

Next-Generation Internet Architectures and Protocols

With ever-increasing demands on capacity, quality of service, speed, and reliability, current Internet systems are under strain and under review. Combining contributions from experts in the field, this book captures the most recent and innovative designs, architectures, protocols, and mechanisms that will enable researchers to successfully build the next-generation Internet. A broad perspective is provided, with topics including innovations at the physical/transmission layer in wired and wireless media, as well as the support for new switching and routing paradigms at the device and sub-system layer. The proposed alternatives to TCP and UDP at the data transport layer for emerging environments are also covered, as are the novel models and theoretical foundations proposed for understanding network complexity. Finally, new approaches for pricing and network economics are discussed, making this ideal for students, researchers, and practitioners who need to know about designing, constructing, and operating the next-generation Internet.

Byrav Ramamurthy is an Associate Professor in the Department of Computer Science and Engineering at the University of Nebraska, Lincoln (UNL). He is a recipient of the UNL CSE Department Student Choice Outstanding Teaching Award, the author of *Design of Optical WDM Networks* (2000), and a co-author of *Secure Group Communications over Data Networks* (2004). His research areas include optical networks, wireless/sensor networks, network security, distributed computing, and telecommunications.

George N. Rouskas is a Professor of Computer Science at North Carolina State University. He has received several research and teaching awards, including an NSF CAREER Award, the NCSU Alumni Outstanding Research Award, and he has been inducted into the NCSU Academy of Outstanding Teachers. He is the author of the book *Internet Tiered Services* (2009) and his research interests are in network architectures and protocols, optical networks and performance evaluation.

Krishna Moorthy Sivalingam is a Professor in the Department of Computer Science and Engineering at the Indian Institute of Technology (IIT), Madras. Prior to this, he was a Professor at the University of Maryland, Baltimore County, and also conducted research at Lucent Technologies and AT&T Bell Labs. He has previously edited three books and holds three patents in wireless networking. His research interests include wireless networks, wireless sensor networks, optical networks, and performance evaluation.

Next-Generation Internet: Architectures and Protocols

Edited by

BYRAV RAMAMURTHY
University of Nebraska-Lincoln

GEORGE ROUSKAS
North Carolina State University

KRISHNA M. SIVALINGAM
Indian Institute of Technology Madras

 **CAMBRIDGE**
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi,
Dubai, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org

Information on this title: www.cambridge.org/9780521113687

© Cambridge University Press 2011

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without the written
permission of Cambridge University Press.

First published 2011

Printed in the United Kingdom at the University Press, Cambridge

A catalog record for this publication is available from the British Library

Library of Congress Cataloging in Publication data

ISBN 978-0-521-11368-7 Hardback

Cambridge University Press has no responsibility for the persistence or
accuracy of URLs for external or third-party internet websites referred to in
this publication, and does not guarantee that any content on such websites is,
or will remain, accurate or appropriate.

To my mother, Mrs. Lalitha Ramamurthy – BR
To Magdalini and Alexander – GNR
To my family – KMS

Contents

<i>Contributors</i>	<i>page</i> xvi
<i>Preface</i>	xix
Part I Enabling technologies	1
1 Optical switching fabrics for terabit packet switches	3
Davide Cuda, Roberto Gaudino, Guido A. Gavilanes Castillo, and Fabio Neri	
1.1 Optical switching fabrics	5
1.1.1 Wavelength-selective (WS) architecture	7
1.1.2 Wavelength-routing (WR) architecture	8
1.1.3 Plane-switching (PS) architecture	9
1.2 Modeling optical devices	10
1.2.1 Physical model	11
1.2.2 Device characterization	12
1.2.3 Multi-plane-specific issues	15
1.3 Scalability analysis	16
1.4 Cost analysis	18
1.5 Results	21
1.5.1 Scalability of the aggregate switching bandwidth	21
1.5.2 CAPEX estimation	23
1.6 Conclusions	24
<i>References</i>	25
2 Broadband access networks: current and future directions	27
Abu (Sayeem) Reaz, Lei Shi, and Biswanath Mukherjee	
2.1 Introduction	27
2.1.1 Current broadband access solutions	27
2.1.2 Passive Optical Network (PON)	28
2.1.3 Extending the reach: Long-Reach PON (LR-PON)	30
2.2 Technologies and demonstrations	32
2.2.1 Enabling technologies	32
2.2.2 Demonstrations of LR-PON	33

2.3	Research challenges in LR-PON	34
2.3.1	Low-cost devices: colorless ONU	34
2.3.2	Resource allocation: DBA with Multi-Thread Polling	34
2.3.3	Traffic management: behavior-aware user assignment	35
2.4	Reaching the end-users: Wireless-Optical Broadband Access Network (WOBAN)	36
2.4.1	WOBAN architecture	36
2.4.2	Motivation of WOBAN	37
2.4.3	Research challenges in WOBAN	38
2.5	Conclusion	39
	<i>References</i>	39
3	The optical control plane and a novel unified control plane architecture for IP/WDM networks	42
	Georgios Ellinas, Antonis Hadjiantonis, Ahmad Khalil, Neophytos Antoniadis, and Mohamed A. Ali	
3.1	Introduction	42
3.2	Overview of optical control plane design	43
3.2.1	Link Management Protocol	44
3.2.2	GMPLS routing protocol	44
3.2.3	GMPLS signaling protocol	46
3.3	IP-over-WDM networking architecture	47
3.3.1	The overlay model	48
3.3.2	The peer and augmented models	48
3.4	A new approach to optical control plane design: an optical layer-based unified control plane architecture	49
3.4.1	Node architecture for the unified control plane	50
3.4.2	Optical layer-based provisioning	51
3.5	Conclusions	68
	<i>References</i>	68
4	Cognitive routing protocols and architecture	72
	Suyang Ju and Joseph B. Evans	
4.1	Introduction	72
4.2	Mobility-aware routing protocol	73
4.2.1	Background	73
4.2.2	Approach	74
4.2.3	Benefits	77
4.2.4	Protocol architecture	78
4.3	Spectrum-aware routing protocol	79
4.3.1	Background	79
4.3.2	Approach	80

4.3.3	Benefits	83
4.3.4	Protocol architecture	84
4.4	Conclusion	84
	<i>References</i>	85
5	Grid networking	88
	Anusha Ravula and Byrav Ramamurthy	
5.1	Introduction	88
5.2	The Grid	89
5.2.1	Grid Computing	89
5.2.2	Lambda Grid networks	90
5.3	Cloud Computing	91
5.4	Resources	92
5.4.1	Grid network resources	92
5.4.2	Optical network testbeds and projects	92
5.4.3	Computational resources	94
5.4.4	Other resources	95
5.5	Scheduling	95
5.6	Optical Circuit Switching and Optical Burst Switching	98
5.6.1	Studies on OCS-based Grids	98
5.6.2	Studies on OBS based Grids	100
5.7	Conclusion	101
	<i>References</i>	102
	Part II Network architectures	105
6	Host identity protocol (HIP): an overview	107
	Pekka Nikander, Andrei Gurtov, and Thomas R. Henderson	
6.1	Introduction	107
6.2	Fundamental problems in the Internet today	108
6.2.1	Loss of universal connectivity	109
6.2.2	Poor support for mobility and multi-homing	109
6.2.3	Unwanted traffic	109
6.2.4	Lack of authentication, privacy, and accountability	110
6.3	The HIP architecture and base exchange	110
6.3.1	Basics	111
6.3.2	HITs and LSIs	112
6.3.3	Protocols and packet formats	113
6.3.4	Detailed layering	117
6.3.5	Functional model	118
6.3.6	Potential drawbacks	120

6.4	Mobility, multi-homing, and connectivity	121
6.4.1	HIP-based basic mobility and multi-homing	121
6.4.2	Facilitating rendezvous	122
6.4.3	Mobility between addressing realms and through NATs	123
6.4.4	Subnetwork mobility	124
6.4.5	Application-level mobility	126
6.5	Privacy, accountability, and unwanted traffic	126
6.5.1	Privacy and accountability	126
6.5.2	Reducing unwanted traffic	127
6.6	Current status of HIP	129
6.7	Summary	131
	<i>References</i>	131
7	Contract-Switching for Managing Inter-Domain Dynamics	136
	Murat Yuksel, Aparna Gupta, Koushik Kar, and Shiv Kalyanaraman	
7.1	Contract-switching paradigm	137
7.2	Architectural issues	138
7.2.1	Dynamic contracting over peering points	139
7.2.2	Contract routing	139
7.3	A contract link: bailouts and forwards	143
7.3.1	Bailout forward contract (BFC)	144
7.3.2	Formalization for pricing a bailout forward contract (BFC)	144
7.3.3	Bailout forward contract (BFC) performance evaluation	147
7.4	Summary	152
	<i>References</i>	153
8	PHAROS: an architecture for optical networks	154
	Iliia Baldine, Alden W. Jackson, John Jacob, Will E. Leland, John H. Lowry, Walker C. Milliken, Partha P. Pal, Subramanian Ramanathan, Kristin Rauschenbach, Cesar A. Santivanez, and Daniel M. Wood	
8.1	Introduction	154
8.2	Background	157
8.3	PHAROS architecture: an overview	157
8.4	Resource allocation	161
8.4.1	Resource management strategies	161
8.4.2	Protection	164
8.4.3	Playbooks	166
8.4.4	Sub-lambda grooming	168
8.5	Signaling system	169
8.5.1	Control plane operation	171
8.5.2	Failure notification	172

8.6	Core node implementation	173
8.7	Performance analysis	175
8.8	Concluding remarks	176
	<i>References</i>	177
9	Customizable in-network services	179
	Tilman Wolf	
9.1	Background	179
9.1.1	Internet architecture	179
9.1.2	Next-generation Internet	180
9.1.3	Data path programmability	180
9.1.4	Technical challenges	181
9.1.5	In-network processing solutions	181
9.2	Network services	182
9.2.1	Concepts	182
9.2.2	System architecture	184
9.3	End-system interface and service specification	186
9.3.1	Service pipeline	186
9.3.2	Service composition	187
9.4	Routing and service placement	188
9.4.1	Problem statement	188
9.4.2	Centralized routing and placement	189
9.4.3	Distributed routing and placement	190
9.5	Runtime resource management	191
9.5.1	Workload and system model	191
9.5.2	Resource management problem	192
9.5.3	Task duplication	192
9.5.4	Task mapping	193
9.6	Summary	194
	<i>References</i>	194
10	Architectural support for continuing Internet evolution and innovation	197
	Rudra Dutta and Ilia Baldine	
10.1	Toward a new Internet architecture	197
10.2	The problems with the current architecture	199
10.3	SILO architecture: design for change	201
10.4	Prior related work	206
10.5	Prototype and case studies	207

10.6	Future work: SDO, stability, virtualization, silo-plexes	208
10.6.1	Virtualization	208
10.6.2	SDO: “software defined optics”	211
10.6.3	Other open problems	212
10.7	Case study	213
	Acknowledgements	214
	<i>References</i>	214
Part III Protocols and practice		217
11	Separating routing policy from mechanism in the network layer	219
	James Griffioen, Kenneth L. Calvert, Onur Ascigil, and Song Yuan	
11.1	Introduction	219
11.2	PoMo design goals	220
11.3	Architecture overview	222
11.3.1	PFRI network structure and addressing	222
11.3.2	PFRI forwarding	223
11.3.3	PFRI routing policies	225
11.3.4	PFRI packet header mechanisms	226
11.4	Scaling the PFRI Architecture	227
11.5	Discussion	230
11.6	Experimental evaluation	232
11.7	Other clean-slate approaches	234
	Acknowledgements	235
	<i>References</i>	235
12	Multi-path BGP: motivations and solutions	238
	Francisca Valera, Iljtsch van Beijnum, Alberto García-Martínez, Marcelo Bagnulo	
12.1	Introduction	238
12.2	Trilogy project	239
12.2.1	Objectives	239
12.2.2	Trilogy technologies	240
12.3	Multi-path routing	241
12.3.1	Higher network capacity	242
12.3.2	Scalable traffic engineering capabilities	242
12.3.3	Improved response to path changes	242
12.3.4	Enhanced security	243
12.3.5	Improved market transparency	243
12.4	Multi-path BGP	244
12.4.1	Intra-domain multi-path routing	244
12.4.2	Inter-domain multi-path routing	245

	12.4.3 Motivations for other solutions	247
	12.4.4 mBGP and MpASS	248
	12.5 Conclusions and future work	253
	<i>References</i>	254
13	Explicit congestion control	257
	Frank Kelly and Gaurav Raina	
	13.1 Fairness	258
	13.1.1 Why proportional fairness?	260
	13.2 Proportionally fair rate control protocol	260
	13.2.1 Sufficient conditions for local stability	263
	13.2.2 Illustrative simulation	264
	13.2.3 Two forms of feedback?	264
	13.2.4 Tatonnement processes	265
	13.3 Admission management	265
	13.3.1 Step-change algorithm	266
	13.3.2 Robustness of the step-change algorithm	267
	13.3.3 Guidelines for network management	268
	13.3.4 Illustrating the utilization–robustness tradeoff	269
	13.3.5 Buffer sizing and the step-change algorithm	270
	13.4 Concluding remarks	272
	<i>References</i>	273
14	KanseiGenie: software infrastructure for resource management and programmability of wireless sensor network fabrics	275
	Mukundan Sridharan, Wenjie Zeng, William Leal, Xi Ju, Rajiv Ramnath, Hongwei Zhang, and Anish Arora	
	14.1 Introduction	275
	14.2 Features of sensing fabrics	278
	14.2.1 Generic Services	278
	14.2.2 Domain-specific services	283
	14.3 KanseiGenie architecture	284
	14.3.1 The fabric model	284
	14.3.2 KanseiGenie Architecture	285
	14.3.3 GENI extension to KanseiGenie	287
	14.3.4 Implementation of KanseiGenie	288
	14.3.5 KanseiGenie federation	290
	14.4 KanseiGenie customization and usage	292
	14.4.1 How to customize KanseiGenie	292
	14.4.2 Vertical APIs and their role in customization	293
	14.4.3 KanseiGenie usage step-by-step run-through	294

14.5	Evolving research issues in next-generation networks	295
14.5.1	Resource specifications for sensor fabrics	295
14.5.2	Resource discovery	296
14.5.3	Resource allocation	296
14.5.4	Data as resource	297
14.5.5	Network virtualization	297
14.6	Conclusion	298
	<i>References</i>	298
Part IV	Theory and models	301
15	Theories for buffering and scheduling in Internet switches	303
	Damon Wischik	
15.1	Introduction	303
15.2	Buffer sizing and end-to-end congestion control	304
15.2.1	Four heuristic arguments about buffer sizing	305
15.2.2	Fluid traffic model and queue model	307
15.2.3	Queueing delay, utilization, and synchronization	309
15.2.4	Traffic burstiness	312
15.3	Queueing theory for switches with scheduling	313
15.3.1	Model for a switched network	313
15.3.2	The capacity region, and virtual queues	314
15.3.3	Performance analysis	315
15.4	A proposed packet-level architecture	320
	<i>References</i>	323
16	Stochastic network utility maximization and wireless scheduling	324
	Yung Yi and Mung Chiang	
16.1	Introduction	324
16.2	LAD (Layering As optimization Decomposition)	326
16.2.1	Background	326
16.2.2	Key ideas and procedures	327
16.3	Stochastic NUM (Network Utility Maximization)	328
16.3.1	Session-level dynamics	328
16.3.2	Packet-level dynamics	332
16.3.3	Constraint-level dynamics	334
16.3.4	Combinations of multiple dynamics	336
16.4	Wireless scheduling	337
16.4.1	Collision-free algorithms	339
16.4.2	Collision-based algorithms	342

	16.4.3 Performance–complexity tradeoff	346
	16.4.4 Future research directions	350
	<i>References</i>	351
17	Network coding in bi-directed and peer-to-peer networks	359
	Zongpeng Li, Hong Xu, and Baochun Li	
	17.1 Network coding background	359
	17.2 Network coding in bi-directed networks	361
	17.2.1 Single multicast in undirected networks	361
	17.2.2 The linear programming perspective	365
	17.2.3 Single multicast in Internet-like bi-directed networks	366
	17.2.4 Towards tighter bounds	367
	17.2.5 Multiple communication sessions	367
	17.2.6 The source independence property of multicast	368
	17.3 Network coding in peer-to-peer networks	369
	17.3.1 Peer-assisted content distribution with network coding	369
	17.3.2 Peer-assisted media streaming with network coding	371
	17.4 Conclusions	374
	<i>References</i>	375
18	Network economics: neutrality, competition, and service differentiation	378
	John Musacchio, Galina Schwartz, and Jean Walrand	
	18.1 Neutrality	380
	18.1.1 Model	381
	18.1.2 The analysis of one- and two-sided pricing	384
	18.1.3 User welfare and social welfare	386
	18.1.4 Comparison	386
	18.1.5 Conclusions	389
	18.2 Competition	390
	18.2.1 Model	392
	18.2.2 Circuit analogy	393
	18.3 Service differentiation	398
	Acknowledgement	400
	<i>References</i>	400
	<i>About the Editors</i>	403
	<i>Index</i>	405

Contributors

Mohamed A. Ali, City University of New York, USA
Neophytos Antoniadis, City University of New York, USA
Anish Arora, The Ohio State University, USA
Onur Ascigil, University of Kentucky, USA
Marcelo Bagnulo, Universidad Carlos III de Madrid, Spain
Ilia Baldine, Renaissance Computing Institute, USA
Kenneth L. Calvert, University of Kentucky, USA
Mung Chiang, Princeton University, USA
Davide Cuda, Politecnico di Torino, Turin, Italy
Rudra Dutta, North Carolina State University, USA
Georgios Ellinas, University of Cyprus, Cyprus
Joseph B. Evans, University of Kansas, USA
Alberto García-Martínez, Universidad Carlos III de Madrid, Spain
Roberto Gaudino, Politecnico di Torino, Turin, Italy
Guido A. Gavilanes Castillo, Politecnico di Torino, Turin, Italy
James Griffioen, University of Kentucky, USA
Aparna Gupta, Rensselaer Polytechnic Institute, USA
Andrei Gurtov, University of Oulu, Finland
Antonis Hadjiantonis, University of Nicosia, Cyprus
Thomas R. Henderson, Boeing, USA
Alden W. Jackson, BBN Technologies, USA
John Jacob, BAE Systems, USA
Suyang Ju, University of Kansas, USA
Xi Ju, The Ohio State University, USA
Shiv Kalyanaraman, IBM Research, India
Koushik Kar, Rensselaer Polytechnic Institute, USA
Frank Kelly, University of Cambridge, UK
Ahmad Khalil, City University of New York, USA
William Leal, The Ohio State University, USA
Will E. Leland, BBN Technologies, USA
Baochun Li, University of Toronto, Canada
Zongpeng Li, University of Calgary, Canada
John H. Lowry, BBN Technologies, USA
Walter C. Milliken, BBN Technologies, USA

Biswanath Mukherjee, University of California Davis, USA
John Musacchio, University of California Santa Cruz, USA
Fabio Neri, Politecnico di Torino, Turin, Italy
Pekka Nikander, Ericsson Research, Finland
P. Pal, BBN Technologies, USA
Gaurav Raina, Indian Institute of Technology Madras, India
Byrav Ramamurthy, University of Nebraska-Lincoln, USA
Subramanian Ramanathan, BBN Technologies, USA
Rajiv Ramnath, The Ohio State University, USA
K. Rauschenbach, BBN Technologies, USA
Anusha Ravula, University of Nebraska-Lincoln, USA
Abu (Sayeem) Reaz, University of California Davis, USA
Ceaser A. Santivanez, BBN Technologies, USA
Galina Schwartz, University of California Berkeley, USA
Lei Shi, University of California Davis, USA
Mukundan Sridharan, The Ohio State University, USA
Franciso Valera, Universidad Carlos III de Madrid, Spain
Iljitsch van Beijnum, IMDEA Networks, Spain
Jean Walrand, University of California Berkeley, USA
Damon Wischik, University College London, UK
Daniel Wood, Verizon Network Systems, USA
Tilman Wolf, University of Massachusetts Amherst, USA
Hong Xu, University of Toronto, Canada
Yung Yi, Princeton University, USA
Song Yuan, University of Kentucky, USA
Murat Yuksel, University of Nevada - Reno, USA
Wenjie Zeng, The Ohio State University, USA
Hongwei Zhang, The Ohio State University, USA

