Wiley Series in Communications Networking & Distributed Systems

Data Center Case Study Included

Miguel Barreiros • Peter Lundqvist

SECOND EDITION

QOS-ENABLED NETWORKS

Tools and Foundations

with a Foreword by **Jeff Doyle**

WILEY

QOS-ENABLED NETWORKS

TOOLS AND FOUNDATIONS

SECOND EDITION

Miguel Barreiros
Juniper Networks, Portugal

Peter Lundqvist
Arista Networks, Sweden

WILEY

This edition first published 2016

First Edition published in 2011

Registered Office
John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ,
United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. It is sold on the understanding that the publisher is not engaged in rendering professional services and neither the publisher nor the author shall be liable for damages arising herefrom. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication data applied for

ISBN: 9781119109105

A catalogue record for this book is available from the British Library.

Set in 11/14pt Times by SPi Global, Pondicherry, India
Printed and bound in Singapore by Markono Print Media Pte Ltd

1 2016

About the Authors

Miguel Barreiros is the Data Center Practice Lead at Juniper Networks responsible for the EMEA region. Previously he was a Senior Solutions Consultant focused on both data centers and IP/MPLS networks. Since he joined Juniper Networks in 2006, he has focused on the creation and development of solutions and has been involved in projects that span all stages of building and expanding networks, from design and testing to implementation and ongoing maintenance. He began his networking career in 2000, when as a hobby he was a network administrator for a British multiplayer gaming website that hosted network servers for various video games. Miguel has a B.Sc. degree in Electronics and Computer Engineering from Instituto Superior Técnico. He holds Juniper Networks Certificate Internet Expert (JNCIE) 193 and is a Juniper Networks Certified Instructor.

Peter Lundqvist is a System Engineer in Arista Networks since 2014, focusing on Datacenter solutions. Previously Peter worked at Juniper Networks in various roles including Juniper Networks Strategic Alliance group with a focus on packet-based mobile networks. Earlier Peter was a member of the Juniper Networks Beta Engineering team, which is responsible for early field testing of new hardware products and software versions. Peter's focus was on routing and control plane protocols. Before joining Juniper in 2000, Peter worked at Cisco Systems as a Consulting Engineer and at Ericsson. Peter has a degree in Systems and Information Technology from Mittuniversitetet (Mid Sweden University). He holds Juniper Networks Certificate Internet Expert (JNCIE) 48 and Cisco Certified Internetwork Expert (CCIE) 3787.

Foreword

Network consolidation has been with us since the 1990s, driven by the simple requirement to reduce the costs of business communication. For IT, it is a matter of controlling CapEx and OpEx. For service providers, it is a matter of offering multiservice solutions at a competitive cost. (Remember when "triple play" was the buzzword of the day?) Consolidation has been so successful that you seldom encounter an organization these days that runs separate data and telephony networks. Voice and video over IP is proven, reliable, and cheap. And modern service providers—whether they got their start as a telephony, cable, long distance, or Internet provider—now run all of their services over an IP core.

Treating all communications as data, and sending it all over a shared IP infrastructure—or series of IP infrastructures—has also revolutionized our modern lives from smart phones to shopping to entertainment to travel. For myself, one of the most interesting impacts of technology has been how different my teenagers' social lives are from my own when I was a teenager. Their activities are more spontaneous, their social groups are larger, and always-available communications make their activities safer.

And consolidation is still evolving. These days the excitement is around virtualization, improving the utilization of our existing communications resources.

From the beginning, one of the biggest challenges of consolidating all communications onto an IP infrastructure stems from the fact that not all data is equal. As users we expect a certain Quality of Experience (QOE) related to the service we're using. So QOE for voice is different than QOE for videoconferencing, both of which are different from high-definition entertainment. Each kind of data stream requires different treatment within the network to meet users' QOE expectations, and that's where Quality of Service (QOS) comes in.

QOS has been around as long as IP has. The IP packet header has a Type of Service (TOS) field for differentiating services, and over the years that field has evolved into the more sophisticated Differentiated Services Code Point (DSCP) field to better fit modern QOS classification strategies. And from the beginning it was understood that although IP provides connectionless best-effort delivery, some applications need reliable, sequenced, connection-oriented delivery. Hence TCP, which "fakes" the behavior of a wired-up point-to-point connection over IP.

QOS is really all about managing limited network resources. You don't get extra bandwidth or faster delivery; you just get to decide what data gets first dibs at the available resources. High-Def video requires very prompt delivery. A web page can wait a bit longer, and e-mail can wait much longer still. Over the years, QOS technologies and strategies have become more and more sophisticated to deal with the diversity of applications using the network. Routers and switches have better and better queues and queuing algorithms, better ingress control mechanisms, and better queue servicing mechanisms. And the advent of Software-Defined Networking (SDN) introduces some new and interesting ways of improving QOE.

All of this growing sophistication brings with it growing complexity for network architects and engineers. There are a lot of choices and a lot of knobs, and if you don't have the understanding to make the right choices and set the right knobs, you can do some serious damage to the overall quality of the network. Or at the least, you can fail to utilize your network's capabilities as well as you should.

That's where this book comes in. My longtime friends Miguel Barreiros and Peter Lundqvist have deep experience designing modern QOS strategies, and they share that experience in this book, from modern QOS building blocks to applied case studies. They'll equip you well for designing the best QOS approach for your own network.

Jeff Doyle

Preface

Five years have elapsed between the original publishing of this book and this second edition, and it is unquestionably interesting to analyze what has changed. The original baseline was that Quality of Service, or QOS, was in the spotlight. Five years have elapsed and QOS prominence has just kept on growing. It has entered in new realms like the Data Center and also spread into new devices. It is no longer just switches and routers—now even servers have at their disposal a complete QOS toolkit to deal, for example, with supporting multiple virtual machines.

This book's focus remains in the roots and foundations of the QOS realm. Knowledge of the foundations of QOS is the key to understanding what benefits it offers and what can be built on top of it. This knowledge will help the reader engage in both the conceptual and actual tasks of designing or implementing QOS systems, thinking in terms of the concepts, rather than thinking of QOS simply as a series of commands that should be pasted into the configuration of the devices. It will also help the reader to troubleshoot a QOS network, to decide whether the undesired results being seen are a result of misconfigured tools that require some fine-tuning or the wrong tools. As Galileo Galilei once said, "Doubt is the father of all invention."

A particular attention is also dedicated to special traffic types and networks, and three case studies are provided where the authors share their experience in terms of practical deployments of QOS.

Although the authors work for two specific vendors, this book is completely vendor agnostic, and we have shied away from showing any CLI output or discussing hardware-specific implementations.

History of This Project

The idea behind this book started to take shape in 2007, when Miguel engaged with British Telecom (BT) in several workshops about QOS. Several other workshops and training initiatives followed, and the material presented matured and stabilized over time. In July 2009, Miguel and Peter, who had also developed various QOS workshop and training guides, joined together to work on this project which led to the creation of the first edition.

In December 2014, both authors agreed that the book needed a revamp to cover the new challenges posed in the Data Center realm, which originated this second edition.

Who Should Read This Book?

The target audience for this book are network professionals from both the enterprise and the service provider space who deal with networks in which QOS is present or in which a QOS deployment is planned. Very little knowledge of other areas of networking is necessary to benefit from this book, because as the reader will soon realize, QOS is indeed a world of its own.

Structure of the Book

This book is split into three different parts following the Julius Caesar approach ("Gallia est omnis divisa in partes tres"):

Part One provides a high-level overview of the QOS tools. It also discusses the challenges within the QOS realm and certain types of special traffic and networks.

Part Two dives more deeply into the internal mechanisms of the important QOS tools. It is here that we analyze the stars of the QOS realm.

Part Three glues back together all the earlier material in the book. We present three case studies consisting of end-to-end deployments: the first focused on VPLS, the second focused on Data Center, and the third one focused on the mobile space.

Have fun.

Miguel Barreiros, *Sintra, Portugal*
Peter Lundqvist, *Tyresö, Sweden*
April 30, 2015

Acknowledgments

Several colleagues of ours have helped us during this project. However, three were absolutely key to the creation of this book:

Patrick Ames led this project in all its nontechnical aspects, allowing Miguel and Peter to focus on writing.

Aviva Garrett played the key role of the editorial review of the entire book and also guided Miguel and Peter in how to improve the book's organization and contents.

Steven Wong (Juniper JTAC) provided technical reviews for much of the book while sharing his immense knowledge of QOS.

The authors would also like to gratefully thank the following people: Antoine Sibout, Bobby Vandalore, Guy Davies, Jeff Doyle, John Peach, Jos Bazelmans, Kannan Kothandaraman, Margarida Correia, and Pedro Marques.

Finally, the authors would like to express their personal thanks:

Miguel: I would like to dedicate this book to Maria Eugénia Barreiros and to my grandparents José Silva and Dores Vilela.

Peter: My work on this book was possible only with the understanding and patience of the most dear ones in my life, my wife Lena and my great kids Ida and Oskar.

Abbreviations

2G	Second Generation
3GPP	Third-Generation Partnership Project
ACK	Acknowledgment
AF	Assured-forwarding
APN	Access Point Name
AUC	Authentication Center
BA	behavior aggregate
BE	best-effort
BHT	Busy Hour Traffic
Bps	bits per second
BSC	Base Station Controller
BSR	Broadband Service Router
BTS	Base Transceiver Station
BU	business
CDMA	Code Division Multiple Access
CEIR	Central Equipment Identity Register
CIR	Committed Information Rate
CLI	Command Line Interface
CNTR	control traffic
CoS	Class of Service
CT	class type
CWND	congestion window
DA	data
DF	Don't Fragment
DHCP	Dynamic Host Configuration Protocol
DiffServ	Differentiated Services

DNS	Domain Name System
DRR	Deficit Round Robin
DSCP	Differentiated Services Code Point
DSL	Digital Subscriber Line
DSLAM	Digital Subscriber Line Access Multiplexer
DWRR	Deficit Weighted Round Robin
EBGP	External Border Gateway Protocol
EF	Expedited-forwarding
EIR	Equipment Identity Register
EPC	Evolved Packet Core
ERO	Explicit Routing Object
eUTRAN	evolved UMTS Terrestrial Radio Access Network
FIFO	First in, first out
FQ	Fair queuing
GBR	Guaranteed Bit Rate
GGSN	Gateway GPRS Support Node
GPRS	General Packet Radio Service
GPS	Generic Processor Sharing
GSM	Global System for Mobile Communications
GTP	GPRS Tunneling Protocol
HLR	Home Location Register
ICMP	Internet Control Message Protocol
IMEI	International Mobile Equipment Identity
IMS	IP Multimedia System
IMSI	International Mobile Subscriber Identity
IntServ	Integrated Services
L2	Layer 2
L3	Layer 3
LBE	lower than that for best-effort
LFI	Link Fragmentation and Interleaving
LSP	label-switched path
LTE	Long-Term Evolution
MAD	dynamic memory allocation
ME	Mobile Equipment
MED	multi-exit discriminator
MF	Multifield
MME	Mobility Management Entity
MPLS	Multiprotocol Label Switching
MPLS-TE	MPLS network with traffic engineering

MS	Mobile System
ms	milliseconds
MSC	Mobile Switching Center
MSS	Maximum Segment Size
MTU	Maximum Transmission Unit
NAS	Non-Access Stratum
NC	Network-control
P2P	point-to-point
PB-DWRR	Priority-based deficit weighted round robin
PCR	Program Clock Reference
PCRF	Policy and Charging Rules Function
PDN	Packet Data Networks
PDN-GW	Packet Data Network Gateway
PDP	Packet Data Protocol
PE	provider edge
PHB	per-hop behavior
PID	Packet ID
PIR	peak information rate
PLMN	Public LAN Mobile Network
PMTU	Path MTU
pps	packets per second
PQ	priority queuing
PSTN	Public Switched Telephone Network
Q0	queue zero
Q1	queue one
Q2	queue two
QCI	QOS Class Identifier
QOS	Quality of Service
RAN	Radio Access Networks
RED	Random Early Discard
RNC	Radio Network Controller
RSVP	Resource Reservation Protocol
RT	real time
RTCP RTP	Control Protocol
RTT	Round Trip Time
SACK	selective acknowledgment
SAE	System Architecture Evolution
SCP	Secure Shell Copy
SCTP	Stream Control Transmission Protocol

SDP	Session Description Protocol
SGSN	Serving GPRS Support Node
S-GW	Serving Gateway
SIM	Subscriber Identity Module
SIP	Session Initiation Protocol
SLA	service-level agreement
SSRC	Synchronization Source Identifier
STP	Spanning Tree Protocols
TCP	Transmission Control Protocol
TE	Traffic Engineering
TOS	Type of Service
TS	Transport Stream
UDP	USER Datagram Protocol
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
UTP	Unshielded Twisted Pair
VLAN	Virtual LAN
VLR	Visitor Location Register
VoD	Video on Demand
VoIP	Voice over IP
VPLS	Virtual Private LAN Service
VPN	Virtual Private Network
WFQ	Weighted Fair Queuing
WRED	Weighted RED
WRR	Weighted Round Robin

Contents