Table of Contents

Introduction 13

Chapter 1. From Multiprocessor Computers to the Clouds 21

1.1. The explosion of demand for computing power 21
1.2. Computer clusters 24
1.2.1. The emergence of computer clusters 24
1.2.2. Anatomy of a computer cluster 24
1.3. Computing grids 26
1.3.1. High-performance computing grids 29
1.3.2. Peer-to-peer computing grids 30
1.4. Computing in a cloud 32
1.5. Conclusion 36

Chapter 2. Utilization of Network Computing Technologies 39

2.1. Anatomy of a distributed computing application 39
2.1.1. Parallelization and distribution of an algorithm . . 41
2.1.1.1. Embarrassingly parallel applications 42
2.1.1.2. Fine-grained parallelism 43
2.1.2. Modeling parallel applications 44
2.1.3. Example of a grid application 44
2.1.4. General classification of distributed applications 47

2.1.4.1. Widely distributed computing 48
2.1.4.2. Loosely coupled computing 49
2.1.4.3. Pipeline computing 50
2.1.4.4. Highly synchronized computing 50
2.1.4.5. Interactive and collaborative computing . . . 51
2.1.4.6. Note 51
2.2. Programming models of distributed parallel applications 52
2.2.1. Main models 52
2.2.2. Constraints of fine-grained-parallelism applications 53
2.2.3. The MPI communication library 54
2.3. Coordination of distributed resources in a grid 57
2.3.1. Submission and execution of a distributed application 57
2.3.2. Grid managers 59
2.4. Conclusion 60

Chapter 3. Specificities of Computing Networks 63

3.1. Typology of computing networks 63
3.1.1. Cluster networks 65
3.1.2. Grid networks 65
3.1.3. Computing cloud networks 67
3.2. Network transparency 68
3.2.1. The advantages of transparency 68
3.2.2. Foundations of network transparency 69
3.2.3. The limits of TCP and IP in clusters 72
3.2.4. Limits of TCP and network transparency in grids 75
3.2.5. TCP in a high bandwidth-delay product network 75
3.2.6. Limits of the absence of communication control 76
3.3. Detailed analysis of characteristics expected from protocols 78
3.3.1. Topological criteria 78

3.3.1.1. Number of sites involved 78
3.3.1.2. Number of users involved 79
3.3.1.3. Resource-localization constraints 79
3.3.2. Performance criteria 80
3.3.2.1. Degree of inter-task coupling 80
3.3.2.2. Sensitivity to latency and throughput 81
3.3.2.3. Sensitivity to throughput and its control . . . 83
3.3.2.4. Sensitivity to confidentiality and security . . 84
3.3.2.5. Summary of requirements 84
3.4. Conclusion 85

Chapter 4. The Challenge of Latency in Computing Clusters 87
4.1. Key principles of high-performance networks for clusters 88
4.2. Software support for high-performance networks 90
4.2.1. Zero-copy transfers 90
4.2.2. OS-bypass 90
4.2.3. Event notification 91
4.2.4. The problem of address translation 93
4.2.5. Non-blocking programming models 95
4.2.5.1. Case 1: message-passing 96
4.2.5.2. Case 2: remote access model 97
4.3. Description of the main high-performance networks . . 99
4.3.1. Dolphins SCI 99
4.3.2. Myricom Myrinet and Myri-10G 100
4.3.3. Quadrics QsNet 104
4.3.4. InfiniBand 105
4.3.5. Synthesis of the characteristics of high-performance networks 107
4.4. Convergence between fast and traditional networks . . 108
4.5. Conclusion 111

Chapter 5. The Challenge of Throughput and Distance . . . 113
5.1. Obstacles to high rate 113
5.2. Operating principle and limits of TCP congestion control 115

5.2.1. Slow Start 116
5.2.2. Congestion avoidance 117
5.2.3. Fast Retransmit 117
5.2.4. Analytical model 119
5.3. Limits of TCP over long distances 120
5.4. Configuration of TCP for high speed 122
5.4.1. Hardware configurations 123
5.4.2. Software configuration 124
5.4.3. Parameters of network card drivers 126
5.5. Alternative congestion-control approaches to that of standard TCP 126
5.5.1. Use of parallel flows 127
5.5.2. TCP modification 129
5.5.2.1. Slow Start modifications 129
5.5.2.2. Methods of congestion detection 130
5.5.2.3. Bandwidth-control methods 131
5.5.3. UDP-based approaches 132
5.6. Exploration of TCP variants for very high rate 133
5.6.1. HighSpeed TCP 133
5.6.2. Scalable 134
5.6.3. BIC-TCP 134
5.6.4. H-TCP 135
5.6.5. CUBIC 135
5.7. Conclusion 136

Chapter 6. Measuring End-to-End Performances 139

6.1. Objectives of network measurement and forecast in a grid 139
6.1.1. Illustrative example: network performance and data replication 140
6.1.2. Objectives of a performance-measurement system in a grid 143
6.2. Problem and methods 144
6.2.1. Terminology 145
6.2.2. Inventory of useful characteristics in a grid 149
6.2.3. Measurement methods 152

6.2.3.1. Active method 152
6.2.3.2. Passive method 152
6.2.3.3. Measurement tools 154
6.3. Grid network-performance measurement systems . . . 155
6.3.1. e2emonit 155
6.3.2. PerfSONAR 155
6.3.3. Architectural considerations 156
6.3.4. Sensor deployment in the grid 160
6.3.5. Measurement coordination 161
6.4. Performance forecast 164
6.4.1. The Network Weather Service tool 164
6.4.2. Network-cost function 166
6.4.3. Formulating the cost function 167
6.4.4. Estimate precision 169
6.5. Conclusion 170

Chapter 7. Optical Technology and Grids 171

7.1. Optical networks and switching paradigms 172
7.1.1. Optical communications 172
7.1.1.1. Wavelength multiplexing 173
7.1.1.2. Optical add-drop multiplexers 174
7.1.1.3. Optical cross-connect 175
7.1.2. Optical switching paradigms 176
7.1.2.1. Optical packet switching 176
7.1.2.2. Optical burst switching 177
7.1.2.3. Optical circuit switching 177
7.1.3. Conclusion 179
7.2. Functional planes of transport networks 179
7.2.1. Data plane 181
7.2.2. Control plane 182
7.2.2.1. Routing 182
7.2.2.2. Signaling 182
7.2.3. Management plane 182
7.2.4. Conclusion 184
7.3. Unified control plane: GMPLS/automatic switched transport networks 184

7.3.1. Label-switching 184
7.3.2. Protocols: OSPF-TE/RSVP-TE/LMP/PCEP . . . 185
7.3.3. GMPLS service models 187
7.3.4. Conclusion 188

Chapter 8. Bandwidth on Demand 189

8.1. Current service model: network neutrality 190
8.1.1. Structure 191
8.1.2. Limits and problems 192
8.1.3. Conclusion 193
8.2. Peer model for bandwidth-delivery services 194
8.2.1. UCLP/Ca*net 194
8.2.2. GLIF 194
8.2.3. Service-oriented peer model 195
8.2.4. Conclusion 196
8.3. Overlay model for bandwidth-providing services 196
8.3.1. GNS-WSI 196
8.3.2. Carriocas 197
8.3.3. StarPlane 198
8.3.4. Phosphorus 198
8.3.5. DRAGON 198
8.3.6. Conclusion 199
8.4. Bandwidth market 200
8.5. Conclusion 201

Chapter 9. Security of Computing Networks 203

9.1. Introductory example 203
9.2. Principles and methods 205
9.2.1. Security principles 206
9.2.2. Controlling access to a resource 207
9.2.3. Limits of the authentication approach 209
9.2.4. Authentication versus authorization 210
9.2.5. Decentralized approaches 211
9.3. Communication security 212
9.4. Network virtualization and security 213

9.4.1. Classic network-virtualization approaches 213
9.4.2. The HIP protocol 215
9.5. Conclusion 216

Chapter 10. Practical Guide for the Configuration of High-speed Networks 217

10.1. Hardware configuration 218
10.1.1. Buffer memory 218
10.1.2. PCI buses 218
10.1.3. Computing power: CPU 219
10.1.4. Random access memory: RAM 220
10.1.5. Disks 220
10.2. Importance of the tuning of TCP parameters 221
10.3. Short practical tuning guide 222
10.3.1. Computing the bandwidth delay product 223
10.3.2. Software configuration 224
10.3.3. Other solutions 225
10.4. Use of multi-flow 226
10.5. Conclusion 228

Conclusion: From Grids to the Future Internet 229

Bibliography 235

Acronyms and Definitions 251

Index 263