the FDL set will be denoted as a general set $\mathcal{A}$.

- **Number of Fibers.** A basic property of an optical buffer is the number of FDLs it has at its disposal; this relates univocally to its size N, defined as the number of FDLs that have non-zero length. Since an FDL set is always required to have a delay line of zero length, it is natural not to include this in the FDL count, even though the amount of fibers amounts to $N + 1$. Even though the number of FDLs is always finite in practice, in two chapters of this work (Chapter 2 and 5), a mathematical model is presented that assumes infinite buffer size. The reason for this assumption is that it leads to a mathematically more elegant solution. However, to derive results that are valid for finite buffer size, a heuristic is presented in Chapter 2, that is able to

convert results for infinite buffer size to results for finite buffer size in an accurate manner. Also, a study of the stability of an infinite-sized buffer is included in Chapter 5. As such, both infinite-sized (in Chapter 2 and 5) and finite-sized buffers (in all chapters) are considered in this work.

- **Feed-forward versus feedback.** In the feedback setting, bursts at the output can be fed back to the input of the buffer again with a feedback line (indicated by (5) in Fig. 1.2), while this is not allowed in a feed-forward setting. Feedback is similar, but not identical to the buffer with recirculations. In the latter case, the entire buffer is placed in the feedback loop, while feedback takes place within the buffer. Feedback allows to improve buffer utilization, but brings about the same issues as buffers with recirculation, notably the accumulation of amplifier noise with each feedback loop, that is different for each burst.

- **Number of stages.** A single-stage buffer like the one in Fig. 1.2 allows a burst to go through exactly one delay line chosen from one set of FDLs, to then leave the buffer (feed-forward) or possibly return to the input (feedback). Such a single-stage buffer is in general easy to control. In a multi-stage buffer, bursts go through as many lines as there are stages, to then leave the buffer (feed-forward) or possibly return to the input (feedback). When compared to a single stage, multiple stages, say m, (potentially) allow for more effective contention resolution, since requested delays can be realized with better approximation, as the sum of m delay line lengths of choice. On the one hand, this leads to smaller voids especially when the buffer is (almost) empty. On the other hand, if many bursts are already present in the buffer, the optimal combination of the m delay line lengths might be impossible, if one of the lines is already reserved by previously arrived bursts. Then, one is to turn to a combination of m delay lines with larger total delay, leading again to larger voids. Further, the implementation cost is considerably higher than that of a single stage buffer, since the cost rises proportionally with the number of switching matrices involved. Finally, note that, if no feedback is applied, the number of delay lines crossed is identical for each burst, which makes the noise accumulation problem less important than in the setting with feedback enabled. A prime example of multi-stage FDL buffering is the SLOB (Switch with Large Optical Buffers) [71]. As such, multi-stage architectures provide a feasible approach from an implementation point-of-view, but do remain more expensive than single-stage settings, and are (probably for that reason) less applied in testbed settings. It is not considered further in this work, but would provide an interesting extension, since it is not improbable that multi-stage buffers provides a competitive alternative to single-stage buffers in the long run, especially if switching

matrix cost reduces over time.

- **Number of wavelengths.** Intrinsically, an FDL buffer can always be used on multiple wavelengths, but naturally, the input and output switching matrix must have switching capabilities on all these wavelengths. Increasing the number of wavelengths allows for a proportional increase in capacity, but increases the price. In this dissertation, both the case of a single wavelength and of multiple wavelengths is studied.

- **Wavelength conversion capability.** Further, as mentioned in Sect. 1.6.2, an FDL buffer might either have no, limited, or full wavelength conversion capability, that is provided by means of tuneable wavelength converters (TWCs). Such TWCs are not depicted in Fig. 1.2, but can be available to buffer control before or just after the input switching matrix. Depending on the availability of TWCs, four possibilities occur.

  – Full wavelength conversion: Buffer control can convert incoming bursts to whatever wavelength, and can choose whatever channel of c channels, given that it is available.

  – Limited number of wavelength converters: A shared pool of TWCs is available. This is studied, for example, in [66]. Buffer control can convert some, but not all of the incoming bursts to output wavelengths, but sometimes runs out of wavelength converters. In that case, an incoming burst with wavelength λ_i is obliged to queue until wavelength i becomes available again, even if another wavelength (different from i) becomes available sooner.

  – Limited wavelength conversion capability: Incoming bursts on wavelength λ_i can only be converted to "near" wavelengths λ_j. Here, the difference between λ_i and λ_j remains smaller than a given small value d, $|\lambda_j - \lambda_i| \leq d$. This results in cheaper TWCs, and is studied, for example, in [78]. Buffer control can convert incoming bursts only to some of the available channels.

  – No wavelength conversion: Buffer control simply lets incoming burst on wavelength λ_i go (and, if necessary, wait) for channel i. As such, behavior of this system is identical to c single-wavelength buffers.

  In this work, whenever the multi-wavelength system is considered, we have adopted full wavelength conversion. We note that, in general, limitations on wavelength conversion capabilities can severely degrade performance. However, it is known that this degradation remains small if the limitations are not too tight, and a sufficient number of wavelength converters [66] combined with sufficient wavelength conversion capability [78] is provided.

- **Number of input channels.** While the number of input channels ((3) in Fig. 1.2) is an important design parameter for optical switch design, it is not taken into account explicitly in this work. Incorporating this design parameter into the system model is possible in principle, and would allow for an even more generic modeling of the optical buffer. However, since all the other design parameters mentioned here play an even larger role, this extension is out of the scope of the current work, and the modeling is done without specific assumptions on the number of inputs. More precisely, it is assumed that the number of inputs is (theoretically) infinite. As a consequence, the performance results obtained in this work are somewhat pessimistic when compared to the results that one would obtain if the number of inputs is taken into account. The discrepancy between both cases (infinite versus finite number of input channels) is however expected to be minor if the number of input channels is not too small when compared to the number of output wavelengths, say for example 2 times or more.

The buffer we consider in this work is a single-stage feed-forward buffer. A a point of reference, note that this is also the setting considered in the OASIS switch of the RACE ATMOS project [55, 79], and in the Broadcast and Select Switch of the ACTS KEOPS project [54, 80]. In most of this work, the buffer is single-wavelength; in Sect. 2.5 and 5.2, the multi-wavelength case is treated, and full wavelength conversion capability is assumed.

1.7.2 The Performance Trade-Off

If a buffer can only realize delays belonging to a limited set,
how does this restriction impact buffer performance? (1.1)

That is, in one sentence, the performance problem that the advent of FDL buffers raised. The fact that the set of possible delays is limited, results in a performance trade-off between two interweaved sources of loss.

Voids On the one hand, given that an FDL buffer provides only a limited number of delays, not all delay values are realizable. One cannot assign the exact delay value needed, but typically a somewhat larger delay, equal to the exact length of the chosen line. Even without taking into account the specific control algorithm (which forms the material of Sect. 1.7.3) this implies that some capacity will inevitably be lost on the outgoing channel: when bursts are present in the buffer, they may not be available yet for transmission, since they are still traversing the delay line. The capacity loss shows up as *voids* on the outgoing channel, defined

as the periods during which the channel remains unused, despite of the fact that there are bursts in the buffer that are to use it. Since the void size is upper-bounded by the difference between fiber delay line lengths, the loss due to voids can be minimized typically by taking short lines, with small differences in length.

Impatience On the other hand, the limited number of delay lines implies that there is a maximum achievable delay, namely the length of the longest line. This creates an effect that is commonly referred to as impatience in queueing literature [81, 82], and results in burst loss, whenever a burst finds that its waiting time is larger than the maximum achievable delay. Note that this is different from the classic case of finite buffer size, where the number of bursts present in the buffer is bounded. The loss due to impatience can be minimized by taking long line lengths, especially for the longest line.

As such, the trade-off consists herein, that one is to take FDL lengths not too long, so as to minimize the effect of voids, and not too short, so as not to lose too many bursts due to impatience.

While impatience is the main source of loss, the occurrence of voids results in increased waiting times, and so indirectly also contributes significantly to the loss probability. Performance loss due to voids is typical for FDL buffers, and has no counterpart in the electronic case with RAM buffering, where packet transmission can start whenever the outgoing channel is available. Not only does this problem generate new challenges in terms of buffer control (see Sect. 1.7.3), it also poses an interesting new performance modeling question. As set out in the first stochastic analysis [24], the problem is new from both a mathematical and a technical point-of-view.

As a mathematical problem, the limitation on the number of delays was never treated in queueing theory (save the pioneering work of Lakatos [20, 21]), and it is clear that the problem cannot be transformed into a known queueing problem—this goes even for the problem's simplest instance, namely an optical M/M/1 queue.

As a performance evaluation problem, it will be shown in this work that it is impossible to optimize the optical buffer parameters, so as to be optimal for any given value of the traffic load. Further, the high sensitivity of buffer performance to the parameter setting forms a second complication, that cannot be ignored when one actually builds an optical switch. As such, the performance evaluation problem received considerable attention in recent years, with many basic optical buffer settings treated either through analysis or through simulation.

Finally, note that [20, 21] did provide the problem with an alternate setting, that is however completely equivalent, as argued in [5]. Lakatos studied a problem connected with the landing of airplanes, where arriving airplanes are obliged to wait for a discrete number of orbits of fixed length before landing. This system has

a degenerate waiting room, which he refers to as cyclic-waiting, while some more recent publications in this tradition [83] speak of a Lakatos-type queueing system. Mykhalevych [83] notes correctly that this system differs fundamentally from a retrial system [84]; a general explanation of the differences can be found in [83]. Note that no impatience was considered for the arriving planes, so the equivalence between the FDL buffering system and the Lakatos-type queueing system is only complete for the case of infinite buffer size.

1.7.3 Control Algorithms

At this point, we highlight some aspects of optical buffering related not to the structure, but to the way in which the buffer is controlled. Most relevant in the context of this doctoral dissertation are the delay-line assignment algorithm (for both single- and multi-wavelength systems) and the wavelength assignment algorithm, that together constitute the contention resolution scheme.

Delay-Line Assignment Algorithm For a single-wavelength system, a delay-line assignment algorithm determines how bursts are scheduled, so as to avoid contention. In this work, we impose a First-In-First-Out (FIFO) scheduling, which corresponds unambiguously to the following delay-line assignment algorithm.

1. Buffer control keeps track of the scheduling horizon, which is the earliest time at which all previous bursts will have left the system.

2. Whenever a new burst arrives, the burst is sent to the smallest delay line that can provide a large enough delay, so as not to overlap with previous bursts.

This delay-line assignment algorithm is consistently assumed throughout this work. Nevertheless, other algorithms exist.

- In [15], Walraevens and his co-authors investigate a “shortest FDL line first” delay-line assignment algorithm, where contention is only partly resolved, leading to an algorithm of minimal complexity. This algorithm does not need to keep track of any system state, but has lowered performance when compared to FIFO scheduling.

- The delay-line assignment algorithm in [14] allows to accommodate extra bursts, by keeping track of several independent queues within one FDL buffer. This algorithm allows to obtain better buffer utilization for a given number of fibers, but is somewhat more complex than FIFO scheduling, and results in increased waiting times.

- In [72], an advanced delay-line assignment algorithm is proposed, that keeps track of the departure instances and burst sizes of all bursts present in the

buffer. Since there is always a finite time between the end of the transmission of a burst and the start of the transmission of the next burst (namely the voids, mentioned in Sect. 1.7.2), this scheduling algorithm aims to fill up these voids, by letting newly arrived bursts be transmitted (if possible) during such void. This method is referred to as void-filling.

Wavelength Assignment Algorithm For a multi-wavelength system, similar principles apply, but with the difference that now, next to delay-line assignment, also the wavelength assignment algorithm plays its part.

For clarity's sake, let us first characterize a situation with contention. When buffer control detects that a number of bursts, say r ($0 \leq r$), need to be switched to the same output port at the same time, contention does not necessarily arise. Some of the wavelengths, say t ($0 \leq t \leq c$), are available immediately, while the other ones ($c - t$ in number) are already reserved by other bursts (some of them being transmitted, others waiting their turn in the buffer). As such, no bursts have to be buffered if $r \leq t$, and all can be transmitted instantaneously, on separate wavelengths. When $r > t$, t bursts can be transmitted directly, and the others ($r - t$ in number) are buffered. With c different wavelengths available to queue for, a wavelength assignment algorithm has to be adopted. In this work (Sect. 2.5 and 5.2), the following three algorithms are considered.

- Join-The-Shortest-Queue (JSQ): Here, in all circumstances, bursts join the shortest of the c queues. Of the three, this algorithm is known to have best performance in terms of loss and delay, especially in case of high traffic load. Main drawback of this algorithm is that it has higher implementation complexity than the other two algorithms.

- Round-Robin (RR): Whenever a burst is sent to queue i, the next one is sent to queue $i + 1$, and so on until queue c, that is followed by queue 1 again. Performance is less than for JSQ, but is sometimes comparable, as will be shown below. Further, hardware implementation complexity is low.

- Random (RND): Each burst is sent to one of the channels in random order. Performance is worst in this case, because the load is not spread intelligently among different queues, and one queue can have an overflow, while another is almost empty. As such, performance is equal to that of c separate single-wavelength buffers, that receive a fraction $1/c$ of the total load. From the implementation point-of-view, this performance weakening occurs when bursts are passed on to the queues in a more or less random order.

Further, note that even with JSQ, this contention resolution scheme is still rather simple, since the delay-line assignment within a wavelength is controlled by keeping track of only the scheduling horizon (just like in the single-wavelength

case). A more advanced method is to keep track of all departure instances and burst sizes on all wavelengths, and apply a void-filling algorithm that first chooses what void to fill, to then assign both the wavelength and the delay line at once. Such algorithms are studied in [72].

1.7.4 Optical Buffers other than FDL Buffers

While the feed-forward FDL buffer we study is probably the most often-assumed type of optical buffer [42], some alternatives have been proposed over the years, mainly with the scope of reducing footprint (physical size) and component count. This being a monograph on optical buffering, we cannot ignore the existence of several optical buffers that do not consist of a set of fiber delay lines.

Miniaturization The main motivation to search for alternatives of FDL-based buffers is miniaturization. An FDL-based buffer, although perfectly feasible with off-the-shelf components, and well-performing in terms of capacity, is indeed large in terms of physical size. For typical OBS specifications (10 Gbit/s link, 100 kbit burst sizes), one needs approximately 2 km of fiber to delay the light for 10 μs, that is, for the duration of a burst. To limit the total footprint, an optical buffer typically consists of only a small number of FDLs, which implies that loss is seldom negligible. Also, the high component count (amounting to the number of fiber delay lines) makes it somewhat impractical for mass production, and motivates the search for a design with either (i) a reduced number of fibers, (ii) reduced length for each fiber, or (iii) a combination of both advantages. As such, the alternative approaches have in common that they aim for miniaturization. Main challenges in the miniaturization however are the high cost, and the difficulty to obtain sufficient capacity for broadband applications.

Recirculation Loop The most viable alternative to the feed-forward buffer we consider is probably the recirculation loop, as studied in for example [36,37,85]. It can be understood in principle as the buffer displayed in Fig. 1.2, with the feedback line (5) enabled. However, the most common implementation, as for example proposed in the LASOR project [57], has the FDL set cut down to a single fiber, placed in a loop. Usually, only a single burst with length shorter than the loop length can enter, and leaves the loop after an integer multiple of recirculations. The entire optical buffer is then a limited set of such fiber loops (typically about ten), able to accustom the same number of bursts.

The stochastic performance of such buffers is somewhat better than that of feed-forward FDL buffers with the same number of lines. This is so, since bursts can exit the loop after every recirculation, whereas bursts in (a long line of) a feed-forward structure are bound to travel through the entire line before being able to

exit. Main drawbacks are however the increased cost of the switching element, and the difficulty to scale such buffers beyond their typically small size.

Slow-Light Devices Although still in their research phase, slow-light devices also provide future candidates to replace FDL buffers in the long run [85]. While FDL-based buffers delay the light by increasing the length of the path the light goes through, slow-light devices delay the light by decreasing its group velocity. Slow-light buffers can be categorized in two types: devices using material-based resonances, and those using coupled resonant structures (CRS). Devices of the first group exploit the resonance between an electromagnetic field and a polarizable medium, either based on electromagnetically-induced transparency or on population oscillation by means of a carrier population grating. The second group, with CRS, uses gratings or photonic crystal defects to lengthen the total lightpath through repeated reflections. More detail can be found in [86, 87], as for now it is important to note that these methods are very expensive when compared with FDL-based buffering, especially when one takes into account the very small capacities, in the order of hundreds to thousands of bits, amounting to typically less than a single burst.

Optical Flip-Flops Another product of recent research is the optical flip-flop. Mentioning just one of many possible designs, the optical flip-flop presented in [88] is based on erbium-ytterbium-doped fiber absorption. Exploiting the effects of fluorescence allows to obtain an optical bistable device, that again has only very small capacity (much smaller than the typical size of a burst), but may provide a competitive rival to the FDL-based buffer setting in the long run.

Small Buffers? Further, note that in the search for possible alternative designs, some performance loss (when compared to that of a feed-forward FDL-based design) is considered acceptable. Given that the buffering performance of a feed-forward FDL buffer is already less than that of a classic buffer, this is far from self-evident. While in [24], it is still argued that optical buffers can achieve very low loss rates if the buffer size is large enough, this argumentation is no longer found in more recent publications. Remarkably, the buffer size is treated in [85] as a secondary requirement, arguing that if access links are slower than the backbone network and the traffic is smoothed, then a capacity of ten packets (or bursts) per output port is needed for 80% throughput, as proposed in [89], propagating an idea that was formulated also in [90], and earlier in [91]. There, it is put forward that although current core routers are provisioned with very large buffer memory (state-of-the-art core routers incorporate line cards operating with more than one GB of RAM), good performance can be obtained with smaller buffer sizes.

As such, whether or not small buffers suffice for broadband applications is a topic of debate, with no definite answer to date. At any rate, of all possible implementations of optical buffers, the one we assume in this work is not only the most feasible for implementation, but also the single one that is not compromised by this debate. Feed-forward FDL buffers can indeed come up to the generally accepted bandwidth requirements in the backbone, and even provide not-so-small-sized buffers. This said, in the rest of this work, the term "optical buffer" is used interchangeably with "FDL buffer".

1.8 This Doctoral Dissertation

1.8.1 Time Setting

In this doctoral dissertation, optical buffer performance will be studied both in discrete-time (DT) and continuous-time (CT) setting. Rather than developing notation for both time settings independently, we chose to denote the DT random variable (rv) and its CT counterpart by the same name, and indicate clearly what time setting is intended. Since the CT stochastic description of optical buffers has a lot in common with that of DT, at some points in Chapter 2, it is possible to present relations that are valid for both time settings at once. Such property will be indicated by placing an asterisk above the relation symbol, for example, "$\overset{*}{=}$" instead of "$=$".

1.8.2 Extended Kendall Notation

As mentioned, the main aim of this doctoral dissertation is to provide an answer to question (1.1). This question is posed for quite different system settings, and as such, it is useful to delimit the set of attributes, needed to uniquely define the systems described in this work. It comes out that we can use the well-known Kendall notation, supplemented with three additional attributes.

Firstly, all systems studied share the assumption on inter-arrival and burst size distribution, namely that both are independent and identically-distributed (iid). Also, the number of wavelengths in an optical buffer system provides a good parallel with the number of servers in a classic queueing system. As such, in the case of CT, the original threefold Kendall notation [92] (of the form A/B/C) provides a good basis to denote the exact assumptions on inter-arrival times (A), burst sizes (B) and the number of wavelengths (C).

As for time setting, in case of DT, a practical modification of the Kendall notation is provided by Bruneel and Kim, in [93]. However, given that we define the arrival process through the inter-arrival times, rather than through the number of arrivals per slot, this notation is not convenient in this work. Therefore, we prefer to add the time setting (DT/CT) explicitly as a first attribute.

extended Kendall notation				section
degenerate	infinite-sized	M/G/1	DT	2.2.3
”	”	”	CT	2.2.4, 2.2.5
”	”	M/M/1	CT	2.2.8.1
”	”	M/D/1	CT	2.2.8.2
”	”	M^X/G/1	DT	2.3
”	”	GI/G/1	DT	2.4
”	”	M/G/c	DT	2.4
non-degenerate	finite-sized	M/G/1	DT	3.1.4
”	”	”	CT	3.1.6
”	”	GI/G/1	DT	3.2
”	”	M/D/1	DT	3.3.3.2
”	”	”	CT	3.3.3.3
degenerate	”	M/M/1	DT	3.4.2.3, 4.5
”	”	”	CT	3.4.2.4, 4.4
non-degenerate	infinite-sized	GI/G/1	CT	5.1
”	”	GI/G/c	CT	5.2

Table 1.1: Systems studied in this doctoral dissertation, denoted with extended Kendall notation, with reference to the relevant section.

For the second attribute, we are to mention the number of FDLs included in the buffer. The well-known fourfold extension of the Kendall notation (of the form GI/G/1/K) is not fit for this purpose, since “K” refers to a limited number of waiting places, whereas we want to express a system with impatience (see Sect. 1.7.2). Therefore, we explicitly mention whether the buffer is finite-sized or infinite-sized.

As third attribute, we indicate whether the FDL set is degenerate or not (see Sect. 1.7.1). Note that the non-degenerate case is the most general case, that encompasses all degenerate settings. As such, in the context of this work, it suffices to use the classic Kendall notation, and add three attributes, namely CT or DT, finite-sized or infinite-sized, and degenerate or non-degenerate, to describe the buffer system intended.

Finally, note that, in a more general context, the single-wavelength case should have an extra attribute to specify the delay-line assignment algorithm (see Sect. 1.7.3), but that this is unnecessary in this work, since it is always FIFO. In the multi-wavelength case, two attributes should be added: the delay-line and the wavelength assignment algorithm. In this work, FIFO is combined with RND, RR, and JSQ.

In the style of Table 1 in [92], we also include here a similar table, namely Table 1.1, that displays the main systems studied here, and the relevant sections. (Some special cases, like a mixed geometric distribution for burst sizes or inter-arrival times, are not explicitly mentioned in the table, but are also treated in this

work.) Hereby, note that M^X denotes a memoryless arrival process with batch arrivals.

1.8.3 Four Main Approaches

As for the structure of this work, the chapters are complementary, in that each represents a different methodological approach: with transform functions (Chapter 2), with Markov Chains (Chapter 3), with impatience (Chapter 4) and with stochastic regenerations (Chapter 5).

The approach with transform functions of Chapter 2 presents results that are obtained in terms of transform functions of the involved random variables (rv's). Here, "transform functions" is an umbrella term that encompasses probability generating functions (pgf's) in case of a DT setting, and Laplace-Stieltjes transforms (lst's) in case of a CT setting. This approach capitalizes on techniques developed for classic (non-optical) buffers, and allows to fit the analysis within the general framework of queueing theory. In both time settings, the FDL lengths are assumed multiples of the granularity D (degenerate buffer setting). The obtained expressions are exact for infinite buffer size, and can be converted into approximative results for finite buffer size by means of a heuristic. For CT, the case of a Poisson arrival process and general iid burst size distribution is studied, for a single output. For DT, we consider the analysis for a Bernoulli arrival process with batch arrivals, for general iid batch and burst size distribution and a single output wavelength. This model is then applied to study the impact of (local) synchronization on buffer performance. Further, also a DT model for general iid inter-arrival times and burst sizes is developed, for a single output wavelength. This is then applied to study the impact of (a mild form of) correlation in the arrival process on buffer performance. Secondly, it is also applied to study the performance of an optical buffer with wavelength conversion.

The second approach for performance evaluation is with Markov chains (Chapter 3). It focuses primarily on the transitions of the imbedded Markov chain, and less on the actual way in which arriving bursts join the queue. The merit hereof is that it allows for a very concise and exact description of optical buffers for finite sizes and general line lengths (not necessarily multiples of D). It yields a numerical method to obtain the exact waiting time probabilities and loss probability, for a memoryless arrival process and general iid burst size distribution, both in DT and CT, and for general iid inter-arrival time and burst size distribution, in DT. Also, results are examined further, to obtain exact closed-form expressions independent of time setting, for a degenerate buffer setting with either bounded burst size, or memoryless burst size distribution.

The approach with impatience (Chapter 4) highlights the similarities between optical buffers and queues with impatience, and exploits this resemblance to show

that exact results for queues with deterministic impatience can lead to good approximations for optical buffer performance. This is verified explicitly in CT, for a Poisson arrival process and negative-exponential burst size distribution.

The fifth chapter sheds light on the stability of infinite-sized optical buffer systems with general FDL line lengths. A general characterization based on stochastic regeneration allows to establish sufficient conditions for the stability of single-wavelength optical buffers in CT, without making any specific assumption on the FDL line lengths, or the distribution of inter-arrival times and burst sizes. This result is then extended to account for the case of multiple wavelengths.

Finally, Chapter 6 presents conclusions, and includes an outlook on possible future work.

1.8.4 List of Publications

The following publications, save two, have provided the material for this doctoral dissertation. The bibliographic data is followed by a reference to the relevant chapter. Note that the contents of Chapter 4 and Sect. 5.2 has not been published nor submitted at the time of writing.

International Journals (A1)

1. W. Rogiest, K. Laevens, D. Fiems and H. Bruneel, A Performance Model for an Asynchronous Optical Buffer, Performance Evaluation 62 Vol. 1-4 (2005), pp. 313–330. [*identical to conference proceedings publication 3*] [→ Chapter 2]

2. W. Rogiest, K. Laevens, J. Walraevens and H. Bruneel, Analyzing a Degenerate Buffer with General Inter-Arrival and Service Times in Discrete Time, Queueing Systems 56 Vol. 3-4 (2007), pp. 203–212. [→ Chapter 2]

3. W. Rogiest, K. Laevens, D. Fiems and H. Bruneel, Modeling the Performance of FDL Buffers with Wavelength Conversion, Submitted to IEEE Transactions on Communications, *currently under revision*. [→ Chapter 2]

4. W. Rogiest, J. Lambert, D. Fiems, B. Van Houdt, H. Bruneel and C. Blondia, A Unified Model for Synchronous and Asynchronous FDL Buffers allowing Closed-Form Solution, Submitted to Performance Evaluation, *currently under revision*. [→ Chapter 3]

International Journals (not A1)

1. W. Rogiest, D. Fiems, K. Laevens and H. Bruneel, Tracing an Optical Buffer's Performance: an Effective Approach, Lecture Notes in Computer Sci-

ence (LNCS), Vol. 4465 (2007), pp. 185–194. [*identical to conference proceedings publication 6*] [→ Chapter 3]

2. W. Rogiest, D. Fiems, K. Laevens and H. Bruneel; A Light-Weight Performance Model for Optical Buffers, International Journal of Communications Networks and Distributed Systems (IJCNDS), Vol. 1(3) (2008), pp. 282–295. [→ Chapter 3]

3. J. Lambert, W. Rogiest, B. Van Houdt, D. Fiems, C. Blondia and H. Bruneel, A Hessenberg Markov Chain for Fast Fibre Delay Line Length Optimization, Lecture Notes in Computer Science (LNCS), Vol. 5055 (2008), pp. 101–113. [*identical to conference proceedings publication 8*]

Conference Proceedings (full paper)

1. W. Rogiest, K. Laevens, D. Fiems and H. Bruneel, Analysis of an Asynchronous Optical Buffer, Proceedings of the COST279 Twelfth Management Committee Meeting, COST279TD(04)034, Antalya, February 2005. [→ Chapter 2]

2. W. Rogiest, K. Laevens, D. Fiems and H. Bruneel, Analysis of an Asynchronous Single-Wavelength FDL Buffer, Proceedings of the Nineteenth International Teletraffic Congress (ITC 19), pp. 1917–1926, Beijing, August 2005. [→ Chapter 2]

3. W. Rogiest, K. Laevens, D. Fiems and H. Bruneel, A Performance Model for an Asynchronous Optical Buffer, Proceedings of the IFIP WG 7.3 International Symposium on Computer Performance, Modeling, Measurements, and Evaluation (PERFORMANCE 2005), Juan-Les-Pins, October 2005. [*identical to A1-journal publication 1*] [→ Chapter 2]

4. W. Rogiest, K. Laevens, D. Fiems and H. Bruneel; Optical Buffers, Batch Arrivals and Synchronization; Proceedings of the second Conference on Next Generation Internet Design and Engineering (NGI 2006), pp. 176–183, Valencia, April 2006. [→ Chapter 2]

5. W. Rogiest, K. Laevens, D. Fiems and H. Bruneel, Quantifying the Impact of Wavelength Conversion on the Performance of Fiber Delay Line Buffers, Proceedings of the Sixth International Workshop on Optical Burst/Packet Switching (WOBS 2006 West), paper 131, pp. 1–10, San Jose, October 2006. [→ Chapter 2]

6. W. Rogiest, D. Fiems, K. Laevens and H. Bruneel, Tracing an Optical Buffer's Performance: an Effective Approach, Proceedings of the First Euro-

FGI International Conference on Network Control and Optimization (NET-COOP 2007), Avignon, June 2007. [*identical to not-A1 journal publication 1*] [→ Chapter 3]

7. W. Rogiest, D. Fiems, K. Laevens and H. Bruneel, Exact Performance Analysis of FDL Buffers with Correlated Arrivals, Proceedings of the Fourth IEEE and IFIP International Conference on Wireless and Optical Communications Networks (WOCN 2007), Singapore, July 2007. [→ Chapter 3]

8. J. Lambert, W. Rogiest, B. Van Houdt, D. Fiems, C. Blondia and H. Bruneel, A Hessenberg Markov Chain for Fast Fibre Delay Line Length Optimization, Proceedings of the 15th International Conference on Analytical and Stochastic Modelling Techniques and Applications (ASMTA'08), Cyprus, June 2008. [*identical to not-A1 journal publication 3*]

Conference Proceedings (abstract)

1. W. Rogiest, K. Laevens, D. Fiems and H. Bruneel, Analysis and performance evaluation of optical buffers, Book of Abstracts sixth FirW PhD Symposium Ghent University, p. 104, Ghent, November 2005. [→ Chapter 2]

2. W. Rogiest, K. Laevens, D. Fiems and H. Bruneel, Analysis of a Lakatos-type queueing system with general service times, the Twentieth Conference on Quantitative Methods for Decision Making, ORBEL 20, pp. 95–97, Ghent, January 2006. [→ Chapter 2]

3. J. Walraevens, B. Van Houdt, J. Lambert, W. Rogiest, D. Fiems, V. Inghelbrecht, C. Blondia, H. Bruneel, Contention resolution in next-generation optical node architectures, Book of Abstracts of the COST 291 / GBOU ONNA Workshop on Design of Next Generation Optical Networks: from the Physical up to the Network Level Perspective, p. 30, Ghent, February 2006. [→ Chapter 2]

4. W. Rogiest, K. Laevens, D. Fiems and H. Bruneel, Modeling the Performance of a Multi-Wavelength Optical Buffer: a Round-Robin Approach, Proceedings of the Third EuroNGI Workshop on New Trends in Modelling, Quantitative Methods and Measurements (WP IA.8.1), pp. 1–4 Torino, June 2006. [→ Chapter 2]

5. W. Rogiest, K. Laevens, J. Walraevens and H. Bruneel, Analyzing a Degenerate Buffer with General Inter-Arrival and Service Times in Discrete Time, Book of Abstracts of the Second Madrid Conference on Queueing Theory (MCQT '06), pp. 49–50, Madrid, July 2006. [→ Chapter 2]

6. W. Rogiest, D. Fiems, K. Laevens and H. Bruneel, An effective modeling for the performance analysis of optical buffers, Book of Abstracts of the Second Benelux Workshop on Performance Analysis of Communication Systems, p. 22, Antwerp, February 2007. [→ Chapter 3]

7. W. Rogiest, D. Fiems, K. Laevens and H. Bruneel, Capturing an optical buffer's performance with a Markov chain, Proceedings of the Euro-FGI Workshop on New Trends in Modelling, Quantitative Methods and Measurements, pp. 90–93, Ghent, June 2007. [→ Chapter 3]

8. W. Rogiest, K. Laevens, D. Fiems and H. Bruneel, Buffering Light: Achilles' Heel of the Next-Generation Backbone Network, Book of Abstracts of the Eighth FirW PhD Symposium Ghent University, p. 23, Ghent, December 2007. [→ Chapter 2 and 3]

Work in Progress

1. W. Rogiest and H. Bruneel, Optical Buffers for Variable Length Packets: Closed-Form Expressions, *submitted to IEEE Communications Letters*. [→ Chapter 3]

2. W. Rogiest, K. Laevens, D. Fiems and H. Bruneel, The Impact of Synchronization on the Performance of FDL Buffers, *submitted to International Journal of Electronics and Communications (AEÜ)*. [→ Chapter 2]

3. W. Rogiest, E. Morozov, D. Fiems, K. Laevens and H. Bruneel, Stability of Single-Wavelength Optical Buffers, *submitted to European Transactions on Telecommunications (ETT)*. [→ Chapter 5]

Awards

1. W. Rogiest, Best Student Paper Award at the Nineteenth International Teletraffic Congress (ITC 19), Beijing, August 2005.

2

Performance Evaluation with Transform Functions

¶ In this chapter, we present a variety of performance models for optical buffers, that were all developed by relying on transform functions throughout the analysis. Here, "transform functions" is an umbrella term that encompasses probability generating functions (pgf's) for a DT setting, and Laplace-Stieltjes transforms (lst's) for a CT setting. This approach we apply to various FDL buffer settings, that all share the property that they are infinite-sized and degenerate.

Pioneering in the stochastic modeling of FDL buffers are [20], in the more general queueing context of cyclic-waiting, and [24], in the specific context of optical networks. In recent years, the number of contributions providing stochastic models of optical buffers has grown impressively [1–37], of which [24] probably still remains the most often cited. Rather than discussing all of these contributions here subsequently, we refer to some of them in more detail only later on, in the chapters where they are most relevant.

While not in terms of transform functions, the analysis provided in [24] also relies on results derived for infinite system size, to then derive results for the finite system. There, and in [25, 26], Callegati provides an approximate analysis of a degenerate M/M/1 buffer in CT. Making use of general notions of queueing theory, he constructs an iterative procedure to provide results for both finite and infinite buffer size. This approach is generalized in [27], where a similar iterative procedure is provided for the multi-wavelength case, for JSQ wavelength assignment.

In [10], Laevens and Bruneel provide the first analysis of FDL buffers in terms

of transform functions. They analyze an infinite-sized degenerate M/G/1 FDL buffer in DT with pgf's, and rely on a heuristic to obtain approximate results for finite buffer size. The contribution of [10] in the context of this chapter is paramount, in that it was this paper that showed that (i) the impact of voids can be traced in an exact manner by means of pgf's, (ii) the effect of impatience can be approximated well by means of a heuristic, and that (iii) the analysis can be done in a mathematically elegant form. While [10] shows that these arguments hold in case of Bernoulli arrivals, for an M/G/1 buffer in DT, this chapter presents the extension of this approach to the more general cases M^X/G/1, also considered in [3], GI/G/1 [6] and M/G/c [4]. Further, the extension toward CT, treated earlier in [2], is also investigated in this chapter, showing that an M/G/1 in CT can still be analyzed, but with somewhat more complex expressions. (The main complication in the CT case is that the solution for general burst size distribution contains an infinite sum.)

Independent of [10] or [2], [32] provided an analysis of a degenerate infinite-sized M/M/1 queue in CT also in terms of transform functions. More precisely, Hong and his co-authors provide an expression for the lst of the modified burst size (there called modified service time), a rv introduced to capture the performance loss due to voids, defined as the sum of the burst size and the void size. The expression for the lst takes into account the correlation between the void sizes and the burst sizes, but makes several simplifying assumptions. Firstly, the correlation between the void sizes and the inter-arrival times is ignored. Secondly, it is not taken into account that, when a burst arrives in an empty system, the void size equals zero. This allows to consider a classic (non-optical) M/G/1 buffer system, with modified burst sizes, as an approximation of the original system. Since the approach of [2] (presented below in Sect. 2.2) is exact for infinite buffer size, it comes as no surprise that the expressions for the lst's in [32] do not correspond to the ones we present in Sect. 2.2.8.1, even though they are framed for exactly the same system setting.

Further, in [18], the topic of optical buffers is treated from a different perspective, as a queue with correlated service and inter-arrival times. After the derivation of general results for such type of queues, a separate section is devoted to the application of this framework to optical buffers with fixed-length bursts. Indeed, focusing on the evolution of the scheduling horizon, an optical buffer can be modeled as a classic buffer model, but with an augmented burst size for bursts that find the system non-empty upon arrival. As such, the model of [18] allows to study a degenerate infinite-sized SM/D/1 in DT (using the notation of Sect. 1.8.2, with SM standing for a discrete-time semi-Markov arrival process). While these assumptions are not more general than for example the degenerate infinite-sized GI/G/1 in DT of [6], [18] currently provides the only analysis in FDL buffering literature that allows to trace the queue length distribution, that is, the distribution

of the number of bursts awaiting transmission in the buffer. Notice that the latter cannot be obtained directly from the steady state distribution of the scheduling horizon, and the approach of this doctoral dissertation would require significant modification to obtain the complete distribution of the queue length[1].

At this point, we refer to Chapter 1 and to the introductory part of Chapter 3-5 for further references, and now move to the results we obtained with transform functions. In the following, we first supply the reader with general notions for the infinite buffer system in Sect. 2.1. Complementary to this, Sect. 2.1.6 presents a general method, that allows to obtain approximate results for finite buffer size from exact results obtained for infinite buffer size. From then on, specific instances of this general framework are considered: Sect. 2.2 highlights the comparison of synchronous and asynchronous single-wavelength optical buffers for memoryless inter-arrival time distribution, while Sect. 2.3 focuses on the impact of synchronization in this context, by means of a batch arrival process. Sect. 2.4 considers general instead of memoryless inter-arrival times. The latter not only allows to study the impact of correlated arrivals in that section, but also of wavelength conversion, in Sect. 2.5.

2.1 Stochastic Model

In this section, the general modeling assumptions of this chapter are introduced, including the buffer setting itself, and the rv's involved in the queueing process. For the single-wavelength system, two alternative system equations are given, that are valid regardless of the specific distribution of the involved rv's, and regardless of time setting. Also, a general approach to the question of stability is introduced, that is again independent of the time setting. As such, this section provides the starting point for the following sections, that consider specific instances of this framework in detail.

2.1.1 Time Setting

In most of this chapter, the model we construct has a DT setting. Events are assumed to take place synchronously, at the beginning of time slots, which is frequently proposed in the context of optical switching. This implies that all time-related variables and measures, including the granularity D, are expressed as multiples of the slot size, that may be arbitrary, and is therefore not mentioned explicitly in this work. In the unslotted case, inter-arrival times and burst sizes may also be continuous (asynchronous). In general it is so that one can approximate the CT

[1] Although Little's theorem does allow to extract the mean queue length from the mean sojourn time, one is to add the buffer occupancy to the system description explicitly in order to obtain the entire queue length distribution. Since tracing either the scheduling horizon or the waiting times equally allows to extract the most important performance measures, we do not elaborate on this in the current work.

case arbitrarily closely if one takes a very small slot size and scales the involved variables and measures accordingly. Such is also done further on, when we treat a basic asynchronous buffer setting with a pgf approach.

2.1.2 Buffer Setting

In this chapter, we initially assume an infinite-sized buffer ($N = \infty$), and only focus on the finite-sized buffer setting in a next step. Utilizing the terms and setting introduced in Chapter 1, the buffer for which we obtain performance results is degenerate and of size N. It is thus a set of $N + 1$ delay lines, with lengths expressed as multiples of the granularity D ($0 \times D$, $1 \times D \ldots N \times D$) and is part of a network node, located at an output port that gives access to a single outgoing fiber. Except in Sect. 2.5, we assume that a single wavelength on the outgoing fiber is accessible. In queueing parlance, this corresponds to a single-server system. Further, recall that the buffer is intended for contention resolution. Whenever a contending burst arrives in a non-empty buffer, it is queued at least until all previous bursts have left the system, and in FIFO order. This amounts to a specific one delay-line assignment algorithm, reflected in the evolution of the scheduling horizon explained next.

2.1.3 Evolution of the Scheduling Horizon

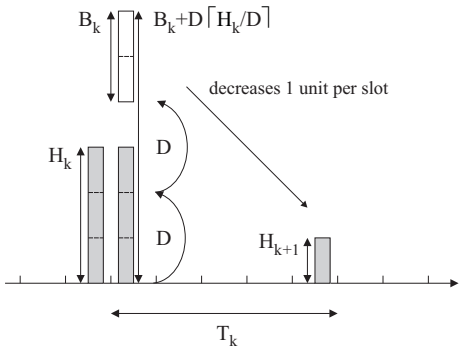


Figure 2.1: The FIFO delay-line assignment in optical buffers corresponds to a scheduling-horizon process. The figure illustrates this for a DT setting, with $D = 2$ slots. A burst of size $B_k = 2$ slots has to wait for at least $H_k = 3$ slots, so as to avoid contention with previously arrived bursts. Since the FDL buffer can only realize delays that are multiples of $D = 2$ slots, burst k has to wait for $D\lceil H_k/D \rceil = 4$ slots.

At this point, we focus on how arriving bursts join the queue of a single-wavelength degenerate buffer, so as to realize a FIFO delay-line assignment algorithm. While this happens independently of time setting, the relation between

the involved rv's is illustrated in Fig. 2.1 for a DT setting, with $D = 2$ slots. Numbering bursts in the order of their arrival, the inter-arrival time T_k (5 slots in Fig. 2.1) captures the time between the kth arrival instant and the next. The burst size B_k is the time it takes to transmit the kth burst. Upon arrival, burst k has to wait for at least an amount H_k, the time needed for all previously arrived bursts to be served. This amount we call the scheduling horizon H_k, as observed by the kth burst upon arrival. Due to the buffer's degeneration, the burst has to wait for a waiting time W_k, that is a multiple of D and is sufficiently long, that is, $W_k \geq H_k$. The evolution of these variables can be captured in one system equation, namely

$$H_{k+1} \stackrel{*}{=} \left[D \cdot \left\lceil \frac{H_k}{D} \right\rceil + B_k - T_k \right]^+ . \tag{2.1}$$

Here, the expression $\lceil x \rceil$ is the ceiling function, equal to the smallest integer greater than or equal to x, while the notation $[x]^+$ is shorthand for max(x, 0). Further, recall that $\stackrel{*}{=}$ denotes that the equation is valid for both time settings, DT and CT. This equation contains two non-linearities. Introducing some additional notation, we can split this equation into two parts. The first non-linearity is

$$H_{k+1} \stackrel{*}{=} [G_k - T_k]^+ , \tag{2.2}$$

with

$$G_k \stackrel{*}{=} B_k + W_k , \tag{2.3}$$

while the second one is

$$W_k \stackrel{*}{=} D \cdot \left\lceil \frac{H_k}{D} \right\rceil . \tag{2.4}$$

In the following, we refer to the non-linearity in (2.2) as the queueing effect, since it occurs frequently in classic queueing analysis. The non-linearity in (2.4) we refer to as the granularity effect, since it reflects the degenerate structure of the FDL buffer.

As mentioned, the buffer's degenerate structure typically results in additional waiting, since $W_k \geq H_k$. This additional waiting results in periods during which the outgoing channel remains unused, despite of bursts still present in the buffer. As mentioned earlier in Sect. 1.7.2, these periods are called voids. Their duration or size, denoted as V_k, relates to W_k and H_k as

$$V_k \stackrel{*}{=} W_k - H_k . \tag{2.5}$$

As said, the occurrence of voids comes at a price, and results in an essential performance gap when compared to classic buffers, for which no void comes about. In this regard, the performance optimization of optical buffers often comes down to minimizing the average void size.

Further, regardless of time setting, we will always assume in this work that the burst sizes B_k and inter-arrival times T_k form two sequences of iid rv's. Furthermore, burst sizes and inter-arrival times are mutually independent.

2.1.4 Evolution of the Waiting Time

Although the system description can be done perfectly well in terms of the scheduling horizon, with (2.2) and (2.4), a description based on the waiting time sometimes proves useful, especially when considering the maximum tolerable load in Sect. 2.2.6. Although not self-evident, the evolution of the waiting times can be captured without considering the scheduling horizon explicitly, and follows directly from the combination of (2.1) and (2.4). Irrespective of time setting,

$$W_{k+1} \stackrel{*}{=} \left[W_k + D \cdot \left\lceil \frac{B_k - T_k}{D} \right\rceil \right]^+ , \tag{2.6}$$

provides a description of the system in its own right, that proves useful in many cases [2]. Note, however, that it only identifies the voids in an implicit manner, and that the void size is still defined in the easiest way through (2.5).

2.1.5 Maximum and Equivalent Load

Maximum Load As mentioned in Sect. 2.1.2, we assume infinite buffer size, and obtain results first for the infinite system. Due to this, stability is not guaranteed, since the waiting times and scheduling horizon can grow unboundedly if on average too much work arrives. The amount of work is traditionally defined through the traffic load ρ, defined as the average amount of arriving work (in the form of bursts) over the average inter-arrival time,

$$\rho \stackrel{*}{=} \frac{\mathrm{E}[B_k]}{\mathrm{E}[T_k]} . \tag{2.7}$$

which is independent of time setting (as indicated by "$*$" in "$\stackrel{*}{=}$") All derivations in the following will assume that the system is stable, which requires the offered load ρ to be below some maximum value ρ_{max},

$$\rho \stackrel{*}{<} \rho_{max} ,$$

that is typically less than unity, unlike conventional queues (where it is unity).

[2] This description also serves well if burst sizes B_k and inter-arrival times T_k were correlated with each other, but independent from arrival to arrival. The latter however goes beyond the scope of this work, and the B_k and T_k are assumed iid, and without correlation between them, as mentioned.

In this regard, it is instructive to compare the evolution of the waiting times, as given by (2.6), to that of a classic buffer system. Independent of time setting, the waiting times in a classic system evolve as

$$W_{k+1} \stackrel{*}{=} [W_k + B_k - T_k]^+ .$$

Clearly, in the classic system, the drift of the waiting time process is brought about by the difference of burst sizes and inter-arrival times, $B_k - T_k$. Demanding the system to be stable then corresponds to demanding *negative drift* (in the sense explained in Chapter 5), so as to assure that the probability that a non-empty system empties again is strictly larger than zero, and that this event takes place within finite time[3]. In terms of the involved rv's, the stability condition for the classic system is that the expected value of the drift is smaller than zero on average,

$$\mathrm{E}[B_k - T_k] \stackrel{*}{<} 0 ,$$

which obviously is just another instance of the well-known condition $\rho \stackrel{*}{<} 1$.

As for the optical system, the drift in the optical waiting time process (2.6) is generated by

$$D \left\lceil \frac{B_k - T_k}{D} \right\rceil ,$$

which is always larger than or equal to $B_k - T_k$, and relates to the occurrence of voids (but only implicitly, as mentioned in Sect. 2.1.4). As a result, also the net drift is larger than in the classic case, and the stability condition now is

$$\mathrm{E}\left[\left\lceil \frac{B_k - T_k}{D} \right\rceil \right] \stackrel{*}{<} 0 , \tag{2.8}$$

which yields a more restrictive condition than in the classic case, that is,

$$\rho \stackrel{*}{<} \rho_{max} \stackrel{*}{\leq} 1 .$$

Notice that the negative drift condition (2.8) is valid only for degenerate buffer structures. Its counterpart for non-degenerate buffer structures is (5.11), as discussed in Chapter 5.

Equivalent Load Given the gap between the maximum tolerable load of a classic system and that of an optical system, we can define an equivalent load, that incorporates the effect of the voids into an altered definition of the load. While an exact tracing of the voids is again only possible by involving the scheduling horizon, we rather extend the notion of *drift*, as mentioned in the previous section,

[3] A more detailed characterization of stability is given in Chapter 5.

since this yields a simple description, and performs well when applied in heuristics for the finite system (see below). More precisely, we define the equivalent load as

$$\rho_{eq} \stackrel{*}{=} 1 + \frac{\mathrm{E}\left[D \cdot \left\lceil \frac{B_k - T_k}{D} \right\rceil\right]}{\mathrm{E}[T_k]}\,, \tag{2.9}$$

which is very similar to the expression of the classic load (2.7), especially when written as

$$\rho \stackrel{*}{=} 1 + \frac{\mathrm{E}[B_k - T_k]}{\mathrm{E}[T_k]}\,.$$

Note however, that the equivalent load does not simply replace the classic load in general, and should not be applied in this manner. For example, while it is well-known for an M/G/1 classic system in CT that the probability of finding an empty system upon arrival equals $1-\rho$, this is not the case for an optical system: neither $(1-\rho)$ nor $(1-\rho_{eq})$ provides an answer, and the probability in question can only be determined after full system analysis. In this regard, the exact definition of ρ_{eq} is prone to discussion: $\Pr[H=0] = 1-\hat{\rho}_{eq}$ can indeed provide a alternative definition for an alternative equivalent load $\hat{\rho}_{eq}$, but leads to impractical expressions, whose form is highly sensitive to the assumptions on the inter-arrival time and burst size distributions. Therefore, such alternative definitions are not explored further, and we maintain (2.9) as definition of the equivalent load throughout this work. As such, the notion of an equivalent load is a useful concept to capture stability and yields good results in the application of heuristics, but does not provide a solid basis for the analysis of the entire system[4]. On the other hand, this provides another illustration of the fact that an optical buffer model differs fundamentally from a classic buffer model, and calls for an analysis in its own right.

Specific instances of the equivalent load are considered in Sect. 2.2.6, and later on also for more general assumptions, in Sect. 2.3.5 and Sect. 2.4.4.

2.1.6 Heuristics for the Loss Probability

In this chapter, the approach will always be to obtain exact results in the transform domain (in terms of either pgf's (DT) or lst's (CT)) first, for a system of infinite size. Only in a second step, a heuristic is applied, and provides us with approximate expressions to calculate the (burst) loss probability (LP), given the buffer size N. Since the same heuristics are applied consistently throughout this chapter, we prefer to present the general approach here, and will refer to it when specific instances are considered. Given that the heuristics capitalize on a dominant pole approximation of tail probabilities, we first specify the method with which we obtain the involved tail probabilities in Sect. 2.1.6.1, to then treat the heuristic itself in Sect. 2.1.6.2.

[4]This is opposed to the situation of for example a classic M/M/1 system in CT, that can be analyzed almost entirely in terms of only the traffic load.

2.1.6.1 Dominant Pole Approximation

For the dominant pole approximation, we first consider the DT case. We assume the system stable, and assume that the analysis yielded an exact expression for the pgf $H(z)$ of the steady-state distribution associated with the scheduling horizon H. The scope of this section is to show how tail probabilities can be extracted from this pgf. To allow for such approximation, we assume that the burst size and inter-arrival time distributions do not possess a heavy tail, and that the associated pgf's have a single dominant pole.

The starting point is the relation

$$\Pr[H > n] = Res\left[\frac{1}{z^{n+1}} \cdot \frac{H(z)-1}{z-1}\right]_{z=0}, \tag{2.10}$$

that implies the calculation of the residue at $z = 0$. If we assume that $H(z)$ has no singularities other than isolated poles, (2.10) becomes

$$\Pr[H > n] = -\sum_{l} Res\left[\frac{1}{z^{n+1}} \cdot \frac{H(z)-1}{z-1}\right]_{z=z_l},$$

where the summation index l runs over all poles z_l of $H(z)$. This relation can now be approximated by keeping only the dominant poles, namely those poles with smallest modulus. As such, a good approximation is

$$\Pr[H > n] \approx -\sum_{k} Res\left[\frac{1}{z^{n+1}} \cdot \frac{H(z)-1}{z-1}\right]_{z=z_k}, \tag{2.11}$$

where the summation index k only runs over the poles z_k of $H(z)$ with smallest modulus. The most common case is that such a dominant pole is unique. However, for the current analysis of a degenerate buffer with granularity D, it will come out that $H(z)$ (and $W(z)$) have D dominant poles, of the form $z_k = z_0\varepsilon_k$ ($k = 0 \ldots D-1$), with z_0 being (by definition) the positive real one, and ε_k the D different complex Dth roots of unity. Without going into detail here, we already mention that this will not constitute a stumble block for a dominant pole approximation, and it will always be possible to write the approximate relation as

$$\Pr[H > n] \quad \approx \quad \frac{C(n)}{z_0^n}, \tag{2.12}$$

where we introduce the notation $C(n)$ to lay emphasis on the quasi-geometrical tail decay, with decay rate z_0. The function $C(n)$ will prove periodical, that is $C(n + mD) = C(n)$ ($m \in \mathbb{N}$), while for $n = ND$, we will even be able to simplify the $C(ND)$ further. Note that this quasi-geometric tail decay can only be assumed because burst size and inter-arrival time distribution are not heavy-tailed, as assumed above.

As for the CT case, a similar approach is possible, and will be explained further in Sect. 2.2.7.

2.1.6.2 Two Heuristics

For conventional queues in CT, fed by a Poisson process of bursts of iid size, a relation exists between (the distributions of) the unfinished work in an infinite system and that in a finite system of, say, capacity M (in time units), see for example [94]. This relation leads to an expression for the loss probability in the finite system of the form

$$\frac{(1-\rho)}{\rho}\frac{\Pr[W_\infty > M]}{1-\Pr[W_\infty > M]}\,, \tag{2.13}$$

where W_∞ denotes the unfinished work in the infinite system (as seen by arrivals), and this independent of time setting. When dealing with degenerate buffers, one can translate this into a heuristic for the loss probability of an arbitrary arriving burst (LP), as

$$\text{LP} \overset{*}{\approx} \frac{(1-\rho_{eq})}{\rho_{eq}}\frac{\Pr[H_\infty > ND]}{1-\Pr[H_\infty > ND]}\,. \tag{2.14}$$

Here, H_∞, the scheduling horizon in an infinite optical buffer, fulfills the role of W_∞, ND is the capacity of the system and ρ_{eq} is the so-called equivalent load, that is, the load on the system taking into account the overhead created by the voids, as introduced in Sect. 2.1.5. Note that formula (2.13) assumes only excess unfinished work is lost, that is, bursts arriving at a nearly full system can still be partially buffered, while in our model, a burst that cannot be delayed sufficiently long due to lack of an appropriate delay line, is dropped entirely.

For small LP, a modified heuristic

$$\text{LP} \overset{*}{\approx} (1-\rho_{eq})\frac{\Pr[H_\infty > ND]}{1-\Pr[H_\infty > ND]}\,, \tag{2.15}$$

(that is, dropping the factor ρ_{eq} in the denominator) turns out to be more accurate. In this chapter, we will refer to (2.14) as "heuristic A" and to (2.15) as "heuristic B" respectively.

It is worth to point out that the same heuristics can also be used to evaluate the LP in overloaded systems, that is, when the equivalent load exceeds 100%. Strictly speaking, no equilibrium distribution then exists for e.g. H_∞. The transform function (a pgf in case of DT, an lst in case of CT) that is used to approximate $\Pr[H_\infty > ND]$, however, remains a proper function. Formally then, one can still compute the quantities $\Pr[H_\infty > ND]$, that will no longer represent probabilities, but still yield a good approximation of the loss probability when plugged into the heuristic, as will be illustrated several times throughout this chapter.

For severely overloaded systems, there is a rather simple, intuitive heuristic. For both classic and optical queues, when $\rho \to \infty$, the loss probability will approximately equal

$$\text{LP} \overset{*}{\approx} \frac{\rho-1}{\rho} \quad \text{for} \quad \rho \gg 1\,.$$

Since such a system will be busy nearly always, the carried load will be close to one. The lost load then equals $\rho - 1$, leading directly to the above approximation. As $\rho \to \infty$, the formal value (not a probability) for $\Pr[H_\infty > ND]$ obtained with the dominant pole approximation goes to infinity, and thus the same limit is retrieved in formula (2.13). Not surprisingly then, for degenerate buffers, heuristic A turns out to be more accurate than heuristic B when $\rho_{eq} \gg 1$. This will be illustrated by numerous numerical examples further in this chapter.

Further, for finite degenerate buffer systems, also the situation where the granularity is much larger than the average inter-arrival time $\mathrm{E}[T_k]$, $D/\mathrm{E}[T_k] \gg 1$ provides an interesting case. In such a system, after a short period of transient behavior, due to the FIFO delay-line assignment algorithm, the system will guide all traffic through the longest line, even in case of low load. Indeed, one can verify that, for $D/\mathrm{E}[T_k] \gg 1$, the equivalent load (2.9) becomes much larger than unity, $\rho_{eq} \gg 1$. As such, the system will always be full. In terms of loss performance, it will behave like a bufferless system, whereas the delay will be fixed, and equal to the length of the longest line, ND. In case of a memoryless arrival process, the LP formula for such a system is extremely simple, namely,

$$\mathrm{LP} \overset{*}{\approx} \frac{\rho}{1+\rho} \quad \text{for} \quad D/\mathrm{E}[T_k] \gg 1\,. \tag{2.16}$$

Also in the case of a more general (not memoryless) arrival process, neither heuristic A nor heuristic B offer accurate results, and one is to treat the system as a (simple) bufferless one to obtain correct results.

Finally, if the granularity of a finite degenerate buffer is much smaller than $\mathrm{E}[T_k]$, $D/\mathrm{E}[T_k] \ll 1$, then also the capacity ND is very small. Therefore, the system's LP also approaches that of a bufferless one, while the waiting time turns zero. As such, in case of a memoryless arrival process and $D/\mathrm{E}[T_k] \ll 1$, Formula (2.16) equally applies.

2.2 Model for Memoryless Arrivals

In this section, we present the results obtained for the case of memoryless arrivals, for both the DT and CT setting. With memoryless arrivals, we intend that the inter-arrival time has a memoryless distribution, being an umbrella term for geometric distribution in DT, and a negative-exponential distribution in CT. For the DT setting, results were obtained earlier by Laevens and his co-authors ([10], later [11, 13]), and are repeated here as a reference. For the CT setting, the results were disseminated through [2]. Within this chapter devoted to the pgf approach, the current section is the only one to treat both DT and CT.

2.2.1 Buffer Setting and Traffic Setting

The buffer model of the current section is exactly as described in Sect. 2.1.2.

To proceed with the analysis of (2.2) and (2.4) under the current assumptions, we introduce further notation, that also reflects the memorylessness of the inter-arrival times T_k. Also, we define appropriate functions in a transform domain: pgf's in DT, lst's in CT.

Traffic Setting for DT In our analysis, we use both the probability mass function (pmf) and the pgf of the involved variables. The burst sizes B_k, for example, have a common pmf

$$b(n) = \Pr[B_k = n] \ , \ n \in \mathbb{N}_0 \, ,$$

that is completely general, except for the conditions $0 \leq b(n)$, and $\sum_{n=1}^{+\infty} b(n) = 1$. The resulting pgf is the z-transform of the pmf, namely

$$B(z) = \mathrm{E}[z^{B_k}] = \sum_{n=1}^{+\infty} z^n \Pr[B_k = n] \ , \ z \in \mathbb{C} \, ,$$

and likewise for the other variables T_k, H_k, G_k and W_k. Throughout this work, we will commonly assume the burst size distribution either geometric or deterministic. In the former case, the pmf is given by

$$b(n) = f \cdot \bar{f}^{n-1} \ , \ n \in \mathbb{N}_0 \, ,$$

with parameter f, $\bar{f} = 1 - f$, which implies for the mean $\mathrm{E}[B_k] = 1/f$; the corresponding pgf is

$$B(z) = \frac{fz}{1 - \bar{f}z} \ , \ z \in \mathbb{C} \, .$$

In the latter case, the pmf is given by

$$b(n) = \delta_{n,B} \ , \ z \in \mathbb{N}_0 \, ,$$

where $\delta_{i,j}$ denotes the Kronecker delta, that is one if $i = j$, and zero everywhere else; the associated pgf is

$$B(z) = z^B \ , \ z \in \mathbb{C} \, .$$

Imposing a memoryless distribution for the inter-arrival times, we define the pmf of the T_k as

$$t(n) = \Pr[T_k = n] = p \cdot \bar{p}^{n-1} \ , \ n \in \mathbb{N}_0 \, , \tag{2.17}$$

with parameter p, where we denoted $\bar{p} = 1 - p$, implying $\mathrm{E}[T_k] = 1/p$. For the pgf, we obtain

$$T(z) = \mathrm{E}[z^{T_k}] = \frac{pz}{1 - \bar{p}z} \ , \ z \in \mathbb{C} \, .$$

Traffic Setting for CT For the CT setting, we use the lst of the cumulative distribution function (cdf) of the involved variables. The burst sizes B_k now have a common cdf

$$F_B(x) = \Pr[B_k \leq x] \ , \ x \in \mathbb{R}^+ \, .$$

The exact form of this distribution is completely general, except for some common assumptions: $F_B(x)$ is a right-continuous non-decreasing function, with $F_B(0) = 0$ and $\lim_{x\to\infty} F_B(x) = 1$ (and thus, $0 \leq F_B(x) \leq 1$). This yields as lst

$$B^{\mathcal{L}}(s) = \mathrm{E}[e^{-sB_k}] = \int_0^\infty e^{-sx} dF_B(x) \ , \ s \in \mathbb{C} \, ,$$

and, again, likewise for the other variables T_k, H_k, G_k and W_k. Note that we use a calligraphic L ("$\mathcal{L}$") here and in the rest of this work, to indicate that the intended transform function is an lst ($B^{\mathcal{L}}(\cdot)$), and not a pgf ($B(\cdot)$). In Sect. 2.2.8, the burst size distribution will be assumed negative-exponential and deterministic, respectively. The cdf for the former is

$$F_B(x) = 1 - e^{-\mu x} \ , \ x \in \mathbb{R}^+ \, ,$$

with μ the parameter, and $\mathrm{E}[B_k] = 1/\mu$. The pgf is then

$$B^{\mathcal{L}}(s) = \frac{\mu}{\mu + s} \ , \ s \in \mathbb{C} \, .$$

The cdf of the latter is

$$F_B(x) = H(x - B) \ , \ x \in \mathbb{R}^+ \, ,$$

with $H(\cdot)$ the Heaviside step function, and B equal to the fixed burst size. The corresponding pgf is

$$B^{\mathcal{L}}(s) = e^{-sB} \ , \ s \in \mathbb{C} \, .$$

The memorylessness of the inter-arrival times implies a negative-exponential distribution, corresponding to the following cdf,

$$F_T(x) = 1 - e^{-\lambda x} \ , \ x \in \mathbb{R}^+ \, ,$$

with parameter λ and resulting mean $\mathrm{E}[T_k] = 1/\lambda$, for which the lst is obtained as

$$T^{\mathcal{L}}(s) = \mathrm{E}[e^{-sT_k}] = \frac{\lambda}{\lambda + s} \, .$$

2.2.2 Equilibrium Distribution

As mentioned first in Sect. 2.1.5, we will assume a maximum load, further specified in Sect. 2.2.6, below which stability is guaranteed. In stable regime, the

distributions of $H_{\rm k}$, W_k and G_k converge, for $k \to \infty$, to a unique stochastic equilibrium distribution, independent of the initial system conditions. The pgf's and lst's obtained are associated with this equilibrium, that is assumed throughout this work for the analysis. By H we will denote a generic rv following that distribution, and likewise for the other rv's, yielding W, G, T and B.

2.2.3 Analysis

In [10], both the queueing effect $[x]^+$ and the granularity effect $\lceil x \rceil$ were studied in a DT setting. The queueing effect (2.2) yielded the following relation between the pgf's of the variables involved:

$$H(z) = \frac{p}{z-\bar{p}} G(z) + K \frac{z-1}{z-\bar{p}}\,, \tag{2.18}$$

where the sum of rv's of (2.3) translates into a product of pgf's, that is,

$$G(z) = B(z)W(z)\,, \tag{2.19}$$

and where K in (2.18) denotes a constant (here, $G(\bar{p})$) of which the exact value can be determined most easily later on, by means of the normalization condition.

The solution to the granularity effect was captured by the following relation:

$$W(z) = \sum_{k=0}^{D-1} \frac{1}{D} \frac{z^D-1}{z\varepsilon_k - 1} H(z\varepsilon_k)\,, \tag{2.20}$$

where the symbols $\varepsilon_k = e^{\,\mathrm{j}\,2\pi k/D}$ represent the D different complex D^{th} roots of unity. In another context, this relation was published earlier in somewhat modified form (floor instead of ceiling function, see formula (13) in [95]). To allow for a CT solution of the granularity effect, we will reorder the terms in the sum as follows,

$$W(z) = \sum_{k} \frac{1}{D} \frac{z^D-1}{z\varepsilon_k - 1} H(z\varepsilon_k)\,, \tag{2.21}$$

where the summation index k now runs over $\lceil -D/2 \rceil < k \leq \lceil D/2 \rceil$.

Using the property that $W(z\varepsilon_k) = W(z)$, which follows directly from the fact that the rv W is always an integer multiple of D, one can combine (2.18), (2.19) and (2.21), to obtain that

$$W(z) = K \left(\sum_k \frac{1}{D} \frac{z^D-1}{z\varepsilon_k - \bar{p}} \right) \cdot \left(1 - \sum_k \frac{1}{D} \cdot \frac{z^D-1}{z\varepsilon_k - 1} \frac{pB(z\varepsilon_k)}{(z\varepsilon_k) - \bar{p}} \right)^{-1}.$$

With the identity

$$\frac{x^{D-1}}{z^D - x^D} = \sum_k \frac{1}{D} \frac{1}{(z\varepsilon_k) - x}\,, \tag{2.22}$$

the expression simplifies to

$$W(z) = K\left(\frac{\bar{p}^{D-1}(z^D-1)}{z^D-\bar{p}^D}\right)\cdot\left(1-\sum_k \frac{1}{D}\cdot\frac{z^D-1}{z\varepsilon_k-1}\frac{pB(z\varepsilon_k)}{(z\varepsilon_k)-\bar{p}}\right)^{-1}. \quad (2.23)$$

The constant K follows from the normalization condition $W(1)=1$, as

$$K = \left(\frac{1}{p} - \mathrm{E}[B] - \frac{D-1}{2} - \sum_{\mathrm{k}\neq 0}\frac{1}{\varepsilon_{\mathrm{k}}-1}\frac{p}{\varepsilon_{\mathrm{k}}-\bar{p}}B(z\varepsilon_{\mathrm{k}})\right)\cdot\left(\frac{D\bar{p}^{D-1}}{1-\bar{p}^D}\right)^{-1}. \quad (2.24)$$

Having determined $W(z)$, $H(z)$ then follows readily from (2.18), as

$$H(z) = \frac{K\cdot p\cdot\bar{p}^{D-1}\cdot(z^D-1)\cdot B(z)}{(z-\bar{p})(z^D-\bar{p}^D)\left(1-\sum_k \frac{1}{D}\cdot\frac{z^D-1}{z\varepsilon_k-1}\frac{pB(z\varepsilon_k)}{(z\varepsilon_k)-\bar{p}}\right)} + K\frac{z-1}{z-\bar{p}}. \quad (2.25)$$

2.2.4 Analysis for CT: Limit Procedure

Our goal is to derive, for the asynchronous system, $H^{\mathcal{L}}(s)$, the lst of the equilibrium distribution of the scheduling horizon H as seen by arrivals. In this subsection, we discuss a limit procedure that allows to retrieve $H^{\mathcal{L}}(s)$ from $H(z)$.

To correctly convert results from DT to CT, one should first observe quantities in the discrete domain. A distinction can be made between time-related quantities and counting-related quantities. The former scale with the slot size Δ, the latter do not. For example, the scheduling horizon H of Sect. 2.2.3 is expressed in slots, and actually represents $H\Delta$ in absolute time. To avoid confusion, we here denote the DT version of the rv H with H_{DT}, and the CT version with H_{CT}; and likewise for D_{DT} and D_{CT}. We can now find a simple relation between the pgf and lst, as

$$H^{\mathcal{L}}(s) = \mathrm{E}[e^{-sH_{\mathrm{CT}}}] = \lim_{\Delta\to 0}\mathrm{E}[(e^{-s\Delta})^{H_{\mathrm{DT}}}] = \lim_{\Delta\to 0} H(e^{-s\Delta}),$$

that is, we need to substitute z by $e^{-s\Delta}$ in the pgf $H(z)$ before taking the limit $\Delta\to 0$. The average inter-arrival time $1/p\cdot\Delta$ scales as $1/\lambda$, the granularity size as $D_{\mathrm{DT}}\Delta = D_{\mathrm{CT}}$. While we limit ourselves to this simple recipe, we refer the interested reader to Appendix B of [96] for its mathematical background.

Applying this limit procedure on the DT solution (2.18) for the queueing effect yields the CT expressions

$$H^{\mathcal{L}}(s) = \frac{\lambda}{\lambda-s}G^{\mathcal{L}}(s) - K\frac{s}{\lambda-s}, \quad (2.26)$$

and

$$G^{\mathcal{L}}(s) = W^{\mathcal{L}}(s)B^{\mathcal{L}}(s).$$

(In taking the limit, here and in the following, the rules of de l'Hôpital need to be applied frequently, to deal with for example indeterminate forms of type $0/0$.)

Concerning the granularity effect, equation (2.21) results in the CT expression

$$W^{\mathcal{L}}(s) = \sum_k \frac{1}{D} \frac{1 - e^{-sD}}{s + \mathrm{j}\, 2\pi k/D} H^{\mathcal{L}}(s + \mathrm{j}\, 2\pi k/D)\,, \tag{2.27}$$

where k now runs from $-\infty$ to $+\infty$, and D_{CT} is again represented by D.

Note that $W^{\mathcal{L}}(s)$ is periodic too, in the sense that

$$W^{\mathcal{L}}(s) = W^{\mathcal{L}}(s + \mathrm{j}\, 2\pi n/D)\,,$$

for any $n \in \mathbb{Z}$. This property now allows combining (2.26) and (2.27) to yield

$$\begin{aligned} W^{\mathcal{L}}(s) \;\; = \;\; & \left(-K \sum_k \frac{1}{D} \frac{1 - e^{-sD}}{\lambda - (s + \mathrm{j}\, 2\pi k/D)} \right) \\ & \cdot \left(1 - \sum_k \frac{1}{D} \frac{1 - e^{-sD}}{t} \frac{\lambda B^{\mathcal{L}}(t)}{\lambda - t} \bigg|_{t = s + \mathrm{j}\, 2\pi k/D} \right)^{-1} . \end{aligned}$$

A further simplification can be made by using

$$\sum_k \frac{1}{D} \frac{1}{(\lambda - s) + \mathrm{j}\, 2\pi k/D} = -\frac{1}{1 - e^{(\lambda - s)d}}\,, \tag{2.28}$$

which follows from applying the limit procedure on (2.22) (where $x = 1$, and we substitute s by $(\lambda - s)$). We then find

$$W^{\mathcal{L}}(s) = \left(K \frac{1 - e^{-sD}}{1 - e^{(\lambda - s)D}} \right) \cdot \left(1 - \sum_k \frac{1}{D} \frac{1 - e^{-sD}}{t} \frac{\lambda B^{\mathcal{L}}(t)}{\lambda - t} \bigg|_{t = s + \mathrm{j}\, 2\pi k/D} \right)^{-1} . \tag{2.29}$$

This is exactly the expression we would have found applying the limit procedure directly to equation (2.23).

The remaining unknown constant K (for CT, as opposed to (2.24), for DT) can be determined, either by applying the limit procedure once more, or by ensuring normalization of $W^{\mathcal{L}}(s)$. The final result reads

$$K = \left(\frac{1}{\lambda} - \mathrm{E}[B] - \frac{D}{2} - \sum_{k \neq 0} \frac{\lambda}{t - \lambda} \frac{B^{\mathcal{L}}(t)}{t} \bigg|_{t = s + \mathrm{j}\, 2\pi k/D} \right) \cdot \left(-\frac{D}{1 - e^{-\lambda D}} \right)^{-1} . \tag{2.30}$$

Equations (2.26), (2.29) and (2.30) together fully specify $H^{\mathcal{L}}(s)$, as

$$\begin{aligned} H^{\mathcal{L}}(s) \;\; = \;\; & \frac{K \cdot \lambda \cdot (1 - e^{-sD}) \cdot B^{\mathcal{L}}(s)}{(\lambda - s)(1 - e^{(\lambda - s)D}) \left(1 - \sum_k \frac{1}{D} \frac{1 - e^{-sD}}{t} \frac{\lambda B^{\mathcal{L}}(t)}{\lambda - t} \Big|_{t = s + \mathrm{j}\, 2\pi k/D} \right)} \\ & - K \frac{s}{\lambda - s}\,, \end{aligned} \tag{2.31}$$

which indeed is similar to the expression of the scheduling horizon's pgf for DT, (2.25).

2.2.5 Analysis for CT: Direct Approach

To consolidate the results of Subsection 2.2.4, we show how they can also be obtained directly. The complexity of the transform-based solution of the queueing effect, mentioned in the above, critically depends on the exact form of the lst of T. (The same goes, in terms of the pgf of T, for the DT case, as discussed in for example [93].) For exponentially-distributed T, the complexity is limited, and results in the CT expression

$$H^{\mathcal{L}}(s) = \frac{\lambda}{\lambda - s} B^{\mathcal{L}}(s) W^{\mathcal{L}}(s) - K \frac{s}{\lambda - s} ,$$

that is, the result we obtained via the limit procedure. A direct proof is rather straightforward. Introducing, for convenience, an auxiliary rv $G = B + W$, with cdf $F_G(t)$ ($t > 0$) and lst $G^{\mathcal{L}}(s) = B^{\mathcal{L}}(s) W^{\mathcal{L}}(s)$, one has, starting from Eq. (2.2),

$$\begin{aligned}
H^{\mathcal{L}}(s) &= \int_0^\infty dF_G(g) \int_0^\infty dF_T(t)\, e^{-s[g-t]^+} \\
&= \int_0^\infty dF_G(g) \int_0^\infty dF_T(t)\, e^{-s(g-t)} \\
&\quad + \int_0^\infty dF_G(g) \int_t^\infty dF_T(t) \left(e^{-s\cdot 0} - e^{-s(g-t)} \right) \\
&= G^{\mathcal{L}}(s) T^{\mathcal{L}}(-s) + \int_0^\infty dF_G(g) \int_g^\infty \lambda e^{-\lambda t} dt \left(1 - e^{-sg} e^{+st} \right) \\
&= G^{\mathcal{L}}(s) \frac{\lambda}{\lambda - s} + \int_0^\infty dF_G(g) \left(e^{-\lambda g} - e^{-sg} \frac{\lambda}{\lambda - s} e^{-(\lambda - s)g} \right) \\
&= G^{\mathcal{L}}(s) \frac{\lambda}{\lambda - s} + G^{\mathcal{L}}(\lambda) \left(1 - \frac{\lambda}{\lambda - s} \right) \\
&= G^{\mathcal{L}}(s) \frac{\lambda}{\lambda - s} - K \frac{s}{\lambda - s} .
\end{aligned}$$

where $K = G^{\mathcal{L}}(\lambda)$ denotes the same constant as in (2.26), of which the exact value can be determined most easily later on, by means of the normalization condition.

The second non-linearity to tackle is the granularity effect, as stated above in Eq. (2.4). The transform-based solution can be obtained by expressing $W^{\mathcal{L}}(s)$ in terms of the probability density function (pdf) $h(x)$ of H. However, H is a "mixed" rv, in that its density function $h(x)$ is continuous for $x > 0$, but possesses a discontinuity at $x = 0$, that accounts for a discrete probability. As such, $h(x)dx = \Pr[x \leq H < x + dx]$ for $x \in \mathbb{R}_0$, while $h(x) = \Pr[H = 0]$ for $x = 0$.

This however poses no further mathematical difficulty, and the expression for (2.4) in terms of lst's is easily obtained as

$$W^{\mathcal{L}}(s) = h(0) + \sum_{\mathrm{k}=0}^{+\infty} \int_{0^+}^{D} h(u+kD) e^{-s(k+1)D} du \,,$$

Rewriting the sum on the right-hand side by introducing the comb function $\sum_{\mathrm{l}} \delta_{t,lD}$ (with $\delta_{i,j}$ still the Kronecker delta) we have

$$\begin{aligned} W^{\mathcal{L}}(s) &= h(0) + \sum_{k=0}^{+\infty} \int_{0^+}^{D} du\, h(u+kD) e^{-s(k+1)D} \\ &\quad \cdot \int_{0^+}^{D} dx\, e^{-s(u-x)} \sum_{l=-\infty}^{+\infty} \delta_{u-x,lD} \,. \end{aligned}$$

Note that the Kronecker delta only has effect when $u - x = 0$, that is, for $l = 0$. Using the Fourier expansion

$$\sum_{k=-\infty}^{+\infty} \delta_{x,kD} = \sum_{k=-\infty}^{+\infty} \frac{1}{D}\, e^{\mathrm{j}\, 2\pi k x/D} \,,$$

and rearranging some terms, we can proceed as

$$\begin{aligned} W^{\mathcal{L}}(s) &= h(0) + \sum_{k=0}^{+\infty} \int_{0^+}^{D} du\, h(u+kD) e^{-s(k+1)D} \\ &\quad \cdot \int_{0^+}^{D} dx\, e^{-s(u-x)} \sum_{l=-\infty}^{+\infty} \frac{1}{D}\, e^{-\mathrm{j}\, 2\pi l(u-x)/D} \\ &= h(0) + e^{-sD} \int_{0^+}^{D} dx \sum_{l=-\infty}^{+\infty} \frac{1}{D}\, e^{(s+\mathrm{j}\, 2\pi l/D)x} \\ &\quad \cdot \sum_{k=0}^{+\infty} \int_{0^+}^{D} du\, h(u+kD) e^{-s(u+kD)-\mathrm{j}\, 2\pi l u/D} \\ &= h(0) + e^{-sD} \sum_{l=-\infty}^{+\infty} \frac{1}{D} \frac{e^{(s+\mathrm{j}\, 2\pi l/D)D} - 1}{s + \mathrm{j}\, 2\pi l/D} \\ &\quad \cdot \sum_{k=0}^{+\infty} \int_{0^+}^{D} du\, h(u+kD) e^{-(s+\mathrm{j}\, 2\pi l/D)\cdot(u+kD)} \,. \end{aligned}$$

In the last step we used the obvious identity

$$e^{-(\mathrm{j}\, 2\pi l/D)kD} = 1 \,,$$

for any l, $k \in \mathbb{Z}$, which allowed us to arrive at an expression in terms of $u + kD$ only in the integral for u. That integration then, combined with the sum over k, amounts to integrating over $\mathbb{R}_0$, that is,

$$\begin{aligned} W^{\mathcal{L}}(s) &= h(0) + e^{-sD} \sum_{l=-\infty}^{+\infty} \frac{1}{D} \frac{e^{(s+\mathrm{j}\,2\pi l/D)D} - 1}{s + \mathrm{j}\,2\pi l/D} \int_{0^+}^{\infty} dt\, h(t) e^{-(s+\mathrm{j}\,2\pi l/D)t} \\ &= h(0) + \sum_{l=-\infty}^{+\infty} \frac{1}{D} \frac{1 - e^{-sD}}{s + \mathrm{j}\,2\pi l/D} \left(H^{\mathcal{L}}(s + \mathrm{j}\,2\pi l/D) - h(0) \right) . \end{aligned}$$

Using identity (2.28) once more, we find that the terms involving $h(0)$ cancel out, yielding

$$W^{\mathcal{L}}(s) = \sum_{l} \frac{1}{D} \frac{1 - e^{-sD}}{s + \mathrm{j}\,2\pi l/D} H^{\mathcal{L}}(s + \mathrm{j}\,2\pi l/D) ,$$

as before in (2.27). As such, the combination of both non-linearities leads to the same solution as the one obtained with the limit procedure, namely (2.31).

2.2.6 Maximum and Equivalent Load

As mentioned in Sect. 2.1.5 and 2.2.2, stability requires the offered load ρ to be below some maximum value ρ_{max}, that is typically less than unity, unlike in conventional queues. The current DT case was treated first in [10], or, more elaborately, in [11]. To derive the stability condition, it can be intuitively understood that, for a buffer with infinite size, the queue length grows unboundedly if the load is too high. For optical buffers, this happens even before the classic load $\rho = \mathrm{E}[B]/\mathrm{E}[T]$ reaches unity. We can characterize a case of unbounded growth by

$$\lim_{k \to \infty} \Pr[H_k = 0] \stackrel{*}{=} 0 ,$$

and this independent of time setting. Taking into account the expression for the scheduling horizon ((2.25) for DT, (2.31) for CT) and waiting time ((2.23) for DT, (2.29) for CT), this occurs when K ((2.24) for DT, (2.30) for CT)) becomes zero.

To capture the instability quantitatively, one can define a maximum tolerable arrival intensity as $1/\mathrm{E}[T]$, that delineates stability. For DT, this is p_{max}, that is the solution to the implicit expression (implicit, as p occurs also on the right-hand side of the equation)

$$\frac{1}{p} = \mathrm{E}[B] + \frac{D-1}{2} + \sum_{k \neq 0} \frac{p}{\varepsilon_k - \bar{p}} \frac{B(\varepsilon_k)}{\varepsilon_k - 1} . \tag{2.32}$$

The symbols ε_k still represent the D different complex D^{th} roots of unity, as in (2.21). The solution is thus function of the FDL granularity D, the complete pgf of the burst-size distribution, and can be found from (2.32) with a simple bisection

algorithm. Note that it is indeed consistent with (2.8), of which it is a specific instance, for DT and geometric inter-arrival times.

Related, we can define an equivalent load, that incorporates the effect of voids into an altered definition of the load. For DT, we obtain

$$\rho_{eq} = p \left(\mathrm{E}[B] + \frac{D-1}{2} + \sum_{k \neq 0} \frac{1}{\varepsilon_k - 1} \frac{pB(\varepsilon_k)}{\varepsilon_k - \bar{p}} \right) ,$$

that plays a key role in the application of the heuristics below, since it allows for results for finite systems. Hereby, note that the exact definition of ρ_{eq} is prone to discussion, as argued in Sect. 2.1.5. Remark that ρ_{eq} indeed equals one when the arrival intensity reaches p_{max}, as should, and is consistent with the general expression for ρ_{eq}, namely (2.9). From here, it is also possible to define an equivalent burst size, B_{eq}, with

$$B_{eq} = \mathrm{E}[B] + \frac{D-1}{2} + \sum_{k \neq 0} \frac{1}{\varepsilon_k - 1} \frac{pB(\varepsilon_k)}{\varepsilon_k - \bar{p}} ,$$

that allows to write $\rho_{eq} = p \cdot B_{eq}$, a counterpart to the classic load $\rho = p \cdot \mathrm{E}[B]$.

Denoting the maximum tolerable arrival intensity with λ_{max} for CT, the limit procedure easily yields that λ_{max} is the solution to

$$\frac{1}{\lambda} = \mathrm{E}[B] + \frac{D}{2} + \sum_{k \neq 0} \frac{\lambda}{t - \lambda} \left. \frac{B^{\mathcal{L}}(t)}{t} \right|_{t = s + \mathrm{j}\, 2\pi k/D} ,$$

that is again an implicit expression, yielding a value for λ_{max} with a bisection algorithm (combined with a fit approximation for the infinite sum). Corresponding, we can define an equivalent load as

$$\rho_{eq} = \lambda \left(\mathrm{E}[B] + \frac{D}{2} + \sum_{k \neq 0} \frac{\lambda}{t - \lambda} \left. \frac{B^{\mathcal{L}}(t)}{t} \right|_{t = s + \mathrm{j}\, 2\pi k/D} \right) ,$$

which is again consistent with the general expression (2.9).

2.2.7 Heuristics for the Loss Probability

Results up to now related to an optical buffer of infinite size. In order to obtain the loss probability (LP) in a finite system, that is, a system with only $(N+1)$ fiber delay lines (realizing delays in the set $\{0, D, \ldots, ND\}$) we rely on the heuristics presented in Sect. 2.1.6. Like there, we assume that the distribution of the burst sizes is not heavy-tailed and that its pgf (in DT) or lst (in CT) has no singularities other than poles.

While the expression for ρ_{eq} is treated in Sect. 2.2.6, we can again combine results of the synchronous FDL buffer [10] with the limit procedure to find expressions for the unknown $\Pr[H_\infty > ND]$ in heuristic A (2.14) en B (2.15). These tail probabilities can indeed be computed by an (approximate) inversion of the pgf $B(z)$, using the dominant pole approximation of (2.12) first in DT, then to move to CT with a limit procedure, and obtain the inversion of the lst $H^{\mathcal{L}}(s)$.

As for (2.12) in DT, z_0 is the dominant pole of $H(z)$ (and of $W(z)$). It is real, positive and larger than 1. This approximate geometric behavior occurs under rather mild conditions on the burst size distribution, a sufficient condition is for example that the pgf of the burst sizes is a rational function. Note, however, that the analysis in the above only requires that $\mathrm{E}[B] < \infty$ (and that the system is stable). When, for instance, the burst size distribution possesses a heavy tail, the distribution of H_∞ would not decay geometrically as above, but would have a heavy tail too. This would require a different approximate inversion formula.

The constant $C(ND)$ (see Sect. 2.12) for DT can be obtained from residue theory and is given by

$$C(ND) = -\frac{1}{z_0}\frac{D}{z_0^D - 1}\left(\lim_{z\to z_0} W(z)(z - z_0)\right) .$$

Applying the limit procedure once more, we find that for asynchronous buffers, in CT,

$$\Pr[H_\infty > ND] \approx \frac{C(ND)}{\gamma^N} ,$$

where we introduced $\gamma = e^{-s_0 D_{\mathrm{CT}}} = \lim_{\Delta\to 0} z_0^{D_{\mathrm{DT}}\Delta}$ for convenience. Here, s_0 denotes the dominant pole of $H^{\mathcal{L}}(s)$ and $W^{\mathcal{L}}(s)$ along the negative real line. In general, a simple bisection algorithm (with possibly an initial search for the appropriate starting interval) suffices to determine γ numerically. In some cases, an explicit expression can also be found, see for example below.

As mentioned in Sect. 2.1.6, the same heuristics can also be used to evaluate the LP in overloaded systems, when the equivalent load exceeds 100%. The transform $H^{\mathcal{L}}(s)$ that is used to approximate $\Pr[H_\infty > ND]$ remains a proper function, and formally, one can still compute the quantities $\Pr[H_\infty > ND]$, the only caveat being that γ is then to be found in the interval $[0, 1)$, that is, $s_0 > 0$. The expression for the constant $C(ND)$ remains the same. (When the equivalent load is exactly 100%, $s_0 = 0$ and $\gamma = 1$. In principle, this requires somewhat modified expressions. Here, we do not pursue this issue further.)

2.2.8 Special Cases

In this section, we take a look at the LP for three special cases for the burst size distribution: exponential, deterministic and a mixture of deterministic burst sizes.

For all three of them, the infinite sum appearing in a.o. equations (2.29), (2.31) and (2.30) can be removed. One obtains closed-form formulas for the lst $H^{\mathcal{L}}(s)$ and for the performance measures derived therefrom. Results given here were obtained via the limit procedure. Formulas for the corresponding DT systems are given in [13].

2.2.8.1 Exponentially-Distributed Burst Sizes

As was the case for the inter-arrival times T, exponentially-distributed burst sizes can be considered as the limit (for slot sizes going to zero) of geometrically distributed burst sizes. Assuming the distribution given in Sect. 2.2.1 for B, expression (2.29) for $W^{\mathcal{L}}(s)$ simplifies significantly to

$$W^{\mathcal{L}}(s) = \frac{1 - e^{-(s+\mu)D}}{1 - e^{-\mu D}} \cdot \frac{\gamma - 1}{\gamma - e^{-sD}} ,$$

with

$$\gamma = \frac{\mu + \lambda}{\mu e^{-\mu D} + \lambda e^{+\lambda D}} .$$

The constant $C(ND)$ appearing in the approximation for the tail distribution becomes

$$C(ND) = \frac{1 - \gamma e^{-\mu D}}{\gamma \left(1 - e^{-\mu D}\right)} .$$

Moreover, we can obtain an explicit expression for the waiting time probabilities. Therefore, it suffices to rewrite $W^{\mathcal{L}}(s)$ as

$$W^{\mathcal{L}}(e^{-sD}) = w(0) + (1 - w(0)) \frac{\bar{\zeta} e^{-sD}}{1 - \zeta e^{-sD}} ,$$

with

$$\begin{aligned} \zeta &= \gamma^{-1} = (\mu e^{-\mu D} + \lambda e^{+\lambda D})/(\lambda + \mu) \; , \; \bar{\zeta} = 1 - \zeta \, , \\ \bar{F} &= e^{-\mu D} \; , \; F = 1 - \bar{F} \, , \\ w(0) &= \Pr[W_k = 0] = \bar{\zeta}/F \, . \end{aligned} \tag{2.33}$$

Here, note that we adopt ζ and F as notation to stress that they are identical to the ζ and F that appear in the results of Sect. 3.4.2.3 and 3.4.2.4, where the same degenerate M/M/1 buffer setting is studied, but then in the case of finite buffer size. Notice how the discrete nature of the W_k is reflected in $W^{\mathcal{L}}(s)$: it suffices to replace e^{-sD} with z^D, to obtain the pgf of the waiting time distribution, expressed as integer multiples of D. From this, one can extract also the other waiting time probabilities explicitly, as

$$w(n) = \Pr[W_k = nD] = (1 - w(0)) \cdot \bar{\zeta} \cdot \zeta^{n-1} \; , \; n \in \mathbb{N}_0 \, . \tag{2.34}$$

As such, the waiting time probabilities $w(n)$ for $n \geq 1$ can be found through a *geometric relation*, but do not follow a *geometric distribution*, since $w(0)$ does not adhere to (2.34). From here, the mean waiting time is easily found as

$$\mathrm{E}[W] = D \cdot \frac{1 - w(0)}{\bar{\zeta}} = D \cdot \left\{ \frac{1}{\bar{\zeta}} - \frac{1}{F} \right\} . \tag{2.35}$$

Finally, note that in the above-mentioned Sect. 3.4.2.3 and 3.4.2.4, the expression for the $w(n)$ and $\mathrm{E}[W_k]$ is derived also in the case of finite buffer size, and countered with similar expressions for DT.

We further obtain

$$\rho_{eq} = 1 + \frac{\lambda D}{\mu + \lambda} \left(\frac{\lambda}{1 - e^{-\mu D}} + \frac{\mu}{1 - e^{+\lambda D}} \right) . \tag{2.36}$$

for the equivalent load.

Note further that, in this case, it is straightforward to verify that

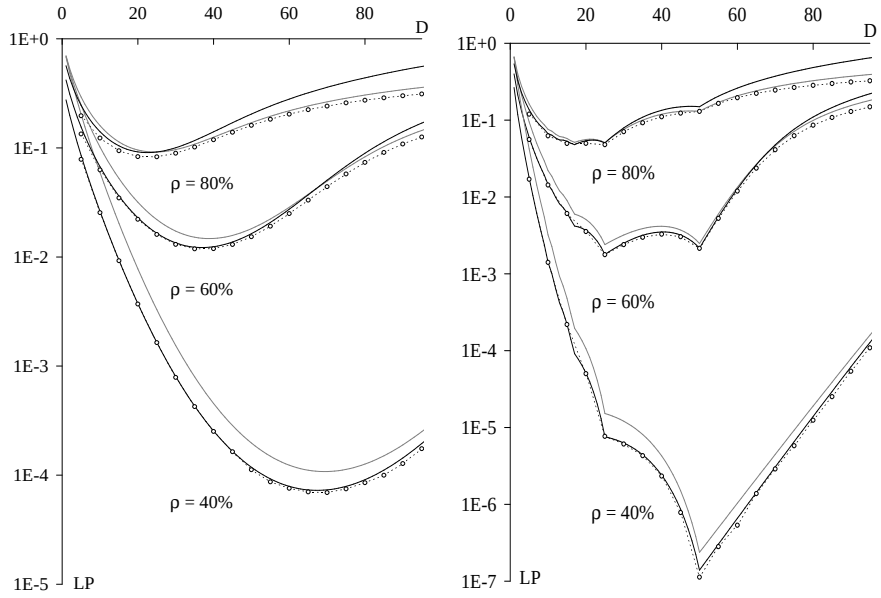
$$\rho_{eq} = 1 \Leftrightarrow \gamma = 1 \Leftrightarrow s_0 = 0 \,,$$

as mentioned above. The condition under which $\rho_{eq} = 1$ fully agrees with the one that can be found by taking the appropriate limit of the condition derived in [11] for the synchronous case.

With these formulas at hand, one can easily calculate the LP via one of the heuristics given above. Some numerical results are shown on the left pane of Fig. 2.2. It compares results from simulation (points connected by dotted lines) with those obtained via heuristic A (solid gray curves) or heuristic B (solid black curves). The mean burst size $\mathrm{E}[B]$ was set to 50 μs, which corresponds to circa 60 kB at 10 Gbit/s. The granularity D varied from 0 to 100 μs (in steps of 5 μs during the simulations). The number of available FDLs was set to $N = 20$. The figure shows results for different input load levels $\rho = \lambda \mathrm{E}[B]$.

The heuristics are a bit pessimistic, that is, they overestimate the LP. Heuristic B is more accurate for low input load levels, but does not converge to the right asymptotic value when $\rho_{eq} \gg 1$, as predicted. (Here, $\rho_{eq} \to \infty$ as $D \to \infty$, while in the actual system, ρ_{eq} plays a minor role when $D \to \infty$, since then LP $\to \rho/(1+\rho)$). For these high loads, heuristic A performs better. Compared to the method of Callegati [26], one gains somewhat in accuracy, especially for low load values or small buffer sizes. Furthermore, since for this special case, explicit formulae were obtained, the numerical complexity involved in our results is close to zero.

There is an optimal granularity D (in terms of LP) , shifting to lower values for higher input load levels, as was the case in synchronous systems, see [10]. As we will illustrate in the next section, the optimal value also depends on the burst size distribution.



(a) Negative-exponential burst size distribution.

(b) Deterministic burst size distribution.

Figure 2.2: The heuristics capture the performance of an asynchronous optical buffer well, both in the case of negative-exponential and deterministic burst size distribution. The solid gray curves represent heuristic A, the solid black curves heuristic B, while the points connected by dotted lines are simulation points. The LP of the CT M/G/1 case is set out as a function of D, for varying D (in μs), $\mathrm{E}[B] = 50$ μs, and $N = 20$.

Further, as a reference, we compare performance results of both a synchronous and asynchronous system setting, as first presented in [1]. In Fig. 2.3(a), the LP is plotted as a function of the granularity, in the case of memoryless burst sizes. The curves now have a different meaning, as both $\rho = 60\%$, and $N = 20$ are fixed, and different curves correspond with different slot sizes Δ. The (mean) burst size is still set to 50 μs, and the granularity D again varies from 0 to 100 μs. In order of increasing LP, the first three sets of curves (lowest LP curves) show results for a DT setting, with a finite slot size Δ, obtained with heuristic B. The first two sets (points connected by dotted lines) are valid for $\Delta = 10$ μs, $\Delta = 5$ μs respectively. The third set has a slot size $\Delta = 1$ μs, and is depicted with solid black curves. (The DT step, although not visible, is $\Delta = 1$ μs.). The last set (highest LP curves) represents the LP for a CT setting, and is also obtained with heuristic B. In DT, the memoryless burst size distribution corresponds to the geometric distribution, in CT to the exponential distribution.

Clearly, higher values for the time slot sizes result in lower LP. For gradually

reducing time slot sizes, curves evolve to the limit distribution of zero slot size, that is, the asynchronous case. For the memoryless burst size distribution considered, the curves mostly preserve their smooth shape.

The above comparison suggests a better performance for synchronized systems in terms of loss probabilities. Note, however, that all curves assumed that at most one arrival can occur per slot. As such, this comparison does not allow to quantify the impact of synchronization, which calls for a separate analysis in Sect. 2.3, by means of a batch arrival model. For the current case, results illustrate however that the assumption on time setting (CT or DT) has a crucial impact on the exact value of the loss probability, but less on the optimum.

2.2.8.2 Deterministic Burst Sizes

In this case, all bursts are of length B. In order to proceed, we need to express B as $aD - b$, where $a \geq 1$ and $0 \leq b < D$. That is, $a = \lceil B/D \rceil$ and $b = aD - B$. With this convention, the limit procedure yields

$$W^{\mathcal{L}}(s) = \frac{-\left(a\left(1-e^{-\lambda D}\right) - e^{-\lambda(D-b)}\right)\left(1-e^{-sD}\right)}{e^{-saD}e^{-\lambda(D-b)}\left(1-e^{-sD}\right) + \left(e^{-\lambda D} - e^{-sD}\right)\left(1-e^{-saD}\right)} \,.$$

The equivalent load is now given by

$$\rho_{eq} = 1 + \lambda D\left(a - \frac{e^{-\lambda(D-b)}}{1-e^{-\lambda D}}\right) \,,$$

and reaches 100% when

$$a = \frac{e^{-\lambda(D-b)}}{1-e^{-\lambda D}} \,,$$

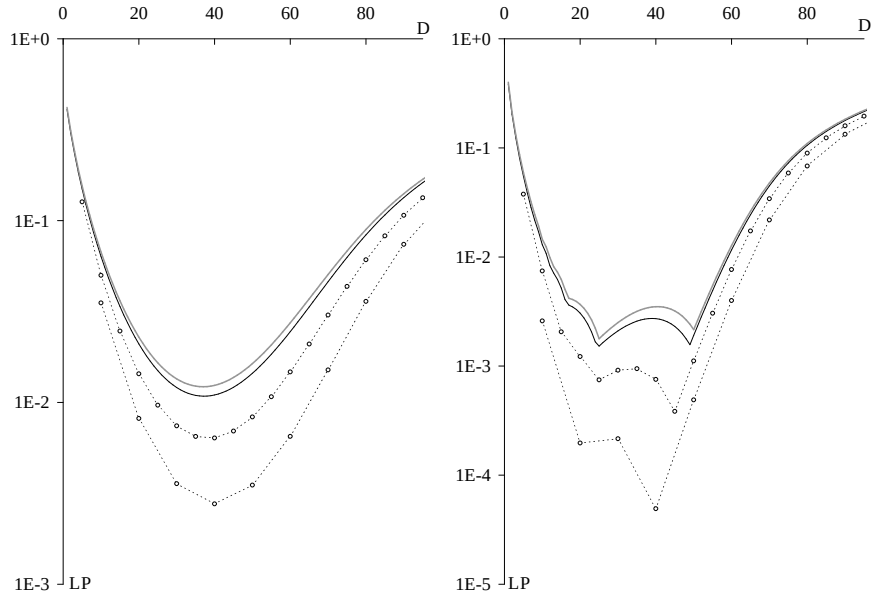
again in agreement with what one would obtain by taking the appropriate limit of the condition given in [11] for this specific case. The dominant pole $\gamma = \mathrm{e}^{-s_0 D}$ now has to be determined as the solution of

$$\gamma^a e^{-\lambda(D-b)}(1-\gamma) + (e^{-\lambda D} - \gamma)(1-\gamma^a) = 0 \,.$$

For $\rho_{eq} < 1$, γ is to be found in $(1, \infty)$. Since $\gamma = \mathrm{e}^{-s_0 D}$, this easily translates into the well-known characterization of stability of lst's, namely $s_0 < 0$. (This is a particular instance of the more general stability condition, requiring all poles to be in the left half-plane.) For $\rho_{eq} > 1$, γ it is situated in $(0, 1)$, and $s_0 > 0$.

Further, the constant needed in the approximation of $\Pr[H_\infty > ND]$ is given by

$$C(ND) = \frac{-\left(a\left(1-e^{-\lambda D}\right) - e^{-\lambda(D-b)}\right)}{a\gamma^a(\gamma - e^{-\lambda D}) - \gamma(1-\gamma^a) - \gamma^a(\gamma + \gamma a - a)e^{-\lambda(D-b)}} \,.$$



(a) Memoryless burst size distribution.

(b) Deterministic burst size distribution.

Figure 2.3: Although for the same traffic and buffer setting, these curves differ completely due to to the different time setting, both in the case of memoryless and deterministic burst size distribution. As the slot length $\Delta \in \{0, 1, 5, 10\}$ (in μs) increases, LP lowers, and this for any value of the granularity D (μs). These figures were obtained for $\mathrm{E}[B] = 50$ μs, $\rho = 60\%$, and $N = 20$.

Finally, note that in the case that $B = D$ (with $a = 1$, $b = 0$), the expression for $W^{\mathcal{L}}(s)$ simplifies to

$$W^{\mathcal{L}}(s) = \frac{\bar{\zeta}}{1 - \zeta e^{-sD}} ,$$

with $\zeta = e^{\lambda D} - 1$, and $\bar{\zeta} = 1 - \zeta$. Given the lst's simple form, it is possible to obtain the waiting time probabilities explicitly, as

$$w(n) = \Pr[W_k = nD] = \bar{\zeta} \cdot \zeta^n \, , \; n \in \mathbb{N} \, , \tag{2.37}$$

and the equivalent load reaches 100% when $\zeta = 1$. That the latter special instance ($B = D$) does indeed provide an interesting case, is argued in the next chapter, in Sect. 3.3, where also the expression for the waiting time probabilities in case of a finite buffer is derived.

Some results for this case are given in Fig. 2.2(b), for similar parameter setting as that of Fig. 2.2(a). Again, the (mean) burst size was set to 50 μs, the granularity D varied from 0 to 100 μs, and the number of available FDLs was set to

$N = 20$. The shape of the curves is substantially different from the ones for negative-exponential burst size distribution, and the LP can be more than an order of magnitude smaller. Again, heuristic B is more accurate for lower values of the LP, but does not converge to the correct limit for $\rho_{eq} \gg 1$. There are now several "notches" in the curves, occuring for those values of D for which B is divisable by D, that is, for $b = 0$. As in the case for exponentially-distributed burst sizes in Sect. 2.2.8.1, the global optimum value of D is sensitive to the load, and decreases when the load increases. However, here, the optimum granularity, say D_0, does not vary in a continuous fashion, but rather jumps from a larger divisor of B to a smaller divisor of B when the load increases. Indeed, as Fig. 2.2(b) illustrates, D_0 is equal to B for $\rho = 40\%$, whereas $D_0 = B/2$ for $\rho = 60\%$, and $D_0 = B/3$ for $\rho = 80\%$. This effect was reported earlier in [10], and is characteristic to a degenerate M/D/1 buffer setting.

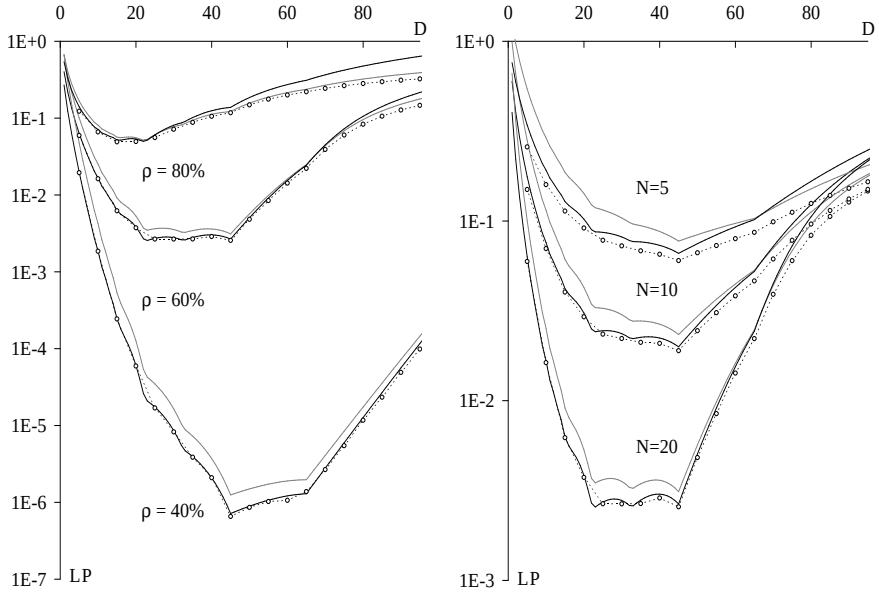
Just like we did for the burst sizes with negative-exponential distribution, we also compare performance results of the synchronous and asynchronous system setting. Figure 2.3(b) presents the LP as a function of the granularity, as obtained with heuristic B, for $\rho = 60\%$, $N = 20$, $\mathrm{E}[B] = B = 50$ μs, D ranging from 0 to 100 μs. Curves are plotted for $\Delta = 10$ μs and $\Delta = 5$ μs (points connected by dotted lines), $\Delta = 1$ μs (solid black curves), and the CT case (solid grey line, the one with highest loss).

Just as it was the case for memoryless burst sizes, also for fixed burst sizes, higher values for the time slot lengths result in lower LP. Here, however, the impact of slot size variation is larger, and the curves all display the local optima ("notches"), typical for the deterministic burst size distribution. Since all curves for DT assumed that at most one arrival can occur per slot, we can again draw no conclusions on the exact impact of synchronization, but rather refer to Sect. 2.3.

2.2.8.3 Mixtures of Deterministic Burst Sizes

It is rather straightforward to extend the above results to mixtures of for example deterministic burst sizes. Burst lengths then take on a limited number of values B_i ($i = 1, \ldots, R$) with probabilities α_i ($\sum \alpha_i = 1$). We again express each B_i as $a_i D - b_i$ as above. The limit procedure results in

$$W^{\mathcal{L}}(s) = \frac{-\sum_{i=1}^{R} \alpha_i \left\{ \left(a_i \left(1 - e^{-\lambda D}\right) - e^{-\lambda(D-b_i)} \right) \left(1 - e^{-sD}\right) \right\}}{\sum_{i=1}^{R} \alpha_i \left\{ e^{-sa_iD} e^{-\lambda(D-b_i)} \left(1 - e^{-sD}\right) + \left(e^{-\lambda D} - e^{-sD}\right) \left(1 - e^{-sa_iD}\right) \right\}},$$



(a) Buffer size fixed to $N = 20$.

(b) Load fixed to $\rho = 60\%$.

Figure 2.4: The analysis provides an accurate approximation of the loss probability also for a mixture of deterministic burst sizes. These figures were obtained for varying D (in μs), $\mathrm{E}[B] = 50$ μs, with $\Pr[B = 45] = 0.75$ and $\Pr[B = 65] = 0.25$.

and

$$\rho_{eq} = 1 + \lambda D \left(\sum_{i=1}^{R} \alpha_i \left\{ a_i - \frac{e^{-\lambda(D-b_i)}}{1 - e^{-\lambda D}} \right\} \right) .$$

The dominant pole $\gamma = e^{-s_0 D}$ has to be determined from

$$\sum_{i=1}^{R} \alpha_i \left\{ \gamma^{a_i} e^{-\lambda(D-b_i)} (1-\gamma) + (e^{-\lambda D} - \gamma)(1 - \gamma^{a_i}) \right\} = 0 .$$

The constant needed in the approximation of $\Pr[H_\infty > ND]$ is given by

$$C(ND) = \frac{-\sum_{i=1}^{R} \alpha_i \left\{ a_i \left(1 - e^{-\lambda D}\right) - e^{-\lambda(D-b_i)} \right\}}{\sum_{i=1}^{R} \alpha_i \left\{ a_i \gamma^{a_i} (\gamma - e^{-\lambda D}) - \gamma(1-\gamma^{a_i}) - \gamma^{a_i}(\gamma + \gamma a_i - a_i) e^{-\lambda(D-b_i)} \right\}} .$$

Results for $R = 2$ are shown in Fig. 2.4(a). Burst sizes are $B_1 = 45$ μs and $B_2 = 65$ μs with probability $\alpha_1 = 0.75$ and $\alpha_2 = 0.25$ respectively, the average burst size being 50 μs again. The shape of the curves clearly resembles those depicted in Fig. 2.3, and the "local optima" induced by the predominant 45 μs

burst sizes can easily be distinguished. However, the optima are not as pronounced as in the (single-valued) deterministic case, due to the presence of the 65 μs burst sizes. The "local optima" induced by the latter can be observed as "notches" in the curves. The example shows that the presence of different burst sizes alleviates the tight connection between average burst size and optimal granularity.

To conclude this section, Fig. 2.4(b) shows results for varying buffer depths N. The overall shape of the curves does not change drastically with N, but the global optimum granularity can. For $N = 5$ and $N = 10$, the optimum is at $D = 45$ μs, while for $N = 20$, it is at $D = 22.5$ μs. Especially in the latter case, the optimum is rather broad, in that a nearly constant LP is observed for all values of D between 22.5 μs and 45 μs.

2.3 Model for Synchronization

While the results in the previous section provide a transform-based modeling for both CT and DT, they do not provide a straight-forward answer to the influence of synchronization on the performance of optical buffers. More particularly, the assumption of a Bernoulli arrival process in DT does not allow to model synchronization, that typically results in multiple arrivals per slot, also known as batch arrivals. As such, the model presented in this section provides an extension of the DT model of the previous section to allow for a batch arrival process. This is then utilized to study the impact of internal synchronization, a process at node level that can be implemented by retiming incoming traffic at the input. In optical network design, synchronization within the network nodes offers several advantages. Although the retiming itself is considered hard to implement (see, for example, [42]), it can reduce the complexity of the node's control logic, and make technologically demanding tasks such as header extraction more feasible. Further, it can better loss performance, in a way that is researched in the current section. Results presented below show that synchronization can improve the loss performance of an optical buffer spectacularly, especially in the case of fixed-sized bursts.

2.3.1 Buffer Setting

The buffer setting is identical to the one treated in the previous section, and is described at the start of this chapter, Sect. 2.1.2.

2.3.2 Evolution of the Scheduling Horizon

Given that the current model allows for the arrival of batches, it is clear that the system equation for single arrivals, (2.1), is to be adapted to this purpose. We again start out with the system of infinite buffer size. We consider arriving batches,

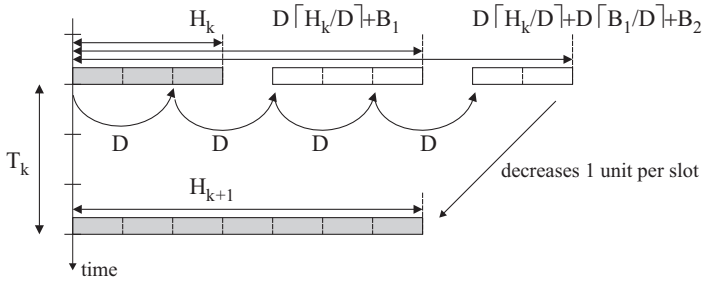


Figure 2.5: The evolution of the scheduling horizon in case of batch arrivals.

which we number in the order of their arrival. We define the scheduling horizon as the time at which all bursts will have left the buffer, and the system will become empty again. Recall that, due to the occurrence of voids, the scheduling horizon is in general larger than the unfinished work, and thus incorporates the effects of granularity. The variable H_k is derived thereof, and is defined as the scheduling horizon as seen by the kth arriving batch, just before its arrival. The arriving batch contains a number of bursts X_k ($X_k \geq 1$), each of which has an associated burst length $B_{k,i}$, $i = 1 \ldots X_k$, which equals the time needed for its transmission. The time between the kth batch arrival and the next is captured by the batch inter-arrival time, denoted by T_k. The evolution of these variables is described by

$$H_{k+1} = \left[D \left\lceil \frac{H_k}{D} \right\rceil + \sum_{i=1}^{X_k - 1} D \left\lceil \frac{B_{k,i}}{D} \right\rceil + B_{k,X_k} - T_k \right]^+ , \qquad (2.38)$$

and is illustrated in Fig. 2.5. When the kth batch sees a non-zero scheduling horizon H_k upon arrival, the first burst of the batch will have to be delayed for at least that amount to avoid contention. As the optical buffer can only realize delays that are a multiple of D, we obtain that the first burst has to be delayed for an amount $D\lceil H_k/D \rceil$. The second burst of the batch (in case there is one, that is, $X_k > 1$) has to be delayed too, in order for it not to overlap with the first. Again, only multiples of D are realizable, so that the second burst needs an additional amount of delay $D\lceil B_{k,1}/D \rceil$. Applying the same reasoning to possible other bursts in the batch, we obtain that delaying and transmitting batch k pushes the scheduling horizon (just after arrival) to $D\lceil H_k/D \rceil + \sum_{i=1}^{X_k-1} D\lceil B_{k,i}/D \rceil + B_{k,X_k}$. Taking into account the batch inter-arrival times T_k, and the possibility that the system becomes empty in between arrivals, one arrives at (2.38). This is illustrated in Fig. 2.5, where the kth batch carries two bursts. For the given setting, with $D = 2$, $H_k = 3$, $B_{k,1} = 3$, $B_{k,2} = 2$ and $T_k = 3$, one can verify that indeed $H_{k+1} = 7$, namely $H_{k+1} = [2\lceil 3/2 \rceil + 2\lceil 3/2 \rceil + 2 - 3]^+ = [4 + 4 + 2 - 3]^+ = 7$.

2.3.3 Traffic Setting

To analyze (2.38), we need to impose restrictions on the involved variables T_k, X_k and $B_{k,i}$, that are similar to the assumptions of the DT model in Sect. 2.2. The T_k have a geometric distribution just as in Sect. 2.2. The batch sizes X_k and burst sizes $B_{k,i}$ also form a sequence of iid rv's, and can have a general distribution, thus being independent of the index k and, in the case of the burst size, also i. Again, we work with the pgf's of the involved rv's, denoted as $T(z)$, $X(z)$ and $B(z)$.

2.3.4 Analysis

We start out with an analysis of the infinite system whereby the system is assumed to be stable. On this condition, to be discussed later on, the distributions of H_k converge, for $k \to \infty$, to a unique stochastic equilibrium distribution, independent of the initial system conditions. Let H denote a generic rv following this equilibrium distribution. Likewise, we will drop the index k for other variables involved. As such, due to (2.38),

$$H = \left[D \left\lceil \frac{H}{D} \right\rceil + \sum_{i=1}^{X-1} D \left\lceil \frac{B_i}{D} \right\rceil + B_X - T \right]^+ . \tag{2.39}$$

Furthermore, we denote

$$F = D \lceil H/D \rceil \ ; \ E_i = D \lceil B_i/D \rceil \ . \tag{2.40}$$

Note that F corresponds to the waiting time of the first burst in a batch; since this however differs from the waiting time of an arbitrary burst, commonly denoted by W throughout this work, we avoid confusion by introducing F as notation, and refer to it as *waiting time of batches*. The other measure, denoted by the rv E, we refer to as *augmented burst size*.

To solve (2.39), we follow the same approach as in the DT model in Sect. 2.2, by solving the two non-linear effects sequentially: the queueing effect, related to the non-negativeness of the buffer content; the granularity effect, related to the finite granularity of the FDLs.

The queueing effect can again be formulated as

$$H = [G - T]^+ ,$$

where

$$G = F + \sum_{i=1}^{X-1} E_i + B_X \, . \tag{2.41}$$

The variable T is the inter-arrival time of batches (and not of distinct bursts, as is the case in the rest of this work). As T is geometrically distributed, and independent of G, the solution in terms of pgf's is exactly the one mentioned in Sect. 2.2,

(2.18):

$$H(z) = \frac{p}{z-\bar{p}} G(z) + K \frac{z-1}{z-\bar{p}}\,, \tag{2.42}$$

with now

$$G(z) = F(z) \frac{X(E(z))}{E(z)} B(z)\,, \tag{2.43}$$

and K a constant that will be determined later on. Note that (2.43) follows directly from (2.41), because the rv's involved are independent.

Also for the granularity effect, we rely on a previously obtained expression (2.21) to obtain

$$\begin{aligned} F(z) &= \sum_{k=0}^{D-1} \frac{1}{D} \frac{z^D-1}{z\varepsilon_k - 1} H(z\varepsilon_k)\,, \\ E(z) &= \sum_{k=0}^{D-1} \frac{1}{D} \frac{z^D-1}{z\varepsilon_k - 1} B(z\varepsilon_k)\,, \end{aligned} \tag{2.44}$$

where $\varepsilon_k = \exp(j2\pi k/D)$ as before, for $k = 0 \ldots D-1$ (the D different complex D^{th} roots of unity).

Combining (2.42) and (2.44), we find after some simplification in the numerator,

$$F(z) = \frac{K \cdot \left[\dfrac{\bar{p}^{D-1}(z^D-1)}{z^D - \bar{p}^D} \right]}{1 - \left[\displaystyle\sum_{k=0}^{D-1} \frac{1}{D} \frac{z^D-1}{z\varepsilon_k - 1} \frac{pB(z\varepsilon_k)}{z\varepsilon_k - \bar{p}} \right] \cdot \dfrac{X(E(z))}{E(z)}}\,. \tag{2.45}$$

This expression is the pgf of F, the waiting time of the first burst of an arriving batch. We further remark that this expression is similar to (2.23). The difference is captured by the factor $X(E(z))/E(z)$, which is the pgf of $\sum_{i=1}^{X-1} E$. This sum is associated with the total amount of work brought about by all bursts that do not arrive last. In the case of single burst arrivals, $X(z) = z$, and this factor simplifies to 1, and we obtain exactly (2.23).

The next task is to determine the constant K occurring in the formula for $F(z)$. This can be done by demanding $F(1)$ to be one, as prescribed by the normalization condition. By applying the rules of de l'Hôpital, we find

$$K = \frac{1-\bar{p}^D}{D\bar{p}^{D-1}} \cdot \left(\frac{1}{p} + \sum_{k=1}^{D-1} \frac{B(\varepsilon_k)}{\varepsilon_k - \bar{p}} - \mathrm{E}[X] \cdot \left[\mathrm{E}[B] + \frac{D-1}{2} + \sum_{k=1}^{D-1} \frac{B(\varepsilon_k)}{\varepsilon_k - 1} \right] \right)\,. \tag{2.46}$$

Now, the expressions for $H(z)$ and $F(z)$ can be obtained as a combination of (2.42), (2.45) and (2.46).

2.3.5 Maximum and Equivalent Load

The probability of finding the system empty upon arrival of a batch, relates to the constant K, as

$$H(0) = \lim_{k\to\infty} \Pr[H_k = 0] = \frac{K}{\bar{p}}\,.$$

Demanding the infinite system to be stable, is equivalent to demanding this probability not to be zero. The condition $H(0) > 0$ defines a maximum tolerable arrival intensity p_{max}, which is the solution of

$$\frac{1}{p} = \mathrm{E}[X] \cdot \left[\mathrm{E}[B] + \frac{D-1}{2} + \sum_{k=1}^{D-1} \frac{B(\varepsilon_k)}{\varepsilon_k - 1}\right] - \sum_{k=1}^{D-1} \frac{B(\varepsilon_k)}{\varepsilon_k - \bar{p}}\,. \tag{2.47}$$

We note that this only involves the mean of the batch size distribution $\mathrm{E}[X]$, and any type of batch size distribution with the same mean thus produces the same p_{max}.

Presently we can also define an equivalent load,

$$\rho_{eq} = p \cdot \left\{\mathrm{E}[X] \cdot \left[\mathrm{E}[B] + \frac{D-1}{2} + \sum_{k=1}^{D-1} \frac{B(\varepsilon_k)}{\varepsilon_k - 1}\right] - \sum_{k=1}^{D-1} \frac{B(\varepsilon_k)}{\varepsilon_k - \bar{p}}\right\}\,, \tag{2.48}$$

that incorporates the effect of voids, and is 100% when $p = p_{max}$.

2.3.6 Maximum and Equivalent Load, Revisited

To allow for more insight in the formulation of ρ_{eq}, it proves useful to also formulate the alternate description of the system evolution here, now in terms of the waiting times of batches. Taking into account (2.38) and (2.40), and denoting $F_k = D\lceil H_k/D\rceil$ and $E_{k,i} = D\lceil B_{k,i}/D\rceil$, one can show that the F_k evolve as

$$F_{k+1} \stackrel{*}{=} \left[F_k + \sum_{i=1}^{X_k-1} E_{k,i} + D\left\lceil \frac{B_{k,X_k} - T_k}{D}\right\rceil\right]^+\,,$$

and, assuming equilibrium for $k \to +\infty$,

$$F \stackrel{*}{=} \left[F + \sum_{i=1}^{X-1} E_i + D\left\lceil \frac{B - T}{D}\right\rceil\right]^+\,.$$

Note that these expressions are independent of time setting, and also are valid for general iid distributions of the inter-arrival times, burst sizes and batch sizes, They allow to follow the same reasoning as in Sect. 2.1.5. Clearly, the term $\sum_{i=1}^{X-1} E_i + D\left\lceil \frac{B-T}{D}\right\rceil$ generates the process of the waiting times of batches, and stability requires thus that

$$(\mathrm{E}[X] - 1) \cdot \mathrm{E}\left[D\left\lceil \frac{B}{D}\right\rceil\right] + \mathrm{E}\left[D\left\lceil \frac{B-T}{D}\right\rceil\right] \stackrel{*}{<} 0\,,$$

which, in the case of geometric inter-arrival times, results in (2.47). In the same manner, one can define a general equivalent load as

$$\rho_{eq} \stackrel{*}{=} 1 + \frac{(\mathrm{E}[X]-1)\cdot \mathrm{E}\left[D\left\lceil\frac{B}{D}\right\rceil\right] + \mathrm{E}\left[D\left\lceil\frac{B-T}{D}\right\rceil\right]}{\mathrm{E}[T]},$$

which is indeed consistent with (2.48).

2.3.7 Heuristics for the Loss Probability

Since the results of the previous section are only valid for infinite buffer size, we rely on the heuristics of Sect. 2.1.6 to obtain the LP for the finite system. Note that the term "LP" still denotes the probability of the loss of an arbitrary burst, but that, when compared to Sect. 2.1.6, an additional approximation is made. More precisely, the LP is approximated by the probability that an arbitrary batch (instead of burst) is lost. Since both H and F are associated with an arriving batch, we thus follow the easiest way to obtain results, that, as we will show, are sufficiently accurate. The possible extension that considers the scheduling horizon of (individual) bursts is discussed at the end of this section.

With the above expression for ρ_{eq} given, we only need to establish expressions for the $\Pr[H_\infty > ND]$ to be able to apply heuristic A (2.14) and B (2.15). These tail probabilities we derive in the same manner as suggested in Sect. 2.1.6, with (2.11),

$$\Pr[H > n] \approx -\sum_k Res\left[\frac{1}{z^{n+1}}\cdot\frac{H(z)-1}{z-1}\right]_{z=z_k},$$

where the summation index k only runs over the poles z_k of $H(z)$ with smallest modulus. Aiming for a dominant pole approximation in the sense of (2.12) in DT, we are to determine the dominant poles of $(H(z)-1)/(z-1)$, that are also the dominant poles of $H(z)$ and, given (2.42), of $G(z)$. The latter being a product of $F(z)$, $B(z)$ and $X(E(z))/E(z)$ (see (2.43)), the assumption made on $H(z)$ has implications for $B(z)$ and $X(z)$. A sufficient condition is to assume that $B(z)$ and $X(z)$ have no singularities (if any) other than poles. This assumption includes the broad class of rational pgf's, and also frequently used others, such as the pgf of the Poisson distribution. Excluded is the class of heavy-tailed distributions. As mentioned earlier in Sect. 2.1.6, this assumption allows us to apply the dominant pole approximation but, even for heavy-tailed distributions, all other derivations from the previous sections remain valid.

With these additional assumptions, we are in the position to use (2.11), retaining the dominant poles of $F(z)$. Note that the poles of $X(E((z))/E(z))$ and $B(z)$ do not interfere, since those of $F(z)$ have smaller absolute value. Their number is D, each has multiplicity one, and is of the form $z_k = z_0\varepsilon_k$ ($k = 0 \ldots D-1$), with

z_0 the positive real one. Applying this allows to write the approximate relation as

$$\begin{aligned}
&\Pr[H > n] \\
&\quad \approx \quad -\frac{\lim_{z\to z_0}[F(z)(z-z_0)]}{z_0^{n+1}} \cdot \sum_{k=0}^{D-1} \frac{1}{\varepsilon_k^n} \frac{B(z_0\varepsilon_k)}{z_0\varepsilon_k - 1} \frac{p}{z_0\varepsilon_k - \bar{p}} \frac{X(E(z_0))}{E(z_0)} \,, \\
&\quad = \quad \frac{C(n)}{z_0^n} \,, \qquad (2.49)
\end{aligned}$$

where we introduced the notation $C(n)$ to lay emphasis on the quasi-geometrical tail decay, with decay rate z_0. The function $C(n)$ is periodical, that is, $C(n + mD) = C(n)$ ($m = 0, 1, 2 \ldots$). The values that will be of interest here are

$$C(ND) = -\frac{1}{z_0} \lim_{z\to z_0}[F(z)(z-z_0)] \frac{D}{z_0^D - 1} \,,$$

which can be found from (2.49), using the fact that z_0 is a pole of F (see (2.45)), that is,

$$1 - \left[\sum_{k=0}^{D-1} \frac{1}{D} \frac{z_0^D - 1}{z_0\varepsilon_k - 1} \frac{pB(z_0\varepsilon_k)}{z_0\varepsilon_k - \bar{p}}\right] \cdot \frac{X(E(z_0))}{E(z_0)} = 0 \,.$$

The latter yields a value for z_0 in the interval $(1, +\infty)$ if $\rho_{eq} < 1$. The same heuristics can also be used to evaluate the LP in overloaded systems ($\rho_{eq} > 1$), by (formally) computing the quantities $\Pr[H_\infty > ND]$, with the same expression for $C(ND)$, for a z_0 that is to be found in $(0, 1)$.

Finally, translating these results to the LP of a random burst is possible by means of pgf's, by quantifying the number of bursts within the same batch that precede the random burst. The associated pgf is obtained for example in [93] (p.20) and can be used as a refinement. Nevertheless, the numerical comparison presented in the next section illustrates that, even with the simpler approach that we applied, the approximation already attains good accuracy.

2.3.8 Numerical Comparison

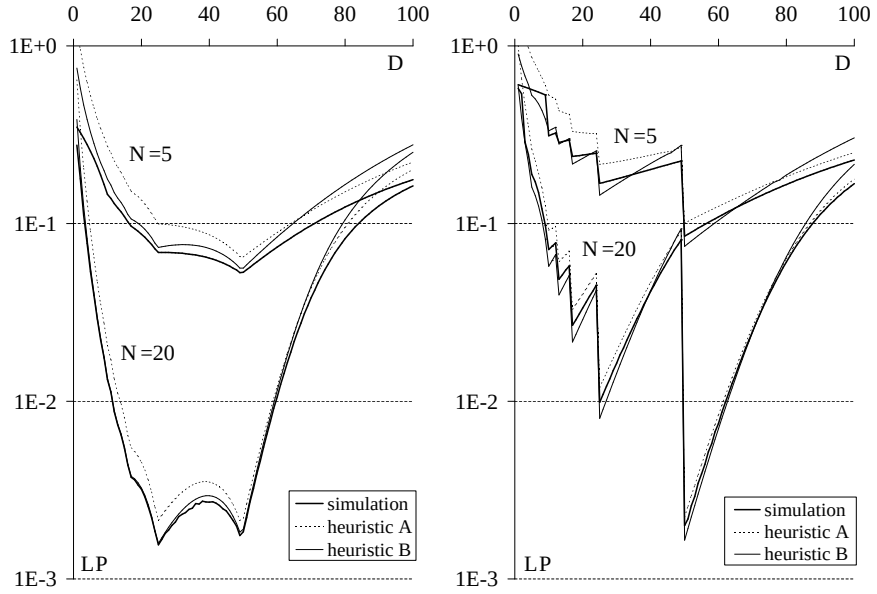
For ease of notation, we introduce an additional rv A, defined as the number of arrivals in a random slot, which is perhaps a more natural way of describing the arrival process in DT. Of course, A, T and X are related. One can easily show that $\Pr[A = 0] = \bar{p}$, and

$$\Pr[A = n \,|\, A > 0] = \Pr[X = n] \,,\ n \in \mathbb{N}_0 \,,$$

and

$$A(z) = \bar{p} + p \cdot X(z) \,,\ z \in \mathbb{C} \,,$$

so that $A(z)$ determines $T(z)$ and $X(z)$.



(a) Poisson batch size distribution.

(b) Deterministic batch size distribution ($X = 2$).

Figure 2.6: Both for classic and more exotic batch size distribution, the heuristics provide an accurate match of the simulation results. These figures were obtained for varying D (in μs), slot size 1μs, deterministic burst size distribution with $\mathrm{E}[B] = B = 50$ *μs, and $\rho = 60\%$.*

We now apply the obtained results to special cases of the burst and batch size distribution. On Fig. 2.6(a), the LP as a function of granularity is considered, for a deterministic burst size distribution, with value $\mathrm{E}[B] = B = 50\mu$s, with slot size fixed to 1μs. The batch size distribution is chosen Poisson. In that case, A is distributed as

$$\Pr[A = k] = e^{-\lambda}\frac{\lambda^k}{k!}\ ,\ k \in \mathbb{N}\,. \tag{2.50}$$

where the parameter λ denotes the average number of arrivals per slot. Further, the load was fixed to $\rho = 60\%$. For buffer sizes $N = 5$ and $N = 20$, results from simulation, heuristic A and heuristic B are compared. As was remarked in Sect. 2.1.6.2 and 2.2.8, heuristic B performs better than A around the optima (small LP), and attains very high accuracy for large buffer sizes. Heuristic A performs better than B only for high values of the granularity. The curves differ only little from the case of Bernoulli arrivals, as the probability of observing more than a single arrival, $\Pr[A > 1]$, is rather small. Results for other values of the load, not included here, reveal similar plots. The accuracy of heuristic B remains high, especially when the LP drops below 10^{-2}.

To further assess the accuracy of our heuristics, we also take a look at a more exotic case of a deterministic batch size distribution, with $\mathrm{E}[X] = X = 2$. Figure 2.6(b) shows curves of the LP as a function of granularity, again for a fixed load of $\rho = 60\%$, with buffer sizes $N = 5$ and $N = 20$. The match between heuristic B and simulations is again very good. As in the case of the single arrivals, local optima occur, but, here, performance is even more sensitive to the value of D.

Having sufficient confidence in the accuracy of the heuristics, we now move on to the study of synchronization.

2.3.9 Synchronization Study

2.3.9.1 Imposing a Slotted Structure

By synchronization we refer to the transformation from an asynchronous (or, unslotted) setting to a setting with a slot length Δ, at node level. This means that, within a network, this synchronization uses a local clock, and only has impact on traffic within the node.

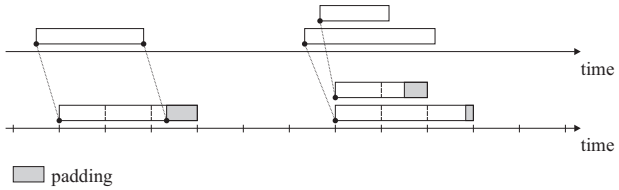


Figure 2.7: Synchronization results in two effects: batch arrivals (multiple arrivals at a slot edge) and padded burst sizes (indicated in grey).

The process is illustrated in Fig. 2.7. In the initial setting, depicted on the upper time axis, arrivals can occur at the input at any instant, and are assumed to happen one at a time. In the synchronized environment, depicted on the lower time axis, bursts are forced to arrive at slot boundaries by imposing a small delay, that can take on values between 0 and Δ. Due to this retiming effect, several bursts can now arrive at the same slot boundary, that is, they can arrive in batches.

For the burst sizes, a different effect comes about. Once synchronized, traffic patterns are captured using discrete variables, that express an integer number of time slots involved. Burst sizes are thus virtually increased to an integer multiple of slots. This effect, that adds to the burst size an amount ranging from 0 to Δ, is denoted padding, as indicated in grey in Fig. 2.7. Remark that the padding results from the interplay between burst size and slot length only, and that the exact arrival instant has no impact on this. Also, note that, if the burst size is already an integer multiple of the slot length, no padding occurs. In other situations, padding can account for serious loss, as it increases the mean value of the burst size, and thus the load. Instances of both situations will be given next.

2.3.9.2 Poisson Batch Sizes and Padded Burst Sizes

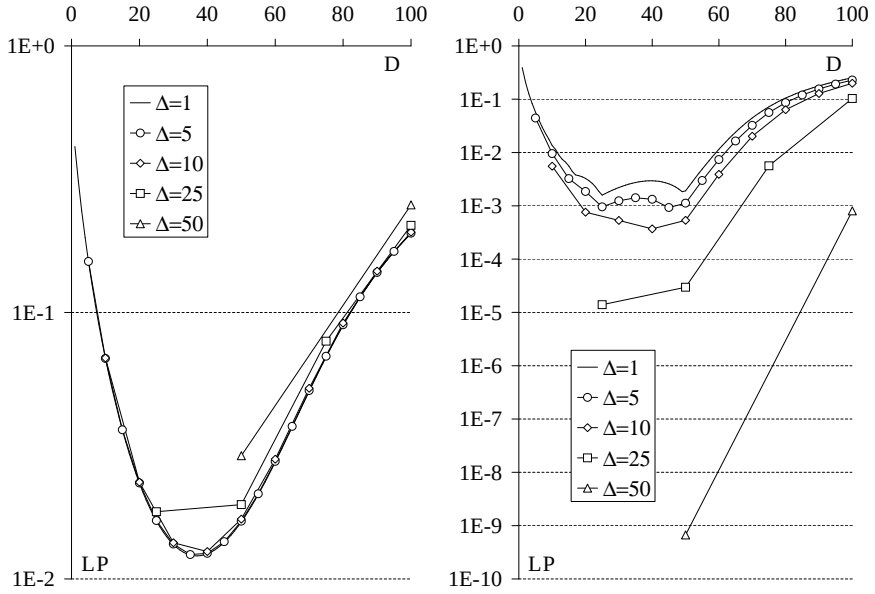
With this definition of synchronization, we can now look at the impact on the number of arrivals per slot A, and the batch size B. The number of arrivals A per slot in a synchronized setting is derived from the number of arrivals A that can be expected to arrive in the asynchronous setting, in a time period of length Δ. In general, this derivation is a non-trivial one, and it is only in some special cases, such as a Poisson arrival process, that synchronization results in a number of arrivals that is distributed iid from slot to slot. In that case, A has a Poisson distribution, with parameter λ, and pmf as in (2.50). We find that A remains Poisson distributed regardless of the slot length, with parameter $\lambda = \Delta \cdot \rho / \mathrm{E}[B]$ scaling linearly with this slot length. This allows us to study the impact of synchronization within the single framework of Poisson batch arrivals. We note that this is not possible if A would have a Bernoulli distribution, as a Bernoulli arrival process is not the result of a synchronization process. For that reason, the comparison made in Sect. 2.2.8 (Fig. 2.3) for geometric and deterministic burst size distribution became less and less accurate with increasing Δ. Also, there, the effect of padding was not taken into account.

2.3.9.3 Numerical Examples

In this section, we take a look at two synchronization settings. Both assume a Poisson arrival process in the asynchronous setting, with single arrivals, that is then synchronized according to a slot length Δ, with the arrival of a batch per slot. As mentioned above, the modeling thereof implies using Poisson batch sizes and padded burst sizes. The load was fixed to $\rho = 60\%$, with a buffer size $N = 20$. Recall that this load applies before any synchronization is done. Since heuristic B yields more accurate results than heuristic A, we will only show heuristic B. Simulation results not included here, confirmed that this heuristic was accurate.

On Fig. 2.8(a), we present the LP for the case of geometric burst sizes, as a function of the granularity. The mean burst size before synchronization is $\mathrm{E}[B] = 50\ \mu$s, which corresponds to approximately 60 kB at 10 Gbit/s. (A rather high value like this one is possible in OBS, where bursts are typically large, as they consist of an aggregation of packets gathered at the edge nodes.) Different slot lengths, ranging from 1 μs to 50 μs, are applied, and the mean burst size increases due to padding. Clearly, synchronization does not have a large impact on LP, and the curves overlap almost completely. Only for large slot lengths ($\Delta = 25\ \mu$s and $\Delta = 50\ \mu$s), the LP alters a little bit. When considering the effects of retiming and padding separately, we found that, while retiming actually betters performance, padding annihilates this benefit.

On Fig. 2.8(b), a similar setting is assumed, now with a deterministic burst size of $B = \mathrm{E}[B] = 50\ \mu$s. Clearly, synchronization now has a huge impact on perfor-



(a) Geometric burst size distribution.

(b) Deterministic burst size distribution ($B = 50\mu$s).

Figure 2.8: How synchronization impacts loss performance depends largely on the burst size distribution. If burst sizes vary, like in case of a geometric burst size distribution, the benefits of synchronization are countered by the effects of padding. If burst sizes are fixed, the LP lowers spectacularly, especially if the threesome of slot length Δ, burst size and granularity D (all three in μs) match. These figures were obtained for $\mathrm{E}[B] = 50\mu$s, $\rho = 60\%$, and $N = 20$. The batch sizes are distributed according to a Poisson distribution, with parameter $\lambda = \Delta\rho/\mathrm{E}[B]$.

mance, especially for large slot lengths. If we consider the loss for $D = B = 50$ μs, we find that LP is reduced more than 10^6 times. Looking for the cause of this drastic reduction, we find that, due to the good choice of slot lengths, burst sizes could always be expressed as a multiple of slots. Therefore, no padding comes about, and B is 50 μs in both the asynchronous and synchronized setting. On the other hand, the benefits of retiming are bigger than in the case of a geometric burst size distribution, and, for a slot length $\Delta = D = B = 50$ μs, voids are even canceled out completely. Especially for $\Delta = D = B$, it can be understood intuitively that the performance of the FDL buffer must benefit from synchronization. Indeed, except for the extra delay due to synchronization, the optical buffer realizes exactly the same buffering behavior as a classic RAM buffer if $\Delta = D = B$.

Taking into account both examples, we conclude that synchronization can lower LP significantly if the burst size is fixed. In that case, we consider retiming

a viable and effective way to mitigate LP drastically, and cost and effectiveness should be weighed to decide wether or not to implement this type of retiming. Further, synchronization is found to be of little merit (with respect to buffer performance) if burst sizes are distributed exponentially. Although not investigated explicitly, the intermediary case where only a limited number of burst sizes plays a role is expected to also allow for a bettering in loss performance. This benefit is expected to be not as drastic as in the case of fixed-sized bursts, yet still significant, if the synchronization slot length is chosen well.

2.4 Model for General Arrivals

While results up to now were obtained under the restrictive assumptions of a memoryless arrival process, the aim of the current section is to alleviate the restrictive assumptions of the M/G/1 buffer model of Sect. 2.2.3, and perform the analysis for general independent arrivals. The latter we intend in the sense of the (extended) Kendall notation: a GI/G/1 model for degenerate infinite-sized optical buffers in DT, with general independent (GI) distribution of the inter-arrival times, with as sole restriction on the inter-arrival times the requirement to have a rational pgf. Such GI arrival process indeed allows to generate a correlated stream of arrivals (with *general* inter-arrival times), as will be treated in this section. Since inter-arrival times still constitute a series of iid rv's, this form of correlation is only a mild variant of correlation, that does not result in the severe performance degradation typical for heavily-correlated input processes, but nevertheless serves the purpose of qualifying the impact of correlation with the lightest-possible model. Since arrival processes in (optical) communication networks are known to be bursty [97], such study of the impact of correlation in the arrival process on performance is crucial. Further, while a generalization of the arrival process naturally gives rise to a study of the impact of correlation, note that the next section will employ the same model for different means, namely to study a buffer with wavelength conversion. Finally, note that the results of this section were disseminated first in [6]. Since the analytic results are obtained from an extension of the DT M/G/1 model of Sect. 2.2, we will rely on previous results in an incremental way, and mainly focus on the way to extend the results of that section.

2.4.1 Buffer Setting and System Equation

As said, the current section treats the extension of previous results in an incremental way, and the buffer setting and system equation are even identical to the one assumed in Sect. 2.1.2 for DT. However, despite the fact that those aspects remains unaltered, the problem now harbors more complexity, since both the T_k and B_k constitute iid rv's with general distribution. This will complicate the analysis in

a non-trivial way, especially in solving the queueing effect, and in the combination of granularity effect and queueing effect.

2.4.2 Traffic Setting

As for the burst sizes B_k and inter-arrival times T_k mentioned in (2.1), we assume both to form a sequence of iid rv's. In our analysis, we will again use their associated pgf, and will consider also the pgf of the other variables H_k, G_k and W_k. Note that, in order to allow for exact analysis, the additional assumption of a rational pgf for the inter-arrival times will be made in the next section. Also, we make the usual assumption that $\mathrm{E}[T_k], \mathrm{E}[B_k] < \infty$.

2.4.3 Analysis

The analysis assumes a system of infinite size, and assumes stable regime. The approach consists in solving the queueing and the granularity effect separately. A crucial third step is then to combine both correctly, which requires the use of Rouché's theorem.

2.4.3.1 Conditions

From this point on, we assume that the system is stable. The condition for stability is in accordance with Sect. 2.1.5, and is investigated below in Sect. 2.4.4. Moreover, different from the case of memoryless arrivals, the combination of general inter-arrival times and general burst sizes might in some "pathological" cases result in a reducible Markov chain for the scheduling horizon. Since these situations are of limited practical relevance, we exclude them in the following, and refer to Sect. 5.1.4.3 for further remarks on this assumption. At any rate, assuming the load is below the tolerable load, and assuming an irreducible and aperiodic Markov chain for the scheduling horizon, the distributions of H_k (and likewise, for the other variables G_k and W_k) converge, for $k \to \infty$, to a unique stochastic equilibrium distribution, that no longer relates to the initial condition of the system. Associated with this distribution is a common rv H, and a pgf $H(z)$, of which we will derive the explicit form below. For the inter-arrival time distribution, we assume that its pgf $T(z)$ is rational. Both the numerator and denominator are thus polynomials of finite degree. This assumption is not very restrictive from a modeling point of view, and includes the distributions treated further in Sect. 2.4.6.

2.4.3.2 The Queueing Effect

Since we assume the system is stable, we analyze the queueing effect as it occurs for the stochastic equilibrium distributions, and (2.3) becomes

$$H =_{st} [G - T]^+ \,, \tag{2.51}$$

where "$=_{st}$" denotes stochastic equivalence. We emphasize that G and T (and, G_k and T_k) are statistically independent, which is essential to our current analysis. Now, the complexity of a transform-based solution to this problem depends critically on the exact form of the distribution, or equivalently, the pgf of T, as discussed in for example [93].

To solve (2.51), we write the rational pgf of T as

$$T(z) = \frac{N(z)}{P(z)} = \frac{\sum_{i=0}^{N} n_i z^i}{\sum_{i=0}^{P} p_i z^i} \,. \tag{2.52}$$

In other words, we label the numerator as a polynomial $N(z)$ of degree N, and the denominator as a polynomial $P(z)$ of degree P. Under these assumptions, Eq. (2.52) lends itself to an exact analysis, involving an auxiliary function $\tilde{T}(z)$. In the first step, we rewrite the pgf of T, so as to facilitate the analysis of the second step, that brings about $\tilde{T}(z)$.

Rewriting the pgf of T The pgf of $T(z)$ mentioned in (2.52) can be rewritten as follows,

$$T(z) = \sum_{i=0}^{R} r_i z^i + \frac{\sum_{i=0}^{Q} q_i z^i}{\prod_{j=1}^{C} (z - \gamma_j)^{m_j}} = \sum_{i=0}^{R} r_i z^i + \sum_{j=1}^{C} \sum_{i=0}^{m_j - 1} \frac{b_{ij}}{(\gamma_j - z)^{i+1}} \,, \tag{2.53}$$

where we introduced R, r_i, Q, q_i and b_{ij} for notational convenience. More precisely, we subsequently isolated the polynomial part of the fraction, to then apply a partial fraction expansion to the second term. The γ_j ($j = 1 \ldots C$) are the C different poles of $T(z)$, each with associated multiplicity m_j ($j = 1 \ldots C$). Utilizing the probability generating property of a pgf, one can express the probabilities t_k ($k \in \mathbb{N}_0$), corresponding with $T(z)$, as

$$t_k = \frac{1}{k!} \frac{\partial^k}{\partial z^k} T(z)|_{z=0} = \sum_{i=0}^{R} r_i \delta_{i,k} + \sum_{j=1}^{C} \sum_{i=0}^{m_j - 1} \binom{i+k}{i} \frac{b_{ij}}{\gamma_j^{i+k+1}} \,,$$

where $\delta_{i,k}$ still denotes the Kronecker delta. Reordering the binomial coefficients $\binom{i+k}{i}$, and introducing coefficients c_{ij}, we arrive at

$$t_k = \sum_{i=0}^{R} r_i \delta_{i,k} + \sum_{j=1}^{C} \sum_{i=0}^{m_j - 1} \frac{k^i}{\gamma_j^k} c_{ij} \,. \tag{2.54}$$

Obtaining the pgf of H With the definition of a pgf as a starting point, we apply (2.52) to obtain

$$\begin{aligned} H(z) &= \sum_{m=0}^{\infty} h_m z^m \\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} g_n t_k z^{[n-k]^+} \\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} g_n t_k z^{n-k} + \sum_{n=0}^{\infty}\sum_{k=n}^{\infty} g_n t_k (z^0 - z^{n-k}) \\ &= G(z)T(z^{-1}) + \sum_{k=0}^{\infty}(1-z^{-k})\sum_{n=0}^{\infty} g_n t_{k+n}\,, \end{aligned} \tag{2.55}$$

where g_n and t_k are the probabilities corresponding with $G(z)$ and $T(z)$ respectively, and $[x]^+$ is shorthand for $\max\{x, 0\}$. We introduce the coefficients $\tilde{t}_k$ ($k \in \mathbb{N}$),

$$\tilde{t}_k = \sum_{n=0}^{\infty} g_n t_{k+n} = \sum_{n=0}^{\infty} g_n \left(\sum_{i=0}^{R} r_i \delta_{k+n,i} + \sum_{j=1}^{C}\sum_{m=0}^{m_j-1} \frac{(k+n)^m}{\gamma_j^{k+n}} c_{mj} \right)\,, \tag{2.56}$$

where we used (2.54) in the last transition. The first term in this expression we can rewrite as

$$\sum_{n=0}^{\infty} g_n \sum_{i=0}^{R} r_i \delta_{k+n,i} = \sum_{i=0}^{min(k,R)} r_i g_{i-k} = \tilde{r}_k = \sum_{i=0}^{R} \tilde{r}_i \delta_{k,i}\,,$$

where we introduced the coefficients $\tilde{r}_k$ for notational convenience. The second term of (2.56) can be reformulated as

$$\begin{aligned} \sum_{n=0}^{\infty} g_n \sum_{j=1}^{C}\sum_{m=0}^{m_j-1} \frac{(k+n)^m}{\gamma_j^{k+n}} c_{mj} &= \sum_{j=1}^{C}\sum_{m=0}^{m_j-1}\sum_{n=0}^{\infty} g_n \sum_{i=0}^{m} \frac{\binom{m}{i} k^i n^{m-i}}{\gamma_j^{k+n}} c_{mj} \\ &= \sum_{j=1}^{C}\sum_{i=0}^{m_j-1} \frac{k^i}{\gamma_j^k} \left(\sum_{n=0}^{\infty} g_n \sum_{m=i}^{m_j-1} \frac{\binom{m}{i} n^{m-i}}{\gamma_j^n} c_{mj} \right) \\ &= \sum_{j=1}^{C}\sum_{i=0}^{m_j-1} \frac{k^i}{\gamma_j^k} \tilde{c}_{ij}\,. \end{aligned}$$

In the second step, we interchanged the sums over m and i, respectively, while in the third step, we introduced other coefficients $\tilde{c}_{ij}$ for notational convenience. Now, the form of the $\tilde{t}_k$ is captured by

$$\tilde{t}_k = \sum_{i=0}^{R} \tilde{r}_i \delta_{k,i} + \sum_{j=1}^{C}\sum_{m=0}^{m_j-1} \frac{k^m}{\gamma_j^k} \tilde{c}_{mj}\,, \tag{2.57}$$

which obviously resembles the form of the t_k (2.54). If we now introduce the auxiliary function $\tilde{T}(z)$,

$$\tilde{T}(z) = \sum_{k=0}^{\infty} \tilde{t}_k z^k \,,$$

the resemblance between (2.54) and (2.57) leads to the conclusion that $T(z)$ and $\tilde{T}(z)$ have a common denominator $P(z)$. Also, we see that the numerator of $\tilde{T}(z)$ has degree N or less. The result of the queueing effect then follows from (2.55), leading to

$$H(z) = G(z)T(z^{-1}) + \tilde{T}(1) - \tilde{T}(z^{-1}) \,, \tag{2.58}$$

that can be seen as an extension of (2.18), obtained in case of a Bernoulli arrival process. As for $\tilde{T}(z)$, we note that this auxiliary function is rational, just like $T(z)$, but not a pgf, as $\tilde{T}(z) \neq 1$. It shares the denominator of $T(z)$, $P(z)$, but differs in its numerator $\tilde{N}(z) \neq N(z)$, that is a polynomial of degree N or less. In accordance with (2.52), we can thus write

$$\tilde{T}(z) = \frac{\tilde{N}(z)}{P(z)} = \frac{\sum_{i=0}^{N} \tilde{n}_i z^i}{\sum_{i=0}^{P} p_i z^i} \,. \tag{2.59}$$

We note that the coefficients $\tilde{n}_i$ are unknown. However, applying Rouché's theorem in the following will enable us to proceed without having to determine them explicitly. We focus now on the second non-linearity.

2.4.3.3 The Granularity Effect

The granularity effect reflects the degeneration of the buffer, that is only capable to realize delays that are a multiple of the granularity D. It is captured by

$$W = D \cdot \left\lceil \frac{H}{D} \right\rceil \,. \tag{2.60}$$

The solution in terms of pgf's is provided by (2.20),

$$W(z) = \sum_{k=0}^{D-1} \frac{1}{D} \frac{z^D - 1}{z\varepsilon_k - 1} H(z\varepsilon_k) \,, \tag{2.61}$$

with the symbols ε_k denoting the D different complex D^{th} roots of unity, that is, $\varepsilon_k = e^{j2\pi k/D}$. A property of this solution is that it displays symmetry in the complex plane. More precisely, $W(z)$ is invariant to rotations over multiples of $2\pi/D$ radians, which is a direct consequence of the fact that $W(z)$ is a function of z^D only, and implies that $W(z\varepsilon_k) = W(z)$, $k = 0 \ldots D-1$.

2.4.3.4 Combining Results

We now start from (2.58), and apply again the property that

$$G(z) = B(z)W(z)\,. \tag{2.62}$$

This is so, because of (2.3), and the fact that the pgf of the sum of two independent rv's is the product of their separate pgf's. We find that

$$H(z) = B(z)W(z)T(z^{-1}) + \tilde{T}(1) - \tilde{T}(z^{-1})\,. \tag{2.63}$$

Further, substituting this expression for $H(z)$ into (2.61), we find a functional equation for $W(z)$,

$$W(z) = \sum_{k=0}^{D-1} \frac{1}{D}\frac{z^D-1}{z\varepsilon_k - 1}\left\{B(z\varepsilon_k)W(z\varepsilon_k)T((z\varepsilon_k)^{-1}) + \tilde{T}(1) - \tilde{T}((z\varepsilon_k)^{-1})\right\}\,.$$

Using $W(z) = W(z\varepsilon_k)$ this results in

$$W(z) = \frac{\displaystyle\sum_{k=0}^{D-1} \frac{1}{D}\frac{z^D-1}{z\varepsilon_k - 1}\left\{\tilde{T}(1) - \tilde{T}((z\varepsilon_k)^{-1})\right\}}{1 - \displaystyle\sum_{k=0}^{D-1} \frac{1}{D}\frac{z^D-1}{z\varepsilon_k - 1}B(z\varepsilon_k)T((z\varepsilon_k)^{-1})}\,. \tag{2.64}$$

This does not represent a complete solution for $W(z)$ yet, as we do not have an exact expression for $\tilde{T}(z)$. However, making use of Rouché's theorem, we can overcome this last obstacle.

2.4.3.5 Applying Rouché's theorem

In order to apply Rouché's theorem, we need to reformulate the denominator of $W(z)$. More precisely, we rewrite the denominator as the sum of two analytic functions of z in the domain $|z| < 1$. Since the factor $T((z\varepsilon_k)^{-1})$ has poles in this domain, we make further assumptions on the form of $T(z) = N(z)/P(z)$ to remove these poles. As for the degree of $N(z)$ and $P(z)$, two possibilities occur: either $N \leq P$, or $N > P$. Derivations for both cases follow a similar line, we confine ourselves here to the case $N \leq P$. Further, we write $P(z)$, the common denominator of $T(z)$ and $\tilde{T}(z)$, as

$$P(z) = \prod_{j=1}^{C}(z-\gamma_j)^{m_j}\,.$$

(This notation is also adopted in (2.53).) The γ_j ($j = 1 \ldots C$) are the C different zeroes of $P(z)$, all outside of the domain $|z| \leq 1$, each with associated multiplicity

m_j ($j = 1 \ldots C$), and thus $P = \sum_{j=1}^{C} m_j$. Now, we are in the position to remove the poles of $T((z\varepsilon_k)^{-1})$, by multiplying both numerator and denominator of $W(z)$ in (2.64) with

$$\prod_{k=0}^{D-1} (z\varepsilon_k)^P P((z\varepsilon_k)^{-1}) \,,$$

which gives us, considering (2.52) and (2.59), for $W(z)$,

$$W(z) = \qquad (2.65)$$
$$\frac{\displaystyle\sum_{k=0}^{D-1} \frac{1}{D}\frac{z^D-1}{z\varepsilon_k - 1} \left\{ \prod_{k=0}^{D-1} (z\varepsilon_k)^P P((z\varepsilon_k)^{-1}) \tilde{T}(1) - \prod_{k=0}^{D-1} (z\varepsilon_k)^P \tilde{T}((z\varepsilon_k)^{-1}) \right\}}{\displaystyle\prod_{k=0}^{D-1} (z\varepsilon_k)^P P((z\varepsilon_k)^{-1}) - \sum_{k=0}^{D-1} \frac{1}{D}\frac{z^D-1}{z\varepsilon_k - 1} B(z\varepsilon_k) \prod_{k=0}^{D-1} (z\varepsilon_k)^P N((z\varepsilon_k)^{-1})} \,.$$

This can be somewhat simplified (not shown here, but applied below), by remarking that

$$\prod_{k=0}^{D-1} (z\varepsilon_k)^P P((z\varepsilon_k)^{-1}) = \prod_{k=0}^{D-1} \prod_{j=1}^{C} (1 - z\varepsilon_k\gamma_j)^{m_j} = \prod_{j=1}^{C} (1 - z^D\gamma_j^D)^{m_j} \,, \quad (2.66)$$

which is an application of the identity $x^D - a^D = \prod_{k=0}^{D-1}(x - \varepsilon_k a)$ at $x = 1$ and $a = z\gamma_j$. A second step to take is to split the denominator of (2.65), denoted $R(z)$, into two functions $R_1(z)$ and $R_2(z)$, as follows,

$$\begin{aligned} R(z) &= R_1(z) + R_2(z) \,, \\ R_1(z) &= \prod_{j=1}^{C} (1 - z^D\gamma_j^D)^{m_j} \,, \\ R_2(z) &= -\sum_{k=0}^{D-1} \frac{1}{D}\frac{z^D-1}{z\varepsilon_k - 1} B(z\varepsilon_k) \prod_{k=0}^{D-1} (z\varepsilon_k)^P N((z\varepsilon_k)^{-1}) \,, \end{aligned}$$

where we applied the simplification of (2.66) to obtain $R_1(z)$. Now, one can verify that $R_1(z)$ and $R_2(z)$ are both analytic in the domain $|z| < 1$: the denominator factor $(z\varepsilon_k - 1)$ does not yield an actual pole due to the numerator factor $(z^D - 1)$, $B(z)$ is analytic in this domain, and also $\prod_{k=0}^{D-1}(z\varepsilon_k)^P N((z\varepsilon_k)^{-1})$ is analytic, since $N \leq P$.

In a third and last step, we remark that

$$R(z\varepsilon_k) = R(z) \quad ; \quad R_1(z\varepsilon_k) = R_1(z) \quad ; \quad R_2(z\varepsilon_k) = R_2(z) \quad (k = 0 \ldots D-1) \,.$$

(For $R_1(z)$, this is obvious; for $R_2(z)$, this can be understood if one compares with the form of $W(z)$ (2.61), which also has this property.) Denoting $y = z^D$,

we introduce equivalent functions $\hat{R}(y)$, $\hat{R}_1(y)$ and $\hat{R}_2(y)$ for which

$$\hat{R}(z^D) = R(z) \quad ; \quad \hat{R}_1(z^D) = R_1(z) \quad ; \quad \hat{R}_2(z^D) = R_2(z)\,.$$

Now, we take up the approach of [98], and apply Rouché's theorem on $\hat{R}(y)$. To do this, it is necessary that $\hat{R}_1(y)$ and $\hat{R}_2(y)$ (i) are analytic functions in $|y| < 1$, (ii) are continuous at the boundary $|y| = 1$, (iii) have derivatives at $y = 1$, (iv) comply with the relation $|\hat{R}_1(y)| > |\hat{R}_2(y)|$ at ($|y| = 1$,$y \neq 1$), (v) comply with the relation $\hat{R}_1(1) = -\hat{R}_2(1) \neq 0$, (vi) comply with the relation

$$\frac{\hat{R}_1'(1) + \hat{R}_2'(1)}{\hat{R}_1(1)} > 0\,, \tag{2.67}$$

where primes denote derivates.

If we assume the involved pgf's to be properly defined, and require the system to be stable, it can be easily checked that these six conditions are fulfilled. (What happens for an unstable system is explained in Sect. 2.4.4.) Invoking the theorem in [98], we conclude that the number of zeroes $N_{\hat{R}}$ of $\hat{R}(y)$ in the domain $|y| < 1$ relates to the number of zeroes $N_{\hat{R}_1}$ of $\hat{R}_1(y)$ in the same domain (which is known to be P) according to the relation

$$N_{\hat{R}} = N_{\hat{R}_1} - 1 = P - 1\,.$$

Including now the zero at $y = 1$, the total number of zeroes for $\hat{R}(y)$ in the domain $|y| \leq 1$ sums up to P. Introducing the notation $\hat{W}(z^D) = W(z)$, we can invoke that $W(z^D)$ (like $W(z)$) is a proper pgf, with no singularities in the mentioned domain. As such, all P zeroes $y = \beta_i$ ($i = 0 \ldots P-1$) of the denominator $\hat{R}(y)$ in the domain $|y| \leq 1$ have to be compensated by the same zeroes $y = \beta_i$ in the numerator. Aware of these zeroes in the numerator, and using the knowledge that the numerator of $\hat{W}(z^D)$ is of degree P (which can be understood if one combines (2.65) and (2.66)), we cast $\hat{W}(z^D)$ in a form

$$W(z) = \frac{K(z^D - 1)}{\hat{R}(z^D)} \prod_{i=1}^{P-1} \frac{z^D - \beta_i}{1 - \beta_i}\,. \tag{2.68}$$

Again, this expression reflects that $W(z)$ is a function of z^D. The constant β_0 is the zero of $\hat{R}(y)$ equal to one, while the other β_i, $i = 1 \ldots P-1$ are zeroes of $\hat{R}(y)$ in the domain $|y| \leq 1$. Since $\hat{R}(y)$ has zeroes β_i, $R(z)$ has DP zeroes α_j, $j = 0 \ldots DP-1$, that correspond to the D-th complex roots of β_i (each β_i, $i = 1 \ldots P$, accounting for D different α_j).

Note that, at this point, we indeed ruled out $\tilde{T}(z)$. The constant K can be determined by demanding that $W(1) = 1$ (normalization condition), and results in

$$K = \prod_{j=1}^{C} \left(1 - \gamma_j^D\right)^{m_j} \cdot \left\{ \frac{\mathrm{E}[T] - \mathrm{E}[B] - \frac{D-1}{2} - \sum_{k=1}^{D-1} \frac{B(\varepsilon_k) T(\varepsilon_k^{-1})}{\varepsilon_k - 1}}{D} \right\}. \tag{2.69}$$

Now, the pgf of the waiting time $W(z)$ is fully known. It then suffices to substitute $W(z)$ in (2.63), to obtain an explicit formula for $H(z)$, which was our aim.

2.4.4 Maximum and Equivalent Load

As for the condition to stability, we can again take

$$\lim_{k\to\infty} \Pr[H_k = 0] = 0\,,$$

as a characterization of instability. Considering (2.63) and (2.68), this occurs when K (given by (2.69)) becomes zero. From this, one can define a maximum tolerable arrival intensity $\hat{p}_{max}$, that puts an upper limit to the arrival intensity $\hat{p}$, defined as $1/\mathrm{E}[T]$. (Note that we use $\hat{p}$ instead of p to avoid confusion with the parameter of the (less general) case of the geometric distribution with parameter p of (2.17).) It follows that $\hat{p}_{max}$ is the solution to the implicit equation (implicit, as $\hat{p}$ also occurs in the expression for $T(z)$)

$$\frac{1}{\hat{p}} = \mathrm{E}[B] + \frac{D-1}{2} + \sum_{k=1}^{D-1} \frac{B(\varepsilon_k)T(\varepsilon_k^{-1})}{\varepsilon_k - 1}\,. \tag{2.70}$$

The symbols ε_k still represent the D different complex D^{th} roots of unity, as in (2.61). The solution is thus function of the FDL granularity D, the (complete) pgf's of both inter-arrival and burst size distribution, again to be found from (2.70) with a simple bisection algorithm. Related, we define the equivalent load as

$$\rho_{eq} = \hat{p} \cdot \left(\mathrm{E}[B] + \frac{D-1}{2} + \sum_{k=1}^{D-1} \frac{B(\varepsilon_k)T(\varepsilon_k^{-1})}{\varepsilon_k - 1} \right)\,. \tag{2.71}$$

Note that the expressions (2.70) and (2.71) are indeed consistent with the general approach presented in Sect. 2.1.5.

Finally, there is also a third method to obtain $\hat{p}_{max}$. One easily verifies that the conditions to apply Rouché's theorem in Sect. 2.4.3.5 are violated when $\hat{p} = \hat{p}_{max}$. More precisely, the assumption that the system is stable then no longer holds, and the numerator of (2.67) turns zero, reflecting the system's instability.

2.4.5 Heuristics for the Loss Probability

To derive the loss probability for a buffer of finite size N, we again invoke the heuristics of Sect. 2.1.6.2, and consider here only the most accurate one, heuristic B. Further, we impose the same assumptions as there: the burst size distribution is not heavy-tailed, and has a single dominant pole. A sufficient condition is that B has a rational pgf like T. We will adopt this convention from here on, and note that it poses no problem for the application of Sect. 2.4.6. Further, the equivalent

load is stated above; the overflow probabilities $\Pr[H_\infty > ND]$ we obtain with the dominant pole approximation of Sect. 2.1.6.1. The tail probabilities have a quasi-geometrical tail decay, with decay rate z_0 as in (2.12). The $C(ND)$ follow from the application of residue theory and are, in their final form, given by

$$C(ND) = -\frac{1}{z_0} \lim_{z \to z_0} \left(W(z) \cdot (z - z_0)\right) \frac{D}{z_0^D - 1} \,.$$

The limit in the above can easily be calculated explicitly. As for the pole z_0, the function $W(z)$ relates to $H(z)$ according to (2.63) and both have the same dominant poles. Since for the latter, we had that $W(z) = W(z\varepsilon_k)$, it is easy to see that there are D dominant poles, of the form $z_k = z_0\varepsilon_k$ ($k = 0 \ldots D-1$), with z_0 being (by definition) the positive real one. That the latter does indeed exist (under the assumptions discussed in Sect. 2.4.2), follows readily by inspecting the denominator of $W(z)$ (see (2.65)) along the positive real axis. Although all formulas were found under the assumption of a stable system, it will be shown that the heuristic again also performs well for overloaded systems, that is, with $\rho_{eq} > 1$.

2.4.6 Numerical Comparison

In this section, we perform a numerical comparison of simulation results against results from the heuristic of Sect. 2.4.5. Our scope is twofold. On the one hand, we want to assess the accuracy of the heuristic. On the other hand, we are particularly interested in the impact of correlation in the arrival process on the performance of an optical buffer, since arrival processes in (optical) communication networks are bursty, as noted before.

The buffer we consider has a fixed size $N = 20$ and feeds into a single channel, for a slot length of 1 μs. Arriving bursts have a fixed burst size B of 20 μs. The load $\rho = \mathrm{E}[B]/\mathrm{E}[T]$ remains fixed at 60 percent, and thus also $\mathrm{E}[T]$ is fixed. Now, we consider a correlation model for the inter-arrival times that uses a probabilistic mix of geometrically distributed rv's T_S (parameter p_1) and T_L (parameter p_2). More precisely, the pgf of the inter-arrival time distribution is given by

$$T(z) = \alpha T_S(z) + (1-\alpha) T_L(z) \,,$$

where α, $0 \leq \alpha \leq 1$, is a weighing factor, and $T_S(z)$ and $T_L(z)$ are the pgf's of T_S and T_L respectively. This model is able to capture an arrival process that we describe with "trains" and "wagons". That is, bursts or "wagons" do not arrive independently (as in the case of a Bernoulli arrival process or a Poisson arrival process), but rather, they arrive in smaller or larger groups or "trains". The time in-between bursts of the same group, or "wagon spacing", generally denoted by T_S, is on average much smaller than the time between two trains or "train spacing",

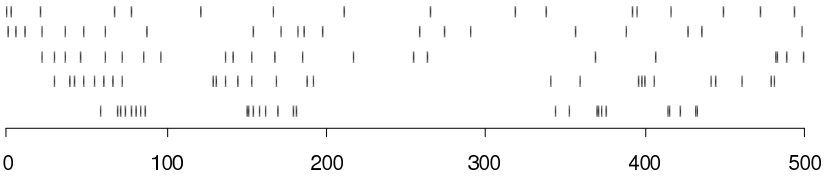


Figure 2.9: Simulation traces of the 5 runs, ordered (1 to 5) from top to bottom.

run	1	2	3	4	5
G	1	2	4	4	4
S	20	20	20	10	5

Table 2.1: Parameter setting used for the five runs.

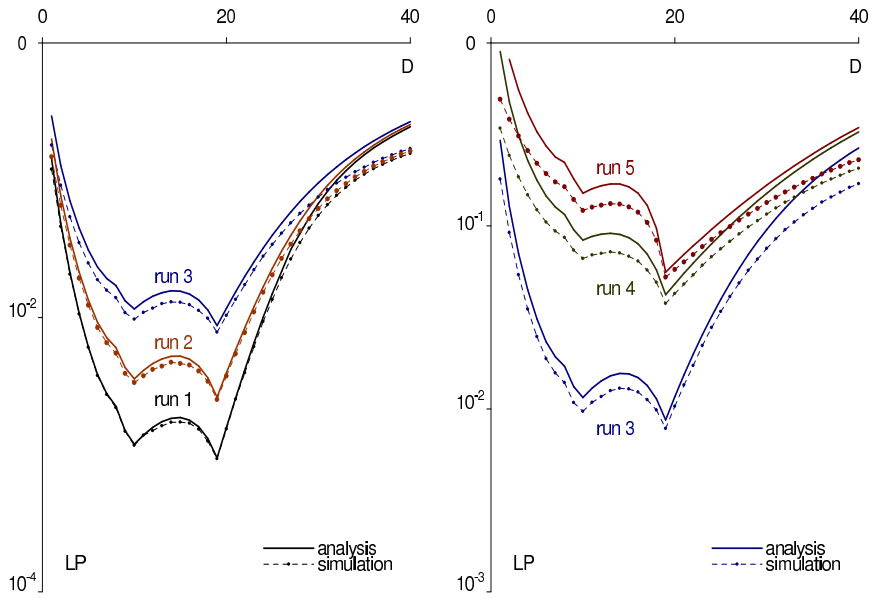
generally denoted by T_L. The average group size is captured by $G = 1/(1-\alpha)$, expressing the average number of bursts per group or "wagons per train". The average wagon spacing is captured by $S = 1/p_1$, the average train spacing by $1/p_2$ (both expressed in slots).

In Fig. 2.9, a sample trace of five different simulation runs is displayed, each with sample length 500 time slots. With each new run (increasing run number) parameter values were altered, so as to go from no correlation between arrivals (run 1) to an increasing amount of correlation (run 2 to 5). The parameter setting for each run is displayed in Table 2.1. As can also be understood from the figure, first we augmented the average group size G (run 1 to 3), then we diminished the wagon spacing S (run 3 to 5). Note that, for all five runs, the load, and thus the overall average inter-arrival time remained fixed.

Figure 2.10 displays the loss probability for varying granularity D (in μs, with each time slot corresponding to 1 μs), for each of the five runs. The results from the analysis are displayed as continuous curves, the simulations as discrete points connected with dotted lines.

As for the performance of the heuristic, it is clear that it attains high accuracy. Indeed, the analytic curve for run 1 matches the simulations very well, while for run 5, the match is less striking. Nevertheless, it is clear that the discrepancy between simulation and analysis is very small around the optimum, where the loss probability is minimal.

Regarding the impact of correlation in the arrival process on the loss performance of a finite-sized degenerate buffer, the observation of increased loss comes rather as expected. Less evident is what happens to the optima. For fixed burst sizes, the results of Sect. 2.2.8.2 (there for CT) showed that the well-pronounced optima around $D_1 = B - 1$ and $D_2 = (B-1)/2$ (in Sect. 2.2.8.2, $D = B$ and $D = B/2$ due to the different time setting) remain optimal for a broad range



(a) Run 1–3. (b) Run 3–5.

Figure 2.10: Although run 1–5 clearly are obtained for a completely different arrival process (with the parameter setting of Table 2.1), the heuristic's output provides good overall accuracy. Remark how the optimal granularity $D = B-1$ stands out for all five runs. These figures were obtained for varying D (in μs), slot length $\Delta = 1$ μs, fixed burst sizes with $\mathrm{E}[B] = 20$ μs, $N = 20$, and $\rho = 60\%$.

of the load ρ if arrivals occur independently. More precisely, it is known that D_1 is optimal for low values of the load (for example 10 percent) and the loss for D_2 is then much higher, while above a threshold load of 0.652, $\rho \geq 0.652$ (for the assumed time setting and burst size), D_2 becomes optimal. Interestingly, Fig. 2.10(a) (together with results not shown here) shows that this balance between D_1 and D_2 is hardly influenced when we increase the average group size in run 1 to 3: although the overall loss increases, the relative position of D_1 and D_2 is hardly influenced. This does not apply in the comparison of run 3, 4 and 5, as inspection of Fig. 2.10(b) illustrates. Apparently, decreasing the wagon spacing has a devastating effect on the optimum at D_2, while the optimum D_1 is much less influenced. As such, the granularity value D_1 that is known to be optimal for memoryless arrivals and low load (smaller than the mentioned threshold load of 0.652, see Sect. 3.3.2 for more detail), also proofs optimal here, for the arrival process of run 5 and a load of 60%.

Although not self-evident, this can be intuitively understood when we think

of a specific situation with $D = D_1 = B - 1 = 19$ μs, with B the fixed burst size of 20 μs. If an arriving burst is queued in delay line jD, and the next burst arrives just 1 slot later, then this next burst is sent to delay line $(j+1)D$ and there will be no time in-between the transmission of both bursts, implying that no void occurs. This situation is indeed likely in run 4 and 5 (Pr$[T = 1]$ is 0.077 and 0.152 respectively), while it is rather rare in run 1-3 (Pr$[T = 1]$ is 0.030, 0.036 and 0.041 respectively). If the arrival would occur not 1, but 2 slots later, then the void size is still just 1 slot. Generalizing, the granularity value D_1 leads to small voids, especially when the case of a very small inter-arrival time is a likely one. And, since minimizing the void size results in better loss performance, one can see how D_1 outperforms D_2 on Fig. 2.10(b), as Pr$[T = 1]$ increases. As a result, the optimum D_1 is more stable than it was under uncorrelated arrivals, and remains optimal also for higher loads. This reveals how special features of the inter-arrival time distribution, such as Pr$[T = 1]$, play an important role in the performance of degenerate buffers.

2.5 Model for Multiple Wavelengths

The current section presents an FDL buffer model that incorporates wavelength conversion, thus considering two types of contention resolution at once. As already mentioned in Sect. 1.5, the combination of FDL buffering and wavelength conversion provides both the most typical and most effective contention resolution scheme, and is therefore of particular practical interest.

As mentioned in the previous section, the optical buffer model presented there can also be applied for different means. Just like in Sect. 2.4, the aim of this section is to provide a generalization of the M/G/1 model in DT of Sect. 2.2.3. However, rather than alleviating the restriction on the arrival process (GI/G/1 in Sect. 2.4), we consider the case of an M/G/c buffer model in DT with wavelength conversion, having access to multiple output wavelengths. The complementarity is complete, in that 2.4 and the current section can be analyzed in an identical way, and both constitute an exact modeling in the case of infinite buffer size. Specific to the current section is however the entirely different buffer setting, and the need for additional assumptions on the wavelength assignment algorithm within the optical network node. As for the latter, we are only able to capture one algorithm in an exact manner, namely round-robin, while others, like join-the-shortest-queue, can (with the method presented) only be captured in an approximate manner, even in the case of infinite buffer size.

2.5.1 Buffer Setting

The current FDL buffer setting differs from the one of Sect. 2.1.2, that is generally considered throughout this chapter, in that it has multiple wavelengths at its disposition. Other assumptions remain identical however: the FDL buffer under consideration can only realize delays that are a multiple of D; a line with index j corresponds to delay jD; the size N of the buffer is the index of the largest delay line. Also in this case voids occur whenever any of the outgoing wavelenghts remains unused, while bursts are present in the buffer, but not yet available. Further, we assume full wavelength conversion capability (see Sect. 1.7.1).

2.5.2 Wavelength Assignment Algorithm

With multiple wavelengths available for contention resolution, both the delay-line and the wavelength assignment algorithm play their part. For each wavelength, the delay-line assignment algorithm is assumed FIFO, just as elsewhere in this work. As for the wavelength assignment algorithm, we will highlight all three possibilities introduced in Sect. 1.7.3: Random (RND), Round-Robin (RR), and Join-the-Shortest-Queue (JSQ). Our analysis will yield analytical results for RR and RND. The crux for the RR model (and also for the RND model) will be to transform the multi-wavelength situation to a single-wavelength one, by studying what happens in a single of the c queues. To make this possible, we will have to transform the arrival process as well (see next subsection). This approach is not feasible for JSQ, that is known to be very hard to analyze in an exact manner, even in the case of classic buffers. Therefore, JSQ performance will be simulated.

2.5.3 Traffic Setting

The burst sizes B_k follow a general distribution, whereas the inter-arrival times T_k follow the distribution that relates directly to the assumed wavelength assignment algorithm. We assume both to form a sequence of iid rv's, thus being independent of the index k. In the analysis, just like in Sect. 2.4, the associated pgf of the rv's is considered.

We assume that bursts arrive in the system according to a Bernoulli arrival process, just like in Sect. 2.2.3. Bursts arrive in the system one by one, with inter-arrival times that are distributed geometrically. Now, the distribution of the inter-arrival time T in a given queue is directly related to the wavelength assignment algorithm.

For RND, the arrivals in the queue are selected from a Bernoulli arrival process in a random manner. Therefore, the same type of arrival process occurs at the level of the queue, and T is also distributed geometrically, with pmf and pgf as introduced in Sect. 2.2.1, with now $\mathrm{E}[T] = c/p$.

For RR, every cth arrival in the system goes to the same queue. As such, the rv T is the sum of c different geometrically distributed rv's. This leads to the so-called negative-binomial distribution (sometimes called Pascal distribution), which is the DT counterpart of the Erlang distribution in CT. The rv T is the sum of c rv's with geometric distribution, each with parameter p. The mean is $\mathrm{E}[T_k] = c/p$, the pmf is

$$\Pr[T_k = n] \quad = \quad \binom{n-1}{c-1} \cdot p^c \cdot \bar{p}^{n-c} \quad n \geq c\,,$$

and the resulting pgf reads

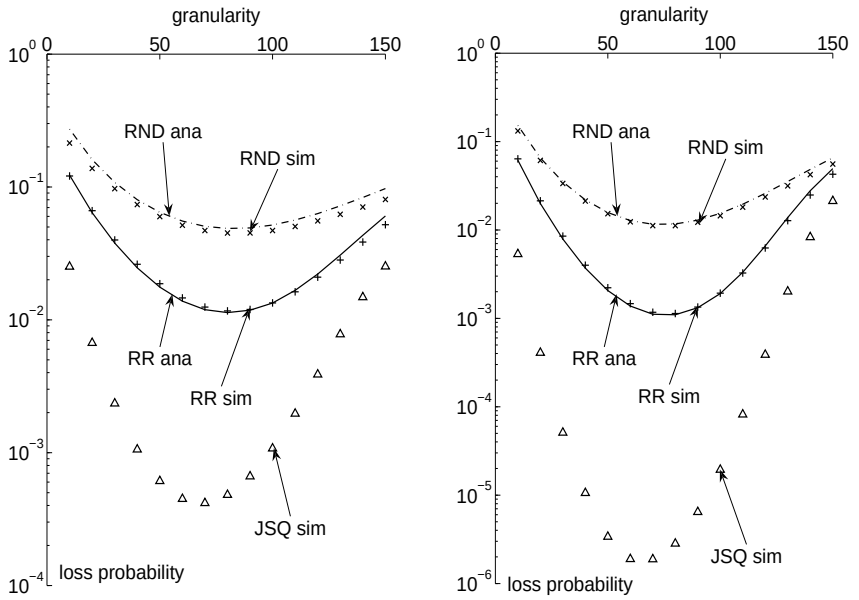
$$T(z) = \left(\frac{pz}{1-\bar{p}z} \right)^c \;,\; z \in \mathbb{C}\,.$$

Note that this approach can in principle be extended also for general (non-geometric) inter-arrival time distribution. Then, the distribution of T would be given by the c-fold convolution of that general distribution. This extension is however not included in this work.

In the following, we will thus no longer consider the arrival process at the entire output port. Instead, we have transformed the system's arrival process to an arrival process at the level of a single queue, with inter-arrival time distribution T. As such, we are able to model RND and RR with a queueing model for a single wavelength. For RR, the price to pay is that the arrival process for the single queue is no longer memoryless. Since we need to utilize the results of Sect. 2.4, notice that both the geometric and the negative-exponential distribution for the inter-arrival times indeed have a rational pgf, as required.

2.5.4 Evolution of the Scheduling Horizon

Having mentioned when bursts enter the buffer, and at what wavelength, we now discuss in what way they join the chosen queue. We now consider a single wavelength i (or, queue i), $i = 1 \ldots c$. (As all queues behave in a similar manner, we will not include the index i in the notation.) Bursts arrive at the given queue, and are numbered in the order in which they arrive at that specific queue by an index k. An arriving burst k has to be buffered for at least H_k, the time needed for all previous bursts in that queue to be transmitted. Instead, it is scheduled to wait for a time period W_k, that is a multiple of D, and is sufficiently long, that is, $W_k \geq H_k$. Depending on the size N, the burst is either queued ($W_k \leq ND$) or dropped ($W_k > ND$). The measure H_k is the so-called scheduling horizon of the given queue, as seen by the kth burst; W_k is the waiting time in the given queue, of the kth burst. For completeness' sake, note that this is different from the approach in Sect. 5.2, where all queues have to be involved in the system processes, in order to be able to analyze the more complex algorithm JSQ.



(a) Buffer size $N = 10$.

(b) Buffer size $N = 20$.

Figure 2.11: For geometrically-distributed burst sizes, the three wavelength assignment algorithms lead to much difference in terms of performance: JSQ outperforms RR, that in its turn outperforms RND. These figures were obtained for $\mathrm{E}[B] = 100$ *slots,* $\rho = 60\%$*, and* $c = 4$*.*

At any rate, here we obtain exactly the same system description as in Sect. 2.1.3, and thus, also the one of Sect. 2.4. Hence, we rely entirely on Sect. 2.4 for the analysis.

2.5.5 Numerical Comparison

At this point, we evaluate the performance of different multi-wavelength buffer settings. Our scope here is twofold. On the one hand, we want to assess the accuracy of our analytic results for RR and RND. On the other hand, we study the relation between RR and JSQ, that will be shown to be intimate for fixed-sized bursts. To do this, we consider all three wavelength assignment algorithms, varying and fixed burst sizes, and two buffer sizes.

2.5.5.1 Geometrically-Distributed Burst Sizes

We first consider varying burst sizes, with a mean length of 100 slots, in Fig. 2.11(a) (for buffer size $N = 10$) and Fig. 2.11(b) ($N = 20$). Assuming a slot size of

100 ns, 100 kbit burst, and a 10 Gbit/s link per wavelength, this corresponds to a mean burst size E$[B] = 10$ μs, and granularity values ranging from 0 to 150 slots (0 to about 3 km of fiber). The load is fixed to 60%. Simulation results (sim) are calculated for multiples of 10 slots, and analytic results (ana) for the whole range of D. The number of wavelengths is four ($c = 4$).

Obviously, the main difference between the figures is in the range of the LP, and they further display a similar behavior. In both figures, the analytic results for RND and RR both match simulation results very well. This asserts the functioning of our heuristic for varying burst sizes.

From the figure, it is obvious that RND is an algorithm that is to be avoided, as both RR and JSQ perform significantly better. Comparing JSQ and RR, the simulation shows that JSQ outperforms RR by far, if burst sizes vary. Therefore, it is understandable that JSQ often is the wavelength assignment algorithm of practical interest. This can be understood intuitively, if one realizes that the next queue (as selected in RR) seldom is the shortest queue (as selected in JSQ) if burst sizes vary. This is different for fixed burst sizes.

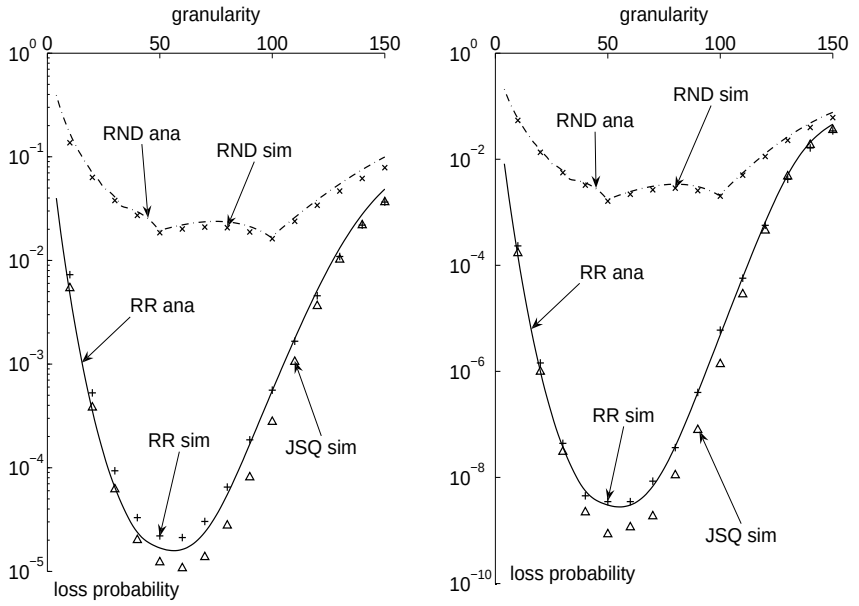
2.5.5.2 Deterministic Burst Sizes

For burst sizes fixed to 10 μs (or, 100 slots), we obtain Fig. 2.12(a) ($N = 10$) and Fig. 2.12(b) ($N = 20$). Again, granularity ranges from 0 to 150 slots, load is 60%, the number of wavelengths is four, simulation results (sim) are calculated for multiples of 10 slots, and analytic results (ana) for the whole range of D.

Similar to the case of varying burst sizes, the impact of the buffer size is mainly in the range of the LP. As for the optimal granularity, the well-pronounced optimum for RND around the burst size (and also, at half the burst size) vanishes for RR and JSQ. Further, simulation results for RND and RR assert the functioning of our heuristic again, now for fixed-sized bursts.

The RND wavelength assignment algorithm is again the one to avoid. The main difference with varying bursts, is that the gap in performance between JSQ and RR is really small, and this for the whole range of D. Because this is so for classic buffers, it comes not as a complete surprise. However, since optical buffers differ from classic buffers in a non-trivial way (one can think of the voids, non-existent for classic buffers), these results show that this is also valid for optical buffers. Thus, the intuition applies that to select the next queue (RR) often comes down to selecting the shortest one (JSQ), if burst sizes are fixed.

Now, the observation that JSQ resembles RR for fixed-sized bursts implies two things. On the one hand, considering also that RR has lower implementation complexity than JSQ, a hardware designer would prefer RR over JSQ. On the other hand, disregarding the choice between RR and JSQ, the analytical model obtained for RR offers a good approximation for both RR and JSQ.



(a) Buffer size $N = 10$.

(b) Buffer size $N = 20$.

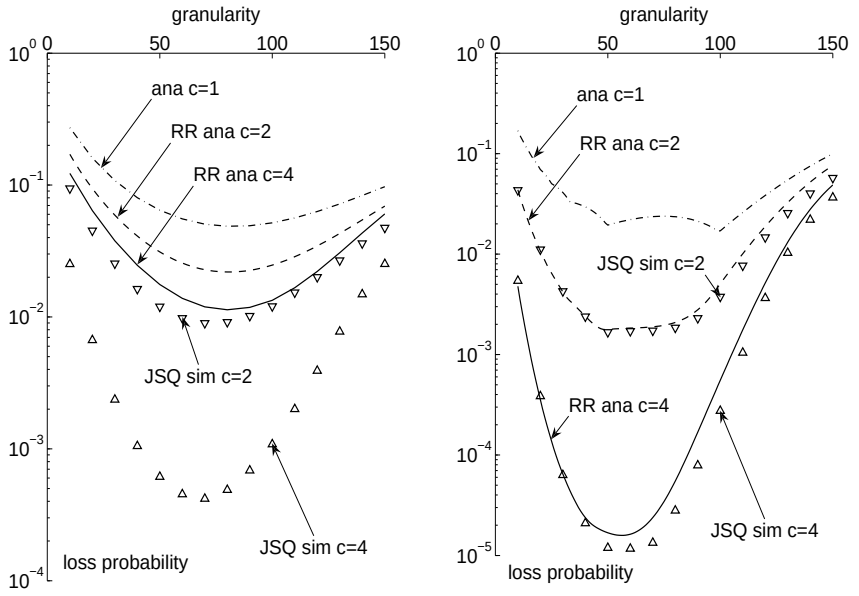
Figure 2.12: For deterministic burst sizes, JSQ and RR yield much better performance than RND. However, the difference between JSQ and RR is much less pronounced than in the case of geometrically-distributed burst sizes (Fig. 2.11). These figures were obtained for $\mathrm{E}[B] = 100$ *slots,* $\rho = 60\%$*, and* $c = 4$*.*

2.5.5.3 Multiplexing Gain

In Fig. 2.13(a) (varying burst sizes) and Fig. 2.13(b) (fixed burst sizes), we focus on the impact of wavelength conversion on loss performance. Therefore, we consider one, two and four wavelengths (c =1, 2 or 4). For a single wavelength, no wavelength assignment algorithm is specified (as there is only one channel to queue for), while for two and four wavelengths, we consider both RR and JSQ. For RND, the results are also shown (albeit implicitly), as they are identical to the results for $c = 1$, independent of the number of available wavelengths. As such, RND does not benefit from multiplexing gain in any way.

We assume a buffer size $N = 10$, a mean burst size $\mathrm{E}[B] = 100$ slots (or, 10 μs at 10 Gbit/s), and a load of 60%. As we have assessed the accuracy of our heuristic in the above, we only show the analytic results for RND and RR (ana), for the whole range of D. (Simulation results for $c = 2$ not included here, display a good match.) Simulation results for JSQ (sim) are calculated for multiples of 10 slots.

As a reference, we also mention the loss in case wavelength conversion is



(a) Geometric burst size distribution.

(b) Deterministic burst size distribution.

Figure 2.13: For both geometrically-distributed and deterministic burst sizes, the benefit from multiplexing gain is largest in case of JSQ. However, for geometrically-distributed burst sizes, RR performs much less than JSQ, while for deterministic burst sizes, the performance of RR almost parallels that of JSQ. These figures were obtained for $\mathrm{E}[B] = 100$ *slots,* $\rho = 60\%$*, and* $N = 10$*.*

applied without FDL buffering (with $N = 0$, corresponding to a single line of length zero). For both varying and fixed burst sizes, the LP is then approximately 0.38 for one wavelength, 0.25 for two, and 0.14 for four wavelengths. These results we obtain from the well-known Erlang-B formula, with 1, 2 and 4 wavelengths respectively. Although this formula by definition only applies in a continuous-time setting, it serves as a good reference.

Figure 2.13(a) shows loss performance for varying burst sizes. For both RR and JSQ, performance does benefit from multiplexing gain, and the curves lower for increasing c. Results not included here, demonstrate a further decrease in loss probability, when one adds even more wavelengths. In this respect, an FDL buffer behaves similarly to a classic multiplexer.

Clearly, the gain is a lot more pronounced for JSQ than for RR: loss for JSQ lowers significantly for two wavelengths, while RR needs four wavelengths to obtain a comparable lowering. For RR and JSQ, we note how the optimum for the granularity (for which the LP is minimal) decreases a little, but not much, as the

number of wavelengths increases. Results not shown here, illustrate that this optimum is more sensitive to the load, than to the number of wavelengths. More precisely, the optimal granularity shifts to lower values as load increases, similar as in the case of a single wavelength.

Figure 2.13(b) displays the loss probability for fixed burst sizes. Again, RND does not benefit at all from multiplexing gain. Loss performance for JSQ and RR is again very similar, and both clearly capitalize on multiplexing gain. Here, the optimal granularity for JSQ and RR drops markedly as the number of wavelengths increases.

2.6 Concluding Remarks

In this chapter, various models for the performance evaluation of optical buffers were presented. Assuming infinite buffer size, their analysis is based on an approach with transform functions, that allows to obtain elegant and exact results in that case. For finite buffer-size results (such as the loss probability), we repeatedly relied on heuristics that yield accurate results.

The first performance model presented was a single-wavelength model for memoryless arrivals in DT. This was extended to CT by taking the limit of the latter for decreasing slot lengths. An analysis directly in CT appeared feasible too, but turned out to be slightly more complex than the limit procedure (at least in the author's opinion). Three special cases of burst size distributions were used to illustrate results, and to establish the accuracy of the heuristics. For these special cases, the resulting formulas turned out to be relatively simple, that is, not involving infinite sums, allowing for easy numerical evaluation. Also, it was shown that the granularity heavily impacts the loss performance, as was noted also for DT.

The second performance model considered batch arrivals, allowing to study the impact of synchronization on performance. It was shown that synchronization benefits loss performance drastically when burst sizes are fixed. Together with other advantages (for example, a simplification of the control logic), this might justify the additional cost of implementing retiming within an optical node, when burst size is fixed. For variable burst size, this benefit is in general much less-pronounced, as is illustrated in the case of memoryless burst size distribution, where synchronization offers no performance benefits at all.

As a third performance model, we relaxed the assumptions on the arrival process, to obtain a single-wavelength model for general inter-arrival times and burst sizes. This model then allowed for a deeper insight in the impact of correlation in the arrival process on loss performance, and the importance of special features of the inter-arrival time distribution in that context.

Employing the latter model in an alternative way provided us with an analytic model for a multi-wavelength optical buffer. We have considered three dif-

ferent wavelength assignment algorithms: Random (RND), Round-Robin (RR), and Join-the-Shortest-Queue (JSQ). These handle either varying- or fixed-sized bursts. For RND and RR, we developed an analytic model, while for JSQ, we relied on simulation. In the case of varying burst sizes, we found that performance is matched accurately in the case of RND and RR. Further, we found that the performance of JSQ can in no case be approximated by that of RR, as they differ too much. We also found that the optimal granularity is not much influenced by the number of wavelengths. For fixed-sized bursts, we found that our model can estimate performance very well for RND and RR. As we found that JSQ performance is close to that of RR, our model is also a good heuristic approximation for JSQ. Finally, we also observed that the optimal granularity drops markedly, for RR and JSQ, as the number of wavelengths increases.

Concluding, this chapter provides a consistent overview of the performance modeling of optical buffers by means of transform functions. Clearly, the single system equation (2.1) posed a challenging problem, that was tackled for either simplified or more general assumptions on the traffic setting. When compared to the Markov chain approach of the next chapter, in the author's opinion, the current method is somewhat more elegant, in that all performance results are captured in very condense mathematical form. Also, the approach with transform functions allows to split up the problem in subproblems. Indeed, the separate treating of granularity effect and queueing effect makes the solution method somewhat more generic, when compared to that of the next chapter. On the other hand, the buffer setting was always chosen degenerate, since the modeling of non-degenerate optical buffers is typically better done with the methods presented in the following chapter. Also, while the current approach yields exact results for infinite system size (with the associated benefit of light-weight formulas), the Markov chain approach of the next chapter is especially well-adapted to finite buffer sizes, and will allow to obtain exact results for some particular and instructive buffer settings.

3

Performance Evaluation with Markov Chains

¶ The current chapter treats the modeling of single-wavelength optical buffers by means of the imbedded Markov chain technique. The system setting is mostly the same as the previous chapter, except that here, the FDL lengths are not limited to multiples of the granularity, but can take on any value. More essential is however the difference in the approach of the analysis, that leads to different but compatible formulations of the performance results. While the previous chapter relied on probability generating functions, the current relies on the probabilities of the involved Markov chain. While the analysis in the previous section was done in the z-domain, the current section shows how to obtain results straight away in the probability domain. Assuming infinite buffer size and degenerate buffer setting in the previous section for the analysis, the current chapter assumes a non-degenerate and finite-sized buffer. While results of the previous chapter took the evolution of the scheduling horizon as a starting point, the current chapter focuses on the evolution of the waiting times.

Apart from the dissimilarities in the assumptions, the results in this chapter are also of a different nature. In the previous chapter, the analysis yielded exact results for infinite-sized buffers in closed-form, but only approximate results for finite-sized buffers. In the current chapter, general results are obtainable only through numerical means, but with the prime advantage of being exact. Also, in some special cases, the Markov chain approach allows to derive an exact closed-form solution for the main performance measures, which clearly constitutes a significant

credential.

Given that the system setting is nearly identical to that of Chapter 2, the references mentioned there remain of interest also to the current chapter, and are not repeated for conciseness' sake, save the most relevant ones, that share either the assumption of non-degenerate buffering, or the solution method.

Firstly, a Markov chain approach is also applied by Lakatos for the cyclic-waiting problem, mentioned earlier in Sect. 1.7.2. His first paper on the topic [20] assumed the most basic system setting in CT: infinite queue size, exponentially-distributed inter-arrival times and burst sizes, with waiting times equal to multiples of the cycle time, or, with the notation of Sect. 1.8.2, a degenerate infinite-sized M/M/1 in CT. While this is exactly the setting of (the first part of) [24], the approach is based on a Markov chain, just like in the current chapter. Main difference however is that Lakatos considered the (Markov chain) evolution of the number of customers in queue, while we consider the evolution of assigned waiting times. Derivation of the number of customers results in an analysis in its own right, with much different expressions, that are somewhat more complicated than the ones we obtain, but prove consistent. As such, [20] is the first to provide the stability condition for this system. This approach is later applied to the DT case in [21], where the stability condition is obtained for the degenerate infinite-sized M/M/1 in DT. The question of stability starting from the waiting time was raised first by [23], and was elaborated in [22], providing a Markov chain approach based on the waiting times, showing that both the CT case and DT case can be solved equally in this manner, and this with somewhat easier calculations and expressions. This is of special interest for the current work, since the approach of [22] is also the one of the current chapter, with the difference that [22] considers the infinite-sized system, whereas we look at the finite-sized, somewhat more complicated variant. The parallel between our work and that of Lakatos is discussed more in detail in [5], and in the introduction of [6].

Most similar to the current approach are the results that Almeida and his co-authors came up with. In [28], they provided a solution method based on the Markov chain of the waiting times, for a non-degenerate finite-sized M/M/1 buffer in CT. Still in CT and for non-degenerate finite-sized buffer setting, they extended this approach to the setting with general burst size distribution (M/GI/1) in [29], to general inter-arrival times (GI/M/1) in [30], and to a GI/GI/1 setting in CT in [31]. Depending on the traffic assumptions, their numerical solution method is either exact or approximate.

Further, especially relevant for the current chapter is [17, 19], where non-degenerate optical buffer settings are considered. This setting, with FDL lengths not necessarily a multiple of the granularity, they refer to as *non-equidistant*. Apart from the naming, the setting is however identical to the one we assume in the current chapter.

In [17], the analysis focuses on the evolution of the scheduling horizon, and follows a Markov chain approach, stated in terms of the involved transition matrix. The arrival process is a general Markovian arrival process, thus yielding a two-dimensional state space. The method is valid for a general set of FDLs, and allows to study both FDL line length optimization and preventive drop mechanisms.

1. FDL line length optimization assumes a non-degenerate buffer; since the line lengths in such structure are not necessarily multiples of D, one can optimize all the line lengths individually, rather than only D. As such, N different parameters are to be optimized for minimal loss (with N equal to the number of fibers of non-zero length). Since exhaustive search of such N-dimensional optimization is unfeasible, three optimization algorithms are introduced: Least Used Elimination (LUE), Smallest Increment Elimination (SIE), and Largest Reduction Addition (LRA). The latter stands out as most practicable; application thereof shows that for fixed burst length and rather high traffic load (ρ larger than about 0.6), non-degenerate buffer structures outperform degenerate ones. (To be precise, the LRA algorithm shortens the length of the longest delay lines to obtain somewhat better loss performance.)

2. Secondly, for degenerate FDL buffers, a preventive drop mechanism based on a Markov decision process points out that preventive dropping of bursts in case of high traffic load can significantly better performance.

Overall, non-degenerate structures obtained with the LRA algorithm yield more performance bettering than preventive dropping. However, preventive dropping has the benefit that it can function optimally on a given set of FDLs regardless of the traffic load, by applying dropping only during periods of high activity, and not applying it during periods of low activity. Non-degenerate structures do not provide this benefit, since reduced length for the longest lines results in optimal performance for a particular (rather high) value of the load, but only in suboptimal performance for low traffic load.

In [19], a joint work in which also the author participated, the LRA algorithm of [17] is investigated further, and a new method is provided for the calculation. More precisely, the method of [7] is adopted as framework, yielding already considerable simplification when compared with [17]. However, the transition matrix involved is stochastically dense, implying that the determination of steady state probabilities is the bottleneck when the LRA algorithm is executed. Therefore, the method is refined further by considering an upper-Hessenberg formulation of the transition matrix. This formulation not only eases the calculation of the steady state probabilities of the Markov chain, but also allows for optimal execution of the LRA algorithm. More precisely, the Hessenberg formulation exploits the similarity of subsequent iterations of the algorithm, by largely reusing the transition

matrix of a given iteration in the next iteration. Comparing execution times of the LRA algorithm of the Hessenberg formulation with those of the stochastically dense description, the approach allows another speed-up of over a 100 times, making the LRA algorithm extremely fast, with computation times of the order of $10 - 100$ ms.

As for the author's contribution, our main assumption was to focus exclusively on the evolution of the waiting times, first applied to the non-degenerate finite-sized M/G/1 buffer in DT in [7]. As such, this method shares elements with the results of Almeida et al. [28–31], but differs in the time setting. A more fundamental difference lies in the state space of the involved Markov chain: in the approach of [7] (and further, also [8, 9]), the size of the state space remains limited to the number of FDLs ($N + 1$), and this independent of the complexity of the inter-arrival or burst size distribution. The latter provides a significant advantage over the approach of Almeida, where the state space size is given by the sum of the number of FDLs ($N + 1$) and the number of blocking states. This number does not pose any problem for the M/M/1- and GI/M/1-setting, for which it equals one [28, 30], but grows with complexity, as it amounts to $N + 1$ for M/GI/1 [29], and to a (possibly infinite) multiple of $N + 1$ blocking states for GI/GI/1 [31]. Opposed to this, we are able to construct a most concise model, that circumvents the need for blocking states, and thus provides a more generic and simpler model, that is much less demanding in terms of computational effort, especially when a more complex inter-arrival or service time distribution is considered.

While [7] considered a Bernoulli arrival process, [8] provides the DT generalization for general independent inter-arrival times, possible without introducing blocking states. Moreover, it proved possible to extend the results of [7] to CT, and so establish a unified model for optical buffers in both DT and CT, in [9]. Also, by considering a specific and rather realistic instance of the latter model, we found that expressions for the loss and waiting time probabilities can be obtained exactly with simple closed-form expressions. Given that even a classic finite-sized M/D/1 buffer system in CT has a more complicated solution [99], the simplicity of the obtained expressions comes as a pleasant surprise.

Finally, an exact closed-form solution is also determined for the non-degenerate finite-sized M/M/1 system in CT, so providing a definite answer to the performance problem first studied in [24].

The current chapter presents the results of [7–9]. Sect. 3.1 treats the exact numerical model for an M/G/1 optical buffer of finite size, for both DT and CT. These results are largely based on [7], for DT, and on [9], for the extension toward CT. Further, Sect. 3.2 provides a generalization to general independent arrivals in DT, as first presented in [8]. Then, we focus on some particular cases of optical buffers, for which a closed-form solution is possible: fixed burst size in Sect. 3.3, and memoryless burst size in Sect. 3.4. The case of fixed burst size was treated

before in CT and DT in [9], whereas memoryless burst sizes were treated first in CT in a contribution currently under review (Sect. 1.8.4, “Work in Progress”, No. 1) for CT, and for the first time here, in the DT case.

3.1 Model for Memoryless Arrivals

In this section, we set out the general performance model of a non-degenerate finite-sized M/G/1 FDL buffer. This is in the spirit of the approach in [7], but with the difference that the latter was only valid for DT, while here we treat CT and DT with one model, as is also done in [9].

3.1.1 Time Setting

To those familiar with queueing theory, it is well-known that different time settings (CT or DT) can give rise to quite different solutions, even when the studied model has a lot in common (see for example [100] for CT, [93] for DT). For the specific case of an M/G/1 FDL buffer, however, it comes out that the results for DT can be converted into results for CT mostly in a plug-and-play fashion. Rather than developing the analysis for both time settings at once, we will present the results subsequently, and this to avoid ambiguity. As such, the DT setting will be adopted as given time setting throughout this chapter, whereas the CT setting will be given only afterward as an extension. Since the notation is especially chosen to fit this purpose, the CT case will be obtainable by mere substitution of certain variables. Note that treating the DT case first is an arbitrary choice, and that it is equally possible the other way around, treating the CT model first, with the DT case as an altered version thereof.

In the DT setting, events take place synchronously, at the beginning of time slots. Therefore, all time-related variables and performance measures are expressed as multiples of the slot length, and for example inter-arrival times and burst sizes take on only strictly positive integer values, contained in $\mathbb{N}_0$. The slot length may be arbitrary, and is therefore not mentioned explicitly in this chapter. In the CT setting, all events take place in an asynchronous fashion, and time-related variables like inter-arrival times and burst sizes can take on any positive real value.

3.1.2 Buffer Setting

The FDL buffer setting of the current section (and chapter) is more general than that of the previous chapter, in that also non-degenerate optical buffers can be studied. In other words, the FDL lengths are in general not multiples of D (degenerate optical buffer), but can be chosen degenerate if desired.

The following assumptions are shared with the previous chapter. Within an optical network, the buffer is located at the output of a backbone switch, and is

dedicated to a single outgoing wavelength. We consider bursts arriving at the buffer randomly, and possibly overlapping in time. Since there is only one wavelength to queue for, overlap during transmission should be prevented. By means of a switching matrix that allows to send any burst to any of the delay lines, buffer control exercises a FIFO (First-In-First-Out) scheduling discipline. Of all lines, it chooses the shortest line with sufficient length, so as to avoid overlap with the one-but-last burst. If the requested delay exceeds the delay provided by the longest line, the burst is dropped.

Different from the previous chapter is that the FDL buffer of size N is now represented by a general finite set, $\mathcal{A} = \{a_0, a_1, a_2, \ldots, a_N\}$ of available delays $a_i \in \mathbb{N}$, $i \in \{0, 1, \ldots, N\}$, with $a_0 = 0$ by definition. As the set of lines is intended to resolve contention for a single outgoing channel, it is necessary that contending bursts undergo different delays, and therefore, a useful FDL set never contains the same length twice, $a_i \neq a_j$ for $i \neq j$. Also, we sort the line lengths ascendingly, $a_0 < a_1 < \ldots < a_N$. The length of the longest line, a_N, is the maximum delay the buffer can provide and is referred to as buffer capacity, while N indicates the buffer size.

The main characteristic of an optical buffer is that it cannot assign the exact delay value needed. When a non-zero delay $x \in \mathbb{N}_0$ is requested ($x > 0$) and is achievable ($x \leq a_N$), a delay a_i is granted from the FDL set $\mathcal{A}$ such that $a_{i-1} < x \leq a_i$, $i = 1 \ldots N$. This assignment procedure can be cast in operator form as

$$a_i = \lceil x \rceil_{\mathcal{A}} = \min\{y \in \mathcal{A}, y \geq x\}\,,\; x \leq a_N\,, \tag{3.1}$$

and will prove useful in the following. Note that negative values for x are also allowed. As a reference we mention that when a degenerate buffer setting $\mathcal{A} = \{0, D, 2D, \ldots, ND\}$ is considered, this operator takes the following form,

$$iD = \lceil x \rceil_{\mathcal{A}} = \left[D \cdot \left\lceil \frac{x}{D} \right\rceil\right]^+ \,,\; x \leq ND\,,$$

where $\lceil x \rceil$ again denotes the smallest integer larger than x, $[x]^+ = \max\{0, x\}$, and i denotes the natural number that adheres to $(i-1)D < x \leq iD$ (so emphasizing that the output of $\lceil x \rceil_{\mathcal{A}}$ is necessarily a multiple of D in this case).

3.1.3 Traffic Setting

Bursts are assumed to arrive one by one, with at most one arrival per slot; upon arrival, a burst is either accepted or dropped. We now number the bursts in the order at which they arrive, but only assign an index to those bursts that are accepted. Notice that this is opposed to the assumption in Chapter 2, that had all arriving bursts numbered.

The distributions of inter-arrival time and burst size distribution are identical to the ones presented in Sect. 2.2.1, therefore we only repeat the main assumptions.

With each accepted burst k, we associate an inter-arrival time $T_k \in \mathbb{N}_0$, that captures the time between the kth arrival and the next, being the arrival of (i) burst $k+1$, if this next burst is accepted or (ii) a burst without number, if this next burst is dropped. In the following, we assume memoryless inter-arrival times T_k, that have (in DT) a geometric distribution, and constitute a Bernoulli arrival process. The inter-arrival times form a sequence of identical and independently distributed (iid) random variables (rv's) with common cumulative distribution function (cdf) $F_T(n)$

$$F_T(n) = \Pr[T_k \leq n] = 1 - \bar{p}^n \quad , \quad n \in \mathbb{N}\,, \tag{3.2}$$

where $\bar{p}$ denotes $1-p$, with $p \in [0,1]$, just like in the previous chapter. The latter probability is also the parameter of the geometric distribution, and gives the probability of having an arrival in an arbitrary slot, and is in tight relation with the mean value, as $\mathrm{E}[T_k] = 1/p$. The inter-arrival times associated with dropped bursts also follow this distribution.

With each accepted burst, we also associate a burst size B_k. The burst sizes also form a sequence of iid rv's with common probability mass function (pmf) $b(n) = \Pr[B_k = n]$ and common cumulative distribution function (cdf) $F_B(n) = \Pr[B_k \leq n]$, $n \in \mathbb{N}$. The latter relates to the pmf as $F_B(n) = \sum_{i=1}^{n} b(i)$, with the same conditions as in the previous chapter, so as to be well-defined.

For notational convenience, we introduce an additional series of rv's $U_k = B_k - T_k$, that enables to express the system's evolution in a more compact way. Their common cdf is denoted by

$$F_U(n) = \Pr[U_k \leq n] = \Pr[B_k - T_k \leq n] \quad , \quad n \in \mathbb{Z}\,.$$

Taking into account the cdf of the inter-arrival times (3.2), we obtain that

$$F_U(n) = \begin{cases} \bar{p}^{-n-1} \sum_{i=1}^{+\infty} b(i) \cdot \bar{p}^i & , n \in \mathbb{Z}^-\,, \\ \bar{p}^{-n-1} \sum_{i=n+1}^{+\infty} b(i) \cdot \bar{p}^i + F_B(n) & , n \in \mathbb{Z}^+\,. \end{cases} \tag{3.3}$$

As the analysis will point out, the input needed for analysis is limited to knowledge of the FDL set $\mathcal{A}$, $F_T(n)$, $F_B(n)$ and $F_U(n)$ (using (3.2) and (3.3), respectively).

3.1.4 Setting for CT

To translate the DT model to a CT setting, only minor changes are involved.

As for the FDL buffer setting (Sect. 3.1.2), it is clear that CT assumes that $a_i \in \mathbb{R}^+$, instead of adopting natural numbers. Similarly, the definition (3.1) for $\lceil x \rceil_{\mathcal{A}}$ remains unaltered except that the domain is to be expanded to $x \in \mathbb{R}$. Further FDL buffer assumptions remain unchanged.

As for the traffic assumptions (Sect. 3.1.3), we adopt the same indexing convention as for DT, and assume that just one arrival per arrival instant occurs. The

inter-arrival times have a memoryless distribution in CT, which constitutes a Poisson arrival process. The inter-arrival times, a sequence of iid rv's, have a common negative-exponential distribution with cdf

$$F_T(x) = \Pr[T_k \leq x] = 1 - e^{-\lambda x}\,,\; x \in \mathbb{R}^+\,, \tag{3.4}$$

where λ denotes the arrival intensity, with $\mathrm{E}[T_k] = 1/\lambda$. The burst sizes again form a sequence of iid rv's, with common cdf $F_B(x) = \Pr[B_k \leq x]$, $x \in \mathbb{R}^+$. The exact form of the cdf is completely general, except for the conditions that any useful cdf has to comply with: $0 \leq F_B(x) \leq 1$, $F_B(0) = 1$, $\lim_{x\to\infty} F_B(x) = 1$, and $F_B(x)$ is non-decreasing. The series $F_U(x)$ in CT has common cdf

$$F_U(x) = \Pr[U_k \leq x] = \Pr[B_k - T_k \leq x] \;\;,\;\; x \in \mathbb{R}^+\,.$$

Using (3.4), we obtain

$$F_U(x) = \begin{cases} e^{\lambda x} \int_0^{+\infty} e^{-\lambda u} dF_B(u) & ,\, x \in \mathbb{R}^-\,, \\ e^{\lambda x} \int_x^{+\infty} e^{-\lambda u} dF_B(u) + F_B(x) & ,\, x \in \mathbb{R}^+\,. \end{cases} \tag{3.5}$$

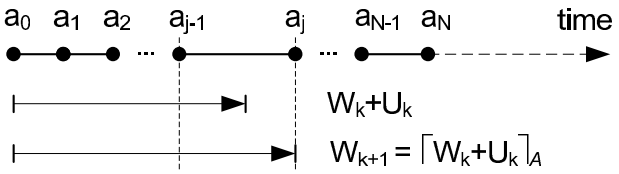
Note that the integral part of this equation does not pose difficulties for typical burst size distributions. For example, if the burst sizes have a common negative-exponential distribution with parameter μ, we have that $U(x) = e^{\lambda x} \cdot \mu/(\lambda + \mu)$ for $x \in \mathbb{R}^-$, and $U(x) = 1 - e^{-\mu x} \cdot \lambda/(\lambda + \mu)$ for $x \in \mathbb{R}^+$. Just like in DT, knowledge of the FDL set $\mathcal{A}$, $F_T(x)$, $F_B(x)$ and $F_U(x)$ (using (3.4) and (3.5), respectively) suffices as starting point for the analysis in CT. The latter will be provided in Sect. 3.1.6, first we focus on the analysis of the equivalent case in DT.

3.1.5 Analysis

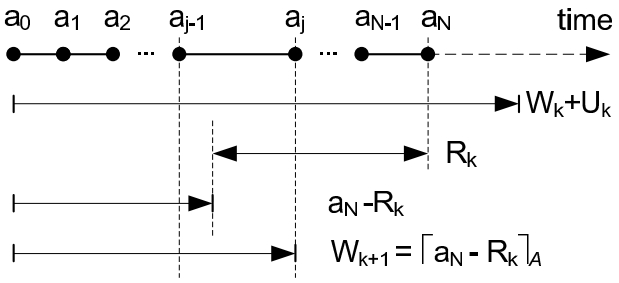
Given the key rv's ($F_U(x)$, $F_T(x)$), together with the parameter set of the FDL lengths $\{a_0, a_1, \ldots, a_N\}$, we are in the position to tackle the analysis. While a system description in terms of the scheduling horizon proved especially useful in the previous chapter, the current chapter is solely concerned with the waiting time evolution. Capturing the system equations (Sect. 3.1.5.1) in a Markov chain of waiting times (Sect. 3.1.5.2) will provide an exact numerical method to obtain the steady-state waiting time probabilities and loss probability of an FDL buffer in DT, under the given assumptions of general burst sizes and a Bernoulli arrival process. The complementary expressions for CT we consider in Sect. 3.1.6.

3.1.5.1 Evolution of the Waiting Time

The main idea of the analysis is that the system's evolution can be captured most condensedly in terms of the waiting time of a burst, as discussed in [7]. However, we still mention the scheduling horizon sporadically in the following, so as to



(a) Lossless transition.



(b) Transition with loss.

Figure 3.1: The waiting time of accepted bursts evolves according to one of two possible scenario's. A lossless transition is governed by the auxiliary function U_k; a transition with loss involves the reactivation time R_k.

make certain system properties more clear. Also, we make use of the term "virtual scheduling horizon", which denotes the scheduling horizon as it would be seen by a hypothetical (and not necessarily actual) arrival. This is opposed to the scheduling horizon itself, that is defined only for actual arrivals.

Still using the same numbering, we associate the waiting time W_k with the kth burst, and define it as the time between the acceptance of burst k, and the start of its transmission. Focusing on the evolution of the waiting time from acceptance to acceptance, we observe two types of transitions, either without or with loss.

Lossless transition. In case of a lossless transition, the burst that arrives just after the kth burst can be provided with a sufficiently large delay, and $W_k + U_k \leq a_N$. Note that, even if burst k pushes the virtual scheduling horizon beyond the maximum delay, that is, $W_k + B_k > a_N$, the transition remains lossless as long as burst $k+1$ arrives "sufficiently late", that is, $W_k + U_k \leq a_N$. Thus, the arriving burst is accepted and is assigned index $k+1$. While the scheduling horizon as seen by burst $k+1$ is $W_k + U_k$, the FDL buffer can only provide delays that are in $\mathcal{A} = \{a_0, a_1, \ldots, a_N\}$, as reflected in operator (3.1). Inferring the waiting time of burst $k+1$ from this, we obtain

$$W_{k+1} = \lceil W_k + U_k \rceil_{\mathcal{A}} \,. \tag{3.6}$$

Transition with loss. In case of a transition with loss, the burst that arrives just after the kth burst cannot be provided with a sufficiently large delay, and $W_k+U_k > a_N$. More precisely, the queueing of burst k pushes the virtual scheduling horizon beyond the maximum delay a_N ($W_k+B_k > a_N$), and the burst that arrives just after burst k arrives "too early" to be accepted, that is, $W_k+U_k > a_N$. As a result, the burst following burst k is dropped, and the virtual scheduling horizon decreases further in time, until some later-arriving burst finds the (actual) scheduling horizon below the maximum achievable delay upon arrival. In this regard, we can call a system with virtual scheduling horizon larger than a_N unavailable, and call it available in the opposite case. Then, notice that a finite time R_k stretches between the moment the system turns available (for virtual scheduling horizon equal to a_N) and the moment that the next arrival occurs (for virtual scheduling horizon equal to $a_N - R_k$. This time period we call the reactivation time, and is associated with burst k. It follows that burst $k+1$ sees $a_N - R_k$ as scheduling horizon. Again, the FDL buffer only provides delays in $\mathcal{A}$, which implies that the waiting time of burst $k+1$ can be captured as

$$W_{k+1} = \lceil a_N - R_k \rceil_{\mathcal{A}} \,. \tag{3.7}$$

Further, in the special case of a memoryless arrival process, as assumed in this section, the cdf of the the inter-arrival time (3.2) is clearly reflected in the cdf of the reactivation time,

$$F_R(n) = \Pr[R_k \leq n] = 1 - \bar{p}^{n+1} \quad , \quad n \in \mathbb{N} \,. \tag{3.8}$$

That R_k and T_k have nearly identical distribution, is a direct implication of the memorylessness of the arrival process. The only difference is an offset of 1, that occurs due to the fact that the minimum of the support of the T_k equals 1. Note that this offset will not come about for the CT setting, as we will see below.

System equations (3.6) and (3.7) describe the waiting time process in a complementary and exhaustive way, and give rise to a uniquely defined Markov chain. Both are illustrated in Fig. 3.1.

3.1.5.2 Markov Chain of Waiting Times

The Markov chain consists of $N+1$ states, that correspond to $N+1$ possible waiting times a_i, $i = 0 \ldots N$. It is characterized by a transition matrix $\mathbf{M}$ with probabilities m_{ij},

$$m_{ij} = \Pr[W_{k+1} = a_j | W_k = a_i] \quad , \quad 0 \leq i, j \leq N \,. \tag{3.9}$$

For ease of notation, we introduce $a_{-1} = -\infty$. We split m_{ij} in two separate contributions, that correspond to the transitions discussed in Sect. 3.1.5.1.

$$\begin{aligned} m_{ij} &= \Pr[a_{j-1} - a_i < U_k \leq a_j - a_i] \\ &\quad + \Pr[U_k > a_N - a_i]\Pr[a_N - a_{j-1} > R_k \geq a_N - a_j]\,. \end{aligned}$$

With the expression for $F_T(x)$ (3.4), $F_U(x)$ (3.5) and $F_R(x)$ (3.8), this can be restated as

$$\begin{aligned} m_{ij} &= F_U(a_j - a_i) - F_U(a_{j-1} - a_i) \\ &\quad + \bar{p}^{a_N}[1 - F_U(a_N - a_i)][\bar{p}^{-a_j} - \bar{p}^{-a_{j-1}}]\,. \end{aligned} \tag{3.10}$$

Due to the memoryless arrival process, the Markov chain of the waiting times is aperiodic, irreducible and has positive-recurrent states. (This is not necessarily so in case of general inter-arrival times, see Sect. 3.2.3.2.) As a result, we obtain a unique equilibrium distribution for $k \to \infty$, and all involved rv's can be associated with a steady-state distribution independent of the k, with $\{W, B, T\}$ denoting generic rv's following the steady-state distributions of $\{W_k, B_k, T_k\}$, respectively. Note that this is very similar to the steady-state assumption in Chapter 2, but with one main difference. To obtain a unique equilibrium distribution in Chapter 2, we were to assume a bound on the traffic load ($\rho < \rho_{max}$), so as to guarantee stability. Here, the nature of the problem implies the existence of a unique stochastic equilibrium distribution, and stability is inherent to the finite buffer setting.

With now the transition probabilities m_{ij} at hand, a simple numerical procedure yields the waiting time probabilities associated with this steady-state. More precisely, the normalized Perron-Frobenius eigenvector of the matrix $\mathbf{M}$ contains the $N + 1$ different steady-state waiting time probabilities

$$\lim_{k\to\infty} \Pr[W_k = a_n] = \Pr[W = a_n] = w(n), \quad 0 \leq n \leq N\,, \tag{3.11}$$

and can easily be obtained numerically, posing no problem for the small N we are interested in. From this, we can also define the mean waiting time $\mathrm{E}[W] = \sum_{n=1}^{N} w(n) \cdot a_n$.

3.1.5.3 Loss Probability

Finally, the loss probability (LP) of an arbitrary burst is also obtainable in a straightforward manner. To find an expression for the LP, we study the unavailable period, associated with an accepted burst k, again distinguishing between the two transitions of Sect. 3.1.5.1. In case of a lossless transition, the arrival of burst k does not push the system into unavailability, and the unavailable period following burst k equals zero. In case of a transition with loss, it takes the system a period of length $W_k + B_k - a_N - 1$ to become available again. (Note that here, the term "-1"

is again a result of the offset on the T_k, as the minimum of its support equals 1.) Combination of both cases yields that the unavailable period, following burst k, is given by $(W_k + B_k - a_N - 1)^+$. Invoking the memoryless nature of the arrival process, the average number $\mathrm{E}[Z_k]$ of bursts lost during the unavailable period following burst k equals $\lambda \cdot \mathrm{E}[(W_k + B_k - a_N - 1)^+]$.

Now, we assume steady state and continue the analysis with the generic rv's associated with it, with Z instead of Z_k. This also enables us to use the $w(n)$ associated with W. Some calculation leads to

$$\mathrm{E}[Z] = p \cdot \left(\mathrm{E}[B] + \mathrm{E}[W] - a_N - 1 + \sum_{n=0}^{N} w(n) \sum_{i=1}^{a_N - a_n} F_B(i) \right) . \qquad (3.12)$$

Now, it suffices to note that, with every accepted burst, a number of $\mathrm{E}[Z]$ bursts on average is dropped, resulting in a LP

$$\mathrm{LP} = \frac{\mathrm{E}[Z]}{1 + \mathrm{E}[Z]} . \qquad (3.13)$$

3.1.6 Analysis for CT

For CT, the system equations are identical to those of Sect. 3.1.5.1:

$$W_{k+1} = \lceil W_k + U_k \rceil_{\mathcal{A}} ,$$

in the case of a lossless transition ($W_k + U_k \leq a_N$), and

$$W_{k+1} = \lceil a_N - R_k \rceil_{\mathcal{A}} ,$$

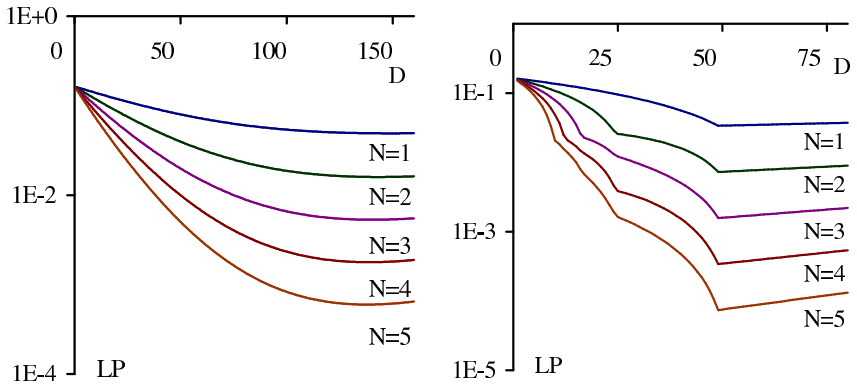
in case of a transition with loss ($W_k + U_k > a_N$). The difference is that the cdf of the reactivation time in CT is given by

$$F_R(x) = \Pr[R_k \leq x] = 1 - e^{-\lambda x} , \; x \in \mathbb{R}^+ ,$$

and thus coincides with that of the T_k (3.4). When compared to the DT case (3.8), the offset of 1 thus disappears, since in CT, the minimum of the support of T_k is zero. Further, the coefficients of the Markov chain (Sect. 3.1.5.2), also coincide, and the expression for the m_{ij} in DT (3.10) (with the $F_U(x)$ of (3.5)) is equally valid for CT (with the $F_U(n)$ of (3.3)), if one makes the simple substitution $\bar{p} = e^{-\lambda}$. The resulting expression for CT reads

$$\begin{aligned} m_{ij} \quad = \quad & F_U(a_j - a_i) - F_U(a_{j-1} - a_i) \\ & + e^{-\lambda a_N} [1 - U(a_N - a_i)][e^{\lambda a_j} - e^{\lambda a_{j-1}}] . \end{aligned}$$

As such, the waiting time probabilities $w(n)$ for steady state can be obtained in exactly the same manner. As for the LP in DT, it suffices to replace p with λ



(a) Geometric burst size distribution.

(b) Deterministic burst size distribution.

Figure 3.2: For low buffer size and geometric burst size distribution (a), the LP curves for varying granularity are not much different from those for large buffer size. The opposite is true for deterministic burst size distribution (b), for which the number of "notches" in the curves matches the buffer size N. These figures were obtained for $\mathrm{E}[B] = 50$ *slots, load* $\rho = 20\%$.

in expression (3.12), substitute the summation by an integration, and take into account the relative offset of 1, to obtain the correct formula for CT,

$$\mathrm{E}[Z] = \lambda \cdot \left(\mathrm{E}[B] + \mathrm{E}[W] - a_N + \sum_{n=0}^{N} w(n) \int_0^{a_N - a_n} F_B(u) du \right) ,$$

that leads to the LP through (3.13), $\mathrm{LP} = \mathrm{E}[Z]/(1 + \mathrm{E}[Z])$.

3.1.7 Numerical Examples

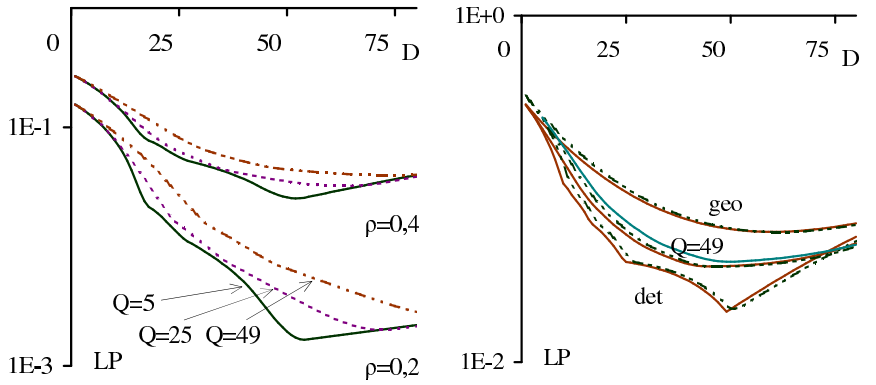
With the above results at hand, one can easily study the impact of the various design parameters on loss performance. More specifically, one wants to determine optima for the granularity, which are values that yield a minimal burst loss probability. While similar curves already occur in the previous chapter for $N = 20$, the approximation applied there lost accuracy for small N. Also, as mentioned, the previous chapter only allowed for a degenerate buffer setting, whereas the current approach makes no special assumptions on the line lengths. As such, the examples given here yield new information, for the case of smaller (more realistic) buffer sizes. We look at four examples for DT, and note that the CT yields similar results.

In both panes of Fig. 3.2, we assess the impact of a small buffer size on loss performance. They show the LP as a function of the granularity D, for a degenerate delay line setting, $\mathrm{E}[B] = 50$, and a load of $\rho = 0.2$. The left pane displays

the situation for geometric burst size distribution, with the expected rise of the loss probability, when the buffer size N decreases. Also, it can be seen that the LP lowers for increasing granularity, and only starts to rise again for D larger than about 150 slots. Considering the five curves together, the figure suggests that the optimal granularity, for a given load, only slightly increases when the buffer size approaches its minimum of $N = 1$. Verification for higher load values confirm that the influence of diminishing buffer size on optimal granularity is but weak, especially if compared to the impact of variations in the load. The latter is illustrated in Fig. 2.2(a) (there for $N = 20$, in CT), and is much stronger than the influence observed here.

On the right pane of Fig. 3.2, we have the same setting, for fixed burst sizes, $\mathrm{E}[B] = B = 50$. Again, a smaller buffer suffers more loss, but now the optima alter in a more surprising way. More precisely, the "notches" at $D = (B-1)/n$, $n = 1, 2, \ldots$, known from previous work, are not uncountable (as was the case in the previous chapter for an infinite-sized buffer, see Fig. 2.2(b) for $N = 20$ in CT), but are limited in number to N. Exactly N notches occur for each curve, at $D = (B-1)/n$, $n = 1, 2 \ldots N$. As such, this relation is only revealed by explicitly assuming finiteness from the start of the analysis, as done in the current chapter. Also, since it is known from the previous chapter that these values correspond to optima (with $D = B-1$ being the optimum for low load, and $D = (B-1)/2$, then $(B-1)/3$,... for increasing load), we verified and found that the set of potential optima, for a load $0 \leq \rho \leq 1$, is indeed limited to (at the maximum) the number of fibers. Further, it was found that the same optimum shift as known from the previous chapter takes place, but now only over the available values: first $D = B-1$ for low load, then $D = (B-1)/2$ for increasing load if $N \geq 2$, and so on if $N \geq 3$...

To verify if the optima for deterministic burst sizes also apply to varying burst sizes, we consider a uniform burst size distribution with radius Q, that has a mean burst size $\mathrm{E}[B] = 50$, and is uniform within the range $[50 - Q, 50 + Q]$. For small Q, this distribution resembles the deterministic distribution. As such, this setting allows to verify what influence variances on the burst size have on loss performance. In Fig. 3.3, the left pane compares the performance of an optical buffer of size $N = 3$ for three uniform distributions, having a narrow range ($Q = 5$, range $[45, 55]$), an intermediate range ($Q = 25$, range $[25, 75]$) and a broad range ($Q = 49$, range $[1, 99]$), and this for load $\rho = 0.2$, $\rho = 0.4$, respectively. For the narrow-ranged one, the curves look very similar to those of the deterministic distribution, and the same optimum around $D = B - 1$ shows. The curves for the intermediate-ranged and broad-ranged case show that increasing Q makes the granularity optimum shift toward higher values, at least for load $\rho = 0.2$ and $\rho = 0.4$. Curves not included here, for higher load, show that the optima for a narrow range (small Q) concentrate around the limited set $D = (B-1)/n$,



(a) Various uniform burst size distributions, degenerate buffer setting, for $N = 3$.

(b) Degenerate (continuous) or non-degenerate (dashed), for $N = 3$, and $\rho = 50\%$

Figure 3.3: For a uniform burst size distribution, the range Q has a paramount impact on LP curves for varying granularity (a). Comparing the performance of a non-degenerate buffer over a degenerate one for deterministic (det), uniform ($Q = 49$) and geometric (geo) burst size distribution, one finds that a slight performance gain that can be realized by choosing a non-degenerate setting (b). These figures were obtained for $\mathrm{E}[B] = 50$ slots.

$n = 1, 2 \ldots N$, known from the deterministic distribution, while for larger Q, the optimum only gradually decreases.

Choosing non-degenerate lengths for the delay lines can in some cases provide better performance. As is shown in [17] for deterministic burst sizes, a non-degenerate set of FDLs can outperform a degenerate set of the same size. However, it turns out that this happens only when the load rises above a certain value (for example 60.17%, for $N = 10$, $B = 20$, $D = 19$). Further, even when the non-degenerate one outperforms the degenerate one, the performance gain is rather small. This said, non-degenerate settings remain interesting, since for more general assumptions (correlated arrivals, multi-wavelength output), the performance gain might be larger. For the right pane of Fig. 3.3, we chose non-degenerate FDL lengths in a way similar to [17], with shortened lengths for the largest lines. A degenerate set (continuous curves) and non-degenerate set (dashed curves) are considered, for a buffer size $N = 5$, load $\rho = 0.5$, and the burst size distributions geometric, deterministic and uniform ($Q = 49$). The non-degenerate set has FDL lengths D, $2D - 2$, $3D - 3$, $4D - 4$, $5D - 8$. The curves show how the non-degenerate set just outperforms the degenerate one, for geometric and uniform burst size distribution, while the opposite is true for a deterministic burst size distribution. Although not included here, figures for the same setting, for a load of $\rho = 0.3$, $\rho = 0.6$ and $\rho = 0.8$ resp., show the same qualitative result, while the

performance difference itself always remains small.

3.2 Model for General Arrivals

In this section, we extend the results of the previous section to account for general inter-arrival times. Just like in the previous chapter, this will allow us to quantify the impact of (a mild form of) correlation, but now in an exact manner. With the pgf approach, such analysis was possible, but only when one assumed a rational pgf for the inter-arrival times. Now, the analysis is possible in an exact manner for finite-sized buffers, but the "price to pay" is that we have to assume that the burst size distribution is upper-bounded by some value B_{max} (defined below in 3.14). The latter is however not a very limiting assumption, since this is in general also assumed in the application.

3.2.1 Time Setting and Buffer Setting

The setting under consideration is very similar to that of the previous one, except that only the DT case is considered. Treating CT is in principle also possible, but is not possible in such plug-and-play fashion as was the case for memoryless arrivals, and is left out of consideration here. The optical buffer structure is simply identical to the one considered in the previous section, in Sect. 3.1.2.

3.2.2 Traffic Setting

The traffic setting is largely the same as in Sect. 3.1.3, but now with a general distribution for the inter-arrival times, and an extra assumption on the distribution of the burst sizes. Also, while we only used the cdf of the involved rv's in Sect. 3.1, we now use both the cdf and probability mass function (pmf). For clarity's sake, we prefer to list all main assumptions here.

We assume a DT setting, with bursts arriving one by one, so either one or no arrival occurs in each slot. Upon arrival, a burst is either accepted or dropped. We number the bursts in the order at which they arrive, but only assign an index to those bursts that are accepted.

With each accepted burst k, we associate an inter-arrival time T_k, that captures the time between the kth arrival and the next, being the arrival of (i) burst $k+1$, if this next burst is accepted or (ii) a burst without number, if this next burst is dropped.

In the following, we assume general (non-zero) inter-arrival times T_k. They form a sequence of iid rv's with some common distribution

$$\Pr[T_k = n] = t(n) \quad , \quad n \in \mathbb{N}_0 \, .$$

The exact form of this distribution is completely general, except for the conditions that any useful probability mass function has to comply with: $0 \leq t(n) \leq 1$, $\sum_n t(n) = 1$. The cdf of the inter-arrival times is $F_T(n) = \Pr[T_k \leq n] = \sum_{i=1}^{n} t(n)$, for $n \in \mathbb{N}_0$. The inter-arrival times associated with dropped bursts also follow this distribution. Since there is also no upper limit on the inter-arrival times, all common distributions (geometric, negative binomial, mix of geometric) can be assumed. One notes that, although the inter-arrival times are uncorrelated, it is nevertheless possible to model correlation in the arrival process, as we will discuss in Sect. 3.2.5.

With each accepted burst, we also associate a burst size B_k. The burst sizes also form a sequence of iid rv's with a common distribution, but with the additional restriction of an essential upper bound on the burst size distribution, B_{max}, defined as

$$B_{max} = \text{esssup}\{B_k\} \equiv \sup(x : \Pr[B_k > x] > 0) \ , \ x \in \mathbb{R}^+ \, , \tag{3.14}$$

so that the distribution itself adheres to

$$\Pr[B_k = n] = b(n) \, ,$$

with $1 \leq n \leq B_{max}$ if B_{max} is finite and $1 \leq n < +\infty$ if B_{max} is infinite, and $0 \leq b(n) \leq 1$, $\sum_n b(n) = 1$.

As in Sect. 3.1, results can be formulated most condensedly in terms of the series of rv's $U_k = B_k - T_k$, with

$$\Pr[U_k = n] = \Pr[B_k - T_k = n] = u(n) \ , \ n \in \mathbb{Z} \, .$$

Taking into account the upper limit to B_k, B_{max}, it follows that

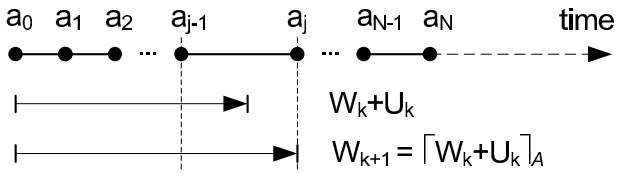
$$u(n) = \begin{cases} \sum_{i=1}^{B_{max}} b(i)t(i + |n|) & , n \in \mathbb{Z}^- \, , \\ \sum_{i=1}^{B_{max}-n} b(i+n)t(i) & , n \in \{1, 2, \ldots, B_{max} - 1\} \, , \\ 0 & , n \in \{B_{max}, B_{max} + 1, \ldots\} \, . \end{cases} \tag{3.15}$$

We will also employ the cdf of U_k in the following, $F_U(n) = \sum_{i=-\infty}^{n} u(n)$, $n \in \mathbb{Z}$. Note that the lower limit for $i \to -\infty$ never poses any numerical issue, since one can always rewrite the expression for $F_U(n)$ as a function of $F_T(n)$ and finite sums.

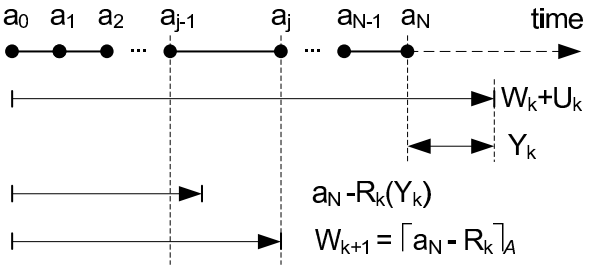
3.2.3 Analysis

3.2.3.1 Evolution of the Waiting Time

In Sect. 3.1.5.1, it was shown that for a memoryless arrival process, the system evolution can be captured in terms of two possible transitions, that correspond to



(a) Lossless transition.



(b) Transition with loss.

Figure 3.4: The waiting time of accepted bursts evolves according to one of two possible scenario's. A lossless transition is governed by the auxiliary function U_k; a transition with loss involves the reactivation time R_k, that is a function of the unavailable period Y_k. Notice the slight difference with the case of memoryless arrivals depicted in Fig. 3.1, where Y_k could be disregarded, whereas here, R_k is a function of Y_k.

two mutually exclusive events. It turns out that, also for general inter-arrival times, it is possible to maintain this kind of description, even without making any change to the form of the related system equations, (3.6) and (3.7). We repeat them here for clarity's sake, and refer to Sect. 3.1.5.1 for the reasoning behind it.

Lossless transition. The condition for this to happen, in terms of waiting times and FDL lengths, is that the burst that arrives just after the kth burst can be provided with a sufficiently long delay, that is,

$$W_k + U_k \leq a_N \,,$$

and results in

$$W_{k+1} = \lceil W_k + U_k \rceil_{\mathcal{A}} \,. \tag{3.16}$$

Transition with loss. Now, the acceptance of burst k pushes the system into unavailability, and the burst that arrives just after the kth burst arrives "too early", and thus can not be provided with a sufficiently long delay:

$$W_k + U_k > a_N \,.$$

The burst following burst k (first drop) is not assigned an index, and possibly, even more bursts are dropped (second drop, third drop,...) before the system turns available again, and, R_k time slots later, a burst arrives and is accepted, with waiting time W_{k+1}.

$$W_{k+1} = \lceil a_N - R_k \rceil_{\mathcal{A}} \,. \tag{3.17}$$

As such, the main difference with the previous section lies not in the system equations, but in the analysis thereof. More particularly, the reactivation time R_k in (3.7) now has to be determined for a general arrival process. In such case, the reactivation time is no longer independent of the system state (as was the case in the previous section) but actually depends on W_k explicitly. This dependence can be characterized by means of an additional rv Y_k, that captures the unavailable period following the acceptance of burst k.

The unavailable period expresses the number of slots between the arrival of the first-dropped burst and the instant that the buffer becomes available again. It is defined as

$$Y_k = W_k + U_k - a_N \,.$$

Conditioning on a given waiting time, $W_k = a_i$, $0 \leq i \leq N$, we can easily express the probabilities for Y_k using (3.15),

$$\Pr[Y_k = n \,|\, W_k = a_i] = u(n - a_i + a_N) \,. \tag{3.18}$$

Now, the moment that the buffer comes available again is fully characterized, since this occurs exactly $T_k + Y_k$ slots after the arrival of burst k. Recall that a burst does not arrive immediately, and in general, it takes burst $k+1$ another R_k slots to arrive, which brings the arrival instant of burst $k+1$ to $T_k + Y_k + R_k$ slots after the arrival of burst k.

The reactivation time R_k depends on W_k only through the unavailable period Y_k: any combination of W_k and U_k with equal Y_k yields the same R_k. More precisely, their relation can be expressed as

$$\Pr[R_k = m \,|\, Y_k = n] = \sum_{l=0}^{n-1} h(l) \cdot t(m + n - l) \,. \tag{3.19}$$

Here, the $t(n)$ are the probabilities of Sect. 3.2.2, while the $h(l)$ are given by the recursive definition

$$h(l) = \sum_{k=0}^{l-1} h(k) \cdot t(l-k) + \delta_l \quad , \quad l \geq 0 \,, \tag{3.20}$$

where δ_l denotes the Kronecker delta, that is one for $l = 0$ and zero elsewhere. Given an arbitrary arrival instant t_{arr} (expressed in slots), $h(l)$, $l \geq 0$, is the

probability that $t_{arr} + l$ (in slots) is an arrival instant too, a probability that often arises in renewal theory. The probabilities can be obtained from (3.20).

Concluding, the characterization of the system behavior falls apart in two types of transitions, (3.16) and (3.17), both illustrated in Fig. 3.4. In the case of a lossless transition (3.16), the analysis coincides with that of the previous section. In case of a transition with loss (3.17), the new waiting time W_{k+1} depends on the reactivation time R_k, that in its turn depends on the unavailable period Y_k, that, finally, depends on the previous waiting time, W_k.

3.2.3.2 Markov Chain of Waiting Times

With the two system equations (3.16) and (3.17) as a starting point, and the expressions for the (conditional) probabilities of the unavailable period (3.18) and the reactivation time (3.19) obtained, we are in the position to consider the Markov chain, associated with the evolution of the waiting times. Similar to (3.9), it consists of $N + 1$ states, that correspond to $N + 1$ possible waiting times a_i, $i = 0 \ldots N$, and is characterized by a transition matrix $\mathbf{M}$ with probabilities m_{ij},

$$m_{ij} = \Pr[W_{k+1} = a_j \,|\, W_k = a_i] \;\;,\;\; 0 \leq i, j \leq N \,. \tag{3.21}$$

As in the previous section, we denote $a_{-1} = -\infty$. We split m_{ij} in two separate contributions, that correspond to the events discussed in Sect. 3.2.3.1.

$$\begin{aligned} m_{ij} &= \Pr\Big[a_i + U_k \leq a_N \,,\, a_j = \lceil a_i + U_k \rceil_{\mathcal{A}}\Big] \\ &\quad + \Pr\Big[a_i + U_k > a_N \,,\, a_j = \lceil a_N - R_k \rceil_{\mathcal{A}}\Big] \\ &= \Pr\Big[a_{j-1} - a_i < U_k \leq a_j - a_i\Big] \\ &\quad + \Pr\Big[Y_k > 0 \,,\, a_N - a_{j-1} > R_k \geq a_N - a_j\Big] \,. \end{aligned} \tag{3.22}$$

With the cdf of U_k, we can rewrite the first term of (3.22) as

$$\Pr[a_{j-1} - a_i < U_k \leq a_j - a_i] = F_U(a_j - a_i) - F_U(a_{j-1} - a_i) \,.$$

Taking into account the upper limit for B_k, B_{max}, the second term of (3.22) can be expressed as

$$\begin{aligned} &\Pr[Y_k > 0 \,,\, a_N - a_{j-1} > R_k \geq a_N - a_j] \\ &= \sum_{n=1}^{B_{max}+a_i-a_N} \Pr[Y_k = n \,|\, W_k = a_i] \cdot \Pr[W_{k+1} = a_j \,|\, Y_k = n] \,. \end{aligned}$$

The conditional probabilities $\Pr[Y_k = n \,|\, W_k = a_i]$ are given by (3.18), while

$$\Pr[W_{k+1} = a_j \,|\, Y_k = n] = \sum_{m=a_N-a_j}^{a_N-a_{j-1}-1} \Pr[R_k = m \,|\, Y_k = n] \,. \tag{3.23}$$

Using (3.19), the latter is, in its final form,

$$\begin{aligned}&\Pr[W_{k+1} = a_j \,|\, Y_k = n] \\ &\quad = \sum_{l=0}^{n-1} h(l) \cdot \{F_T(n-l+a_N-a_{j-1}-1) - F_T(n-l+a_N-a_j-1)\}\,,\end{aligned}$$

where $F_T(n)$ denotes the cdf of the inter-arrival times.

Filling in the last four expressions in (3.22) yields an explicit expression for the coefficients m_{ij}, in terms of the (given) a_i, $t(n)$ and $b(n)$.

With all m_{ij} determined, next step is to move to the steady-state distribution of the waiting times, like in Sect. 3.1.5.2. There, the steady-state distribution of the waiting times was unique, since the involved Markov chain was aperiodic, irreducible, and composed of positive-recurrent states. This is no longer so in the present case, due to the fact that both inter-arrival times and burst sizes are general. More precisely, in some "pathological cases", it is possible that the Markov states fall apart in multiple irreducible classes. In that case, if we assume zero initial conditions (that is, an empty buffer upon arrival of the first burst), only the states belonging to the same class as the zero waiting time state can be reached, and the steady-state distribution as such depends on the initial state. Further, note that also periodicity can arise. Since these pathological cases are however of minor practical relevance, we exclude them in the following for simplicity's sake. For a mathematically sound approach to this problem, we refer the interested reader to Sect. 5.1.4.3. For now, assuming an irreducible and aperiodic Markov chain with positive-recurrent states allows to obtain a unique equilibrium distribution for $k \to \infty$, with all involved rv's associated with a steady-state distribution, and $\{W, B, T, Y\}$ denoting generic rv's following the steady-state distributions of $\{W_k, B_k, T_k, Y_k\}$, respectively. Now, the normalized Perron-Frobenius eigenvector of the matrix $\mathbf{M}$ contains the $N+1$ different steady-state waiting time probabilities

$$\lim_{k\to\infty} \Pr[W_k = a_n] = \Pr[W = a_n] = w(n) \;\;,\;\; 0 \le n \le N\,, \tag{3.24}$$

and can easily be obtained with standard (numerical) methods, posing no problem for the small N we are interested in. From this, we can also define a mean waiting time $\mathrm{E}[W] = \sum_{i=1}^{N} w(i) a_i$.

3.2.3.3 Loss Probability

Up to now, dropped bursts remained unnumbered, and the assigned waiting times are those of the accepted burst. We now focus on the LP, defined as the probability that an arbitrary arriving burst is lost. To this end, we study the number of dropped

bursts between arrival k and $k+1$, denoted Z_k, and its mean, $\mathrm{E}[Z_k]$. Again, we consider the two transitions of Sect. 3.2.3.1. In case of a lossless transition, $\mathrm{E}[Z_k]$ trivially equals zero. In case of a transition with loss, at least one burst (the first-dropped one) is lost, so $Z_k \geq 1$. Moreover, during the unavailable period Y_k, additional loss occurs. Although not self-evident, it turns out that the mean number of losses (including the first-dropped one), given an unavailable period $Y_k = n$, can be expressed as

$$\mathrm{E}[Z_k \,|\, Y_k = n] = \sum_{l=0}^{n-1} h(l)\,,$$

with the $h(l)$ still given by (3.20). Employing (3.24) and (3.18), and assuming steady state (with thus Z instead of Z_k), we obtain that

$$\begin{aligned}
\mathrm{E}[Z] &= \sum_{i,n} \{\, \Pr[W = a_i] \cdot \Pr[Y = n \,|\, W = a_i] \cdot \mathrm{E}[Z \,|\, Y = n] \,\} \\
&= \sum_{i=0}^{N} w(i) \sum_{n=1}^{B_{max}+a_i-a_N} u(n - a_i + a_N) \sum_{l=0}^{n-1} h(l) \\
&= \sum_{i=0}^{N} w(i) \sum_{l=0}^{B_{max}+a_i-a_N} h(l)\{1 - F_U(l - a_i + a_N)\}\,.
\end{aligned}$$

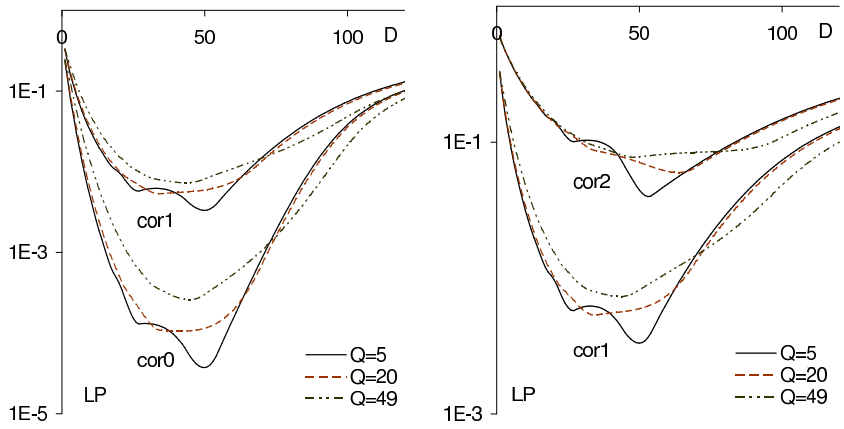
Since with every accepted burst, a number of $\mathrm{E}[Z]$ bursts on average is dropped, we obtain as LP, through (3.13), that $\mathrm{LP} = \mathrm{E}[Z]/(1 + \mathrm{E}[Z])$.

3.2.4 Light-Weight

The algorithm proposed in Sect. 3.2.3.2 and 3.2.3.3 is light-weight in that it poses minimal computational burden, also in comparison with [31]. This is the result of two distinct effects: state space reduction, and the use of proper auxiliary variables, here F_U and h.

The first effect is the main difference with [31], from which all other differences stem. Since the assigned waiting times W_k (as they appear in system equations (3.16) and (3.17)) are bound to be in the FDL set $\mathcal{A}$, the number of delay lines comes about as the minimal number of states, necessary to be able to describe the system in an exact way. Fixing the number of states to $N+1$ thus not only makes the problem more generic (yielding considerable numerical benefits), but also minimizes the state space, reducing the necessary numerical inversion of the transition matrix to a simple problem.

The first effect imposing an aim rather than being a solution itself, the main task in this work was to express the transition probabilities m_{ij} of (3.21) in such way that they can actually be calculated, in finite time. To this end, it proved useful to impose an upper limit on the burst size, and introduce the auxiliary rv h in (3.20) (where the variable F_U was already introduced, first in [20], to the same end).



(a) Correlation level 0 and 1 (cor0, cor1). (b) Correlation level 1 and 2 (cor1, cor2).

Figure 3.5: Maintaining a fixed traffic load of $\rho = 0.5$, different combinations of uniform burst size distributions ($Q \in \{5, 20, 49\}$) and inter-arrival time distributions (cor0, cor1, cor2) yield nine different LP curves in total (with three of them repeated across figures as a reference). These figures were obtained for a degenerate FDL set, $\mathrm{E}[B] = 50$ slots, and $N = 20$.

Note that the algorithm as presented here also displays a clear distinction between traffic pattern and buffer setting: given a certain distribution for B and T, one can first calculate all needed values F_U and h at once. As such, it is only in a second step that the specific values of the FDL set $\mathcal{A}$ are taken into account. The second step uses the output of the first step (F_U and h) to obtain the m_{ij} of (3.21), which are essentially just linear combinations of values of F_U, h and F_T. Also, note that the influence of a load variation is strictly limited to the first step. This two-step calculation process is especially fit for optimizations of the FDL set for a given load, since one can cut down the optimization process to only a repetition of the second step. For the case of an M/G/1 in DT, this optimization is the subject of [19], in which the author is involved.

3.2.5 Numerical Examples

With the formulas at hand, a simple numerical program suffices to study the complex interplay between all involved variables, yielding results fast for any combination. Since the model allows to obtain exact results instantaneously for any combination of inter-arrival time and burst size distribution (and thus any load), any set of FDLs (not necessarily degenerate) and any buffer size, it can provide the framework for an extensive optimization study. Here, we confine ourselves to an

	$Q = 5$	$Q = 25$	$Q = 49$
range	$[45, 55]$	$[25, 75]$	$[1, 99]$

	cor0	cor1	cor2
G	1	4	4
S	50	50	50/4

Table 3.1: Parameter settings of three uniform burst size distributions with different range (left), for three levels of correlation in the arrival process (right).

instructive example, that focuses on the loss optimization of the granularity value given that burst sizes are uniform, with variable range, and correlation occurring on three different levels. Note that this is also the example considered in the previous chapter, where it was treated only approximatively, whereas now, exact results are obtained.

In both panes of Fig. 3.5, we quantify the loss performance of an FDL buffer with size $N = 20$ and degenerate fiber lengths $a_i = iD$, $i = 0 \dots 20$. The burst loss probability is displayed as a function of the granularity D, that varies from 1 slot (line lengths $1, 2 \dots 20$) to 120 slots (line lengths $120, 240 \dots 2400$). We consider three different burst size distributions, all uniform with $\mathrm{E}[B] = 50$ but with different radius Q and associated range [$50 - Q$,$50 + Q$]: a narrow range ($Q = 5$, range $[45, 55]$), an intermediate range ($Q = 25$, range $[25, 75]$) and a broad range ($Q = 49$, range $[1, 99]$). This is also set out in Table 3.1. The load ρ, with common definition $\mathrm{E}[B]/\mathrm{E}[T]$, is fixed to $\rho = 0.5$, and thus also the mean inter-arrival time $\mathrm{E}[T]$, to 100 slots.

For these three burst size distributions, we study the impact of correlation on three different levels, by considering three different parameter settings for the inter-arrival time distribution, a mix of geometric distributions,

$$t(n) = \alpha p_S (1 - p_S)^{n-1} + (1 - \alpha) p_L (1 - p_L)^{n-1} \ , \ n \geq 1 \, ,$$

with $0 \leq \alpha \leq 1$, and $0 < p_L < p_S < 1$. This distribution is indeed able to capture correlation, through the arrival of groups, and was used earlier for that scope in Sect. 2.4.6, for Fig. 2.10. The time between arrivals of members of the same group is distributed geometrically with parameter p_S and short mean value $S = 1/p_S$. This occurs with probability α. The time between arrivals of two groups, that terminates the previous group and initiates a new one, is distributed geometrically with parameter p_L and long mean value $1/p_L$. This occurs with probability $1 - \alpha$. Resulting, one obtains that the mean number of arrivals within one group is $G = 1/(1 - \alpha)$. In Fig. 3.5, pane (a) considers correlation level 0 and 1 (cor0, cor1), while pane (b) repeats correlation level 1, in order to compare it with level 2 (cor1, cor2). All three levels have $\mathrm{E}[T] = 100$ slots, but different G and S. Correlation level 0 has $G = 1$ and $S = 50$ slots, which corresponds to a memoryless geometric distribution, with no correlation. At level 1, correlation is augmented by increasing the group size, $G = 4$, while S remains 50 slots. At

level 2, correlation rises further as S diminishes to $50/4$ slots, while G remains 4. The parameter settings for the three correlation levels are tabulated in Table 3.1. Notice that exactly the same setting was considered in Sect. 2.4.6, with cor0, cor1, cor2 there called run 1, run 3 and run 5, respectively. Further, note that the results in Sect. 2.10 were approximate, whereas we obtain exact results here.

A first interpretation of the results comes as expected: increased correlation leads to increased loss, for all three radii $Q = 5, 20, 49$. Further verification showed that this is so for any load $0 < \rho \leq 1$ and any radius $0 < Q < 50$. Also, the optimal granularity is always between 0 and B_{max}. This optimum is a result of the trade-off between the granularity D (that is best kept small) and the maximum achievable delay ND (that is best kept large). However, the evolution of this optimum under a correlation increase is far from self-evident and calls for a separate discussion.

For a small radius ($Q = 5$), we observe a pronounced optimum for D around $\mathrm{E}[B + Q]$, that was already reported for the uncorrelated case in [7], and is very similar to the case of deterministic burst sizes, first discussed in [10]. As Fig. 3.5 suggests, this value remains optimal also under a correlation increase. Results for other loads point out that the optimum is even more stable in case of higher correlation. Specifically, $D \lesssim \mathrm{E}[B + Q]$ is optimal for $\rho < 0.58$ (cor0), $\rho < 0.59$ (cor1) and $\rho < 0.96$ (cor2). For an intermediate ($Q = 20$) and large radius ($Q = 49$), Fig. 3.5 displays how the optimum for D, associated with cor0, still remains stable under a first correlation increase, cor1. However, going from cor1 to cor2, the optimum for $Q = 20$ shifts to the higher value $D \lesssim \mathrm{E}[B + Q]$, while the optimum for $Q = 49$ also increases, but to a lesser extent. Verification for $0 < \rho \leq 1$ confirmed this for any Q: under a highly correlated arrival process (cor2) the optimum either remains the same, or shifts towards a higher value.

3.3 Solution for Upper-Bounded Burst Size Distribution

In the current and the next section, we focus on two special instances of the general model discussed in Sect. 3.1. More precisely, we narrow down the assumptions to consider only a degenerate buffer setting with granularity D, but however maintain generality for the time setting. As in the rest of this chapter, we maintain the (arbitrary) convention to first study the system in DT, and then in CT. For upper-bounded burst size distribution (this section), direct application of the model in Sect. 3.1 enables a closed-form solution when the granularity is larger than or equal to the mentioned upper bound. For memoryless burst size distribution (Sect. 3.4), a modified model based on that of Sect. 3.1 allows for a closed-form solution of the problem for any value of the granularity.

We note that the main motivation for a closed-form solution is not the reduction in computation time, since the method of Sect. 3.1 can be easily implemented in software, yielding results instantly (order of μs) for any parameter setting. Rather, we feel that closed-form expressions are extremely easy to use, especially when their form is as simple as that of the formulas we obtain here. Also, their simple form allows for more insight in the functioning of an FDL buffer in general.

3.3.1 Implications of an Upper Bound

When burst sizes are upper-bounded, and a single-wavelength buffer with FIFO scheduling is considered, an important notion is that the optimal granularity D, say D_0, will never be larger than $B_{max} - T_{min}$, with B_{max} defined in Eq. (3.14), and T_{min} defined as the essential lower bound of the inter-arrival time distribution,

$$T_{min} = \text{essinf}\{T_k\} \equiv \inf(x : \Pr[T_k < x] > 0) \ , \ x \in \mathbb{R}^+ \ . \tag{3.25}$$

This result is generally valid, and is independent of the time setting, the inter-arrival-time distribution, and the specifics of the burst size distribution. The result even holds if batch arrivals are considered (as in Sect. 2.3), with the understanding that then, $T_{min} = 0$. The only assumption is thus the upper-boundedness of the burst size distribution, for a single-wavelength buffer with FIFO scheduling.

Note that $D_0 \leq B_{max} - T_{min}$ is even valid for non-iid burst sizes or service times. It can easily be shown that it is also valid for an infinite-sized buffer setting. It can be stated somewhat more general as $D_0 \leq U_{max}$, where U_{max} denotes the maximum of the support of the rv U_k that is the difference of B_k and T_k, that may be correlated to each other.

Returning now to $D_0 \leq B_{max} - T_{min}$, this can be understood through intuitive reasoning, by making use of the scheduling horizon, a key rv in the previous chapter, that is only sporadically mentioned in the current chapter. We consider two mutually exclusive cases.

1. Imagine that an arbitrary burst k was assigned waiting time $W_k = iD$, with $i \leq N - 1$. This corresponds to the situation where burst k used any but the longest delay line, and there is always a delay line left that is longer than iD, and available for contention resolution. Burst $k+1$ arrives T_k *time units* later (*seconds* for CT, *time slots* for DT), and has to wait for at least $H_{k+1} = [iD + B_k - T_k]^+$ time units to avoid contention, with H_{k+1} is the scheduling horizon of burst $k + 1$, just like in the previous chapter. In a worst case scenario, this amount of time is maximized, when B_k equals B_{max}, and T_k equals T_{min}. In this regard, even in a worst case scenario (with clearly $H_{k+1} > 0$), burst $k + 1$ only needs to be delayed for $H_{k+1} = iD + B_{max} - T_{min}$. As such, the delay line with index $i + 1$ can always provide burst $k + 1$ with sufficient delay, as long as $B_{max} - T_{min} \leq D$,

since then, $H_{k+1} \leq (i+1)D$, showing that $(i+1)D$ is sufficiently long. The latter being the maximally required line length for line $i+1$, there is no reason to make the line any longer than $B_{max} - T_{min}$, since this would only result in extra delay for burst $k+1$, increase the average waiting time, and so even increase the average loss probability.

2. Imagine that an arbitrary burst k was assigned waiting time $W_k = ND$. Now, burst k is delayed with the longest delay line. Whether or not burst $k+1$ can be accepted depends entirely on B_k and T_k: if $B_k - T_k \leq 0$, burst $k+1$ is accepted; if $B_k - T_k > 0$, burst $k+1$ is lost. As such, in this case, D is not to be optimized, since it plays no part in the acceptance of burst $k+1$.

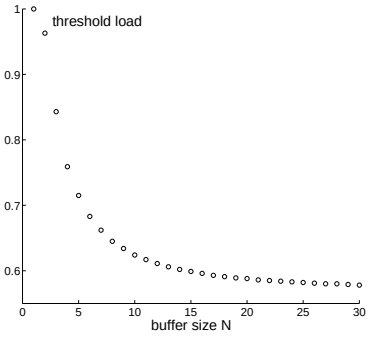
Concluding, we see how upper-bounded burst sizes imply that $D_0 \leq B_{max} - T_{min}$. In other words, increasing the granularity beyond $B_{max} - T_{min}$ is useless[1], since it does not yield better contention resolution, and results in performance degradation. The bound $B_{max} - T_{min}$ is reflected in for example the optima for low load and deterministic burst size distribution: $B_{max} - 1$ for DT, B_{max} for CT (as mentioned frequently in the previous chapter, and also in the current, see for example Sect. 3.1.7).

3.3.2 Rather General Assumptions

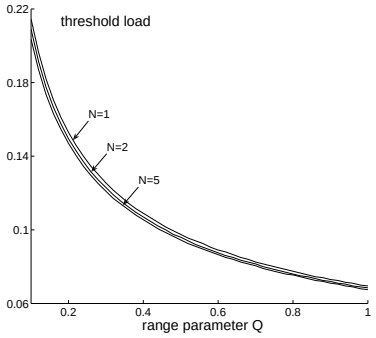
In the following, we will add three assumptions to the ones made in Sect. 3.1. We first treat the DT case, for which we adopt the following: (i) we assume that the burst sizes B_k are upper-bounded by some B_{max}, that is, $B_k \leq B_{max}$; (ii) the buffer is degenerate, with FDL set $\mathcal{A} = \{0, D, 2D, \ldots, ND\}$, with granularity D; (iii) we assume that the granularity (nearly) matches the maximum burst size B_{max}, that is, $D = B_{max} - 1$. As known from Sect. 3.3.1, the value $B_{max} - 1$ also constitutes the largest granularity value that still can be called useful. The CT case also assumes (i) and (ii) but not (iii) since for CT, the offset of one vanishes, and D is chosen equal to B_{max}. This is explained further in Sect. 3.3.3.3. An explicit formula for a more general class of FDL buffers is presented in [9], where we prove that the minimum loss rate within this class is realized using the (three) above-mentioned assumptions.

While these assumptions are helpful from a mathematical point of view, the main motivation to adopt these three assumptions comes from the application. Especially in the case of fixed burst sizes (often considered in OBS and OPS), this

[1] Strictly speaking, $B_{max} - T_{min}$ can only be called "the largest useful granularity value" in a performance modeling context. In an actual implementation, the value still plays a similar role, but the granularity value can be chosen somewhat larger for other reasons, for example to compensate for finite switching and processing times.



(a) Deterministic burst size distribution.



(b) Uniform burst size distribution, with increasing range.

Figure 3.6: Characterization of the threshold load of a CT degenerate M/D/1 buffer for increasing size N (a), and of a CT degenerate M/U/1 buffer for increasing burst size range Q (b). Below the threshold load, the degenerate buffer setting, with $D = B_{max}$, is optimal. Note that results are independent of the numerical value of B_{max} (expressed in μs).

setting is a very plausible one. The first assumption trivially arises from the application. As for the second and third, it is widely accepted that, in the case of fixed-sized bursts, it is natural to choose D equal to the size of the bursts , as reported for CT in for example [24]. In the DT case, more recent work confirms this [10], showing that D is best chosen equal to $B_{max} - 1$ (offset of one when compared with the CT case), given that this is optimal in terms of loss as long as the load is smaller than some threshold load of about 60%. (When the load exceeds this threshold, the optimum for the granularity jumps to lower granularity values.) Also in DT, this is further confirmed for a more general arrival process [6]. In the case of non-degenerate buffers, [17] illustrates that even when non-degenerate buffers are considered, the degenerate case comes out as the optimal one for low loads, again for loads up to a threshold of about 60%. To validate this qualitative result for a CT setting, we traced this threshold load in an exact manner, by means of the method of Sect. 3.1.5, for degenerate fiber lengths. Note that we prefer to consider the CT setting rather than the DT setting, since the CT is somewhat simpler (no influence of the slot length comes about), but is equally instructive. Further numerical experimenting for the DT case showed that increasing the slot lengths results in a minor increase of the threshold load.

The left pane of Fig. 3.6, valid for any value of B_{max}, shows that this threshold load depends substantially on the buffer size if this buffer size is small, but flattens out for larger buffer size. This shows that, even for large buffer sizes, the current setting is the optimal one for a load smaller than about 58%. While this is valid

for CT, numerical verification confirmed that this also goes for DT. As such, it is a prime candidate for implementation in an actual OBS/OPS switch.

The current setting with $D = B_{max}$ can also be the optimal one in case of variable burst sizes, albeit only for low traffic load. While the exact value of such threshold load depends considerably on the specific burst size distribution, it is instructive to verify the basic case of a uniform burst size distribution. Without loss of generality (at least for CT), we normalize the mean burst size to $\mathrm{E}[B] = 1$, and obtain a tuneable range $[1 - Q, 1 + Q]$ by varying the range parameter Q between 0 and 1. This is applied on the right pane of Fig. 3.6 to study the threshold load, for varying Q, and three different buffer sizes, $N \in \{1, 2, 5\}$. Firstly, the figure shows that the impact of the buffer size is only minor, since the three lines nearly coincide. Secondly, the value of the threshold load is already low for only small variations of the burst size: the threshold load is about 21% for range $[0.9, 1.1]$ ($Q = 0.1$), which is much lower than the 60% mentioned for fixed burst size. Thirdly, note that, even for the widest possible range, $[0, 2]$ ($Q = 1$), the threshold load remains larger than zero.

Summarizing, the current setting is the optimal one for any range of a uniform burst size distribution, but only for (really) low loads. Note that the performance model of Sect. 3.1.5 allows to determine the optimal buffer setting for any given burst size distribution. However, numerical trials for load values above the threshold load pointed out that the optimum is exceedingly dependent on the traffic load and burst size distribution range, making it impossible to identify a "best design choice" for the granularity for general traffic load. As such, the setting assumed here is also an interesting point of reference in case burst lengths vary.

3.3.3 Analysis

3.3.3.1 Markov Chain of Waiting Times

For the traffic setting, the assumption that $D = B_{max} - 1$ allows to specify the $F_U(n)$ further, as

$$F_U(n) \;=\; \begin{cases} \bar{p}^{-n} \cdot \sum_{i=1}^{B_{max}} b(i) \cdot \bar{p}^{i-1} & , n \in \mathbb{Z}^- \,, \\ \bar{p}^{-n} \sum_{i=n+1}^{B_{max}} b(i) \cdot \bar{p}^{i-n-1} + F_B(n) & , n \in \{1, 2, \ldots, B_{max} - 2\} \,, \\ 1 & n \in \{B_{max} - 1, B_{max}, \ldots\} \,. \end{cases}$$

For conciseness' sake, we introduce additional notation, that directly relates to the parameters introduced in Sect. 3.1:

$$\bar{P} = \bar{p}^D \;,\; P = 1 - \bar{P} \;\;;\;\; \bar{g} = F_U(0) = \sum_{n=1}^{B_{max}} b(n) \bar{p}^{n-1} \;,\; g = 1 - \bar{g} \,. \quad (3.26)$$

All four parameters have range in $[0, 1]$ and account for probabilities. Note that $F_U(0)$ can be obtained as a function of the pgf of the burst size distribution, $B(z)$,

evaluated in $z = \bar{p}$: $F_U(0) = B(\bar{p})/\bar{p}$. With $w(n)$ the steady-state probability that an accepted burst is delayed with a_n (like above), it is required that $\sum_{i=0}^{N} w(i) \cdot m_{in} = w(n)$, for $n \in \{0, 1, \ldots, N\}$. Since the evolution of the waiting time is exactly as in Sect. 3.1.5, we utilize (3.10), to obtain after simplification that

$$\mathbf{M} = \begin{bmatrix} \bar{g} & g & 0 & \ldots & & 0 \\ \bar{g}\bar{P} & \bar{g}P & g & & & \\ \bar{g}\bar{P}^2 & \bar{g}P\bar{P} & gP & & & \vdots \\ \vdots & \vdots & & \ddots & 0 & 0 \\ & & & & g & 0 \\ \bar{g}\bar{P}^{N-1} & \bar{g}P\bar{P}^{N-2} & \ldots & & \bar{g}P & g \\ \bar{P}^N & P\bar{P}^{N-1} & \ldots & & P\bar{P} & P \end{bmatrix}, \tag{3.27}$$

where $m_{ij} = 0$ if $j \geq i + 2$, for $0 \leq i \leq N - 2$. Already of very simple form, we remark that $\mathbf{M}$ further simplifies if burst sizes are fixed to $B = B_{max}$, since then $\bar{g} = \bar{P}$, $g = P$, and the last two rows coincide.

3.3.3.2 Closed-Form Solution

The symmetry of $\mathbf{M}$ (3.27) confirms that the Markov chain formulation indeed is fit for the specific problem, since it allows us to obtain a closed-form solution for the steady-state waiting time probabilities and the LP.

Like in Sect. 3.1.5.2, the Markov chain of waiting times, now defined by $\mathbf{M}$ (3.27), is aperiodic, irreducible and has positive-recurrent states, and thus allows to obtain a unique equilibrium distribution for $k \to \infty$, with all involved rv's associated with the steady-state distribution, and $\{W, B, T\}$ denoting generic rv's following the steady-state distributions of $\{W_k, B_k, T_k, Y_k\}$, respectively.

For this steady state, we evaluate $\sum_{i=0}^{N} w(i) \cdot m_{in} = w(n)$, which provides us with $N + 1$ conditions for the $w(n)$ ($1 \leq n \leq N - 1$):

$$gw(n-1) + \bar{g}P \sum_{i=0}^{N-1-n} \bar{P}^i w(n+i) + P\bar{P}^{N-n} w(N) = w(n) \;\; , \;\; 1 \leq n \leq N-1 \, .$$

Adding the evaluation of the matrix expression for $w(N)$, $w(N) = w(N-1) \cdot g/Q$, and the normalization condition, $\sum_{n=0}^{N} w(n) = 1$, provides us with $N + 1$ conditions for the $w(n)$, sufficient for a unique solution. Some calculation shows that the waiting time probabilities $w(n)$ for accepted bursts have a truncated (shifted) geometric distribution,

$$w(n) = \zeta^n \cdot \frac{1 - \zeta}{1 - \zeta^{N+1}} \;, \; 0 \leq n \leq N \, , \tag{3.28}$$

with $\zeta = g/\bar{P}$. The mean waiting time $\mathrm{E}[W_k]$ of an arbitrary accepted burst k is directly derived from this, as

$$\mathrm{E}[W_k] = D \cdot \left(\frac{\zeta}{1-\zeta} - \frac{(N+1)\zeta^{N+1}}{1-\zeta^{N+1}} \right) . \tag{3.29}$$

This formula is valid for $\zeta \neq 1$; in the case that $\zeta = 1$, the $w(n)$ are distributed uniformly, $w(n) = 1/(N+1)$, $0 \leq n \leq N$, with mean waiting time $\mathrm{E}[W_k] = ND/2$.

To obtain the LP, like in Sect. 3.1.5.3, we consider again $\mathrm{E}[Z_k]$, the average number of lost bursts during the unavailable period following burst k, that is captured by the expression $\mathrm{E}[Z_k] = p(\mathrm{E}[B_k] - 1)w(N)$. The same reasoning as in Sect. 3.1.5.3, now assuming steady state, leads to

$$\mathrm{LP} = \frac{p(\mathrm{E}[B]-1)\zeta^N(1-\zeta)}{p(\mathrm{E}[B]-1)\zeta^N(1-\zeta)+1-\zeta^{N+1}} . \tag{3.30}$$

As such, the closed-form solution comprises (3.28) and (3.30), and is of particularly simple form. For deterministic burst size distribution, $\mathrm{E}[B] = B_{max}$, $\mathrm{E}[B] - 1 = B_{max} - 1 = D$, and $\zeta = P/\bar{P}$.

3.3.3.3 Closed-Form Solution for CT

To account for CT, a slight change in the third assumption of Sect. 3.3.2 comes about: it suffices to assume that the granularity equals B_{max}, that is, $D = B_{max}$, and this (as in Sect. 3.1.6) due to the omission of the relative offset in the minimum of the support of T_k in CT ($T_{min} = 0$) when compared to DT ($T_{min} = 1$). We adopt the notation of Sect. 3.1.4, and add the following four parameters (complementary to (3.26)),

$$\bar{P} = e^{-\lambda D} \ , \ P = 1 - \bar{P} \ ; \ \bar{g} = F_U(0) = \int_0^{B_{max}} e^{-\lambda u} dF_B(u) \ , \ g = 1 - \bar{g} \ . \tag{3.31}$$

Substituting these parameters, the matrix $\mathbf{M}$ for CT is equally valid for DT, and yields the same formulas for the steady-state waiting times probabilities $w(n)$ and the mean waiting time, (3.11) and (3.29), respectively. In the case that $\zeta < 1$, one can compare (3.29) to the expression of the mean waiting time for infinite buffer size. As was derived in the previous chapter (Sect. 2.2.8.2, Eq. (2.37)), then, the waiting times have a (shifted) geometric distribution with parameter $\bar{\zeta}$, from which it follows that $\mathrm{E}[W] = D \cdot \zeta/\bar{\zeta}$ for $N \to \infty$. For the finite system considered here, expression (3.29) for the mean waiting time remains valid also for $\zeta > 1$, while for $\zeta = 1$, the $w(n)$ are distributed uniformly, $w(n) = 1/(N+1)$, $0 \leq n \leq N$, with mean waiting time $\mathrm{E}[W] = ND/2$.

As for the LP, the CT formula somewhat differs from (3.30), again due to the aforementioned relative offset. Assuming steady state, the mean number of lost bursts during the unavailable period following an arbitrary burst is now $\mathrm{E}[Z] = \lambda \mathrm{E}[B] w(N)$, which results in the following expression for the LP,

$$\mathrm{LP} = \frac{\rho \zeta^N (1-\zeta)}{\rho \zeta^N (1-\zeta) + 1 - \zeta^{N+1}} \,, \tag{3.32}$$

with $\rho = \lambda \mathrm{E}[B]$ the traffic load.

3.3.3.4 Comparison to Classic System

If burst sizes are fixed to B_{max}, (3.32) can be written solely in terms of only the traffic load ρ, since then $\zeta = P/\bar{P} = e^{\rho} - 1$. When considering the formula it comes out that this is even somewhat simpler than the solution for a classic M/D/1 buffer of size $N+1$ in CT [99], for which the loss probability LP_c (c for classic) is expressed by

$$\mathrm{LP}_c = \frac{1 + (\rho - 1) F_N}{2 + (2\rho - 1) F_N} \,, \tag{3.33}$$

with $F_i = \sum_{k=0}^{i} \frac{(-1)^k}{k!} (i-k)^k e^{(i-k)\rho} \rho^k$. The same goes for the expressions for the mean waiting time, since the mean waiting $\mathrm{E}[W^c]$ for a classic M/D/1 buffer of size $N+1$ [99], with burst sizes fixed to D, is given by

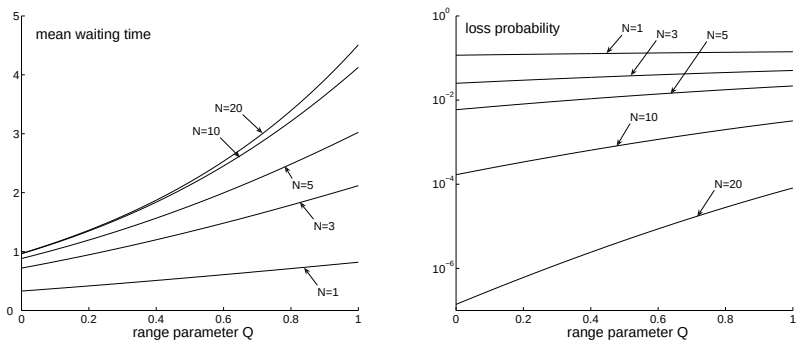
$$\mathrm{E}[W^c] = D \cdot \left(N - \frac{\sum_{i=0}^{N} F_i - N - 1}{\rho F_N} \right) \,, \tag{3.34}$$

with the same F_i as above, which is clearly somewhat more involving than (3.29). A comparison of the performance of both systems is given in Sect. 3.3.4.

3.3.4 Numerical Examples

While the obtained results allow to trace both the DT and CT setting for a wide variety of settings, we will focus here on just two examples of the CT setting, and also provide an instructive comparison of an optical and a classic buffer system.

Uniform Burst Size Distribution In a first example, displayed in Fig. 3.7, we consider burst sizes with a uniform distribution. The traffic load is fixed to 40%; the burst size distribution has the same (normalized) characteristics as the one considered in Sect. 3.3.2: $\mathrm{E}[B] = 1$ μs, with a range $[1-Q, 1+Q]$ that can be tuned by varying the range parameter Q in $[0, 1]$. The left pane of Fig. 3.7 shows the average waiting time of accepted bursts, for five different buffer sizes, $N \in \{1, 3, 5, 10, 20\}$, ($N$ also being the number of lines with non-zero length) as a function of the range parameter Q; the values are obtained from (3.29). Firstly, it



(a) How Q impacts E[W]. (b) How Q impacts the LP.

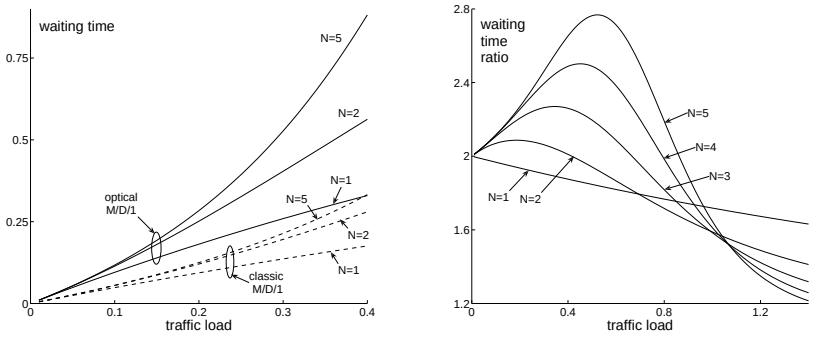
Figure 3.7: Impact of an increase of the range of the uniform burst size distribution on the mean waiting time and LP of a degenerate M/D/1 buffer in CT. The (normalized) range $[1-Q, 1+Q]$ (in μs) of the burst size distribution increases along with Q. These figures were obtained for $\rho = 40\%$.

comes as no surprise that the average waiting time increases when the buffer size augments. This can be immediately learned from (3.29), and it is also the case for classic buffers. Note that the average waiting time for an infinite-sized buffer system is also on display, since for the parameter settings considered, the difference with the curve for $N = 20$ is negligible. Most relevant now is the observation that performance degrades as the burst size range increases, a performance loss that is inherent to the FDL buffer system.

The right pane of Fig. 3.7 displays the LP for the same setting, and is obtained from (3.30). Although increasing the buffer size indeed lowers the loss for any burst size range, it is clear that loss is mitigated much more effectively when the burst size range is limited. This again confirms that performance worsens as the burst size range increases. Relying also on further numerical trials, we conclude that an FDL buffer in general yields better performance when burst sizes are fixed than in the case where the latter vary.

Deterministic Burst Size Distribution In a second numerical example, we focus on a case with better performance, with burst sizes fixed. More particularly, we compare its performance with that of a classic M/D/1 buffer of size $N+1$ (N places available for waiting, 1 for serving), with burst sizes fixed to $B_{max} = D$.

Note that this is indeed a fair comparison, which is not necessarily so in the case of general burst sizes. More precisely, the FDL buffer suffers loss from so-called balking: bursts are lost, whenever the requested waiting time exceeds the buffer capacity ND. In a classic buffer of size $N+1$, loss occurs when all $N+1$



(a) Absolute values of $\mathrm{E}[W]$.

(b) Ratio of $\mathrm{E}[W]$ and $\mathrm{E}[W^c]$.

Figure 3.8: Mean waiting time of the classic and the degenerate M/D/1 buffer of finite size in CT, set out for increasing traffic load and various buffer sizes. These figures were obtained for $\mathrm{E}[B] = 1$ *μs.*

places (N for waiting, 1 for serving) are occupied. As such, the loss process is determined by either the waiting time characteristics, or the number of bursts in the queue, respectively. However, due to the fact that burst size is fixed, limiting the waiting time to ND yields the same loss condition as limiting the number of places available for waiting to N (classic buffer case).

The left pane of Fig. 3.8 displays the mean waiting time, for three different buffer sizes, $N \in \{1, 2, 5\}$, with $\mathrm{E}[B] = B_{max} = 1$, and varying traffic load $\rho = \lambda B_{max}$. The continuous curves are valid for an FDL buffer, and are obtained from (3.29) (with $\zeta = e^{\rho} - 1$); the dotted curves account for the classic buffer case, and are calculated using (3.34). As can be seen, the performance gap between the classic and the FDL case is considerable, with the discrepancy growing for increasing traffic load. On the other hand, for low load, the curves for the classic buffer converge to one waiting time value. The same can be said about the curves for the FDL buffer, that also converge to one waiting time value when the load approaches zero, as should. Note, however, that even for small loads, the curves of the FDL buffer always display higher waiting time values than those of the classic buffer.

Inspecting this, the right pane of Fig. 3.8 shows the mean waiting time ratio of the FDL case and the classic case, $\mathrm{E}[W]/\mathrm{E}[W^c]$, for five different buffer sizes, $N \in \{1, 2, 3, 4, 5\}$. Most interestingly, it comes out that, in the limit of the load approaching zero, this ratio is exactly 2, and this independent of the buffer size. Although not self-evident, this observation comes with an intuitive explanation. More precisely, the case of very low load implies that the buffer is almost always empty. If a burst has to wait (and thus is buffered), it will nearly always be because

exactly one burst (and not more) is receiving service. On the one hand, in a classic system with a Poisson arrival process, the waiting time of such burst (also, the residual service time of the previous burst) is half of the burst size on average, $D/2$. On the other hand, in an FDL buffer system with a Poisson arrival process, the residual service time of the previous burst is also $D/2$ if the load is near to zero. However, given the FDL buffer's functioning, its waiting time has to be a multiple of D, and such burst always gets assigned D in the FDL buffer system. Concluding, FDL M/D/1 buffers see waiting times doubled when compared to their classic counterpart, at least for low load.

For larger loads, the right pane of Fig. 3.8 shows that the performance gap is largest for a traffic load between 40% and 80%. This gap then decreases for augmenting load, and becomes minimal when the system is in overload ($> 100\%$). The latter being a less interesting regime in practice, we conclude that, for loads $< 80\%$, the waiting times in M/D/1 FDL buffers are usually more than doubled (except for $N = 1$), when compared to those of the classic M/D/1 buffer.

3.4 Solution for Memoryless Burst Size Distribution

With the previous section devoted to the case of deterministic burst size distribution, the current one assumes a memoryless burst size distribution, but maintains the assumption of memoryless arrivals, and all the other system assumptions of the general model discussed in Sect. 3.1. The aim is again to derive a closed-form solution for the waiting time distribution and loss probability, in the case of a degenerate buffer setting with granularity D. The main difference with the approach of the previous section is that we do not base the analysis directly on the method of Sect. 3.1.5, but consider a Markov chain with an additional state to account for loss. Since the latter allows to exploit the memorylessness of the burst size distribution, we obtain a simpler transition matrix, which eases the extraction of closed-form expressions from there. Note that the "all-round" approach of Sect. 3.1.5 also leads to the correct expressions in principle, but poses more mathematical difficulty than the light-weight approach presented here, yielding closed-form expressions for any value of the granularity. As in the previous section, we first consider the DT case, then the CT case.

3.4.1 Traffic Setting

In this section, we follow an approach that is different from that of Sect. 3.1, by introducing an additional Markov state to account for loss (see further). This also corresponds to a different numbering for the arriving bursts: rather than numbering only accepted bursts (like in Sect. 3.1), we number the bursts in the order of their arrival, regardless of their acceptance, like in the previous chapter. As such, index

k indicates that an arbitrary burst k either experiences a waiting time W_k in the buffer, $W_k \in \{0, D, 2D, \ldots, ND\}$, or is discarded, without assigning a waiting time, if the requested delay exceeds the maximum delay. The latter we denote with D_M in the following, instead of using a_N, to stress that D_M is indeed a multiple of D.

As for the burst sizes and inter-arrival times, both constitute a series of iid rv's with memoryless distribution, which is a geometric distribution in the current case of DT. With burst k, we associate a burst size B_k with mean value $1/f$, whereas the time between the kth burst and the $(k+1)$th is captured by the inter-arrival time T_k, with mean $1/p$, corresponding to the cdf's

$$F_T(n) = \Pr[T_k \leq n] = 1 - \bar{p}^n \ , \ F_B(n) = 1 - \bar{f}^n \ , \ n \in \mathbb{N} \, ,$$

with $\bar{f} = 1 - f$, $\bar{p} = 1 - p$. Again considering $U_k = B_k - T_k$, this traffic setting results in the following expression for the cdf of U_k,

$$F_U(n) \quad = \quad \begin{cases} \bar{g} \cdot \bar{p}^{-n} & , n \in \mathbb{Z}^- \, , \\ \bar{g} + g \cdot (1 - \bar{f}^n) & , n \in \mathbb{Z}_0^+ \, , \end{cases}$$

with

$$\bar{g} = F_U(0) = \frac{f}{1 - \bar{f} \cdot \bar{p}} \, ,$$

where $\bar{g} = 1 - g$.

3.4.2 Analysis

3.4.2.1 Evolution of the Waiting Time

Numbering each arriving burst, the system's evolution can be captured completely in terms of the rv's U_k and T_k in a way that is similar but not identical to Sect. 3.1.5.1, due to the different numbering convention. By considering the acceptance or loss of burst $k+1$, conditioned on the acceptance or loss of burst k, we obtain four mutually exclusive events, with the corresponding waiting time transition.

1. Burst $k+1$ accepted, given that burst k was accepted. The buffer is never full, and both burst k and $k+1$ can be accepted. This occurs if

$$W_k + D \cdot \left\lceil \frac{U_k}{D} \right\rceil \leq D_M \, ,$$

and leads to

$$W_{k+1} = \left[W_k + D \cdot \left\lceil \frac{U_k}{D} \right\rceil \right]^+ \, .$$

2. Burst $k+1$ lost, given that burst k was accepted. Accepting burst k filled up the buffer, so that burst $k+1$ is blocked upon arrival. This occurs if

$$W_k + D \cdot \left\lceil \frac{U_k}{D} \right\rceil > D_M \,,$$

and, given that burst $k+1$ is lost, no waiting time is to be associated with it.

3. Burst $k+1$ lost, given that burst k was lost. Both burst k and $k+1$ cannot be provided with sufficient delay, and are blocked for this reason. Thanks to the memoryless nature of the burst size distribution, it is of no importance when the blocking started, and the residual time needed for the latest-accepted burst to enter the buffer, B, has the same distribution as B_k. Loss occurs whenever the inter-arrival time is smaller than B:

$$T_k < B \,,$$

and again, since burst $k+1$ is lost, no waiting time is assigned to it.

4. Burst $k+1$ accepted, given that burst k was lost. Now, the system blocked burst k due to fullness, but the time until arrival of burst $k+1$ is sufficiently large when compared to B, that is,

$$T_k \geq B \,.$$

To obtain the new waiting time, that is necessarily smaller or equal than D_M, we utilize the reactivation time R_k, that plays exactly the same role as in Sect. 3.1.5.1, Eq. (3.7) and Sect. 3.2.3.1, Eq. (3.17), to obtain obtain

$$W_{k+1} = \left[D \cdot \left\lceil \frac{D_M - R_k}{D} \right\rceil \right]^+ \,. \tag{3.35}$$

Just like in Sect. 3.1.5.1, the R_k and T_k have nearly identical distribution due to the memoryless nature of the arrival process, and the cdf of R_k reads

$$F_R(n) = \Pr[R_k \leq n] = 1 - \bar{p}^{n+1} \quad , \quad n \in \mathbb{N} \,,$$

which coincides with (3.8). The only difference between $F_T(n)$ and $F_R(n)$ is again the offset of 1, that reflects the minimum of the support of the T_k, namely 1.

3.4.2.2 Markov Chain of Waiting Times

Now we associate a Markov state i to the case that $W_k = i \times D$, $i = 0, 1 \ldots N$, and associate the case of loss of burst k with state $i = N+1$ for notational convenience. Then, upon each arrival, the Markov state i is updated to a new state j, according to one of the four events mentioned above. Denoting the probability

of such transition by m_{ij}, $0 \leq i, j \leq N+1$, calculations show that the transition matrix $\mathbf{M}$ is as follows:

$$\mathbf{M} = \begin{bmatrix} \bar{g} & gF & gFG & \dots & gFG^{N-1} & gG^N \\ \bar{g}Q & \bar{g}P & gF & \dots & gFG^{N-2} & gG^{N-1} \\ \bar{g}Q^2 & \bar{g}PQ & \bar{g}P & \dots & gFG^{N-3} & gG^{N-2} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \bar{g}Q^{N-1} & \bar{g}PQ^{N-2} & & \dots & gF & gG \\ \bar{g}Q^N & \bar{g}PQ^{N-1} & & \dots & \bar{g}P & g \\ \bar{g}Q^N & \bar{g}PQ^{N-1} & & \dots & \bar{g}P & g \end{bmatrix} \tag{3.36}$$

with

$$\begin{aligned} \bar{P} &= \Pr[T_k > D] = \bar{p}^D \,,\; P = 1 - \bar{P} \,, \\ \bar{F} &= \Pr[B_k > D] = \bar{f}^D \,,\; F = 1 - \bar{F} \,. \end{aligned}$$

3.4.2.3 Closed-Form Solution

Just like in the previous sections, we assume steady state, possible since the Markov chain defined by (3.36) is aperiodic, irreducible and has positive-recurrent states. While it is possible to obtain the steady-state probabilities by numerical means like above, in this special case, we are able to tackle the problem analytically. The steady-state probabilities $\pi(i)$ of the Markov states i are to fulfill the relation

$$\pi(i) = \sum_{j=0}^{N+1} \pi(j) \cdot m_{ji} \;\;;\;\; 0 \leq i \leq N+1 \,. \tag{3.37}$$

Now, exploiting the symmetry in (3.37) when evaluated for $i = 1, 2 \dots N$, some calculation allows to establish a geometric relation between all but the first ($\pi(0)$) and last ($\pi(N+1)$) entry of π:

$$\pi(i) = \pi(1) \cdot \zeta^{i-1} \;\;;\;\; 1 \leq i \leq N \,, \tag{3.38}$$

with

$$\zeta = \bar{g}\bar{F} + g/\bar{P} = (f \cdot \bar{f}^D + \bar{f} \cdot p \cdot \bar{p}^{-D})/(1 - \bar{f} \cdot \bar{p}) \,. \tag{3.39}$$

Given (3.38), we have $N-1$ linear equations (excluding the trivial instance $i = 1$) at disposition to determine the $N+2$ entries of π. As such, we need only three additional conditions to determine all $\pi(i)$, which we obtain by evaluating (3.37) in $i = 0$ and $i = N+1$, and invoking the normalization condition, $\sum_{i=0}^{N+1} \pi(i) = 1$. This allows to write down the final expressions for the $\pi(i)$,

$$\begin{aligned} \pi(0) &= K \cdot \bar{P} \cdot \left(\bar{g}/g \cdot \chi_P + (\bar{F}\bar{P})^{N-1} \cdot \chi_F\right) \,, \\ \pi(N+1) &= K \cdot \bar{F}^{N-1} \cdot \left(\bar{F} \cdot \bar{P} \cdot \chi_P + g/\bar{g} \cdot \chi_F\right) \,, \\ \pi(1) &= K \cdot \left(1 - (\bar{F}\bar{P})^N\right) \,, \end{aligned} \tag{3.40}$$

where we introduced

$$\begin{aligned}
\chi_1 &= (1-\zeta^N)/(1-\zeta)\,, \\
\chi_P &= (1-(\zeta\bar{P})^N)/(1-\zeta\bar{P})\,, \\
\chi_F &= (1-(\zeta/\bar{F})^N)/(1-\zeta/\bar{F})\,, \\
K^{-1} &= (1-(\bar{F}\bar{P})^N)\cdot\chi_1 + \bar{P}(\bar{F}^N+\bar{g}/g)\cdot\chi_P \\
&\quad +\bar{F}^{N-1}(\bar{P}^N+g/\bar{g})\cdot\chi_F\,.
\end{aligned}$$

Note that other $\pi(i)$, $2 \leq i \leq N$, are readily obtained from (3.38). Now, the probability that an arriving burst is lost simply equals $\pi(N+1)$. The steady-state probabilities of the waiting time of accepted bursts, $w(i)$, $i = 0 \ldots N$, can be obtained as

$$w(i) = \Pr[W_k = iD \,|\, \text{burst } k \text{ accepted}] = \frac{\pi(i)}{1-\pi(N+1)}\,,\; i \in \{0,1,\ldots,N\}\,.$$

Together with (3.38), this yields that

$$w(i) = w(1)\cdot\zeta^{i-1}\,,\; i \in \{1,2,\ldots,N\}\,, \tag{3.41}$$

and the coefficient ζ plays a similar role as in Sect. 3.3.3.2, but with the minor difference that the waiting times do not follow a simple truncated geometric distribution like in Sect. 3.3.3.2, since $w(0)$ does not obey the geometric relation (3.41). The mean waiting time of accepted bursts however easily follows, as

$$\mathrm{E}[W] = D\frac{\sum_{i=1}^{N} i\pi(i)}{1-\pi(N+1)} = D\pi(1)\frac{N\zeta^{N+1}-(N+1)\zeta^N+1}{(1-\pi(N+1))(1-\zeta)^2}\,. \tag{3.42}$$

It is interesting to place this result in juxtaposition with the expressions for infinite-sized buffers. While the current work only treats the CT case for that (in Sect. 2.2.8.1), [13] accounts for the DT case, reporting amongst others the pgf of the waiting times. Inversion of the latter yields that, for the infinite-sized case, the waiting time probabilities adhere to

$$\begin{aligned}
w(0) &= \bar{\zeta}/F\,, \\
w(i) &= (1-w(0))\cdot\bar{\zeta}\cdot\zeta^{i-1}\,,\; i \in \mathbb{N}_0\,,
\end{aligned} \tag{3.43}$$

and thus prove similar to the current result, in that the decay rate for increasing i is ζ in both cases. The mean waiting time for the DT case and infinite buffer size is also mentioned in [13] (with notation γ^D for ζ, $\bar{\mu}^D$ for $\bar{F}$), and can be simplified to

$$\mathrm{E}[W] = D\cdot\{\frac{1}{\bar{\zeta}} - \frac{1}{F}\}\,, \tag{3.44}$$

that is in general larger than the mean waiting time for finite buffer size (3.42), and provides an upper bound, that yields very good approximations for large N.

Finally, since the first purpose of the performance model is to determine the optimal granularity D_0, one might attempt to derive an analytic expression for the latter. Such approach is possible in principle, as the expression for the LP, $\pi(N+1)$, can be written as an explicit function of D, $\pi(N+1) = \mathcal{F}(D)$. Further, as discussed earlier in Sect. 2.1.6.2, $\mathcal{F}(D)$ takes on the loss value of a bufferless system, both for $D \to 0$ and $D \to \infty$. As such, (2.16) (obtained for a memoryless arrival process) applies also here,

$$\lim_{D\to 0} \mathcal{F}(D) = \lim_{D\to\infty} \mathcal{F}(D) = \rho/(1+\rho)\,.$$

Between these extremes, the curve displays only one extremum, namely the optimum D_0, that can be found by solving the integer optimization problem by numerical means.

3.4.2.4 Closed-Form Solution for CT

Just like in the previous sections, the analysis for CT follows almost directly from the analysis in DT, and requires but little modification. Now, the iid inter-arrival times and burst sizes both have negative-exponential distribution, with cdf

$$F_T(x) = \Pr[T_k \le x] = 1 - e^{-\lambda x}\ ,\ F_B(x) = 1 - e^{-\mu x}\ , x \in \mathbb{R}^+\ ,$$

with $\mathrm{E}[T_k] = 1/\lambda$, and $\mathrm{E}[B_k] = 1/\mu$. This corresponds to the following expression for the cdf of U_k ,

$$F_U(x) \quad = \quad \begin{cases} \bar{g} \cdot e^{\lambda x} & , x \in \mathbb{R}^-\ , \\ \bar{g} + g \cdot (1 - e^{-\mu x}) & , x \in \mathbb{R}_0^+\ , \end{cases}$$

with

$$\bar{g} = F_U(0) = \frac{\mu}{\lambda + \mu}\,.$$

Other aspects (degenerate buffer with waiting times in $\{0, D, 2D, \ldots, ND\}$, with $D_M = ND$) remain the same, except that the granularity now takes on any positive real value, $D \in \mathbb{R}_0^+$.

As for the waiting time transitions, the four events mentioned go unaltered, but now with

$$F_R(x) = \Pr[R_k \le x] = 1 - e^{-\lambda x}\ , x \in \mathbb{R}^+\ ,$$

that coincides with $F_T(x)$. The matrix **M** (3.36) for DT is valid equally for CT if one assumes

$$\bar{P} = e^{-\lambda D}\ ,\ P = 1 - \bar{P}\ \ ;\ \ \bar{F} = e^{-\mu D}\ ,\ F = 1 - \bar{F}\,. \tag{3.45}$$

(Due to the identical arrival process, this substitution is similar to the one for upper-bounded burst sizes, see (3.31).) Since the transition matrix is formally

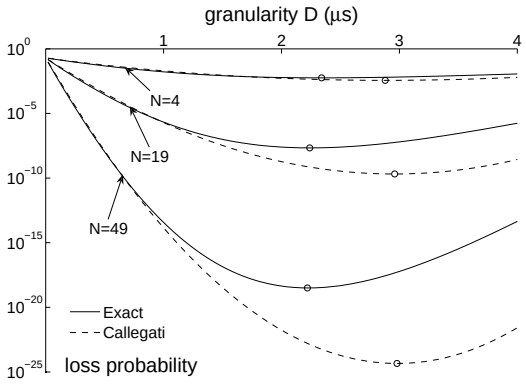


Figure 3.9: Loss probability as a function of the granularity D, for different values of the buffer size N, for fixed traffic load $\rho = 0.25$, and $\mathrm{E}[B] = 1/\mu = 1$ μs (although results are independent of time scale). Results from Callegati's model are compared to the exact results presented in this section. In both cases, the resulting optimal granularity value D_0 is indicated with "o".

identical, the expressions for steady state go unaltered, and the loss probability $\pi(N+1)$ (3.40), waiting time probabilities $w(i)$ (3.41) and $\mathrm{E}[W_k]$ (3.42) maintain their form, and ζ (3.39) now equals

$$\zeta = \bar{g}\bar{F} + g/\bar{P} = (\mu e^{-\mu D} + \lambda e^{\lambda D})/(\lambda + \mu)\,. \tag{3.46}$$

Again, it is elucidating to contrast these results with those for infinite-sized buffers. Indeed, the expressions for the $w(i)$ (3.43) and the mean waiting time (3.44) for DT are equally valid for CT if one substitutes $\{\zeta, F\}$ according to (3.45) and (3.46). Further, this very expression for ζ was encountered earlier in the case of infinite-sized buffers, see Sect. 2.2.8.1, Eq. (2.33). Also, Eq. (2.34) for $w(i)$ and Eq. (2.35) for $\mathrm{E}[W]$ indeed provide identical results, as should.

For the determination of D_0, numerical minimization of $\pi(N+1)$ for varying D again provides the easiest method. Although $\partial\mathcal{F}(D)/\partial D = 0$, with $\pi(N+1) = \mathcal{F}(D)$, yields a condition for D_0, it does not allow for a solution in closed form.

3.4.3 Numerical Examples

The current case of a memoryless burst size distribution was considered earlier by Callegati in [24, 26], in CT. As mentioned in the introduction of Chapter 2, [24] presents the first stochastic model for optical buffers, and is probably the best-known stochastic model for optical buffers to date. Therefore, it provides an interesting point of reference to readers acquainted with optical buffers. Since the model of Callegati was approximate, we find it appropriate to compare its output

to that of our exact model. Also, we focus on the optimal granularity for a given load and buffer size, also in CT. (The choice for CT is opportune, in that it is simpler than the DT case, where also the influence of the slot size is to be taken into account. Numerical verification for DT showed that this influence is minor as long as Δ is small when compared to $\mathrm{E}[B]$.)

As reflected in [24,26], Callegati's approach showed that it yields very accurate results for the cases considered there: $\rho = 0.7$ and $\rho = 0.8$, for $N = 254$ and $N = 510$. As such, the somewhat more complicated solution presented here would be of lesser significance, if its only merit were that results are obtained in closed form. Therefore, we extensively explored the output of the numerical model in [26] (that is more accurate than [24]), to find that it is very accurate for $0.4 < \rho < 0.7$. However, for $\rho > 0.8$, accuracy drops especially for small N, whereas for $\rho < 0.4$, accuracy decreases for any N. The latter is illustrated in Fig. 3.9, where it is shown that for $\rho = 0.25$ and $\mathrm{E}[B] = 1/\mu = 1$ μs both the loss probability, and the related optimum D_0, are inexact when determined with the approximate approach of Callegati, for different buffer sizes[2] $N \in \{4, 19, 49\}$. In this respect, the current approach relieves the network engineer of questions of applicability, since it is exact for any load or buffer size, and also yields the correct loss limit for large granularity, namely $\lim_{D\to+\infty} \pi(N+1) = \rho/(1+\rho) = g$, that also goes for any bufferless system, as should. We note that the accurateness for small N is a considerable merit, as debate in recent years proposes small backbone buffers for optical networks, see Sect. 1.7.4 and [89].

At this point, we focus on the dimensioning problem itself: given a certain buffer size N and load ρ, what is the optimal granularity D_0? The answer follows from (3.40), by evaluating for different values of D. While Fig. 3.9 provides the situation for just one load value, an evaluation for any load $\rho \in [0, 1]$ and for five different buffer sizes $N \in \{1, 2, 4, 19, 49\}$ is displayed in Fig. 3.10. Clearly, the optimal value for low load (< 0.25) is extremely large, which seemingly compromises the feasibility of implementation in that case. However, as discussed in Sect. 3.3.1, it suffices to assume that the burst size is limited (which is in practice always so), to show that the optimal granularity D_0 is never larger than the maximum burst size. More important is the observation that also for higher loads, the optimal granularity is considerably impacted by the value of the load, whereas the number of FDLs (buffer size) plays only a secondary role. This constitutes a fundamental issue in optical buffer design, implying that a granularity value chosen for implementation is only optimal for one value of the load, while the actual network evidently has a fluctuating load. Note, however, that other important requirements will probably result in a granularity smaller than the mean burst size, since shorter

[2]The reason to consider $N \in \{4, 19, 49\}$ instead of more obvious values (like multiples of 5) is to ease the comparison with the work of Callegati, that had buffer size captured not by N, but by a parameter that is equal to $N + 1$ (yielding $\{5, 10, 50\}$).

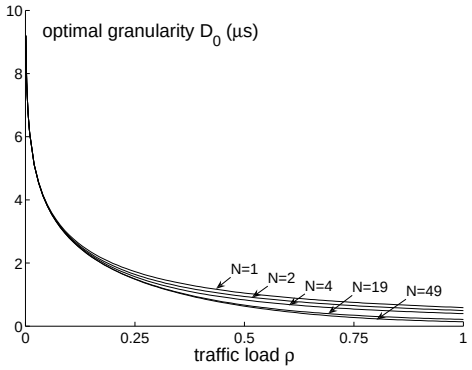


Figure 3.10: Optimal granularity D_0 as a function of the traffic load ρ, for different values of the buffer size N.

FDLs result in a smaller footprint for the switch and reduced cost, and also in a reduced mean waiting time. This contrasts with the case of fixed-length bursts, where matching the granularity with the burst size often yields optimal results, as discussed in Sect. 3.3.2.

3.5 Concluding Remarks

In this chapter, we provided (i) an exact numerical method based on a Markov chain approach, that allows to obtain exact performance results by means of numerical calculation; and (ii) closed-form expressions for some special cases, that can be obtained directly or indirectly from the results of (i).

The exact numerical method was done first for memoryless arrivals, both in CT and DT. By focusing on the evolution of assigned waiting times only, we obtained a very concise system description, with $N + 1$ possible waiting times corresponding to an equal number of system states. Transition probabilities between two states proved expressible by considering either the case of acceptance or loss. The steady-state waiting time probabilities could then easily be extracted from the transition matrix, allowing exact results also for the loss probability, and this with minimal computational burden. This was applied to some numerical examples, that focused on the performance of (very) small buffers, as especially the latter could not be traced in an accurate manner with the pgf approach of the previous chapter.

Further, the exact numerical method was extended to a general arrival process, in DT. Under the single restriction of an upper bound on the burst size, this approach is applicable for any inter-arrival and burst size distribution. By tracing the loss process in a more exact manner, we were able to maintain the same simple

system description of $N + 1$ states, and so retain the same advantages of minimal complexity and computational burden.

In a second part, we focused on the closed-form solution of two particular system settings with memoryless arrivals. Valid for both DT and CT, the first solution was obtained for upper-bounded burst size, and is valid for a particular parameter setting, where the granularity matches the maximum burst size. The derivation proved straightforward, and took the general numerical model as a starting point. The output of the latter was illustrated with some numerical examples, confirming the severe impact of burst size variation on buffer performance. Also, numerical comparison reveals that waiting times in the FDL M/D/1 buffer system are (more than) doubled, when compared to those of the classic M/D/1 buffer.

The second closed-form solution is a general solution for the case of a memoryless burst size distribution. Again valid for both DT and CT, we constructed a somewhat different system description with $N + 2$ states, which enabled an exact solution. Numerical examples showed that the granularity can be tuned so as to yield minimal loss, but only for a given value of the traffic load.

When compared with the previous chapter, apart from providing exact results, the material of this chapter also provides additional insight in the functioning of optical buffers. In particular, we quantified the influence of the small buffer size, and showed that especially for fixed burst size, the optimal granularity is indeed impacted by this. Also, for varying burst sizes (memoryless burst size distribution), it was shown how engineering an optical buffer is indeed a delicate task: setting a large value for the granularity will make the buffer well-adapted to low traffic load, whereas a small granularity value makes the buffer better adapted to high traffic load. Taking into account the smaller resulting footprint of the switch, and the reduced average waiting time resulting from a smaller granularity, it seems credible to the author that optical buffers for variable length bursts will typically have short line lengths, with the granularity significantly smaller than the mean burst size. The difference with the case of fixed-sized bursts is striking, in that then, for any buffer size, matching the granularity with the burst size provides an optimal design, yielding minimal loss up to a certain threshold load. For the case in-between, we considered a burst size distribution with tuneable range, showing that the optimization problem is highly impacted by the burst size range, and has the optimal granularity lowering as the burst size range increases. However, independent of the range, an intuitive reasoning yielded that the optimal granularity is never to be sought beyond the maximum burst size, but approximates the latter when the load is low.

Concluding, the Markov chain approach of the current chapter yields exact results for a wide range of traffic settings, and provides a general-purpose tool on the one hand, while yielding deeper insight on the other hand, through the closed-form solutions that were obtained.

4

Performance Evaluation with Impatience

¶ In this chapter, we briefly consider a simple alternative approach to study the loss probability (LP) of an optical buffer. The model is based on an exact description of a system with impatience, that however does not include any granularity effect. It only yields approximate results, that could already be obtained with somewhat more accurateness with the pgf approach of Chapter 2, and even in an exact manner with the Markov chain approach of Chapter 3. However, since the method is simple, instructive, and might prove useful in cases that are not investigated to date, we consider it too useful not to mention. Also, since the method bases itself merely on the results of a paper by Barrer [82], and does not follow the approach of either Chapter 2 nor 3, it seems most appropriate to present it in a separate chapter.

4.1 Optical Buffers: Voids and Impatience

As mentioned in Sect. 1.7.2, the fact that an FDL buffer can only realize delays belonging to a limited set implies both the occurrence of voids, and impatience. The loss process associated with impatience is different from that of a classic buffer model, in that impatience assumes an upper bound on assignable waiting times, whereas finite waiting room puts an upper bound on the maximum number of bursts in the queue. This difference in loss process results in fundamentally differ-

ent queueing behavior. Even if two systems differ only in the type of loss process, the results can be completely different, with the solution typically somewhat more involving in the case of impatience. Only in some special cases, for example when service times have a deterministic distribution, both types of loss coincide.

In the previous chapters, the combination of the effects of voids and impatience was treated in two different ways. In Chapter 2, the effect of voids was traced in an exact manner, but only for a degenerate buffer setting. Impatience was modeled only in an approximate manner, by means of a heuristic. The latter took into account the tail probabilities of the waiting time distribution, which indeed reflects impatience, but nevertheless yields only approximate results. In Chapter 3, the combined effect of voids and impatience was modeled in an exact manner. However, both features were implied by the system description in an entwined manner, not allowing for specific insight in the influence of both characteristics separately.

In the following, the approach will be complementary to the one of Chapter 2: rather than focusing on the effect of voids, we first trace the effect of impatience in an exact manner, and only introduce the effect of voids in a second step, through an approximation that makes use of the equivalent load. We will focus only on the CT M/M/1 optical buffer system, with the focus rather on insight than on performance modeling, since we already obtained exact results for this system in Sect. 3.4.

4.2 Impatience in Literature

Impatience goes by many names in queueing literature, and further specification of the type of impatience is necessary so as to avoid confusion. The type of impatience we are interested in has *assigned* waiting times bounded by a fixed value D_M, but puts no bound on the scheduling horizon, that evolves unboundedly. Loss occurs whenever the scheduling horizon is larger than D_M.

To the best of the author's knowledge, this type of impatience occurs first in literature in two papers of Barrer: [81] (for random order of service) and [82] (first-come-first-serve). The latter case is especially relevant for us, as it assumes arriving customers (or bursts) to be served in the order in which they arrive, just like we assume in the current work. The model is in CT, and considers negative-exponential inter-arrival and service times, and an arbitrary number of servers c: an M/M/c model.

As for the impatience of the customers, Barrer considers two types of customer behavior in [82]. Quoting the latter paper (except for the notation D_M), the following two cases occur: (1) if a customer is accepted for service before he has waited a time D_M, he remains in the queue until served irrespective of whether or not his total waiting time exceeds D_M. Only those customers who wait for a time D_M without being accepted for service become "lost" customers; and, (2) a customer whose total waiting time is D_M becomes a lost customer irrespective of

whether he is accepted for service or not.

As for the application of optical buffers, clearly, the first type of customer behavior is almost the one that we are interested in, since optical buffers can only provide a maximum delay D_M equal to the delay time of the longest FDL. One difference lies in the fact that in this system setting, customers (or bursts) that eventually renege wait in line for a time D_M, whereas in the optical buffer model, such customers do not enter the queue at all. Apart from the queue content, this does however not influence the mean waiting time of accepted customers, nor the loss ratio. This is also mentioned in [101], where the same type of impatience is studied for a GI/G/1 model with generalized impatience: rather than a deterministic parameter D_M, [101] considers a rv for impatience. In [102], first results were obtained for the same system and the same generalized notion of impatience, but with impatience there described as *reneging phenomena*. More general system settings that all comply with our notion of impatience are considered in [103–105]. Finally, results in [106, 107] are derived for a DT setting, for both the case of limited sojourn time and limited waiting time, with deterministic patience distribution and general patience distribution. The case of general patience distribution, combined with Markovian arrival process, is considered in [108].

Inspired by [104], the type of impatience we consider can be unambiguously described as *deterministic customer impatience, with impatience until the beginning of service.* We note however that, apart from the type of impatience mentioned, many others occur in literature, and, as mentioned by Stanford in [109], one is to beware of confusion due to different terminology. For example, the term "reneging" is often used interchangeably with "impatience" in literature, but one is to check the entire system description to be sure that it is so. Another often-recurring term for impatience is "balking", that is mentioned notably in [24], in the context of optical buffers. Further, in [110], a complete section is devoted to the non-related case of *uniformly bounded actual waiting time*, that assumes service times to evolve dynamically, so as to ensure that customers awaiting service endure waiting times never larger than some maximum waiting time. Gavish considers *bounded waiting time* in [111], but actually studies a limit on sojourn times (waiting time plus service time).

4.3 Barrer's Result and Granularity

In this section, we will show that the model that Barrer investigated upon shares characteristics of the optical buffer model. More precisely, it turns out that Barrer's model coincides with a CT M/M/1 optical buffer model of finite size, if the granularity D of the latter is considered infinitely small. As such, we come to study an optical buffer that still is degenerate (FDLs have lengths iD), that still has a finite maximum waiting time ($D_M = ND$), but that has infinitely small

granularity value D	0	0.2	0.5	1	2
N for $D_M = 4$	$N \to \infty$	20	8	4	2
N for $D_M = 20$	$N \to \infty$	100	40	20	10

Table 4.1: Parameter setting used for the comparison presented in Fig. 4.1.

granularity, $D \to 0$, and an infinitely large number of FDLs, $N \to \infty$. Clearly, this system cannot actually be implemented, but serves as an instructive point of reference, that can be analyzed with the simple formula presented in [82].

We note that the simplification of $N \to \infty$ is indeed shared with the pgf approach of Sect. 2, but with the difference that here, we maintain an exact description of the maximum delay D_M (still a "finite-sized" buffer), whereas D_M was assumed infinite in the model of Sect. 2, to obtain an infinite-sized buffer. By assuming the correct value for D_M right from the start, the model allows for a better tracing of the loss process in case of high load.

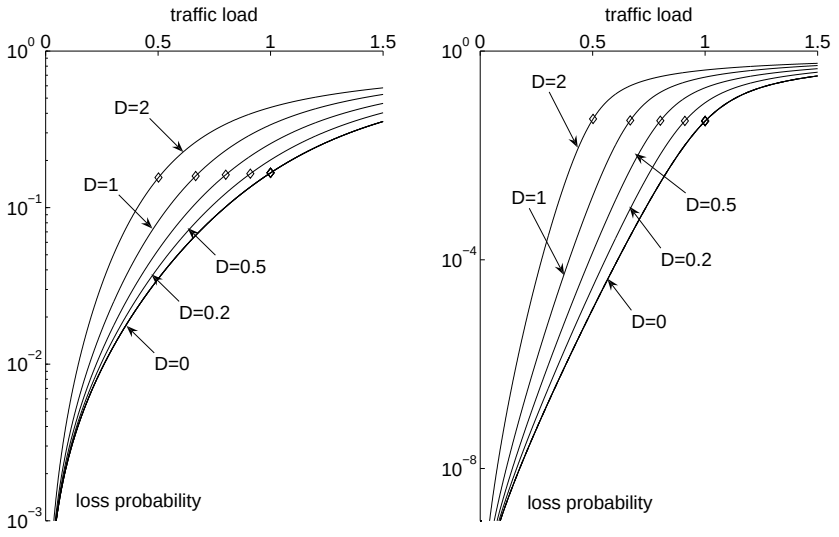
The purpose now is to use the exact result of Barrer, formulated for classic (non-optical) buffers, to evaluate the performance of optical buffers. The results of Sect. 3.4 serve the purpose of reference, and allow to easily evaluate the accuracy of the approximation.

Barrer obtains a closed-form expression for the LP of an M/M/1 system with impatience, namely

$$\mathrm{LP} = \begin{cases} \dfrac{(1-\rho)\rho}{e^{\mu D_M(1-\rho)} - \rho^2} & \rho \neq 1\,, \\ \dfrac{1}{\mu D_M + 2} & \rho = 1\,, \end{cases} \tag{4.1}$$

where we adopted the notations μ^{-1} for the mean service time, $\rho = \lambda/\mu$ for the traffic load, and λ^{-1} for the mean inter-arrival time, as in Sect. 3.4. This formula is extremely simple, and tightly relates to an optical buffer's performance with the same impatience (or maximum delay), especially when D is small (or N large).

The application being a degenerate optical buffer system, in Fig. 4.1, we compare the LP obtained from (4.1), for $D = 0$, to the results of Sect. 3.4, with various values for the granularity. The optical burst sizes have mean burst size $\mu^{-1} = 1\ \mu$s, and a Poisson arrival process with varying intensity λ, equal to the traffic load ρ (since $\mu^{-1} = 1$). Fixing the maximum achievable amount of delay $D_M = ND$ to $4\ \mu$s (left pane) and to $20\ \mu$s (right pane), the granularity values considered lead either to a finite amount of FDLs, or to the limit of an infinite number of FDLs, as displayed in Table 4.1. As a reference, all curves carry a diamond ($\diamond$) to indicate the load level at which the equivalent load, ρ_{eq}, reaches one. In the current case of an optical M/M/1 system in CT, the expression for the latter is obtained earlier in



(a) Maximum delay D_M fixed to 4 μs. (b) Maximum delay D_M fixed to 20 μs.

Figure 4.1: Comparison of the LP of CT M/M/1 optical buffer systems with different granularity D, but with the same maximum delay D_M. For $D = 0$, the curve corresponds to a classic (non-optical) M/M/1 system with deterministic patience fixed to D_M. The diamond ($\diamond$) on each curve indicates the point where $\rho_{eq} = 1$. These figures were obtained with $\mathrm{E}[B]$ normalized to 1 μs.

Sect. 2.2.8.1, Eq. (2.36), and reads

$$\rho_{eq} = 1 + \frac{\lambda D}{\mu + \lambda} \left(\frac{\lambda}{1 - e^{-\mu D}} + \frac{\mu}{1 - e^{+\lambda D}} \right) . \tag{4.2}$$

As mentioned before, the latter defines the load in the infinite system (that is, the system with fixed granularity D but both $N \to \infty$ and $D_M \to \infty$), and turns one when overload in the infinite system is reached. As such, it provides a characterization of the load without being dependent on the buffer size.

On the left pane of Fig. 4.1, with $D_M = 4$ μs, it comes as no surprise that decreasing granularity (more FDLs) leads to significant performance bettering. Further, it shows that, even with D_M fixed, the granularity has a paramount impact on performance, that remains visible even for granularity values as small as 0.2. Also, the point where ρ_{eq} turns one ($\diamond$) comes about as a reference point: for $0 < \rho < \rho_{eq}$, the LP grows fast with the traffic load (as reflected in quasi-linear curves on the log-lin scale applied), while for $\rho > \rho_{eq}$, loss grows slowly, with an asymptote at $(\rho - 1)/\rho$ for $\rho \to \infty$, with the effect of granularity gradually fading.

This role of reference point comes about even more distinctly when we consider the curves for a larger achievable delay, $D_M = 20\,\mu$s, on the right pane of Fig. 4.1. The latter curves further confirm the major role of the granularity in performance evaluation, since even the case of $D = 0.2$ significantly differs from the limiting case with $D = 0$. However, for D even smaller, ($D \ll \mu^{-1}$, here, $D < 0.1$), the curves nearly overlap, as should.

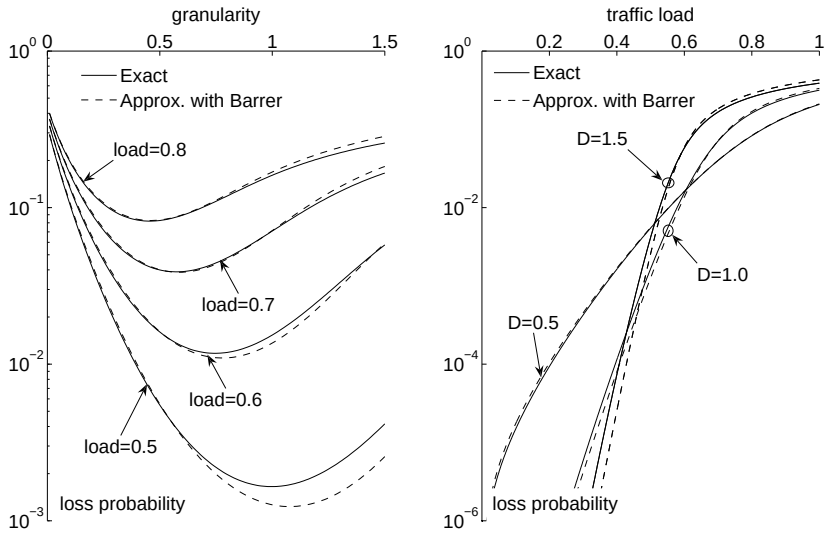
4.4 Approximate Model

In the previous section, we showed how the expression of Barrer (4.1) is consistent with the results of Sect. 3.4, and even provides a good approximation for the latter, but only if $D \ll \mu^{-1}$. Further, the curves in Fig. 4.1 showed how the equivalent load can serve as a reference, especially for $\rho_{eq} = 1$. Now, question is whether the role of the equivalent load can be expanded, so as to allow a simple approximation of an optical buffer's performance with Barrer's model.

Given that the equivalent load is immediately derived from the drift of the waiting time process of the infinite system (see Sect. 2.1.5), the load indeed incorporates the impact of the granularity on the waiting time process, without considering the buffer size. As such, it is the counterpart of Barrer's model, that traces the performance impact of finiteness (finite patience), with a correct modeling of the loss process, without considering any FDL effect in the waiting times. This feeds the intuition that the combination of both might yield a good approximation to trace an M/M/1 optical buffer's performance, by simply substituting ρ in (4.1) by ρ_{eq} (4.2).

In Fig. 4.2, the latter approximation is put to test, by comparing its output to results from the exact analysis presented in Sect. 3.4, all for buffer size $N = 20$, and mean burst size $\mu^{-1} = 1\,\mu$s. On the left pane, the LP is plotted as function of the granularity, for four cases of the traffic load. We recognize the typical U-shape of the granularity optimization curves, as studied earlier, notably in Fig. 2.2, where the curves for $\rho = 0.6$ and $\rho = 0.8$ are also on display (up to a scaling factor, since $\mu^{-1} = 50\,\mu s$ there). The approximation displays good accuracy, especially for high traffic load, and not too large a value of the granularity. However, to assess the impact of traffic load and granularity, we consider the plots of the right pane, where the LP is displayed as function of the load, for three different values of the granularity. This further confirms that the approximation works best for small granularity, since the curves for $D = 0.5\,\mu s$ nearly coincide with the exact ones. For large granularity, the approximation is accurate only for high traffic load. Further, results not shown here confirm these results also for other buffer sizes N, showing that accuracy is not influenced much by the latter.

As an approximation, we note that the mentioned approach shares some of the properties of Callegati's approximation, that also performed best for not too large a



(a) The LP in function of the granularity (in μs), for load $\rho \in \{0.5, 0.6, 0.7, 0.8\}$.

(b) The LP in function of the traffic load, for $D \in \{0.5, 1.0, 1.5\}$ μs.

Figure 4.2: Comparison of exact results for an M/M/1 optical buffer in CT with the results of the approximation with Barrer's model. The approximation allows for accurate results, especially for high traffic load, and not too large values for the granularity ($D \leq \mu^{-1}$). These figures were obtained for $N = 20$, and with $\mathrm{E}[B]$ normalized to 1 μs.

granularity (see Sect. 3.4), and has comparable accuracy. However, there, an iterative procedure was needed to obtain results. Opposed to this, the current approach only involves Barrer's simple formula, and the notion of equivalent load, that can be derived (also in cases much more general than M/M/1) without analyzing the entire system, thanks to its simple definition (Sect. 2.1.5).

4.5 Approximate Model for DT

While results up to now only cover the CT case, we now move to the DT case, that is especially interesting in the current context. More precisely, the relation between queues with impatience and optical buffer modeling is tighter than in the CT case. While in CT a queue with impatience was obtainable from the optical buffer by letting $D \to 0$, it suffices to set $D = 1$ in an optical buffer model to obtain a DT model for queues with impatience. In the present context, it suffices to set $D = 1$ in all DT results of Chapter 3, to obtain exact results for a general GI/G/1

system with deterministic impatience in DT. Contrasting this with results for CT, it is remarkable that even the less general M/G/1 system in DT of [105] requires numerical approximations in order to obtain results. Apparently, assuming a DT setting somewhat simplifies the queueing problem with impatience.

Further, DT is also the time setting studied in [107], and as such, some of the results of Chapter 3 are also covered there, allowing to verify consistency. In [107], a general impatience distribution with upper bound r is assumed, and thus allows to handle a deterministic patience distribution, by setting $r = D_M + 1$. (The term +1 is to be introduced to be compatible with the definition of r in [107]).

An instructive comparison is the special case of M/M/1 in DT, treated in both Sect. 3.4 and [107]. Setting $D = 1$, the results of Sect. 3.4.2.3 can be simplified, with $D_M = N$, to obtain that the LP equals

$$\text{LP} = \begin{cases} \dfrac{(1-\rho)\cdot\rho}{(\bar{p}/\bar{f})^{N+1} - \rho^2} & \rho \neq 1\,, \\ \dfrac{1}{(N+1)q/\bar{f} + 2} & \rho = 1\,, \end{cases} \tag{4.3}$$

with $\mathrm{E}[T_k] = 1/p$, $\bar{p} = 1-p$, $\mathrm{E}[B_k] = 1/f$, $\bar{f} = 1-f$ and $\rho = p/f$. Assuming this notation and applying the result of Sect. 3 of [107] yields that the probability that the age of the customer in service is zero (there denoted $\hat{\pi}_0$) equals

$$\hat{\pi}_0 = \frac{\bar{p}^r f(p-f)}{p^2\bar{f}^r - \bar{p}^r f^2}\,.$$

With now $\text{LP} = 1 - (1-\hat{\pi}_0)/\rho$ and $r = N+1$, one easily obtains (4.3), proving consistency.

Finally, we note that the expression for the LP in DT (4.3) is tightly related to the one of CT (4.1). To see this, we rewrite the CT expression for $\rho \neq 1$ as

$$\text{LP} = \frac{(1-\rho)\rho}{e^{\mu D_M} e^{-\lambda D_M} - \rho^2} \quad , \quad \rho \neq 1\,.$$

Now, a substitution similar to the one of Sect. 3.4 (Eq. (3.45)) is needed to yield correspondence,

$$\bar{p} = e^{-\lambda}\,,\ p = 1-\bar{p}\ ;\ \ \bar{f} = e^{-\mu}\,,\ f = 1-\bar{f}\,,$$

completed with $D_M = N+1$, to indeed obtain (4.3). The expression for $\rho = p/f = 1$ follows by taking the limit for $p \to f$. The link between DT and CT is less intuitive at one point, since D_M is "virtually expanded" to $N+1$ in DT, rather than N. The latter however forms no stumbling block: just like before (for example in Sect. 3.1.6 and Sect. 3.3.3.3) it can again be understood as (an indirect) result of the difference in the offset in the inter-arrival times in DT ($T_{min} = 1$ in DT, $T_{min} = 0$ in CT).

Together with the appropriate expression for ρ_{eq} in DT, (4.3) can be used for an approximate modeling of optical buffers. This is not treated further here, since results are similar to the CT case.

4.6 Concluding Remarks

In this chapter, we highlighted the link between optical buffer modeling and queues with impatience. While the most characteristic feature of optical buffers is the limited set of possible waiting times, a second feature is the finiteness of the achievable delay. The latter property corresponds to an often-studied type of impatience, and allows for a comparison between optical buffers and queues with impatience. Given that the loss process is shared, this also allowed for an approximate modeling of optical buffers, where the effects of denumerability of waiting times is incorporated in the equivalent load. For M/M/1 optical buffers in CT, the approximation was compared to exact results, showing that it attains high accuracy especially for not too large granularities and high traffic load. Also, the complementary case for DT is mentioned, and allows a similar approximation.

While the focus of the current section was solely on the M/M/1 optical buffer system of finite size (both for CT and DT), other cases would certainly deserve further exploration. More precisely, there is no reason why an exact model for impatience, combined with the appropriate expression for the equivalent load, would not yield accurate results in more general cases. For example, results not included showed that good accuracy is also obtained for the M/D/1 system, by introducing the equivalent load in the result of Brun and Garcia [99] (also, Eq. (3.33)). Other cases remain unexplored to date, and might prove particularly useful when the exact problem becomes intractable, for example when multiple wavelengths are available for service.

5

Stability Analysis with Regenerations

¶ In the previous chapters, the performance analysis of optical buffers was sometimes based on the analysis of an infinite buffer system. More precisely, in Chapter 2, (exact) results were obtained for an infinite system, and only applied in a second step for finite optical buffer systems. As usual when working with infinite-sized queueing systems, stability is to be taken into account explicitly. In Chapter 2 it was obtained that optical buffers display instability even when the average burst size (or service time) does not exceed the average inter-arrival time. To be precise, an equivalent load ρ_{eq} was introduced and provides a characterization of stability ($\rho_{eq} < 1$, see (2.9)), but only for the case of degenerate optical buffers, with a single outgoing wavelength. However, note that this equivalent load was introduced as a modeling assumption, necessary for the model to be applicable, rather than as an actual stability guarantee. Further, neither the case of multiple wavelengths nor the case of a non-degenerate FDL set was studied before in terms of stability.

In this chapter, both cases are studied, and a proof of sufficient stability conditions is presented for general non-degenerate FDL buffers of infinite size, general arrival process and general burst sizes, and either one, or multiple wavelengths. While the emphasis of previous chapters was on performance modeling, this chapter foregrounds the stochastic model itself. Therefore, the stochastic processes that produced the performance results of previous chapters are characterized here in much more detail, which allows for additional insight into the functioning of FDL buffers.

The stability conditions we present below are valid for a setting as general as

a non-degenerate GI/G/c system. This is considerably more general than previous results in the field of optical networking: [11,16] are the sole contributions treating stability explicitly, stating an exact value for the sustainable load of a degenerate M/G/1 buffer in DT [11], and of a degenerate setting with Markovian arrivals and possibly correlation between successive burst sizes, in DT [16]. In a different context, Lakatos already obtained the first stability result in [20,21] for an M/M/1 setting, and casted the DT and CT case recently in one model in [22], as mentioned in Chapter 2.

Further, the stability analysis presented in this chapter accounts for the first application of the regenerative approach (in the spirit of [112–114]) to the FDL buffer problem. This is complementary to all previous work, that could rely on more common notions of stability (for example, by considering the exact transition probabilities of the Markov chain [20–22]) because the assumptions there were less general. The conditions we obtain are simple and general, and therefore easily applicable to any asynchronous optical buffer setting with iid inter-arrival times and burst sizes. Moreover, these conditions only refer to the line lengths and involve only the first moments of inter-arrival time and burst size distribution. The other side of the coin is that the conditions are sufficient, but (save some "pathologic cases") not necessary. As such, it is possible that an FDL buffer setting does not adhere to the conditions given in this chapter, but nevertheless is stable. This difference between "sufficient" and "necessary" is discussed in Sect. 5.1.5.

As for the analysis itself, the regenerative property of the waiting-time process plays a key role. Given that the system is described by a modified form of the Lindley equation, the analysis shares elements with the stability analysis of a classic GI/G/1 queue, respectively GI/G/c queue, especially at the start. However, the analysis follows completely different lines when the discrete nature of the waiting times is taken into account, necessary to establish positive recurrence of the corresponding regenerative process. The case of a single wavelength and that of multiple wavelengths are treated subsequently, in Sect. 5.1 and 5.2 respectively.

Further, the time setting is CT throughout the entire chapter, and the stability analysis in case of DT is not discussed further. However, such analysis should follow from the analysis given here in a straight-forward manner, since the time setting plays only an implicit role in the analysis, and no specific assumptions are made for the distribution of inter-arrival times and burst sizes.

5.1 Single-Wavelength System

In this section, the single-wavelength GI/G/1 optical buffer system is treated, for non-degenerate line lengths. First, the specifics of the stochastic model are presented together with an instructive example, that points out the difference between a classic waiting-time process and an optical one. In Sect. 5.1.4, the proof of the

main stability result using regenerative stability analysis is presented. Further, in Sect. 5.1.5, a comparison is made between the stability conditions obtained here and the ones of Chapter 2, for the case of degenerate buffering.

5.1.1 Buffer Setting

In the current case of a single wavelength, the buffer is dedicated to a single wavelength of a fiber channel and hence, no more than one burst can be transmitted at a time. As in the rest of this work, we assume a FIFO delay-line assignment algorithm. The buffer setting is non-degenerate like in Chapter 3, but with the difference that the buffer size is assumed infinite here, instead of finite. We mainly follow the notation as coined in Chapter 3: an FDL set $\mathcal{A} = \{a_0, a_1, a_2, \ldots\}$ of available delays $a_i \in \mathbb{R}^+$, $i \in \mathbb{N}$, with $a_0 = 0$ by definition, $a_i \neq a_j$ for $i \neq j$, $a_0 < a_1 < \ldots$ Further, we assume that no impatience is involved: $\lim_{i\to\infty} a_i = \infty$, so assuring that a suitable delay line can always be found.

The effect of voids is also captured by means of the operator introduced in Chapter 3: when a delay x is requested, the actual delay a_i is chosen from the FDL set $\mathcal{A}$ such that $a_{i-1} < x \leq a_i$. The resulting delay-line assignment procedure (select a_i given x) is captured in operator form as

$$\lceil x \rceil_{\mathcal{A}} = \inf\{a_i \in \mathcal{A} : a_i \geq x\},\ x \in \mathbb{R}^+ .$$

5.1.2 Traffic Setting

The traffic setting is shared with the rest of this work: bursts arrive in the optical buffer system one by one. Upon arrival, a burst is always accepted since the buffer has an infinite size. Numbering the bursts in the order in which they arrive, let t_k be the arrival instant of burst k, $k \geq 1$. Also denote $T_k = t_{k+1} - t_k$, the inter-arrival time between the kth and $(k+1)$th arrival, and assume that T_k are iid random variables (rv's) and follow a general distribution (renewal input). With burst k we associate a burst size B_k, with $\{B_k\}$ also constituting an iid sequence with general distribution. For notational convenience, we also introduce the iid rv $U_k = B_k - T_k$, $k \geq 1$.

Further, we will assume a CT time setting. This implies that the rv's can take on arbitrary values in $\mathbb{R}^+$ (t_k, T_k, B_k) or in $\mathbb{R}$ (U_k).

5.1.3 Evolution of the Waiting Time

As mentioned earlier in this work (Sect. 2.1), for the analysis of a single-wavelength FDL buffer, one can either focus on the evolution of the scheduling horizon (as is done in Chapter 2), or on the evolution of the waiting times (as is done in Chapter 3). While the waiting time evolution allows for the most concise description

of a single-wavelength buffer, the scheduling horizon evolution is somewhat more informative, since it shows explicitly how bursts join the buffer queue. With the emphasis in this chapter on the description of all stochastic processes involved, we will treat both the waiting-time and the scheduling-horizon process, and also make the comparison with the waiting-time process in a classic buffer system. Further, even though the role of the scheduling horizon is minimal in the analysis for a single wavelength, it will prove indispensable for the analysis of the case with multiple wavelengths. Therefore, both processes merit discussion also at this point. Also, we will treat the involved processes not only upon arrival instants, as was done in the rest of this work, but also on arbitrary instants. Note that knowledge of the latter is not a strict necessity for the stability analysis, and that it is included here only to provide the reader with additional insight in the difference between classic and optical buffers.

First of all we introduce a left-continuous *virtual waiting-time process* $\hat{W} = \{W(t)\ ,\ t \geq 0\}$ (with both $\hat{W}$ and $W(t)$ thus indicating the same process.) For each instant t, $W(t)$ equals the waiting time a burst gets, if it would arrive at instant t. By the left-continuity, the *actual* waiting time of burst k is defined as $W_k = W(t_k)\ ,\ k \geq 1$. In accordance with the FIFO delay-line assignment, the consecutive waiting times $W_k \in \mathcal{A}$ satisfy the following recursion,

$$W_{k+1} = \lceil W_k + U_k \rceil_{\mathcal{A}},\ \ k \geq 1\,. \tag{5.1}$$

To gain more insight in the waiting-time process, we introduce the underlying (left-continuous) process $\hat{H} = \{H(t),\, t \geq 0\}$ that describes the (virtual) scheduling horizon at every time instant. Let now $H_k = H(t_k),\, k \geq 1$. The *scheduling-horizon process* $H = \{H_k\,,\, k \geq 1\}$ is associated with the optical system and describes the delay that an arriving burst *requires*, with $H_k \in \mathbb{R}^+$. The evolution is described by the following recursion:

$$H_{k+1} = [\lceil H_k \rceil_{\mathcal{A}} + U_k]^+,\ \ k \geq 1\,. \tag{5.2}$$

The equations (5.1) and (5.2) relate through

$$W_k = \lceil H_k \rceil_{\mathcal{A}},\ k \geq 1\,,$$

and thus in general $H_k \leq W_k$ for each k.

The third process is the *classic virtual waiting-time process* $\hat{W}_c = \{W_c(t), t \geq 0\}$, that traces the virtual waiting time in a classic single-server system with infinite buffer, the same input and burst sizes allowing any possible delay (continuous waiting room). From this, one can derive the *classic waiting-time process* $W_c = \{W_c^k,\, k \geq 1\}$, through $W_c^k = W_c(t_k),\, k \geq 1$. The waiting-time process in a classic system evolves according to the well-known Lindley equation

$$W_c^{k+1} = [W_c^k + U_k]^+,\ k \geq 1\,.$$

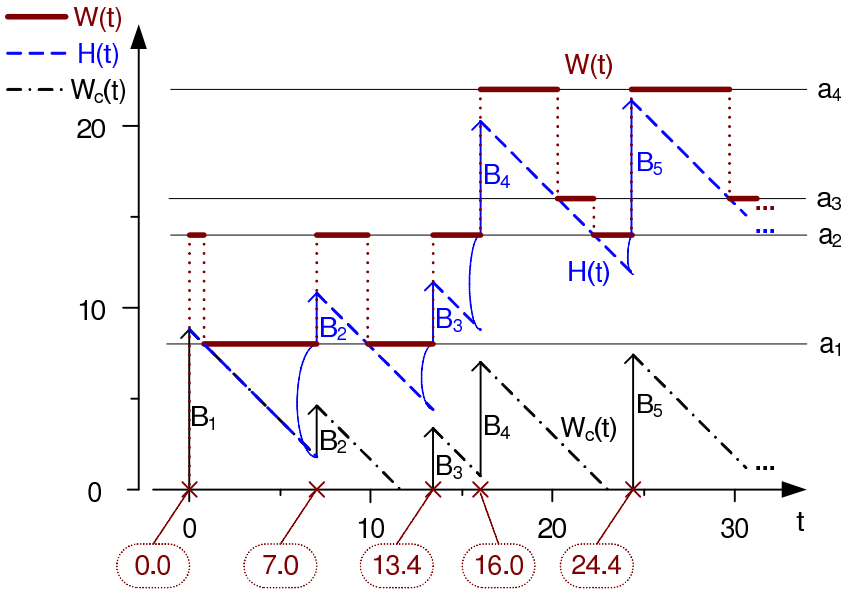


Figure 5.1: The virtual processes $W(t)$ and $H(t)$ associated with the FDL buffer clearly illustrate the impact of voids, since they attain higher levels in general than the classic virtual waiting-time process $W_c(t)$. The crosses on the (horizontal) time-axis represent arrivals $k = 1 \ldots 5$. Both the classic and optical process have the same burst sizes B_k and arrival instants t_k, $k \geq 1$, respectively. These values are summed up in Table 5.1 together with the values of other relevant rv's, whereas the FDL lengths a_i are lined up in Eq. (5.3).

k	1	2	3	4	5
B_k	8.8	2.8	3.4	6.6	7.4
t_k	0.0	7.0	13.4	16.0	24.4

Table 5.1: Values of the variables corresponding to Fig. 5.1.

Given that the evolution of waiting times in the system described by (5.1) is similar but not identical to this form, we can refer to (5.1) as *modified Lindley equation*. Note that, for the classic system, each arriving burst k *always* gets the delay it requires. As opposed to the optical buffer, the classic system never wastes time by additionally delaying bursts and therefore, the classic virtual waiting-time process $\hat{W}_c$ also denotes the virtual workload of the classic system.

As an instructive example, a typical evolution of the three processes is consid-

ered in Fig. 5.1; specifics of the used trace are given in Table 5.1, where

$$\mathcal{A} = \{a_0, a_1, a_2, a_3, a_4, \ldots\} = \{0, 8, 14, 16, 22, \ldots\}\,, \tag{5.3}$$

with a_i, $i \geq 5$, not further mentioned since those values do not impact the figure.

In both the optical and classic system, bursts are served one by one over time. Moreover, $H(t)$ and $W_c(t)$ diminish at rate one in-between arrivals, whereas the evolution of the process $W(t)$ is described by a step function reflecting the non-work-conserving discipline in the optical system.

The difference between both comes from the behavior upon arrivals. Arriving burst k brings an amount of work B_k. In a classic system, governed by the process $\hat{W}_c$, this amount sums up with the work that was already in queue, and yields a virtual waiting time $W_c^k + B_k$. In the optical system, the scheduling horizon $\hat{H}$ is first augmented with an extra delay (indicated by the curbed lines in Fig. 5.1) upon arrival of burst k, because this burst can only commence service after spending some time $W_k = \lceil H_k \rceil_{\mathcal{A}}$. This gives the virtual waiting time $\lceil W_k + B_k \rceil_{\mathcal{A}}$ upon arrival, and results in a fundamental difference between $\hat{W}_c$ and $\hat{W}$ as shown in Fig. 5.1.

With now a broader look on the involved processes, we narrow down the analysis to only the discrete-time process $W_k = W(t_k)$, $k \geq 1$, that suffices as input to the single-wavelength stability analysis. Notice that this is complementary to the case of multiple wavelengths (Sect. 5.2), where the scheduling horizon will be the prime variable of interest.

5.1.4 Regenerative Stability Analysis

With the system equation (5.1) at hand, we now move to the actual stability analysis of the single-wavelength system. As mentioned, our method will exploit the regenerative technique in the sense of [113]. There, and in the stability analysis of Markov chains in general [115], concepts of both renewal theory and Markov chain theory are put to use. This gives rise to a specific parlance that was not highlighted earlier in this work, and therefore deserves some explanation here.

In terms of Markov chain theory, it is clear that the the waiting time process is a Markov chain, and thus, with each possible waiting time a_i, we can associate a Markov state. Each arrival then corresponds to a transition in the infinite-sized Markov chain. Of these states, the state with zero waiting time plays a special role, since it can be used to characterize a renewal process, with each transition to the renewal state referred to as a regeneration. Further, this state, like any state in the Markov chain, is either transient or recurrent. In the case of a transient state, the probability that the hitting time (defined below by (5.6)) is infinite is larger than zero, whereas this probability is zero for a recurrent state. For recurrent states, a further distinction is made between the case of finite expectancy for the hitting

time, referred to as positive recurrence, and the case of infinite expectancy for the hitting time, which is referred to as null-recurrence.

In terms of renewal theory, the busy period of the waiting time process is taken as renewal interval, and each arrival instant that finds the buffer empty accounts for a renewal instant or regeneration. The time between renewal instants, in general referred to as the renewal interval, is called the regeneration time in this context, and corresponds to the hitting time of the zero waiting time state. Further, the renewal process constitutes a regenerative process, on the condition that the expected value of the regeneration time is finite. As such, relating now Markov chain theory and renewal theory, the renewal process of the waiting times is called regenerative if and only if the Markov state associated with zero waiting time is positive-recurrent. This is especially relevant for stability analysis, because this provides us with a way to prove stability under zero initial conditions: it suffices to prove that the zero waiting time state is positive-recurrent, to show that, still under zero initial conditions, the system is stable. This is the main part of the proof, and is presented in Sect. 5.1.4.2. The second part consists in relaxing the assumption of a zero initial state, to show that the system is stable regardless of the initial state, if a minor additional assumption is made. This is presented in Sect. 5.1.4.3. Further, still in Sect. 5.1.4.3, we show that that same set of stability conditions not only guarantees stability, but also provides a sufficient condition for the existence of a stationary regime for the waiting time process. Finally, the difference between sufficient and necessary conditions for the single-wavelength case is highlighted in Sect. 5.1.5.

5.1.4.1 Notations and Characterization

Before moving to the main stability result, we introduce the notation that is needed for the proof. For the waiting-time process $W = \{W_k, k \geq 1\}$, we denote the differences as

$$\Delta_W(k) = W_{k+1} - W_k,\ k \geq 1\,,$$

whereas, for the buffer setting, we introduce

$$\begin{aligned} g_n \equiv a_{n+1} - a_n\ ,\ n \geq 0 \quad ; \quad \delta^* \equiv \inf_{n\geq 0} g_n\,; \\ \Delta^* \equiv \sup_{n\geq 0} g_n \quad ; \quad \Delta_0 \equiv \limsup_{n\to\infty} g_n\,, \end{aligned} \tag{5.4}$$

where "$\equiv$" denotes that the relation concerns a definition. Hereby, we note that Δ^* and Δ_0 are identical in many common cases. One example is a degenerate buffer setting, where $g_n = D$, $\forall n \geq 0$, and both Δ^* and Δ_0 are equal to the granularity D. Another example is the case where the FDL set of a non-degenerate buffer is periodical in the g_n, that is, $g_{n+jP} = g_n$ for some period P, $\forall j \in \mathbb{N}$. In that case, Δ^* and Δ_0 are equal to $\sup_{n\in\{0,1,\ldots,P-1\}} g_n$. On the other hand, note that the

difference between Δ^* and Δ_0 is fundamental in the context of stability analysis, and that the case with $\Delta^* > \Delta_0 > 0$ also corresponds to a possible system setting, for example in the situation where only a finite number of the g_n is larger than some $\hat{D}$, and $g_n = \hat{D}$ for all n larger than some system parameter $n_{\hat{D}}$. Summarizing, all the system settings mentioned here are included in the current description, by considering both Δ^* and Δ_0.

Further, introduce the distribution $F_U(x) = \Pr[U \leq x]$, where $U = B - T$, and B, T are generic variables for B_k and T_k, respectively. Also, for each $x \geq 0$ define the overshoot

$$\lceil x \rceil_{\mathcal{A}} - x = \Delta(x)\,. \tag{5.5}$$

Note that $\Delta(x) \geq 0$ and $\Delta(a_k) = 0$ for any $k \geq 0$.

Now, we define the regeneration times $\{\beta_n,\, n \geq 1\}$ for the waiting-time process W in the following (conventional) way: let $\beta_0 = 0$ and

$$\beta_{n+1} = \inf(k > \beta_n : W_k = 0),\ n \geq 0\,, \tag{5.6}$$

where we put $\inf \emptyset = \infty$. As such, the regeneration times, also called regenerations or hitting times, capture the indices of those arrivals that get assigned zero as waiting time. As outlined above, our main purpose, treated in Sect. 5.1.4.2, is to establish conditions which imply positive recurrence of the zero waiting time state, or, equivalently, of the process β. The latter means that

$$\beta_1 < \infty \text{ with probability 1 (w.p.1) and } \mathrm{E}[\beta_2 - \beta_1] \equiv \alpha < \infty\,. \tag{5.7}$$

In what follows we assume a zero-delayed process β when $\beta_1 \sim \beta_2 - \beta_1$, where $\sim$ stands for stochastic equivalence. In other words, the process W starts at regeneration instant, $W_1 = 0$ and $\mathrm{E}[\beta_1] = \alpha$. Note that the zero-delayed process $\beta = \{\beta_n\}$ is well-defined since $a_0 = 0 \in \mathcal{A}$, allowing to assume zero initial conditions. As mentioned, this assumption will be relaxed in Sect. 5.1.4.3, where we also consider the case of non-zero initial conditions

Further, the recurrence property of the renewal process β can be characterized in several ways. In this chapter, we choose a formulation via the limiting behavior of the forward regeneration time at instant n, which is defined as

$$\beta(n) = \inf_k(\beta_k - n : \beta_k - n > 0),\ n \geq 1\,. \tag{5.8}$$

In words, given an arbitrary arrival with index n, $\beta(n)$ is the index of the first arrival following arrival n that is assigned zero waiting time. Finally, as mentioned above, positive recurrence of the zero waiting time state demands that its hitting time (the regeneration time) has finite expectancy. As known from [116], this relates to the forward regeneration time through

$$\alpha = \infty \text{ if and only if } \beta(n) \to \infty \text{ in probability as } n \to \infty\,. \tag{5.9}$$

Since $\mathrm{E}[\beta_1] = \alpha$ by assumption, $\alpha < \infty$ implies $\beta_1 < \infty$ w.p.1, and positive recurrence follows. Summarizing, showing that

$$\beta(n) \not\to \infty \text{ as } n \to \infty \text{ in probability}\,,$$

implies positive recurrence of the zero waiting time state, which in its turn guarantees stability under zero initial conditions. This marks the approach of the next section.

5.1.4.2 Main Stability Result

Now we formulate the assumptions which are adopted throughout this chapter. First we assume that $\mathrm{E}[B] < \infty$, $\mathrm{E}[T] < \infty$ and thus, $\mathrm{E}[U] \in (-\infty, +\infty)$. Furthermore, we sum up the three stability conditions, sufficient to guarantee stability for any single-wavelength FDL buffer adhering to them.

The first condition is that the FDL buffer set is properly defined,

$$\delta^* > 0, \quad \Delta^* < \infty\,, \tag{5.10}$$

or, in other words, that all delay lines have non-identical and finite length. Secondly, we assume as negative drift condition that

$$\Delta_0 + \mathrm{E}[U] < 0 \tag{5.11}$$

holds. This condition defines a bound on the maximum tolerable load. As such, it is a generalized instance of the stability characterization for single-wavelength FDL buffers with degenerate structure, as presented in Chapter 2 (Eq. (2.8)). Of the three conditions, it is this one that results in sufficient stability conditions, rather than necessary conditions, as discussed in Sect. 5.1.5.

Thirdly, we assume as regeneration condition that

$$\Pr[T > \Delta^* + B] = \delta_1 \tag{5.12}$$

holds for some $\delta_1 > 0$. This condition is specific to the regenerative approach, and guarantees that even the largest difference between line lengths, Δ^*, forms no obstacle to return to the zero waiting time state with non-zero probability. Rather than an actual necessity for stability, this condition is auxiliary to our approach, allowing us to characterize stability as positive recurrence of the regeneration-time process.

At this point, we prove that these three conditions guarantee stability for zero initial conditions. Hereby, recall that positive recurrence of the regeneration-time process β (5.7) implies positive recurrence of the zero waiting time state, which in its turn implies stability under zero initial conditions.

Theorem 5.1.1. *Under assumptions (5.10)-(5.12), the zero-delayed renewal process β satisfying (5.6) is positive-recurrent, that is, (5.7) holds.*

Proof. Because of the independence between U_k and W_k, we can write

$$\begin{aligned}\mathrm{E}[\Delta_W(k)] &= \mathrm{E}\Big[\lceil W_k + U_k \rceil_{\mathcal{A}} - W_k\Big] = \sum_{y\in\mathcal{A}} \mathrm{E}\Big[\lceil W_k + U_k \rceil_{\mathcal{A}} - W_k; W_k = y\Big] \\ &= \sum_{y\in\mathcal{A}} \mathrm{E}[\lceil y + U \rceil_{\mathcal{A}} - y]\mathrm{Pr}[W_k = y]\ ,\ k \geq 1\,.\end{aligned}$$

Here we used the fact that $W_k \in \mathcal{A}$. If $U \leq -y$ then $\mathrm{E}[\lceil y + U \rceil_{\mathcal{A}} - y] = -y$. Thus, we have

$$\begin{aligned}&\mathrm{E}[\lceil y + U \rceil_{\mathcal{A}} - y] \\ &= -y\mathrm{Pr}[U \leq -y] + \int_{z\geq -y} (\lceil y + z \rceil_{\mathcal{A}} - y) dF_U(z) \\ &= -y\mathrm{Pr}[U \leq -y] + \int_{z\geq -y} (\lceil y + z \rceil_{\mathcal{A}} - (y+z) + z) dF_U(z) \\ &= -y\mathrm{Pr}[U \leq -y] + \int_{z\geq -y} \Delta(y+z) dF_U(z) \\ &\quad + \int_{z\geq -y} z dF_U(z)\,. \end{aligned} \tag{5.13}$$

We now provide an upper bound for the former expression. We consider the three terms in (5.13) separately. As for the first term, since $\mathrm{E}[U] > -\infty$, we find

$$0 > -y\mathrm{Pr}[U \leq -y] \geq \int_{-\infty}^{-y} x dF_U(x) \uparrow 0 \ \text{ as } \ y \to \infty\,, \tag{5.14}$$

while for the third term we obtain

$$\int_{z\geq -y} z dF_U(z) \downarrow \mathrm{E}[U] \ \text{ as } \ y \to \infty\,, \tag{5.15}$$

(since $y \geq 0$). Rewriting the second term of the right-hand side of equation (5.13) yields the following form:

$$\begin{aligned}\int_{-y}^{\infty} \Delta(y+z) dF_U(z) &= \int_{-y}^{-y/2} \Delta(y+z) dF_U(z) \\ &+ \int_{-y/2}^{\infty} \Delta(y+z) dF_U(z) \equiv I_1 + I_2\,,\end{aligned}$$

where we introduced I_1 and I_2 for notational convenience. Considering the latter two terms subsequently, we find for I_2 that

$$I_2 \leq \sup_{x\geq y/2} \Delta(x)\mathrm{Pr}[U > -y/2]\,,$$

while for I_1, it is clear that

$$I_1 \leq \sup_{0\leq x\leq y/2} \Delta(x)\Pr[-y < U \leq -y/2] \leq \Delta^* \Pr[U \leq -y/2]\,.$$

Fix now an arbitrary $\varepsilon > 0$ and choose y_0 such that following three equations hold for $y \geq y_0$,

$$\Pr[U \leq -y/2] \leq \frac{\varepsilon}{4\Delta^*}\,, \tag{5.16}$$

$$\int_{z\geq -y} z dF_U(z) \leq \mathrm{E}[U] + \frac{\varepsilon}{4}\,, \tag{5.17}$$

$$\sup_{x\geq y/2} \Delta(x) \leq \Delta_0 + \frac{\varepsilon}{4}\,. \tag{5.18}$$

Such choice of y_0 is possible: for (5.16), due to (5.14); for (5.17), due to (5.15). For (5.18), we apply (5.4) and (5.5). More precisely, one can find a value n_0, sufficiently large, such that

$$\sup_{n\geq n_0} g_n \leq \Delta_0 + \frac{\varepsilon}{4}\,.$$

Now we take into account that $\lceil x \rceil_{\mathcal{A}} \geq a_{n_0}$ for $x \geq a_{n_0}$. Then

$$\Delta(x) \leq \sup_{n\geq n_0} g_n \leq \Delta_0 + \frac{\varepsilon}{4},\ \ x \geq a_{n_0}\,, \tag{5.19}$$

and (5.18) follows if $y_0/2 \geq a_{n_0}$. Considering equations (5.15) to (5.19) and taking into account the negativity of (5.14), we find the following upper bound for (5.13),

$$\mathrm{E}[\lceil y+U \rceil_{\mathcal{A}} - y] \leq \frac{\varepsilon}{4} + \Delta_0 + \frac{\varepsilon}{4} + \mathrm{E}[U] + \frac{\varepsilon}{4} < \Delta_0 + \mathrm{E}[U] + \varepsilon,\ y \geq y_0\,. \tag{5.20}$$

Exploiting the fact that $\varepsilon > 0$ was arbitrary up to now, we choose it to be

$$\Delta_0 + \mathrm{E}[U] = -2\varepsilon\,.$$

Notice that the left-hand side of the former expression is indeed negative in view of (5.11). It then follows from (5.20) that

$$\mathrm{E}[\Delta_W(k) \,|\, W_k = y] = \mathrm{E}[\lceil y+U \rceil_{\mathcal{A}} - y] \leq -\varepsilon\ ,\ y \geq y_0\,. \tag{5.21}$$

In other words, we obtain a negative drift of the waiting-time process W outside the compact set $[0, y_0]$. Further, let C denote the following upper bound for the overshoot,

$$\begin{aligned} \mathrm{E}[\lceil y+U \rceil_{\mathcal{A}} - y] &\leq \mathrm{E}[\lceil y+B \rceil_{\mathcal{A}} - y] = \mathrm{E}[\lceil y+B \rceil_{\mathcal{A}} - y - B] + \mathrm{E}[B] \\ &\leq \Delta^* + \mathrm{E}[B] \leq \max(\varepsilon,\, \Delta^* + \mathrm{E}[B]) \equiv C < \infty\,. \end{aligned}$$

Here ε is introduced in the definition of C to ensure $\varepsilon/C \leq 1$ in (5.25). In view of the former expressions, we also find

$$\begin{aligned} &\mathrm{E}[\Delta_W(k) \,|\, W_k \leq y_0] \\ &\quad = \sum_{y \in \mathcal{A}:\ y \leq y_0} \frac{\mathrm{E}[\Delta_W(k) \,|\, W_k = y]\mathrm{Pr}[W_k = y]}{\mathrm{Pr}[W_k \leq y_0]} \leq C\,. \end{aligned} \tag{5.22}$$

For any k, write now

$$\begin{aligned} \mathrm{E}[\Delta_W(k)] \;&=\; \mathrm{E}[\Delta_W(k) \,|\, W_k \leq y_0]\mathrm{Pr}[W_k \leq y_0] \\ &\quad + \sum_{y \in \mathcal{A}:\ y > y_0} \mathrm{E}[\Delta_W(k) \,|\, W_k = y]\mathrm{Pr}[W_k = y]\,. \end{aligned}$$

such that

$$\mathrm{E}[\Delta_W(k)] \leq C\mathrm{Pr}[W_k \leq y_0] - \varepsilon\mathrm{Pr}[W_k > y_0]\,, \tag{5.23}$$

in view of the inequalities (5.21) and (5.22).

Now we prove that $W_k \not\to \infty$ in probability as $k \to \infty$. We use a proof by contradiction. Hence, assume that we have,

$$W_k \to \infty \text{ in probability as } \; k \to \infty\,. \tag{5.24}$$

Therefore, one can find a value k_0 such that

$$\mathrm{Pr}[W_k \leq y_0] \leq \frac{\varepsilon}{4C}\ ,\ k \geq k_0\,,$$

which allows us to rewrite (5.23) as

$$\mathrm{E}[\Delta_W(k)] \leq \frac{\varepsilon}{4} - \varepsilon\Big(1 - \frac{\varepsilon}{4C}\Big) \leq -\frac{\varepsilon}{2},\ k \geq k_0\,. \tag{5.25}$$

Thus, the process W has negative drift under assumption (5.24).

Now we show that $\mathrm{E}[W_{k_0}] < \infty$. Clearly, we have $\lceil x \rceil_{\mathcal{A}} \leq \lceil y \rceil_{\mathcal{A}}$ for $x \leq y$. Further, recall that $W_1 = 0$. Therefore, we find

$$\begin{aligned} \mathrm{E}[W_2] \;&=\; \mathrm{E}[\lceil U_1 \rceil_{\mathcal{A}}] \leq \mathrm{E}[\lceil B_1 \rceil_{\mathcal{A}}] = \int_0^\infty \lceil z \rceil_{\mathcal{A}} dF_B(z) \\ &=\; \int_0^\infty \Delta(z) dF_B(z) + \int_0^\infty z dF_B(z) \\ &\leq\; \Delta^* + \mathrm{E}[B] \equiv C^*\,, \end{aligned} \tag{5.26}$$

where $F_B(x) = \mathrm{Pr}[B \leq x]$. Expanding this expression iteratively, we obtain

$$\begin{aligned} \mathrm{E}[W_3] \;&\leq\; \mathrm{E}[\lceil W_2 + B_2 \rceil_{\mathcal{A}}] \\ &=\; \int_{(x,y)\in[0,+\infty)^2} (\lceil x + y \rceil_{\mathcal{A}} - (x+y) + (x+y)) dF_{W_2}(x) dF_B(y) \\ &=\; \int_{(x,y)\in[0,+\infty)^2} \Delta(x+y) dF_{W_2}(x) dF_B(y) + \mathrm{E}[W_2] + \mathrm{E}[B] \\ &\leq\; \Delta^* + C^* + \mathrm{E}[B] = 2C^*\,, \end{aligned}$$

where $F_{W_2} = \Pr[W_2 \leq x]$, and $[0,+\infty)^2$ denotes the two-dimensional domain $[0,+\infty) \times [0,+\infty)$. By induction on the index k, we obtain

$$\mathrm{E}[W_k] \leq (k-1)C^*,\ k \geq 1\,,$$

and thus,

$$\mathrm{E}[W_{k_0}] < k_0 C^* \equiv D_0 < \infty\,. \tag{5.27}$$

We finally conclude from (5.25) that for any $k \geq k_0$ we have,

$$\mathrm{E}[W_k] < D_0\,.$$

This then contradicts the assumption (5.24).

Since (5.24) leads to a contradiction, there must exist a non-random (sub)sequence $z_k \to \infty$, a $\delta' > 0$ and a finite constant R such that

$$\inf_k \Pr[W_{z_k} \leq R] \geq \delta'\,. \tag{5.28}$$

First assume that the r.v. T is unbounded. In this case we have, $\Pr[T > x] > 0$ for any $x \geq 0$. Further, in view of (5.12), the event $\{W_{z_k} \leq R,\, T_{z_k} > R + B_{z_k} + \Delta^*\}$ implies that burst $z_k + 1$ meets an empty system. For z_k belonging to the sequence $\{z_k\}$ which satisfies (5.28)), a regeneration occurs (for burst $z_k + 1$) with probability,

$$\begin{aligned}
\Pr[\beta(z_k) = 1] &\geq \Pr[W_{z_k} \leq R\,,\, T_{z_k} > R + \Delta^* + B_{z_k}] \\
&\geq \Pr[W_{z_k} \leq R]\Pr[T > R + a + \Delta^*,\, B \leq a] \\
&\geq \delta' \Pr[T > R + a + \Delta^*]\Pr[B \leq a] > 0\,.
\end{aligned}$$

Here, independence between W_{z_k}, B_{z_k} and T_{z_k} is used, and the constant $a > 0$ is chosen such that $\Pr[B \leq a] > 0$.

Assume now that T is bounded, then it is possible that $\Pr[T > R+a+\Delta^*] = 0$. For each fixed $x \geq 0$, define $N(x) = \min\{k : a_k \geq x\}$. By (5.10), we find,

$$N(R) \leq \left\lceil \frac{R + \Delta^*}{\delta^*} \right\rceil < \infty\,.$$

Moreover, on the event $\mathcal{E}(z_k) = \{W_{z_k} \leq R\}$, the following inequality holds,

$$W_{z_k} \leq \lceil W_{z_k} \rceil_{\mathcal{A}} \leq \lceil R \rceil_{\mathcal{A}} \leq a_{N(R)}\,.$$

Let us now introduce the following events:

$$\mathcal{D}(z_k+i) = \{T_{z_k+i} > \Delta^* + B_{z_k+i}\} = \{B_{z_k+i} - T_{z_k+i} < -\Delta^*\},\ i \geq 0\,. \tag{5.29}$$

Then, on the event $\mathcal{E}(z_k) \cap \mathcal{D}(z_k)$, we have that

$$\begin{aligned}
W_{z_k+1} &= \lceil W_{z_k} + B_{z_k} - T_{z_k} \rceil_{\mathcal{A}} \leq \lceil a_{N(R)} - \Delta^* \rceil_{\mathcal{A}} \\
&\leq \lceil a_{N(R)} - g_{N(R)} \rceil_{\mathcal{A}} = a_{N(R)-1}\,.
\end{aligned}$$

and $\Pr[\mathcal{E}(z_k) \cap \mathcal{D}(z_k)] \geq \delta'\delta_1$ in view of (5.12), (5.28) and (5.29). Analogously, on the event $\mathcal{E}(z_k) \cap \mathcal{D}(z_k) \cap \mathcal{D}(z_k + 1)$, we obtain,

$$W_{z_k+2} \leq a_{[N(R)-2]^+} \,,$$

and by the independence of the components we have $\Pr[\mathcal{E}(z_k) \cap \mathcal{D}(z_k) \cap \mathcal{D}(z_k + 1)] \geq \delta'\delta_1^2$. It is now easy to see that with a probability

$$\Pr\Big[\mathcal{E}(z_k) \cap \bigcap_{i=0}^{N(R)} \mathcal{D}(z_k + i)\Big] \geq \delta'\delta_1^{N(R)} > 0 \,,$$

there is a burst that enters an empty system between (and including) bursts z_k and $z_k + N(R)$; that is in the interval $[z_k,\, z_k + N(R)]$. As such, a regeneration occurs within this interval with non-zero probability, since

$$\Pr[\beta(z_k) \leq N(R)] \geq \delta'\delta_1^{N(R)} \,. \tag{5.30}$$

Invoking that the arrival z_k is arbitrary, (5.30) holds for all z_k satisfying (5.28). Because the sequence $\{z_k\}$ is non-random and the length of the interval is a constant $N(R)$, we find that

$$\beta(n) \not\to \infty \text{ as } n \to \infty \text{ in probability} \,,$$

and it follows from (5.9) that $\alpha < \infty$, which was to be proved. □

As outlined at the beginning of this section (Sect. 5.1.4), this proves that the FDL buffer is stable under zero initial conditions, since it implies that the zero waiting time state is positive-recurrent. In the next section, we will extend the initial conditions to show that conditions (5.10)-(5.12) guarantee stability for any initial state, so allowing for a complete proof.

Finally, to contrast this result with those obtained in Chapter 2, we consider the degenerate case, where $g_n \equiv D$ for all n and for granularity $D \in \mathbb{R}^+$. In that case, we have $\Delta^* = \Delta_0 = D$, which yields the following result.

Corollary 5.1.1. *The waiting-time process of the initially empty degenerate optical buffer is positive-recurrent regenerative if the negative drift condition*

$$\mathrm{E}[U] + D < 0 \,, \tag{5.31}$$

holds.

Proof. Indeed, in this case $0 < \mathrm{E}[-U - D] = \mathrm{E}[T - B - D]$ and hence the regeneration condition $\Pr[T > D + B] > 0$ holds automatically. Also, note that a degenerate buffer setting is indeed well-defined, implying that also condition (5.10) is fulfilled. □

Notice that condition (5.31) is not identical to the one obtained in Chapter 2 because it is a sufficient but not necessary condition. This is argued further below in Sect. 5.1.5.

5.1.4.3 Extension to Non-Zero Initial Conditions

In this section, we show that conditions (5.10)-(5.12) not only suffice to guarantee stability under zero initial conditions (as proved in the previous section), but, if extended with a minor condition, provide sufficient conditions regardless of the initial state. At the end of this section, we further show that the same set of conditions suffices for the existence of a stationary regime.

In terms of Markov chain theory, the proof we obtained for zero initial conditions is valid immediately for any initial condition if we assume that the Markov chain of the waiting times is irreducible. In that case, any state $W_k = a_j$ can be reached from the state $W_1 = a_0$ with finite probability, and all Markov states belong to the single communicating class $\mathcal{A}$. From a practical point of view, assuming irreducibility is not a very limiting assumption, and as such, it was not much of a restriction when we assumed it earlier in this work in the case of a non-degenerate finite-sized GI/G/1 buffer, in Sect. 3.2.3.2. However, we remark that in principle, it is possible to establish stability also for Markov chains that are reducible. In the following, we will however limit the analysis to aperiodic irreducible Markov chains, and provide a sufficient (but not necessary) condition for irreducibility. Also, we will show that aperiodicity is already implied by the negative drift condition (5.31).

We now consider non-zero initial conditions, assuming thus that the waiting time (state) of the first burst is non-zero, $W_1 > 0$. We show that any initial state $W_1 = a_k$ in the Markov chain can be used that can be reached from the zero waiting time state state $W_1 = 0$ with a positive probability, in a finite number of Markov chain transitions. In case of a reducible Markov chain, all Markov states belonging to the same communicating class as the zero waiting time state can be reached; in case of an irreducible Markov chain, all states can be reached. Let $\Pr_0[\cdot]$ denote the probability operator for the zero initial state. Notice that the expectation operator $\mathrm{E}[\cdot]$ and probability operator $\Pr[\cdot]$ in Sect. 5.1.4.2 in fact referred to the zero initial state too.

Under the assumptions (5.10)-(5.12) of Theorem 5.1.1, we have $\Pr_0[\beta_1 < \infty] = 1$. Assume that a state $W_n = a_k$, for some k, is reached (at instant n) with a positive probability $\pi_k > 0$. Now, as a proof by contradiction, assume that the unfinished regeneration time since instant n is infinite with probability $\pi_\infty > 0$. Then we obtain $\Pr_0[\beta_1 = \infty] \geq \pi_k \pi_\infty > 0$ which contradicts the finiteness of β_1 w.p.1 for the zero initial state. Hence it is proved by contradiction that, if a state a_k can be reached from state $a_0 = 0$ with a positive probability then, under initial state $W_1 = a_k$, the first regeneration period β_1 is finite w.p.1. As such, the stability conditions (5.10)-(5.12) of Theorem 5.1.1 are also sufficient for stability if the state a_k (that can be reached from the zero state with a positive probability) is taken as initial state.

At this point, we propose sufficient assumptions to obtain an irreducible Markov

chain. In that case, all states a_k are reachable from state a_0, so allowing to extend the class of initial states to all Markov states at once. Let T_{min} denote the essential lower bound of the inter-arrival time distribution, defined like in Chapter 3 (Eq. (3.25)). Since $\mathrm{E}[T] < \infty$, it follows that $T_{min} < \infty$. For the negative-exponential distribution, that is often assumed throughout this work, $T_{min} = 0$. By definition, for any $\delta_0 > 0$ we have,

$$\Pr[T_{min} \leq T < T_{min} + \delta_0] \equiv \varepsilon_0 > 0\,.$$

Finally, assume that the pdf of the burst size distribution, $b(x) \equiv F'_B(x)$, exists in a (right) neighborhood of T_{min} and let $F_T(x) = \Pr[T \leq x]$ denote the cdf of T, like in Sect. 3.1.4.

Theorem 5.1.2. *If the assumptions of Theorem 5.1.1 hold and moreover, if the condition*

$$\inf_{x \in [T_{min},\, T_{min}+\Delta^*+\delta_0)} b(x) \equiv b_0 > 0\,, \tag{5.32}$$

is satisfied for some $\delta_0 > 0$, then the regenerative waiting-time process W is positive-recurrent under any initial state $W_1 = a_k \in \mathcal{A}$.

Proof. As a starting point, we assume that $W_1 = a_0 = 0$ like in the previous section, and we recall that the notation $\Pr_0[\cdot]$ relates to this zero initial condition. Also, we note that

$$F_U(x) = \int_{y \in [0,\infty)} \Pr[B \leq x + y] dF_T(y),\ x \in \mathbb{R}\,.$$

Considering now the Markov chain evolution over time, following inequality can be established,

$$\begin{aligned}
\Pr_0[W_2 = a_1] &= \Pr[0 < U_1 \leq a_1] = F_U(a_1) - F_U(0) \\
&= \int_{y \in [0,\infty)} \Pr[y < B \leq a_1 + y] dF_T(y) \\
&\geq \int_{y \in [T_{min},\, T_{min}+\delta_0)} \Pr[y < B \leq a_1 + y] dF_T(y) \\
&= \int_{y \in [T_{min},\, T_{min}+\delta_0)} \int_{z \in [y,\, y+a_1)} b(z) dz dF_T(y) \\
&\geq \inf_{z \in [T_{min},\, T_{min}+\delta_0+a_1)} b(z) \Pr[T_{min} \leq T < T_{min} + \delta_0] \cdot a_1 \\
&\geq g_0 b_0 \varepsilon_0\,.
\end{aligned} \tag{5.33}$$

In the last step, we used the fact that $a_1 = g_0 \leq \Delta^*$, and also the fact that

$$b_0 = \inf_{z \in [T_{min},\, T_{min}+\Delta^*+\delta_0)} b(z) \leq \inf_{z \in [T_{min},\, T_{min}+a_1+\delta_0)} b(z)\,.$$

Similarly,

$$\begin{aligned} \Pr_0[W_3 = a_2] &\geq \Pr[W_3 = a_2 \,|\, W_2 = a_1]\Pr_0[W_2 = a_1] \\ &= \Pr[\lceil a_1 + U \rceil_{\mathcal{A}} = a_2]\Pr_0[W_2 = a_1] \\ &\geq g_0 g_1 (b_0\varepsilon_0)^2 \geq (\delta^* b_0 \varepsilon_0)^2 \,. \end{aligned} \tag{5.34}$$

Continuing in a similar way, we obtain

$$\begin{aligned} \Pr_0[W_{n+1} = a_n] &\geq \Pr_0[W_n = a_{n-1}]\Pr[W_{n+1} = a_n \,|\, W_n = a_{n-1}] \\ &\geq (\delta^* b_0 \varepsilon_0)^n > 0 \,,\ n \geq 1 \,. \end{aligned}$$

Thus, each state a_n can be reached from the zero initial state $W_1 = 0$ (in n arrivals) with positive probability. As proved by contradiction in the above, this implies that each state a_n can be assumed as initial state for stability. Hence, the statement of Theorem 5.1.1 holds for any initial state $W_1 = a_k$. □

As such, the proposed assumption (5.32) is a sufficient condition to extend the class of possible initial states to $\mathcal{A}$. Note that the condition is not necessary, and that it is equally possible to treat alternate cases. Rather, (5.32) provides a sufficient condition for irreducibility of the Markov chain. On the other hand, the Markov chain is aperiodic due to the regeneration assumption. More precisely, the regeneration-time process is aperiodic in general, and this independent of condition (5.32), as

$$\Pr[\beta_2 - \beta_1 = 1] = \Pr[W_{k+1} = 0 \,|\, W_k = 0] = \Pr[U \leq 0] = \Pr[B \leq T] > 0 \,, \tag{5.35}$$

and hence poses no additional restrictions, since it follows from the negative drift condition (5.11) that $\mathrm{E}[U] = \mathrm{E}[B - T] < 0$ and thus $\Pr[B \leq T] > 0$.

Therefore, given that the Markov chain of waiting times is irreducible, aperiodic and composed of positive-recurrent states, we can conclude that the process W_k weakly converges to W_∞, $W_k \to W_\infty$, with associated stationary distribution $\Pr_{W_\infty}[\cdot] = \Pr[W_\infty \in \cdot]$.

Theorem 5.1.3. *Under assumptions (5.10)-(5.12) and (5.32), the regenerative waiting-time process W is positive-recurrent with respect to the renewal process of regenerations (5.6) and has a stationary distribution $Pr_{W_\infty}[\cdot]$ under any initial state $W_1 = a_k \in \mathcal{A}$.*

Let $\Pr_k[\cdot] = \Pr[\cdot \,|\, W_1 = a_k]$. Hence, under the assumptions of Theorem 5.1.3, $\Pr_k[\beta_1 < \infty] = 1$, $k \geq 0$.

5.1.5 Sufficient Condition

Juxtaposing (5.31) with the characterization of stability of GI/G/1 degenerate buffers in Chapter 2, it is clear that (5.31), even though obtained for the same setting,

is not as tight a bound as the one provided by (2.8), or, equivalently, by $\rho_{eq} < 1$, characterized by (2.9).

More precisely, if one assumes (5.31), one will find that stability is guaranteed only for ρ smaller than some value ρ_1, while (2.9) guarantees stability for ρ smaller than some ρ_2, with in general $\rho_1 \leq \rho_2$. In other words, the conditions established here are sufficient but not necessary for stability. As such, the approach of this chapter is not able to guarantee stability for $\rho \in [\rho_1, \rho_2)$, while the method of Chapter 2 can, but has the drawback that it is only applicable to degenerate buffer settings.

Seemingly a drawback to the current approach, it should however be understood as a natural consequence of the fact that in this chapter, only the first moments of the inter-arrival time and burst size distribution were taken into account. Also for the FDL structure, we only capture the FDL lengths by means of Δ_0, rather than involving all FDL lengths. This is opposed to the situation for degenerate buffers, where condition (2.8) took into account not only the entire distribution of the inter-arrival times and burst sizes, but also the lengths of all FDLs, possible with only the granularity D. Taking into account all the FDL line lengths of an arbitrary infinite-sized FDL set is infeasible, and as such, coming up with *necessary* stability conditions for general non-degenerate line lengths is not practicable. Instead, the approach given here captures the non-degenerate structure in one parameter, Δ_0, which results in very simple *sufficient* stability conditions, valid also for non-degenerate buffer structures, that fall beyond the reach of the modeling in Chapter 2. Further, note that, for some more "pathological cases" of the inter-arrival time and burst size distributions, the bound provided by (5.31) is indeed a necessary condition, like (2.9). The simplest instance of such a "pathological case" would be a degenerate system with deterministic inter-arrival times (length T_p), and deterministic burst sizes (length B_p) in CT. One can verify that, for such system, stability is characterized by $\mathrm{E}[B-T]+D < 0$ (or, $T_p > B_p + D$), making condition (5.31) an exact characterization of stability in that case.

5.2 Multi-Wavelength System

The purpose of this section is to extend the results of Theorem 5.1.1 to a multi-wavelength optical buffer system, for general line lengths. In Sect. 5.2.1, the stochastic model is generalized to multiple wavelengths. A notation is introduced that takes into account the assumed wavelength assignment algorithm, namely Join-the-Shortest-Queue (JSQ), previously studied in Sect. 2.5. An example illustrates the system's evolution over time, and highlights the differences with the single-wavelength case.

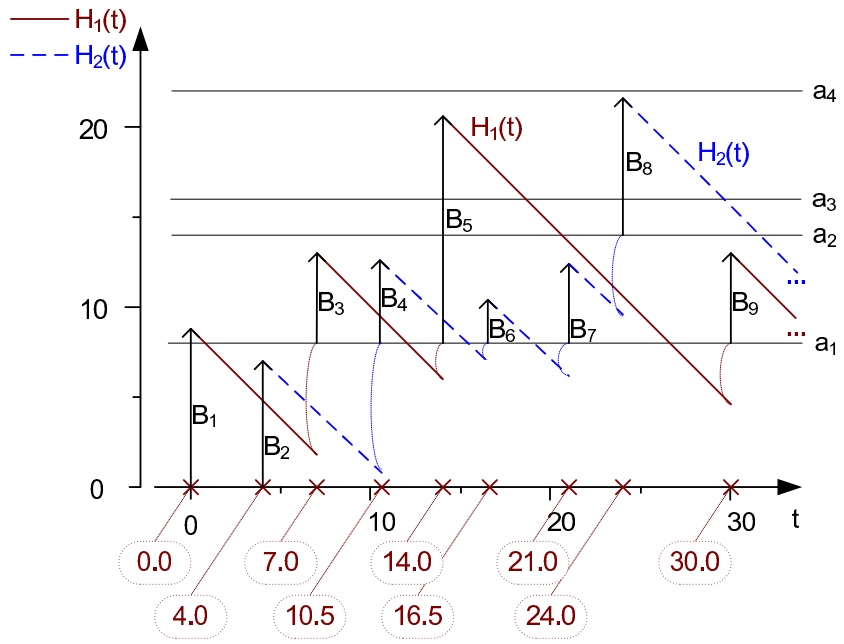


Figure 5.2: The evolution over time of the virtual scheduling-horizon process $\hat{H}$ in case of two queues, queue 1 ($\hat{H}_1$) and queue 2 ($\hat{H}_2$). The crosses on the (horizontal) time-axis represent arrivals $k = 1 \ldots 9$, and the FDL lengths a_i are the ones in (5.37). Upon arrival, burst k joins the shortest queue (JSQ) (lowest of both lines), generating a void (curbed line) and adding B_k to the virtual scheduling horizon. All values are given in Table 5.2.

5.2.1 Traffic Setting and Buffer Setting

Firstly, the traffic assumptions remain the same as in the previous section: iid burst sizes B_k, iid inter-arrival times T_k, and a general FDL set $\mathcal{A}$. However, now, we treat multiple wavelengths, amounting to a non-degenerate GI/G/c buffer model. Further, as introduced in Sect. 2.5, the multi-wavelength system is governed by a wavelength assignment algorithm, exercised by buffer control to determine which wavelength is chosen for the transmission of burst k. More precisely, in this section we consider a JSQ wavelength assignment algorithm, exactly in the same sense as it was introduced in Sect. 2.5.2. However, in Sect. 2.5, the mathematical description of the involved stochastic processes was stripped down to the bare minimum, needed to allow for performance evaluation. For the purpose of stability analysis, here, we will provide a complete mathematical description of the processes involved in multi-wavelength optical buffering. While these processes were defined

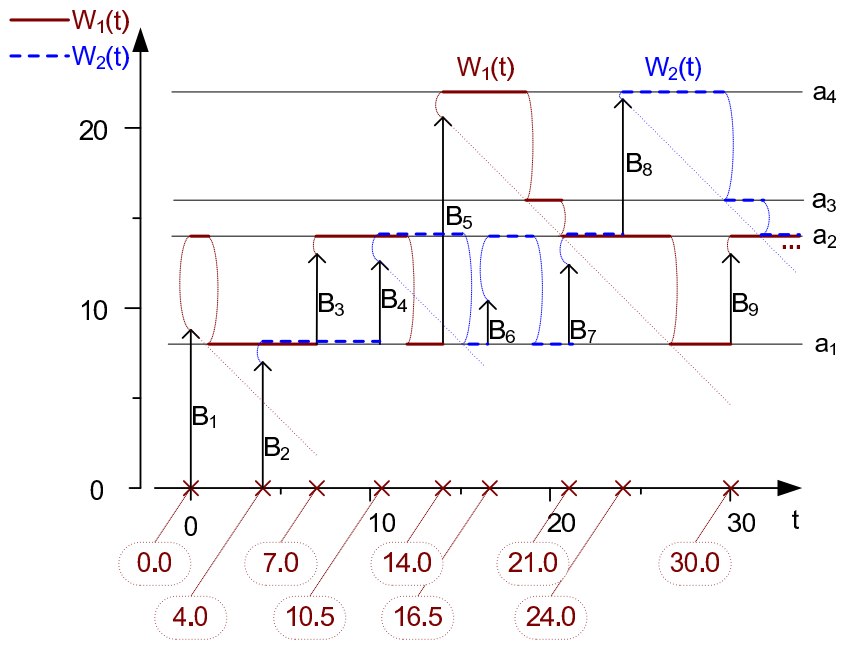


Figure 5.3: The same system setting and traffic pattern as in Fig. 5.2, now with the virtual waiting-time process on display.

k	1	2	3	4	5	6	7	8	9
B_k	8.8	7.0	5.0	4.5	12.5	2.4	4.4	7.6	4.8
t_k	0.0	4.0	7.0	10.5	14.0	16.5	21.0	24.0	30.0

Table 5.2: Values of the variables displayed in Fig. 5.2 and 5.3.

only implicitly in the analysis of Sect. 2.5, note that these are identical to the ones discussed here.

5.2.2 Evolution of the Scheduling Horizon

Independent of the wavelength assignment algorithm (see below), in the c-wavelength system, there are c different wavelengths to queue for. Like in the single-wavelength case, we study the involved processes not only upon arrival instants, but also on arbitrary instants. Again, knowledge of the latter is not strictly necessary for the stability analysis, and that it is included here only to provide the reader with additional insight on the difference between the waiting-time and the

scheduling-horizon process.

As usual, we number the bursts in the order of arrival by means of an index k. Upon arrival of burst k, we can associate with each queue j, $j = 1 \dots c$, a scheduling horizon $H_{j,k}$ as seen by arrival k, that captures the earliest time at which all previous bursts on wavelength j have left the system. Similar as in the single-wavelength system (see Sect. 5.1.3), each of these c processes $\bar{H}_j = \{H_{j,k}, k \geq 1\}$ can be derived from a left-continuous process $\hat{H}_j = \{H_j(t),\, t \geq 0\}$, that describes the *virtual scheduling-horizon process* of each physical queue independently, through the relation $H_{j,k} = H_j(t_k)$, $k \geq 1$. Further, for each wavelength j the *waiting time as seen upon arrival* of burst k is described by $W_{j,k}$, and as a process $\bar{W}_j = \{W_{j,k}, k \geq 1\}$ can be derived from the left-continuous process $\hat{W}_j = \{W_j(t),\, t \geq 0\}$ that describes the *virtual waiting-time process* for each queue i, $j = 1 \dots c$, through the relation $W_{j,k} = W_j(t_k)$, $k \geq 1$. The relation between the scheduling horizon (the delay that is needed) and the waiting time (the delay that is assignable) is given by

$$W_{j,k} = \lceil H_{j,k} \rceil_{\mathcal{A}} \,,\ j = 1 \dots c \,.$$

Note that the discrete-time processes $\bar{H}_j$ and $\bar{W}_j$ are now *virtual* processes (just like $\hat{H}_j$ and $\hat{W}_j$), in the sense that they do not coincide with the *actual* scheduling horizon and waiting time experienced by arrival k, that is determined by the wavelength assignment algorithm.

Up to this point, the stochastic processes are formulated independently of the chosen wavelength assignment algorithm. However, since we perform the stability analysis solely for JSQ, it is possible to exploit the symmetry of the problem in that case, with a specific notation. Instead of identifying the c different queues one by one independently, we sort the scheduling-horizon components $H_{i,k}$ in increasing order, and denote them as $H_k^{(i)}$, $i = 1 \dots c$. As such, the c-dimensional scheduling-horizon process $H_k = \{H_k^{(i)}\}$ consists of c components

$$H_k = (H_k^{(1)}, \dots, H_k^{(c)}),\ k \geq 1 \,,$$

with

$$H_k^{(1)} \leq \dots \leq H_k^{(c)} \,.$$

When the kth burst arrives, the wavelength assignment algorithm JSQ will assign the burst to some queue j, that is not necessarily the queue with $j = 1$, but is indeed the shortest, with thus always $i = 1$, and with $H_k^{(1)}$ as scheduling horizon. The corresponding waiting time reads

$$W_k^{(1)} = \lceil H_k^{(1)} \rceil_{\mathcal{A}} \,,$$

whereas the scheduling horizon of the other wavelengths $i = 2 \dots c$ (as long as positive) will decrease by an amount T_k. As such, the overshoot operator $\Delta(x)$

of (5.5) applied to $H_k^{(1)}$ indeed corresponds to the void V_k created by the arrival of burst k, as defined earlier in this work, in (5.5). Therefore, in the following we denote

$$V_k = \Delta(H_k^{(1)}) \ , \ k \geq 1 \, .$$

Note that the evolution of the other components of H ($H_k^{(i)}$, $2 \leq i \leq c$) cannot be captured with the overshoot operator, since their evolution does not relate to the set of FDLs. As such, the waiting time alone does not suffice to capture the system state, and one has to rely on the scheduling-horizon process. This is opposed to the single-wavelength case, where the system description could be done either with the waiting time or the scheduling horizon.

As such, the evolution of the system is to be traced in terms of the scheduling horizon, and is formed by a Kiefer-Wolfowitz-type recursion [117], namely

$$H_{k+1} = \mathrm{R}\Big([\lceil H_k^{(1)} \rceil_{\mathcal{A}} + U_k]^+, \, [H_k^{(2)} - T_k]^+, \, \ldots, [H_k^{(c)} - T_k]^+ \Big) \, , \qquad (5.36)$$

where the operator $\mathrm{R}(\cdot)$ ensures that the components are put in increasing order. Notice how the operator $\lceil \cdot \rceil_{\mathcal{A}}$ occurs in exactly one component.

We give an example of the evolution of the scheduling-horizon process for an initially empty two-wavelength system, that is $H_1 = (0, 0)$. Then

$$H_2 = \mathrm{R}(\lceil 0 \rceil_{\mathcal{A}} + [U_1]^+ \, , \, 0) = (H_2^{(1)}, \, H_2^{(2)}) \, ,$$

(note that $H_2^{(1)} = 0$) and

$$H_3 = \mathrm{R}([\lceil H_2^{(1)} \rceil_{\mathcal{A}} + U_2]^+, \, [H_2^{(2)} - T_2]^+) \, ,$$

and so on. In the following, given that $\{B_k\}$, $\{T_k\}$ and $\{U_k\}$ are sequences of iid rv's, we can omit the index k unless it is relevant to the context.

Finally, as an instructive example, a typical evolution of the involved virtual processes is considered in Fig. 5.2 (scheduling horizon) and 5.3 (waiting time), where

$$\mathcal{A} = \{a_0, a_1, a_2, a_3, a_4, \ldots\} = \{0, 8, 14, 16, 22, \ldots\} \, , \qquad (5.37)$$

identical to the FDL set considered earlier in (5.3), with only the first five fiber lengths relevant to the example. Specifics of the used trace, including the actually assigned scheduling horizon and waiting time, are given in Table 5.2.

While the above provides insight in the interaction between the waiting-time and the scheduling-horizon process, we now formulate the analysis entirely in terms of the process H_k, $k \geq 1$, that is the only necessary input to the multi-wavelength stability analysis.

5.2.3 Regenerative Stability Analysis

Again adopting a regenerative approach, we will now focus on the positive recurrence of the zero scheduling horizon state. Notice that the Markov chain of the scheduling horizon now has a continuous state space, as opposed to the discrete state space of the waiting time evolution, as it was studied for the single-wavelength case. This however forms no stumbling block to characterize stability. As such, the renewal process of the scheduling horizon has regenerations whenever a Markov chain transition to the state with zero scheduling horizon is made. Regeneration times $\{\beta_n\}$ for this c-dimensional process $H = \{H_k\}$ are constructed as in the single-wavelength case (see (5.6)), namely, we put $\beta_0 = 0$ and

$$\beta_{n+1} = \inf\Big(k > \beta_n : H_k = \bar{0}\Big),\ n \geq 0\,, \tag{5.38}$$

with $\bar{0}$ corresponding to c components, or $(0, \ldots, 0)$. Also the unfinished renewal/regeneration time $\beta(n)$ at instant n is defined as in (5.8).

Similar as in the single-wavelength case, the main stability result will be to prove that, under zero initial conditions, a set of three stability conditions guarantees stability. Of these three conditions, two are exactly the same as in the single-wavelength case: the assumption that the FDL buffer is properly-defined (5.10), and the regeneration assumption (5.12). The only condition that alters is the negative drift condition (5.11), that is adapted to the case of multiple wavelengths in Sect. 5.2.3.1. A second step consists in establishing the tightness of the drift of the different components of the scheduling-horizon process. This step of the proof has no counterpart in the single-wavelength case, and constitutes an additional challenge, typical to the multi-wavelength case. It is treated in Sect. 5.2.3.2, and is applied subsequently in Sect. 5.2.3.3, for the main stability result, showing that the three conditions guarantee stability under zero initial conditions. Although the extension for non-zero initial conditions is not included in this work, a remark is also devoted to it at the end of Sect. 5.2.3.3.

5.2.3.1 Negative Drift

Before moving to the main stability result, we present a first lemma to describe the negative drift of the (total) scheduling-horizon process H_k, and a second lemma to characterize the tightness of individual components of H_k. Further, recall from (5.4) that $\Delta_0 \equiv \limsup_{n\to\infty} g_n$.

Lemma 5.2.1. *Denote*

$$\Delta_H(k) = \sum_{i=1}^{c} (H_{k+1}^{(i)} - H_k^{(i)})\,,$$

and estimate the mean increment $\mathrm{E}[\Delta_H(k)]$ *by assuming* $H_k = y = (y_1, \ldots, y_c)$. *The drift of the scheduling-horizon process* H_k *adheres to*

$$\limsup_{y\to\infty} \mathrm{E}[\Delta_H(k) \mid H_k = y] \le \Delta_0 + \mathrm{E}[B] - c\mathrm{E}[T]\,, \tag{5.39}$$

where the limit $y \to \infty$ *is to be applied to all components, that is,* $y_i \to \infty$.

Proof. As mentioned, the vector process H_k is a Markov chain, and by using the independence between the burst sizes of the arriving bursts and the inter-arrival times, we obtain

$$\mathrm{E}\Big[\Delta_H(k) \mid H_k = y\Big] = \mathrm{E}\Big[[\lceil y_1\rceil_{\mathcal{A}} + U_k]^+ - y_1 + \sum_{i=2}^{c}([y_i - T_k]^+ - y_i)\Big]\,, \tag{5.40}$$

where the operator R could be omitted, since we evaluate the sum of the components.

As for the sum involved in (5.40), as $y_i \to \infty$,

$$\mathrm{E}[[y_i - T_k]^+ - y_i] \downarrow -\mathrm{E}[T],\ \ i = 2 \ldots c\,.$$

Consider now the rest of (5.40), that can be rewritten as

$$\begin{aligned}
&\mathrm{E}\Big[[\lceil y_1\rceil_{\mathcal{A}} + U_k]^+ - y_1\Big] \\
&\quad = \int_{\mathbb{R}} ([\Delta(y_1) + y_1 + z]^+ - y_1) dF_U(z) \\
&\quad = -y_1 \mathrm{Pr}[U \le -y_1 - \Delta(y_1)] + \Delta(y_1)\mathrm{Pr}[U > -y_1 - \Delta(y_1)] \\
&\qquad + \int_{z \ge -y_1 - \Delta(y_1)} z dF_U(z) \\
&\quad \equiv I_1 + I_2 + I_3\,,
\end{aligned}$$

where we introduced I_1, I_2 and I_3 for notational convenience. We now consider these three terms subsequently. As above (see (5.14)), because $\mathrm{E}[U] > -\infty$, we have,

$$-y_1\mathrm{Pr}[U \le -y_1 - \Delta^*] \ge -y_1\mathrm{Pr}[U \le -y_1 - \Delta(y_1)] \ge -y_1\mathrm{Pr}[U \le -y_1] \uparrow 0\,,$$

as $y_1 \to \infty$, and hence, $I_1 \uparrow 0$. For I_2, by definition, like in (5.18),

$$\limsup_{y_1\to\infty} \Delta(y_1)\mathrm{Pr}[U > -y_1 - \Delta(y_1)] \le \Delta_0\,,$$

while for I_3, as $y_1 \to \infty$, we have

$$\int_{z \ge -y_1 - \Delta(y_1)} z dF_U(z) \downarrow \mathrm{E}[U]\,.$$

Given the above limits, it follows that (5.39) holds. Note that convergence $y \to \infty$ in (5.39) is component-wise. □

As a result, the negative drift condition in the multi-wavelength case takes on the form

$$\Delta_0 + \mathrm{E}[U] - (c-1)\mathrm{E}[T] < 0\,,$$

or

$$\Delta_0 + \mathrm{E}[B] < c\,\mathrm{E}[T]\,. \tag{5.41}$$

This condition, together with the conditions known from the single-wavelength case ((5.10) and (5.10)), will prove sufficient to guarantee stability. However, for the proof of the main stability result, an auxiliary lemma concerning tightness is needed, and is therefore presented beforehand.

5.2.3.2 Tightness

For the second lemma, a further characterization of the scheduling-horizon process is required, that allows to relate the drift of the largest component of H_k to that of the smallest. To this end, we introduce the differences

$$d_k = H_k^{(c)} - H_k^{(1)},\ k \geq 1\,. \tag{5.42}$$

The tightness of the sequence $\{d_k\}$ plays a crucial role in the following stability analysis.

Lemma 5.2.2. *The sequence $\{d_k\}$ is tight, that is for any $\varepsilon > 0$ there exists a finite constant N such that*

$$\inf_{k\geq 1} \Pr[d_k \leq N] \geq 1 - \varepsilon\,.$$

Proof. Denote the difference

$$e_k \equiv \sum_{i-1}^{c-1} \left(H_k^{(c)} - H_k^{(i)}\right) = (c-1)H_k^{(c)} - \sum_{i=1}^{c-1} H_k^{(i)}\,, \tag{5.43}$$

which is larger than or equal to zero for all k. Further, obviously, $d_k \leq e_k$ for all k. Hence, tightness of the sequence $\{e_k\}$ implies tightness of the sequence $\{d_k\}$. We rewrite the basic relation (5.36) as follows:

$$H_{k+1} = \mathrm{R}\Big([H_k^{(1)} + V_k + B_k - T_k]^+,\, [H_k^{(i)} - T_k]^+,\, i = 2\ldots c\Big),\ k \geq 1\,. \tag{5.44}$$

In the proof, we mainly follow [117] and also provide a detailed analysis based on an extended Kiefer-Wolfowitz recursion from [112, 118]. Upon the arrival of an arbitrary burst k, the scheduling horizon of the shortest queue increases. Depending on the size of the burst B_k and of the associated void V_k, the queue chosen by burst k will either be the longest queue as seen by burst $k+1$, or not. These two complementary cases are treated subsequently.

1. Firstly, assume that the queue chosen by burst k (the shortest queue) evolves to become the longest queue as seen by burst $k+1$, that is,

$$B_k + V_k \geq d_k = H_k^{(c)} - H_k^{(1)}\,, \tag{5.45}$$

which implies, given (5.44), that

$$H_{k+1}^{(c)} = [H_k^{(1)} + V_k + B_k - T_k]^+\,.$$

If assumption (5.45) holds then we obtain for the $\{e_k\}$ of (5.43) that

$$e_{k+1} = (c-1)\Big[H_k^{(1)} + V_k + B_k - T_k\Big]^+ - \sum_{i=2}^{c}[H_k^{(i)} - T_k]^+\,.$$

Now, excluding the trivial case where $H_{k+1}^{(i)} = 0$ for all $i = 1 \ldots c$ (implying $e_{k+1} = 0$), we assume the opposite case, namely that $H_k^{(1)} + V_k + B_k \geq T_k$. In that case, e_{k+1} has the following upper bound,

$$\begin{aligned} e_{k+1} &= (c-1)\Big(V_k + B_k\Big) - (c-1)T_k + (c-1)H_k^{(1)} \\ &\quad - \sum_{i=2}^{c}[H_k^{(i)} - T_k] - \sum_{i=2}^{c}[H_k^{(i)} - T_k]^- \\ &\leq (c-1)\Big(V_k + B_k\Big) + (c-1)H_k^{(1)} - \sum_{i=2}^{c} H_k^{(i)} \\ &\leq (c-1)\Big(V_k + B_k\Big) \\ &\leq (c-1)(\Delta^* + B_k)\,. \end{aligned} \tag{5.46}$$

To establish the first inequality, we used the identity $[\cdot]^+ = [\cdot] + [\cdot]^-$, with $[\cdot]^- = -\min(0,\cdot) \geq 0$, and omitted the negative term $-\sum_{i=2}^{c}[H_k^{(i)} - T_k]^-$. For the second inequality, we invoked $H_k^{(1)} \leq H_k^{(i)}$, for $i = 2 \ldots c$. The last inequality follows from $V_k \leq \Delta^*$.

2. Secondly, assume that the queue chosen by burst k is not the longest queue as seen by burst $k+1$, that is, complementary to (5.45) we assume,

$$V_k + B_k \leq d_k = H_k^{(c)} - H_k^{(1)}\,, \tag{5.47}$$

which, given (5.44), results in

$$H_{k+1}^{(c)} = [H_k^{(c)} - T_k]^+\,.$$

This then implies that

$$e_{k+1} = (c-1)\Big[H_k^{(c)} - T_k\Big]^+ - \sum_{i=2}^{c-1}[H_k^{(i)} - T_k]^+ - \Big[H_k^{(1)} + V_k + B_k - T_k\Big]^+ ,$$

where the sum disappears if $c = 2$. Excluding again the trivial case where $H_{k+1}^{(i)} = 0$ for all $i = 1 \ldots c$ (implying $e_{k+1} = 0$), we obtain that $H_k^{(c)} \geq T_k$. From this, it follows that

$$\begin{aligned} e_{k+1} &= (c-1)(H_k^{(c)} - T_k) - \sum_{i=2}^{c-1}(H_k^{(i)} - T_k) \\ &\quad - \Big(H_k^{(1)} + V_k + B_k - T_k\Big) \\ &\quad - \sum_{i=2}^{c-1}\Big[H_k^{(i)} - T_k\Big]^- - \Big[H_k^{(1)} + V_k + B_k - T_k\Big]^- . \end{aligned}$$

Omitting two non-positive terms we obtain as upper bound

$$e_{k+1} \leq e_k - V_k - B_k \leq \; e_k - B_k \, . \tag{5.48}$$

To summarize both cases, we introduce the notation $D_k = \Delta^* + B_k$. Note that $\mathrm{E}[D_k] = C^*$, see (5.26). Now, taking into account (5.45) and (5.47), it always goes that

$$e_{k+1} \leq \max\{e_k - B_k, \, (c-1)D_k\} \, .$$

Developing this expression recursively, we obtain

$$e_{k+1} \leq \max\Big(e_{k-1} - B_{k-1} - B_k, \, (c-1)D_{k-1} - B_k, \, (c-1)D_k\Big) \, .$$

After k iterations we obtain (assuming for a moment that $k \geq 3$)

$$\begin{aligned} e_{k+1} \quad \leq \quad & \max\Big(e_1 - B_1 - \cdots - B_k, \, (c-1)D_1 - B_2 - \cdots - B_k, \\ & \ldots, (c-1)D_{k-1} - B_k, \, (c-1)D_k\Big), \; k \geq 3 \, . \end{aligned} \tag{5.49}$$

Note that in our case, $e_1 = 0$ but the proof holds also for any initial state H_1, see [118]. Denote now

$$\begin{aligned} Y_k \quad &\equiv \quad \max\Big(e_1 - \sum_{i=1}^{k} B_i, \, (c-1)D_i - \sum_{n=i+1}^{k} B_n, \; i = 1 \ldots k\Big) \\ &= \quad \max\Big((c-1)D_i - \sum_{n=i+1}^{k} B_n, \; i = 1 \ldots k\Big) \, , \end{aligned} \tag{5.50}$$

with $\sum_{n=k+1}^{k}(\cdot) = 0$, using the non-negativity of Y_k in the last step. Taking into account (5.49), we have that

$$e_{k+1} \leq Y_k, \;\; k \geq 1\,,$$

and therefore, proving that the sequence of the $\{Y_k\}$ is tight, suffices to prove the tightness of the $\{e_k\}$, and, in its turn, of the $\{d_k\}$. This will be the approach in the remainder of this proof. Because the B_n are iid we may interchange indexes in (5.50) such that

$$Y_k = \max\Big((c-1)D_i - \sum_{n=1}^{i-1} B_n,\; i = 0 \ldots k-1\Big)\,,$$

keeping stochastic equivalence. By the strong law of large numbers (SLLN), w.p.1 as $i \to \infty$,

$$\frac{\sum_{n=1}^{i-1} B_n}{i} \to \mathrm{E}[B] > 0, \quad \frac{D_i}{i} \to 0\,.$$

Hence, for any fixed $\varepsilon \in (0,\, 1)$, there exists an i_0 such that

$$\Pr\Big[(c-1)D_i - \sum_{n=1}^{i-1} B_n < 0,\; i \geq i_0\Big] \geq 1 - \varepsilon/2\,.$$

Now we take $x_0 = x_0(\varepsilon)$ such that, for $x \geq x_0$,

$$\Pr\Big[(c-1)D_i \leq x,\; i = 1 \ldots i_0\Big] \geq 1 - \varepsilon/2\,.$$

Note that then

$$\Pr\Big[(c-1)D_i - \sum_{n=1}^{i-1} B_n \leq x,\, n \geq i_0\Big] \equiv \Pr[\mathcal{G}_x] \geq 1 - \varepsilon/2,\; x \geq 0\,,$$

where we denoted with $\mathcal{G}$ the event that the argument of $\Pr[\cdot]$ in the above is true.

$$\Pr\Big[(c-1)D_1 \leq x,\, (c-1)D_i \leq x + B_1 + \cdots + B_{i-1},\; i = 2 \ldots i_0\Big] \\ \equiv\; \Pr[\mathcal{B}_x] \geq 1 - \varepsilon/2\,,\; x \geq x_0\,,$$

where we introduced the event $\mathcal{B}_x$. Since

$$\Pr[\mathcal{G}_x \cap \mathcal{B}_x] = \Pr[\mathcal{G}_x] - \Pr[\mathcal{G}_x \cap \bar{\mathcal{B}}_x] \geq \Pr[\mathcal{G}_x] - \Pr[\bar{\mathcal{B}}_x]\,,$$

we obtain the lower bound

$$\Pr[Y_k \leq x] \geq \Pr[\mathcal{G}_x \cap \mathcal{B}_x] \geq 1 - \varepsilon/2 - \varepsilon/2 = 1 - \varepsilon\,.$$

Note that the event $\mathcal{G}_x$ contains an infinite sequence of (sub)events indexed by k, while in the event $\{Y_k \leq x\}$, the range of the index k is limited by i. Since ε is arbitrary small it follows that the sequence $\{Y_k\}$ (and hence $\{d_k\}$) is tight. □

We are now ready to establish the main stability result for the multi-wavelength system.

5.2.3.3 Main Stability Result

Theorem 5.2.1. *Assume that conditions (5.10) (assuming a properly-defined FDL buffer), (5.12) (regeneration assumption), and (5.41) (negative drift for multiple wavelengths) hold. Also, assume that a JSQ wavelength assignment algorithm is applied, as reflected in the Kiefer-Wolfowitz-type recursion (5.36). Then the zero-delayed regenerative process H is positive-recurrent with respect to regenerations (5.38).*

Proof. It follows easily from (5.39) that again (as in the single-wavelength case)

$$\sup_{y\geq 0} \mathrm{E}[\Delta_H(k) \,|\, H_k = y] \leq C^* \,, \tag{5.51}$$

where the mean increment $\mathrm{E}[\Delta_H(k)]$ is estimated by assuming $H_k = y = (y_1, \ldots, y_c)$ like in (5.39), and C^* denotes $\Delta^* + \mathrm{E}[B]$, as in (5.26). Take now any (fixed)

$$\varepsilon \in \Big(0,\, \min(C^*,\, c\,\mathrm{E}[T] - \Delta_0 - \mathrm{E}[B])\Big) \,,$$

and denote with S^c the c-dimensional domain$[x_0,\, \infty) \times \ldots \times [x_0,\, \infty)$, where we take $x_0 \geq 0$ such that (see (5.41))

$$\sup_{y\in S^c} \mathrm{E}[\Delta_H(k) \,|\, H_k = y] \leq -\frac{\varepsilon}{2} \,.$$

Thus, we can write (see (5.51))

$$\mathrm{E}[\Delta_H(k)] \leq C^* \mathrm{Pr}[H_k \notin S^c] - \frac{\varepsilon}{2} \mathrm{Pr}[H_k \in S^c] \,.$$

By means of a proof by contradiction, we show that $H_k^{(1)} \not\to \infty$ in probability as $k \to \infty$. Therefore, assume that as $k \to \infty$,

$$H_k^{(1)} \to \infty \;\; \text{in probability} \,. \tag{5.52}$$

Then, of course, all $H_k^{(i)} \to \infty$, $i = 2 \ldots c$, or $H_k \to \infty$. Therefore, we can find k_0 such that

$$\mathrm{Pr}[H_k \in S^c] \geq 1 - \frac{\varepsilon}{4C^*} \,,\;\; k \geq k_0 \,,$$

and thus,

$$\mathrm{Pr}[H_k \notin S^c] \leq \frac{\varepsilon}{4C^*} \,,\;\; k \geq k_0 \,.$$

Hence, for $k \geq k_0$,

$$\mathrm{E}[\Delta_H(k)] \leq \frac{\varepsilon}{4} - \frac{\varepsilon}{2}\Big(1 - \frac{\varepsilon}{4C^*}\Big) \leq -\frac{\varepsilon}{8} \,.$$

Using the expressions in the proof of (5.41) one can show (exactly as in the single-wavelength case, see (5.27)) that

$$\mathrm{E}\Big[\sum_{i=1}^{c} H_{k_0}^{(i)}\Big] < k_0 c\, C^* \,,$$

which contradicts assumption (5.52), as desired. It implies that $H_k^{(1)} \not\to \infty$, or there exist $\varepsilon_0 > 0$ a finite constant C and non-random (sub)sequence $z_k \to \infty$ such that

$$\Pr[H_{z_k}^{(1)} \le C] \ge \varepsilon_0 \,. \tag{5.53}$$

Because of the tightness of increments $\{d_k = H_k^{(c)} - H_k^{(1)},\, k \ge 1\}$, one can find a constant R^* (generally $\ge C$) such that

$$\Pr\Big[H_{z_k} \in [0,\, R^*]^c\Big] \ge \frac{\varepsilon_0}{2} \,.$$

Recall that t_k is the arrival instant of burst k, ($t_1 = 0$), with $T_k = t_{k+1} - t_k$, $k \ge 1$. Fix any z_k (belonging to the sequence $\{z_k\}$ in (5.53)), and denote the event

$$\mathcal{E}_k = \{H_{z_k} \in [0,\, R^*]^c\} \,.$$

With condition (5.12) assumed, it follows that also the condition

$$\Pr\Big[c\,T > B + \Delta^* + \delta_0\Big] = \Pr\Big[T > \frac{B + \Delta^* + \delta_0}{c}\Big] = \delta' \,,$$

holds for some $\delta_0 > 0$ and $\delta' > 0$. Assume that all wavelengths are busy ($H_{z_k}^{(i)} > 0$ for all $i = 1 \ldots c$) within the interval $[t_{z_k},\, t_{z_{k+1}})$, and that the event

$$\mathcal{G}(z_k) \equiv \Big\{T_{z_k} > \frac{B_{z_k} + \Delta^* + \delta_0}{m}\Big\} \,,$$

is realized. Then

$$\begin{aligned}
\sum_{i=1}^{c} H_{z_k+1}^{(i)} &= H_{z_k}^{(1)} + \Delta(H_{z_k}^{(1)}) + B_{z_k} - T_{z_k} + \sum_{i=2}^{c} (H_{z_k}^{(i)} - T_{z_k}) \\
&= \sum_{i=1}^{c} H_{z_k}^{(i)} + \Delta(H_{z_k}^{(1)}) + B_{z_k} - c\,T_{z_k} \\
&\le \sum_{i=1}^{c} H_{z_k}^{(i)} - \delta_0 \,.
\end{aligned}$$

Thus, on the event $\mathcal{G}(z_k)$, if all wavelengths are occupied within the (continuous-time) inter-arrival time $[t_{z_k},\, t_{z_k+1})$, the components of the scheduling horizon decrease by δ_0 (within this inter-arrival time). It is easy to see that this holds

for each inter-arrival time as long as all wavelengths are busy if we use at most $\lceil R^*/\delta_0 \rceil$ events $\mathcal{G}(z_k+i)$, $i \in \{0,\ldots,\lceil R^*/\delta_0\rceil - 1\}$. In other words, with a probability

$$\Pr\Big[\mathcal{E}_k \cap \bigcap_{i=0}^{\lceil R^*/\delta_0\rceil-1} \mathcal{G}(z_k+i)\Big] \geq \frac{\varepsilon_0}{2}(\delta')^{\lceil R^*/\delta_0\rceil} > 0\,,$$

a burst $\xi(n)$ say, arrives within (discrete-time) interval $[z_k, z_k+\lceil R^*/\delta_0\rceil)$ and sees at least one idle wavelength, that is

$$W^{(1)}_{\xi(n)} = \Big\lceil H^{(1)}_{\xi(n)} \Big\rceil_{\mathcal{A}} = 0\,.$$

Now, since instant $\xi(n) \in [z_k,\, z_k+\lceil R^*/\delta_0\rceil)$, we realize at most $\lceil R^*/\delta_0\rceil$ (independent) events

$$\mathcal{D}_{\xi(n)+i} \equiv \{T_{\xi(n)+i} > B_{\xi(n)+i} + \delta_0\}\,,\ i \in \{0,\ldots,\lceil R^*/\delta_0\rceil - 1\}\,,$$

to obtain a burst which finds the buffer completely empty. Indeed, for each burst $\xi(n)+i$, $i \in \{1,\ldots,\lceil R^*/\delta_0\rceil - 1\}$, we may use the same idle wavelength, allowing to transmit these bursts immediately. As such, the scheduling horizon of the other $c-1$ queues decreases continuously, and all wavelengths become idle within the interval $[\xi(n),\, \xi(n)+\lceil R^*/\delta_0\rceil)$. This is possible, since the regeneration assumption (5.12) implies that for each i,

$$\Pr[\mathcal{D}_{\xi(n)+i}] \geq \delta_1 > 0\,.$$

Finally, a regeneration occurs within interval $[z_k,\, z_k + 2\,\lceil R^*/\delta_0\rceil)$ with a constant length $D_0 = 2\,\lceil R^*/\delta_0\rceil < \infty$ with a probability $\geq \varepsilon_0/2(\delta')^{2\lceil R^*/\delta_0\rceil} > 0$. In other words, the unfinished regeneration time $\beta(z_k)$ at instant z_k is limited by D_0 with a probability

$$\Pr\Big[\beta(z_k) \leq D_0\Big] \geq \frac{\varepsilon_0}{2}(\delta')^{2\lceil R^*/\delta_0\rceil} > 0\,, \tag{5.54}$$

which is strictly positive, so implying positive recurrence. □

Remark 5.2.1. It follows from the proof of Theorem 5.2.1 that instead of the regeneration assumption (5.12), one can use a set of two assumptions:

$$\Pr\Big[T > \frac{B+\Delta^*}{c}\Big] > 0\,, \tag{5.55}$$

and

$$\Pr\Big[T > B\Big] > 0\,, \tag{5.56}$$

which both follow from (5.12). If $\Pr[B \geq \Delta^*] = 1$ then the latter one implies the former one, since, for $c \geq 2$,

$$0 < \Pr[T > B] = \Pr[c\,T > c\,B] \leq \Pr[c\,T > B + \Delta^*]\,.$$

Corollary 5.2.1. *Under the assumptions of Theorem 5.2.1, the process H has stationary distribution $Pr_{H_\infty}[\cdot]$, that is,* $\Pr[H_k \in \cdot] \to Pr_{H_\infty}[\cdot]$ *as* $k \to \infty$.

Proof. Similar to the single-wavelength case (see (5.35)), the regeneration assumption (5.12) guarantees aperiodicity of the renewal process β of regenerations since

$$\Pr[H_{k+1} = 0 \,|\, H_k = 0] = \Pr[T_k \geq B_k] \geq \Pr[T > B + \Delta^*] > 0\,,$$

showing that the stationary regime exists. □

The proof of the following result is obvious.

Corollary 5.2.2. *For the degenerate case $g_n \equiv D$, the zero-delayed process H is positive-recurrent if conditions (5.41) and* $\Pr[T > D + B] > 0$ *hold.*

Remark 5.2.2. The results above assume zero initial conditions. While we do not address the extension to non-zero initial conditions explicitly in this work, it should be possible in principle to follow an approach similar to that of the recent contribution [119]. However, the main problem is that the system under consideration evolves in a state-dependent way: the exact increment or decrement from arrival to arrival depends on the (state-dependent) size of the void. Therefore, we cannot use the known monotonicity properties which typically are used in the classic approach to stability. As such, a complete proof requires a more detailed and refined analysis.

5.3 Concluding Remarks

In this chapter, sufficient conditions were given for the stability of both single-wavelength and multiple-wavelength optical buffers. The key element of analysis is the exploitation of the regenerative property of the main system process and a characterization of the limiting forward renewal time process.

In case of one wavelength, the main system process is that of the waiting times. We first established stability under zero initial conditions, that we then extended to non-zero initial conditions.

For the multi-wavelength system, the main system process is that of the scheduling horizon. In this case, we were able to obtain a stability guarantee that assumes zero initial conditions, under a JSQ wavelength assignment algorithm.

The obtained stability guarantee is valid for the wide class of GI/G/c optical buffers, and poses no restriction on the FDL lengths. For now, the stability conditions for the multi-wavelength system still assume zero initial conditions; extending results also for non-zero initial conditions is a challenging problem that is still unsolved at the time of writing.

6
Concluding Remarks and Outlook

¶ Nearing the end of this doctoral dissertation, we return to the central question in this work (1.1):

If a buffer can only realize delays belonging to a limited set,
how does this restriction impact buffer performance?

We now summarize the answers provided in this doctoral dissertation. Also, we draw up possible new challenges in the stochastic modeling of optical buffers, and close with a general remark.

6.1 Main Results for Designers

At this point, we gather the conclusions of this work in one listing, with each item providing part of the answer to the central question of this work.

- As first discussed in Sect. 1.7.2, an FDL buffers differs fundamentally from a classic RAM buffer due to the occurrence of voids, and the effects of impatience. Both effects are directly implied by the limited number of possible delays.

- When compared to a classic RAM buffer, the effect of voids is strictly a performance degradation. The voids are implied by the FDL buffer structure, and can be mitigated, but not annihilated.

- When compared to a classic RAM buffer, the effect of impatience is not necessarily a performance degradation. Rather, it provides a characterization of buffer finiteness that is complementary to the typical limit on buffer size, namely an upper bound on the number of waiting places.

- The main way to optimize the performance of FDL buffers is by fine-tuning the fiber lengths. In case of degenerate buffering, this comes down to fine-tuning the granularity.

- In the fine-tuning of the fiber lengths, the most important factor to take into account is probably the burst size distribution. More particularly, in case of a burst size distribution with bounded support (with related supremum B_{max}), the optimal granularity for the degenerate buffer setting is necessarily smaller than B_{max}.

  Other major factors include the inter-arrival time distribution, and the number of wavelengths. If multiple wavelengths are considered, the wavelength assignment algorithm also plays a major role.

  Although a major factor in the LP, the buffer size is only a minor factor when it comes to fine-tuning the fiber lengths, and only impacts this process if very few lines (less than about 5) are considered.

- In the case that burst sizes are fixed, for low load values ($\rho < 0.6$) and memoryless arrivals, one obtains best results with a degenerate buffer setting and with a granularity equal to the burst size. For high load values ($\rho > 0.6$), the optimal granularity shifts to lower values. Also, in Sect. 3.3.4, for a degenerate M/D/1 setting and for low load values, it was pointed out that the average waiting time in an FDL buffer is (more than) doubled, when compared to a RAM buffer with identical maximum waiting time.

- In the case that burst sizes are varying, it equally goes that an increase in load results in a lower value for the optimal granularity. However, the latter value is much more sensitive to the load value than in case of fixed-sized bursts, as was illustrated for the case of memoryless burst sizes.

- All the numerical examples throughout this work considering the question of optimal granularity share a common trend: the optimal granularity is strictly non-increasing for increasing load. More precisely, numerous figures illustrated that, for increasing load, the optimal granularity either remains unaltered, or that it decreases.

  - In a degenerate M/G/1 system with upper-bounded burst size distribution, the optimal granularity remains invariant if the load remains below a certain value called the threshold load. Curves of the threshold load are displayed on Figure 3.6, for the case of M/D/1 and M/U/1.

– Above the threshold load, as the load increases, the optimal granularity in a degenerate M/G/1 system shifts to lower values in either a continuous manner (as is the case for a degenerate M/M/1) or in a non-continuous manner (as is the case for a degenerate M/D/1).

Note that there is no theoretical framework to substantiate this claim, but that it seems plausible to the author that it can be proven for a degenerate M/G/1 buffer setting. In the more general case of GI/G/1, however, several pathological cases exist for which this rule simply does not hold.

6.2 Main Results for Queueing Theorists

Within the broader context of optical buffer research, our angle of incidence in this monograph is stochastic modeling. Following conclusions can be drawn.

- The modeling of infinite-sized FDL buffers pointed out that these buffers are instable for lower load values than those that apply for classic buffers. For the single-wavelength degenerate case, the maximum tolerable load is strictly smaller than one, and typically (but not always) decreases with increasing granularity. In Chapter 2, we provided exact stability bounds for a GI/G/1 degenerate buffer setting. In Chapter 5, we constructed sufficient conditions to guarantee stability for the wider class of a non-degenerate GI/G/1 buffer setting. Also in Chapter 5, assuming zero initial conditions, we obtained stability guarantees for a non-degenerate GI/G/c buffer setting with JSQ wavelength assignment algorithm.

- Throughout this work, the modeling of single-wavelength FDL buffers was done in terms of both the waiting time and the scheduling horizon. Both in the non-degenerate and degenerate case, the description based on the waiting times is probably the most suitable of the two. However, in the multi-wavelength case, one necessarily has to include the scheduling horizon in the analysis.

- From a modeling point of view, probably the simplest FDL buffer system is given by the degenerate M/D/1 setting. In case of infinite buffer size, the analysis is already quite basic, and in case of finite buffer size and granularity equal (CT) or nearly equal (DT) to the burst size, the analysis is even simpler than that of a classic M/D/1 buffer. Further, exact analysis was also feasible for the degenerate M/M/1 setting, for both infinite and finite buffer setting. In this work, all the cases mentioned are provided with a closed-form solution, and this regardless of the time setting.

- When compared to classic buffer models, the occurrence of voids leads to additional complexity in the buffer model. Nevertheless, the results of Chapter 2 point out that exact analysis of the single-wavelength case is feasible for a class as wide as GI/G/1. On the other hand, for the multi-wavelength case, exact analysis is only possible for some special cases. As demonstrated in Chapter 2, assuming a simple wavelength assignment algorithm such as RR allows for exact analysis. Such model can provide a useful approximation also in case of JSQ, but only if one assumes in addition that burst sizes are fixed.

- For finite single-wavelength optical buffers, a powerful technique is provided in Chapter 3. More precisely, loss and waiting time probabilities for a class as broad as GI/G/1 can be computed exactly, with minimal numerical effort. This approach exploited the limited number of possible waiting times, by associating each possible waiting time with exactly one Markov state. Although not treated in Chapter 3, it can be expected that such numerically efficient approach cannot be translated to the multi-wavelength case. The main problem lies therein, that one can still trace the first wavelength by means of the waiting time, but that one is to trace the scheduling horizon of each additional wavelength. Since the scheduling horizon can take on a wider (and possibly infinite) range of values, this quickly leads to a state space explosion. More precisely, in DT, the scheduling horizon can take on any value in $\{0, 1, \ldots a_N\}$, with a_N the length of the longest line. As such, while the single-wavelength case could be modeled with $N+1$ states, the c-wavelength case requires a state space of dimension $(N+1) \times (a_N+1)^{c-1}$, which is much larger than $(N+1)^c$.

- As for the aspect of impatience involved in FDL buffering, application of impatience modeling to the degenerate M/M/1 buffer model pointed out that this is a promising candidate for an approximate modeling of more complicated FDL buffer systems. A possible extension is treated in the next section.

6.3 Outlook and Possible Future Work

With Transform Functions or With Markov Chains While Chapter 2 and 3 provide solutions to a rather broad class of FDL buffering problems, many problems remain open. Most relevant is beyond doubt the multi-wavelength case, where many possible wavelength assignment algorithms remain unstudied. Also interesting is the single-wavelength case with void filling, that was not considered in this work but nevertheless is promising as alternative.

Whether these problems are best tackled by means of transform functions or Markov chains, depends mainly on the researcher's willingness to explore approximations. Given the irregularity of the stochastic processes produced by advanced delay-line or wavelength assignment algorithms, it can be expected that an approach with transform functions will not enable an exact tracing of this problem. On the other hand, a Markov chain approach may (or may not) allow an exact tracing. However, in the case that an exact tracing is possible, it is not said that such tracing with Markov chains allows for a feasible numerical computation. More particularly, tracing the scheduling horizon of multiple queues in the multi-wavelength case is bound to quickly result in a state space explosion. Nevertheless, appropriate techniques, such as the formulation of state transitions in terms of Hessenberg matrices (as done in for example [19]) might provide a remedy to such problems.

If one opts for an approach with transform functions, it is probably best to first introduce a variety of simplifications of the system. After verification of the accuracy of each approximation through simulation, one could then turn to a solution in terms of transform functions of the most accurate simplified system. A possible simplified system, that probably allows for solution in terms of transform functions, was discussed in Chapter 4.

Impatience As for the approach with impatience of Chapter 4, it is clear that we have merely lifted the tip of the veil. It follows logically from the results of Chapter 4 how one could apply the method with impatience to degenerate finite-sized GI/G/1 buffers. The necessary ingredients to follow such approach are the equivalent load in that case (that is, Eq. (2.71)), combined with the appropriate model for impatience (with the analysis for D-MAP/PH/1 of [106] as a possible starting point).

As for the approach with impatience applied to the multi-wavelength system, the situation is less self-evident. Firstly, one is to propose an appropriate equivalent load, since no general expression is available for that case. Also, note that such expression will most likely be valid only for one particular wavelength assignment algorithm, since variation of the latter greatly impacts the buffer's loss performance (see Sect. 2.5).

Finally, given what we know at this point about FDL buffer behavior, it is very likely that a further exploration of the impatience model will provide even more accurate approximations than the simple approach of Chapter 4. One possible extension would be to model the impact of voids by means of a model with exceptional first service. In such models, arrivals that find an empty system upon arrival get a different service (or, a different burst size) than arrivals that find a non-empty system upon arrival. In case of FDL buffers, an approximative performance model with exceptional first service could assign a normal burst size B (with B a rv) to

arrivals that find an empty system, and an augmented burst size $B + V$ (with B and V two rv's) to arrivals that find a non-empty system. The additional work, represented by the rv V, could then account for the effect of voids in the actual FDL buffer system.

With Regenerations As for the stability results obtained in Chapter 5, a challenging open problem remains the extension of stability guarantees to non-zero initial conditions for the multi-wavelength case with JSQ wavelength assignment. Further, note that no stability guarantees have been obtained for other wavelength assignment algorithms. Also of interest is the question whether a necessary stability bound, in terms of a maximum tolerable load, can also be formulated for a non-degenerate GI/G/1 buffer, as it was presented in Chapter 2 for a degenerate GI/G/1 buffer.

6.4 Let There Be Light...

If one is to name a single reason for the interest of the research community in FDL buffers, it is beyond doubt their key role in the next-generation optical network. That FDL buffers have been a hot topic over the last decennium is an effect of their inclusion in numerous OPS/OBS testbeds across the world. On the other hand, when and how the final step to commercialization of optical switching will be made is hard to tell. At any rate, as El-Bawab states it [39], the Internet does grow, and the yearly incremental rise in traffic volume may even be increasing. Obviously, this urges for additional bandwidth not only in the access network, but also in the backbone. The exact moment at which the shift from electrical switching to optical switching will take place might not be predictable, but the technical and operational drivers for the vision of the optical layer are genuine: the network architecture needs simplification, flexibility, intelligence, and scalability. Optical networking comes up to these needs, and optical buffering, although but a small piece of this optically-switched network, surely provides a vital piece. Given the vital role of performance evaluation in the propagation of new technologies, the author expresses his hope that this dissertation both directly and indirectly stimulates the deployment of optical buffers in actual networks, and the deployment of optical networking in general.

Bibliography

[1] W. Rogiest, K. Laevens, D. Fiems, and H. Bruneel. *Analysis of an Asynchronous Single-Wavelength FDL Buffer*. Proceedings of the Nineteenth International Teletraffic Congress, ITC 19 (Beijing), pages 1917–1926, 2005.

[2] W. Rogiest, K. Laevens, D. Fiems, and H. Bruneel. *A Performance Model for an Asynchronous Optical Buffer*. Performance Evaluation, 62(1-4):313–330, 2005.

[3] W. Rogiest, K. Laevens, D. Fiems, and H. Bruneel. *Optical Buffers, Batch Arrivals and Synchronization*. Proceedings of the 2nd Conference on Next Generation Internet Design and Engineering, NGI 2006 (València), 2006.

[4] W. Rogiest, K. Laevens, D. Fiems, and H. Bruneel. *Quantifying the Impact of Wavelength Conversion on the Performance of Fiber Delay Line Buffers*. Proceedings of the 2006 International Workshop on Optical Burst/Packet Switching, WOBS 2006 (San José, CA), 2006.

[5] W. Rogiest, K. Laevens, D. Fiems, and H. Bruneel. *Analysis of a Lakatos-Type Queueing System with General Service Times*. Book of Abstracts of the Twentieth Conference on Quantitative Methods for Decision Making, ORBEL 20 (Ghent), pages 95–97, 2006.

[6] W. Rogiest, K. Laevens, J. Walraevens, and H. Bruneel. *Analyzing a Degenerate Buffer with General Inter-Arrival and Service Times in Discrete Time*. Queueing Systems, 56 (3-4):203–212, 2007.

[7] W. Rogiest, D. Fiems, K. Laevens, and H. Bruneel. *Tracing an Optical Buffer's Performance: An Effective Approach*. Lecture Notes in Computer Science, NET-COOP 2007 Special Issue, 4465:185–194, 2007.

[8] W. Rogiest, D. Fiems, K. Laevens, and H. Bruneel. *A Light-Weigth Performance Model for Optical Buffers*. Accepted for publication in the International Journal of Communications Networks and Distributed Systems (IJCNDS).

[9] W. Rogiest, J. Lambert, D. Fiems, B. Van Houdt, H. Bruneel, and C. Blondia. *A Unified Model for Synchronous and Asynchronous FDL Buffers.* submitted to Performance Evaluation, 2008.

[10] K. Laevens and H. Bruneel. *Analysis of a Single-Wavelength Optical Buffer*. Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2003 (San Francisco, CA), 2003.

[11] K. Laevens, B. Van Houdt, H. Bruneel, and C. Blondia. *On the Sustainable Load of Fibre Delay Line Buffers.* IEE Electronics Letters, 40(2):137–138, 2004.

[12] K. Laevens and H. Bruneel. *Analysis of Fiber Delay Line Optical Buffers*. Workshop on Next Generation Networks: Architecture, Protocols, Performance (Belize), 2005.

[13] K. Laevens, M. Moeneclaey, and H. Bruneel. *Queueing Analysis of a Single-Wavelength Fiber-Delay-Line Buffer*. Telecommunication Systems, 31(2-3):259–287, 2006.

[14] D. Fiems, J. Walraevens, and H. Bruneel. *Discrete Time Queueing Analysis of a Fibre Delay Line Structure with Granularity*. Proceedings of the Eighth IFIP Working Conference on Optical Network Design and Modelling, ONDM 2004 (Ghent), 2004.

[15] 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 (AEÜ), 57(4):270–276, 2003.

[16] 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*. Proceedings of the 2004 IEEE Global Telecommunications Conference, Globecom 2004 (Dallas), pages 23–33, paper OC05–07, 2004.

[17] J. Lambert, B. Van Houdt, and C. Blondia. *Single-Wavelength Optical Buffers: Non-Equidistant Structures and Preventive Drop Mechanisms*. Proceedings of the 2005 Networking and Electronic Commerce Research Conference, NAEC 2005 (Riva del Garda), pages 545–555, 2005.

[18] J. Lambert, B. Van Houdt, and C. Blondia. *Queues with Correlated Inter-Arrival and Service Times and Its Application to Optical Buffers.* Stochastic Models, 22(2):233–251, 2006.

[19] J. Lambert, W. Rogiest, B. Van Houdt, D. Fiems, C. Blondia, and H. Bruneel. *A Hessenberg Markov chain for Fast Fibre Delay Line Length Optimization*. Lecture Notes in Computer Science, ASMTA 2008 Special Issue, 2008.

[20] L. Lakatos. *On a Simple Continuous Cyclic-Waiting Problem*. Annales Univ. Sci. Budapest., Sect. Comp. 14, pages 105–113, 1994.

[21] L. Lakatos. *On a Simple Discrete Cyclic-Waiting Queueing Problem*. Journal of Mathematical Sciences (New York), 4(92):4031–4034, 1998.

[22] L. Lakatos and G. Zbaganu. *Waiting Time in Cyclic-Waiting Systems*. Annales Univ. Sci. Budapest., Sect. Comp. 27, pages 217–228, 2007.

[23] E.V. Koba. *On a GI/G/1 Queueing System with Repetition of Requests and Service in the Order of Arrivals (written in Russian)*. Dopovidi NAN Ukraini, 6:101–103, 2000.

[24] F. Callegati. *Optical Buffers for Variable Length Packets*. IEEE Communications Letters, 4(9):292–294, 2000.

[25] F. Callegati. *On the Design of Optical Buffers for Variable Length Packet Traffic*. Proceedings of the Ninth IEEE International Conference on Computer Communications and Networks, ICCCN 2000 (Las Vegas), pages 448–452, 2000.

[26] F. Callegati. *Approximate Modeling of Optical Buffers for Variable Length Packets*. Photonic Network Communications, 3(4):383–390, 2001.

[27] M. Murata and K. Kitayama. *Ultrafast Photonic Label Switch for Asynchronous Packets of Variable Length*. Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2002 (New York), 2002.

[28] H. Waldman, R. C. Almeida, and J. U. Pelegrini. *An Infinite Granularity Bound on the Performance of Delay-Line Buffering*. Proceedings of the 7th IFIP Working Conference on Optical Network Design and Modelling, ONDM 2003 (Budapest), pages 1187–1200, 2003.

[29] R. C. Almeida, J. U. Pelegrini, and H. Waldman. *Delay-Line Buffering for Asynchronous Optical Networks*. Proceedings of the 2003 SPIE/IEEE Optical Networking and Communications Conference, OptiComm 2003 (Dallas), pages 381–391, 2003.

[30] R. C. Almeida, J. U. Pelegrini, and H. Waldman. *Optical Buffer Modelling for Performance Evaluation Considering Any Packet Inter-Arrival*

Time Distribution. Proceedings of the 2004 IEEE International Conference on Communications, ICC 2004 (Paris), 3:1771–1775, 2004.

[31] R. C. Almeida, J. U. Pelegrini, and H. Waldman. *A Generic-Traffic Optical Buffer Modeling for Asynchronous Optical Switching Networks*. IEEE Communications Letters, 9(2):175–177, 2005.

[32] D. Hong, P. Poppe, J. Reynier, F. Baccelli, and G. Petit. *The Impact of Burstification on TCP Throughput in Optical Burst Switching Networks*. Proceedings of the Eighteenth International Teletraffic Congress, ITC 18 (Berlin), 2003.

[33] T. Zhang, K. Lu, and J. P. Jue. *Architectures and Performance of Fiber Delay Line Buffers in Packet-Based Multifiber Optical Networks*. Proceedings of the 2005 Optical Fiber Communication Conference – Technical Digest, OFC/NFOEC 2005 (San Diego, CA), 2005.

[34] T. Zhang, K. Lu, and J.P. Jue. *An Analytical Model for Shared Fiber-Delay-Line Buffers in Asynchronous Optical Packet and Burst Switches*. Proceedings of the 2005 IEEE International Conference on Communications, ICC2005 (Seoul), 3:1636–1640, 2005.

[35] T. Zhang, K. Lu, and J.P. Jue. *Shared Fiber Delay Line Buffers in Asynchronous Optical Packet Switches*. IEEE Journal On Selected Areas in Communications, 24(4):118–127, 2006.

[36] A. Rostami and S.S. Chakraborty. *On Performance of Optical Buffers with Fixed Length Packets*. Proceedings of the Second IFIP International Conference on Wireless and Optical Communications Networks, WOCN 2005 (Dubai), 2005.

[37] A. Rostami and S.S. Chakraborty. *On Performance of Optical Buffers With Specific Number of Circulations*. IEEE Photonics Technology Letters, 17(7):1570–1572, 2007.

[38] *Alcatel-Lucent Announces New Optical Transmission Records at ECOC 2007, Paving the Way for the Future Introduction of 100 Gigabit/s Ethernet*. Alcatel-Lucent News Releases, 19 September 2007 (http://usstock1.jrj.com.cn/news/2007-09-20/000002701271.html), 2007.

[39] T. S. El-Bawab. *Optical Switching*. Springer, 2006.

[40] D. Blumenthal. *Photonic Packet Switching and Optical Label Swapping*. Optical Networks Magazine, 2(6):54–65, 2001.

[41] P. Green. *Progress in Optical Networking*. IEEE Communications Magazine, 39(1):54–61, 2001.

[42] G.N. Rouskas and H.G. Perros. *A Tutorial on Optical Networks*. Networking 2002 Tutorials (LNCS 2497), pages 155–194, 2002.

[43] H. Perros. *Connection-Oriented Networks: SONET/SDH, ATM, MPLS and Optical Networks*. Wiley, 2005.

[44] E. Van Breusegem. *Study and Design of Optical Network Architectures*. PhD thesis, Ghent University (UGent), 2003.

[45] P. De Dobbelaere, K. Falta, and S. Gloeckner. *Advances in Integrated 2D MEMS-Based Solutions for Optical Network Applications*. IEEE COmmunications Magazine, 41(5):16–23, 2003.

[46] M. Listanti, V. Eramo, and R. Sabella. *Architectural and Technological Issues for Future Optical Internet Networks*. IEEE Communications Magazine, 38(9):82–92, 2000.

[47] A. Jourdan, D. Chiaroni, E. Dotaro, G.J. Eilenberger, F. Masetti, and M. Renaud. *The Perspective of Optical Packet Switching in IP-Dominant Backbone and Metropolitan Networks*. IEEE Communications Magazine, 39(3):136–141, 2001.

[48] T. El-Bawab and J.-D. Shin. *Optical Packet Switching in Core Networks: Between Vision and Reality*. IEEE Communications Magazine, 40:60–65, 2002.

[49] F. Ramos, E. Kehayas, J. M. Martinez, R. Clavero, J. Marti, L. Stampoulidis et al. *IST-LASAGNE: Towards All-Optical Label Swapping Employing Optical Logic Gates and Optical Flip-Flops*. IEEE Journal of Lightwave Technology, 23(10):2993–3011, 2005.

[50] L. Xu, H.G. Perros, and G.N. Rouskas. *Techniques for Optical Packet Switching and Optical Burst Switching*. IEEE Communications Magazine, 39(1):136–142, 2001.

[51] J. Turner. *Terabit Burst Switching*. Journal on High-Speed Networks, 8:3–16, 1999.

[52] I. Chlamtac et al. *CORD: Contention Resolution by Delay Lines*. IEEE Journal on Selected Areas in Communications, 14(5):1014–1029, 1996.

[53] D.K. Hunter et al. *WASPNET: a Wavelength Switched Packet Network*. IEEE Communications Magazine, 37(3):120–129, 1999.

[54] C. Guillemot, M. Renaud, P. Gambini, C. Janz et al. *Transparent Optical Packet Switching: the European ACTS KEOPS Project Approach*. IEEE/OSA Journal of Lightwave Technology, 12(16):2117–2134, 1998.

[55] F. Masetti, J. Benoit, F.F. Brillouet, J.M. Gabriagues et al. *High Speed, High Capacity ATM Optical Switches for Future Telecommunication Transport Networks*. IEEE Journal on Selected Areas in Communications, 14(5):979–998, 1996.

[56] L. Dittmann, C. Develder, D. Chiaroni, F. Neri et al. *The European IST Project DAVID: A Viable Approach Toward Optical Packet Switching*. IEEE Journal on Selected Areas in Communications, 7(21):1026–1040, 2003.

[57] D. J. Blumenthal and M. Masanovic. *LASOR (Label Switched Optical Router): Architecture and Underlying Integration Technologies*. Proceedings of the 2005 European Conference of Optical Communication, ECOC 2005 (Glasgow), 2005.

[58] C. Qiao and M. Yoo. *Optical Burst Switching–a New Paradigm for an Optical Internet*. Journal on High-Speed Networks, 8:69–84, 1999.

[59] Y. Chen, C. Qiao, and X. Yu. *Optical Burst Switching: A New Area in Optical Networking Research*. IEEE Network, 18(3):16–23, 2004.

[60] V. Inghelbrecht, B. Steyaert, and H. Bruneel. *Study of the Burstification Mechanism of an OBS Edge Router*. Proceedings of the Seventh IFIP Working Conference on Optical Network Design and Modelling, ONDM 2003 (Budapest), II:1221–1239, 2003.

[61] Jung Yul Choi, H. L. Vu, and Minho Kang. *On Achieving the Optimal Performance of FDL Buffers Using Burst Assembly*. IEEE Communications Letters, 11:895–897, 2007.

[62] Y. Xiong, M. Vandenhoute, and H. Cankaya. *Control Architecture in Optical Burst-Switched WDM Networks*. IEEE Journal on Selected Areas in Communications, 18(10):1838–1851, 2000.

[63] S. Yao, B. Mukherjee, and S. Dixit. *Advances in Photonic Packet Switching: An Overview*. IEEE Communications Magazine, 38(2):84–94, 2000.

[64] Shun Yao, B. Mukherjee, S. J. B. Yoo, and S. Dixit. *A Unified Study of Contention-Resolution Schemes in Optical Packet-Switched Networks*. Journal of Lightwave Technology, 21(3):672–683, 2003.

[65] M. Yoo, C. Qiao, and S. Dixit. *The Effect of Limited Fibre Delay Lines on QoS Performance of Optical Burst Switched WDM Networks*. Proceedings

of the 2000 IEEE International Conference on Communications, ICC 2000 (New Orleans), pages 974–979, 2000.

[66] Christoph M. Gauger. *Optimized Combination of Converter Pools and FDL Buffers for Contention Resolution in Optical Burst Switching*. Photonic Network Communications, 8(2):139–148, 2004.

[67] D.K. Hunter, M. Chia, and I. Andonovic. *Buffering in Optical Packet Switches*. IEEE/OSA Journal of Lightwave Technology, 16(12):2081–2096, 1998.

[68] M.J. Karol and M.G. Hluchyj. *Input Versus Output Queueing on a Space-Division Packet Switch*. IEEE Transactions on Communications, 35(12):1347–1356, 1987.

[69] K. Laevens. *Stochastic Modelling of ATM Switching Elements with Queues at the Input*. PhD thesis, Ghent University (UGent), 1999.

[70] C. Develder. *Design and Analysis of Optical Packet Switching Networks*. PhD thesis, Ghent University (UGent), 2003.

[71] D.K. Hunter, W.D. Cornwell, T.H. Gilfedder, and I. Andonovic. *SLOB: a Switch with Large Optical Buffers for Packet Switching*. IEEE/OSA Journal of Lightwave Technology, 16(10):1725–1736, 1998.

[72] L. Tancevski, S. Yegnanarayanan, G. Castanon, L. Tamil, F. Masetti, and T. McDermott. *Optical Routing of Asynchronous, Variable Length Packets*. IEEE Journal on Selected Areas in Communications, 18(10):2084–2093, 2000.

[73] L. Tancevski, A. Ge, and G. Castanon. *Optical Packet Switch with Partially Shared Buffers: Design Principles*. Proceedings of the 2001 Optical Fiber Communication Conference and Exhibit, OFC 2001 (Anaheim), 2:TuK3–1–TuK3–3, 2001.

[74] Z. Haas. *The Staggering Switch: an Electronically Controlled Optical Packet Switch*. IEEE/OSA Journal of Lightwave Technology, 11(5-6):925–936, 1993.

[75] R. Van Caenegem, D. Colle, M. Pickavet, P. Demeester, K. Christodoulopoulos, K. Vlachos et al. *The Design of an All-Optical Packet Switching Network*. IEEE Communications Magazine, 45(11):52–61, 2007.

[76] A. Jajszczyk. *Optical Networks–the Electro-Optic Reality*. Optical Switching and Networking, 1:3–18, 2005.

[77] L. Tancevski, S. Tamil, and F. Callegati. *Non-Degenerate Buffers: A Paradigm for Building Large Optical Memories*. IEEE Photonic Technology Letters, 11(8):1072–1074, 1999.

[78] V. Puttasubbappa and H.G. Perros. *Performance Analysis of Limited-Range Wavelength Conversion in an OBS Switch*. Telecommunication Systems, 31(2-3):227–246, 2006.

[79] J.M. Gabriagues and J.B. Jacob. *OASIS: A High-Speed Photonic ATM Switch Results and Perspectives*. Proceedings of the 15th International Switching Symposium, ISS 1995 (Berlin), pages 457–461, 1995.

[80] F. Masetti, M. Sotom, D. De Bouard, D. Chiaroni, P. Parmentier, F. Callegati, G. Corazza, C. Rafaelli, S.L. Danielsen, and K.E. Stubkjaer. *Design and Performance of a Broadcast and Select Photonic Packet Switching Architecture*. Proceedings of the 1996 European Conference of Optical Communication, ECOC 1996 (Oslo), pages 15–19, 1996.

[81] D. Y. Barrer. *Queuing with Impatient Customers and Indifferent Clerks*. Operations Research, 5(5):644–649, 1957.

[82] D. Y. Barrer. *Queuing with Impatient Customers and Ordered Service*. Operations Research, 5(5):650–656, 1957.

[83] K.V. Mykhalevych. *A Comparison of a Classical Retrial M/G/1 Queueing System and a Lakatos-Type M/G/1 Cyclic-Waiting Time Queueing System*. Annales Univ. Sci. Budapest., Sect. Comp. 23, 2003.

[84] J.R. Artalejo. *A Classical Bibliography of Research on Retrial Queues: Progress in 1990–1999*. TOP 7, pages 187–211, 1999.

[85] E.F. Burmeister, D.J. Blumenthal, and J.E. Bowers. *A Comparison of Optical Buffering Technologies*. Optical Switching and Networking (OSN), 5(1)(1):10–18, 2008.

[86] J. Khurgin. *Optical Buffers Based on Slow Light in Electromagnetically Induced Transparent Media and Coupled Resonator Structures: Comparative Analysis*. Journal of the Optical Society of America B, 22:1062–1074, 2005.

[87] R. Tucker, P.-C. Ku, and C. Chang-Hasnain. *Slow-Light Optical Buffers: Capabilities and Fundamental Limitations*. Journal of Lightwave Technology, 23:4046–4066, 2005.

[88] A. Malacarne, A. Bogoni, and L. Poti. *ErbiumYtterbium-Doped Fiber-Based Optical Flip-Flop*. IEEE Photonics Technology Letters, 19(12):904–906, 2007.

[89] N. Beheshti, Y. Ganjali, R. Rajaduray, D. Blumenthal, and N. McKeown. *Buffer Sizing in All-Optical Packet Switches*. Proceedings of the 2006 Optical Fiber Communication Conference, OFC 2006 (Anaheim), 2006.

[90] M. Enachescu, Y. Ganjali, A. Goel, N. McKeown, and T. Roughgarden. *Routers with Very Small Buffers*. Proceedings of the 25th IEEE International Conference on Computer Communications, INFOCOM 2006 (Barcelona), 2006.

[91] G. Appenzeller, I. Keslassy, and N. McKeown. *Sizing Router Buffers*. Proceedings of the 2004 ACM Conference of the Special Interest Group on Data Communication, SIGCOMM 2004 (New York), pages 281–292, 2004.

[92] D.G. Kendall. *Stochastic Processes Occurring in the Theory of Queues and Their Analysis by the Method of the Imbedded Markov Chain*. Ann. Math. Statist., 24 (3):338–354, 1953.

[93] H. Bruneel and B. G. Kim. *Discrete-Time Models for Communication Systems Including ATM*. Boston: Kluwer Academic Publishers, 1993.

[94] J. C. Van Ommeren. *Asymptotic Analysis of Queueing Systems*. PhD thesis, Vrije Universiteit te Amsterdam, 1989.

[95] H. Bruneel. *Message Delay in TDMA Channels with Contiguous Output*. IEEE Transactions on Communications, 34(7):681–684, 1986.

[96] D. Fiems. *Analysis of Discrete-Time Queueing Systems with Vacations*. PhD thesis, Ghent University (UGent), 2004.

[97] W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson. *On the Self-Similar Nature of Ethernet Traffic (Extended Version)*. IEEE/ACM Transactions on Networking, 2(1):1–15, 1994.

[98] V.I. Klimenok. *On the Modification of Rouche's Theorem for the Queueing Theory Problems*. Queueing Systems, 38(4):431–434, 2001.

[99] O. Brun and J.-M. Garcia. *Analytic Solution of Finite Capacity M/D/1 Queues*. Applied Probability Trust 37, pages 1092–1098, 2000.

[100] L. Kleinrock. *Queueing Systems, Volume 1: Theory*. John Wiley & Sons, 1975.

[101] F. Baccelli, P. Boyer, and G. Hebuterne. *Single-Server Queues with Impatient Customers*. Advances in Applied Probability, 16(4):887–905, 1984.

[102] R. E. Stanford. *Reneging Phenomena in Single Channel Queues*. Mathematics of Operations Research, 4(2):162–178, 1979.

[103] A. R. Swensen. *On a GI/M/c Queue with Bounded Waiting Times*. Operations Research, 34(6):895–908, 1986.

[104] A. Movaghar. *On Queueing with Customer Impatience until the Beginning of Service*. Queueing Systems, 29:337–350, 1998.

[105] W. Xiong, D. Jagerman, and T. Altiok. *M/G/1 Queue with Deterministic Reneging Times*. Performance Evaluation, 65:308–316, 2008.

[106] B. Van Houdt, R.B. Lenin, and C. Blondia. *Delay Distribution of (Im)Patient Customers in a Discrete Time D-MAP/PH/1 Queue with Age-Dependent Service Times*. Queueing Systems, 45:59–73, 2003.

[107] J. Van Velthoven, B. Van Houdt, and C. Blondia. *On the Probability of Abandonment in Queues with Limited Sojourn and Waiting Times*. Operations Research Letters, 34(3):333–338, 2006.

[108] J. Van Velthoven, B. Van Houdt, and C. Blondia. *Response Time Distribution in a D-MAP/PH/1 Queue with General Customer Impatience*. Stochastic Models, 21:745–765, 2005.

[109] R. E. Stanford. *On Queues with Impatience*. Advances in Applied Probability, 22(3):768–769, 1990.

[110] J.W. Cohen. *The Single Server Queue*. North-Holland Publishers, 1969.

[111] B. Gavish. *The Markovian Queue with Bounded Waiting Time*. Management Science, 23(12):1349–1357, 1977.

[112] E. Morozov. *The Tightness in the Ergodic Analysis of Regenerative Queueing Processes*. Queueing Systems, 27:179–203, 1997.

[113] E. Morozov. *Weak Regeneration in Modeling of Queueing Processes*. Queueing Systems, 46:295–315, 2004.

[114] E. Morozov. *A Multiserver Retrial Queue: Regenerative Stability Analysis*. Queueing Systems, 56 (3-4):157–168, 2007.

[115] S.P. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability*. Springer-Verlag, London. Available at probability.ca/MT, 1993.

[116] W. Feller. *An Introduction to Probability Theory and Its Applications, Volume II*. John Wiley & Sons, 1971.

[117] J. Kiefer and J. Wolfowitz. *On the Theory of Queues with Many Servers*. Trans. Amer. Math. Soc., 78:01–18, 1955.

[118] E. Morozov. *Stochastic Boundness of Some Queueing Processes*. Sci.Report R-95-2022, Dept.of Math.and Computer Science, Aalborg Univ., Denmark, 1995.

[119] E. Morozov and R. Delgado. *Stability Analysis of Regenerative Queues*. Sci. Report Centre de Recerca Matemàtica (CRM), 812:1–33, 2008.