

Contents

1. Introduction	1
1.1. Structure of this Thesis	3
1.2. Contributions	5
2. Background	9
2.1. Optical Networks	9
2.2. Software-Defined Networking	14
2.3. Network Characteristics	17
I. Circuit Switching	21
3. A Preemptive Routing and Wavelength Allocation Algorithm	23
3.1. Introduction	24
3.2. Related Work	25
3.3. Generic Graph model	27
3.4. Routing and wavelength assignment	28
3.5. Simulation	35
3.6. Conclusion	40
4. Selfish Virtual Topology Reconfiguration in IP-over-WDM	43
4.1. Introduction	44
4.2. Related Work	45
4.3. Software Defined IP-over-WDM Networks	46
4.4. Selfish Virtual Topology Reconfiguration	48
4.5. Simulation	49
4.6. Conclusion	56
II. Packet Switching	57
5. A Low-Cost, Software-Defined Data-Center Network	59
5.1. Introduction	59

5.2. Related Work	61
5.3. HybridTE	62
5.4. Evaluation	66
5.5. Conclusion	76
6. Elephant Detection	77
6.1. Introduction	77
6.2. OpenFlow Packet Sampling Extension	78
6.3. Using Packet Sampling for Elephant Detection	80
6.4. Evaluation	81
6.5. Conclusion	82
III. Emulation Tools	85
7. MaxiNet: A Scalable Network Emulator	87
7.1. Introduction	87
7.2. Related Work	89
7.3. MaxiNet	90
7.4. Performance Evaluation of MaxiNet	92
7.5. NetSLS: A MaxiNet Application	101
7.6. Conclusion	105
8. DCT²Gen: A Traffic Generator for Data Centers	107
8.1. Introduction	107
8.2. Related Work	110
8.3. Traffic Properties	112
8.4. Architecture of DCT ² Gen	115
8.5. Deconvolving Traffic Matrix Entries	118
8.6. Generating Traffic Matrices	122
8.7. Generating TCP Flows	128
8.8. Empirical Evaluation	133
8.9. Conclusion	140
9. Conclusion and Future Work	141
9.1. Conclusion	141
9.2. Future Work	142
Bibliography	145

List of Figures

2.1.	Model of an optical WDM network with three established lightpaths.	10
2.2.	Simplified working principle of a ROADM.	11
2.3.	The offline routing and wavelength allocation problem.	12
2.4.	The online routing and wavelength allocation problem.	13
2.5.	The physical and logical topology of an optical WDM network.	13
2.6.	Topology of the janos-us network.	17
2.7.	Schematic view of a fat tree topology.	18
2.8.	Schematic view of a clos-like topology.	19
3.1.	Workflow of our proposed RWA algorithm.	25
3.2.	Layered graph with W wavelengths.	28
3.3.	Flowchart of the proposed RWA algorithm.	30
3.4.	Construction of the conflict graph G_C from a lightpath configuration.	31
3.5.	Rerouting lightpaths.	33
3.6.	Unsuccessful cleanup.	33
3.7.	Influence of c on the blocking probability.	37
3.8.	Influence of cleanup on the blocking probability.	37
3.9.	Influence of the holding time on the blocking probability.	38
3.10.	Average hop count of established lightpaths.	38
3.11.	Success rate of the rerouting operation.	39
3.12.	Cardinality of each set of rerouting candidates.	39
3.13.	Blocking probability of different configurations.	40
4.1.	SDN-over-WDM Network.	46
4.2.	Germany50 network with 50 nodes and 88 links.	49
4.3.	Total traffic demand over a 24-hour period.	50
4.4.	Exemplary flow schedule for a pair of nodes.	51
4.5.	Throughput over different load levels.	53
4.6.	Maximum used physical interfaces at a single switch.	53
4.7.	Average flow entries per SDN switch.	54
4.8.	Maximum flow entries used at a single SDN switch.	54
4.9.	Throughput over different flow size limits.	54
4.10.	Hop Count over different load levels.	55

4.11. Achieved throughput over iterations.	55
5.1. Clos-like data-center topology with an exemplary forwarding tree. . .	63
5.2. Average flow-completion time over all ten flow traces on load level 1. .	68
5.3. Average flow-completion time over all ten flow traces on load level 1.5.	69
5.4. Average flow-completion time over all ten flow traces on load level 1.75.	70
5.5. Average flow-completion time over all ten flow traces on load level 2. .	70
5.6. Average flow completion time created by HybridTE ₁₀₀	71
5.7. Reduction of the average flow completion time created by HybridTE.	72
5.8. Maximum number of elephants on load level 1.	74
5.9. Maximum number of elephants on load level 1.5.	74
5.10. Maximum number of elephants on load level 1.75.	75
5.11. Maximum number of elephants on load level 2.	75
6.1. <code>ofp_flow_sampling_request</code> message.	79
6.2. <code>ofp_sample_flow</code> message.	79
6.3. The token bucket algorithm for packet sampling.	80
6.4. Throughput over different sampling rates	80
6.5. Rates of false negatives for different parameter sets.	83
6.6. Reporting delay for different parameter sets.	83
6.7. Rates of false positives for different parameter sets	83
7.1. Schematic view of MaxiNet.	91
7.2. Minimal example experiment using the MaxiNet Python API.	92
7.3. One switch pair consisting of 2 switches and 2 hosts connected in a line.	93
7.4. Required dilation factor when emulating on a single physical machine.	93
7.5. Speedup gained from an increasing number of workers.	94
7.6. Distribution of forwarding delays.	95
7.7. Sketch of the emulated Clos-like topology.	96
7.8. Distribution of flow completion times.	96
7.9. QQ-plot of flow completion times.	97
7.10. Aggregate traffic demand used in the large-scale experiment.	98
7.11. CPU utilization and network usage of the OpenFlow controller.	99
7.12. Average CPU usage of the workers during the experiment.	100
7.13. Average network usage of the workers during the experiment.	100
7.14. Schematic view of SLS.	102
7.15. Schematic view on NetSLS.	104
8.1. High-level overview of DCT ² Gens work flow.	108
8.2. Relationship between Layer 2 flows, ACKs, and payload flows.	114
8.3. Architectural overview of DCT ² Gen.	118
8.4. Payload bytes per ACK packet over different flow sizes.	120
8.5. Result of the deconvolution operation.	122

8.6. Run time of ILP 1.	126
8.7. Distribution of node degrees.	128
8.8. Algorithm for selecting s-d pairs for flows.	131
8.9. Topsøe distance of flow assignment methods.	132
8.10. Sketch of the Clos-like topology that was used in our experiments. . .	134
8.11. Comparison between S^{gen} and S^{obs}	136
8.12. Comparison between $N_{\text{inter}}^{\text{gen}}$ and $N_{\text{inter}}^{\text{obs}}$	136
8.13. QQ-plot of $N_{\text{inter}}^{\text{gen}}$ and $N_{\text{inter}}^{\text{obs}}$	137
8.14. Comparison between $N_{\text{intra}}^{\text{obs}}$ and $N_{\text{intra}}^{\text{gen}}$	137
8.15. QQ-plot of $N_{\text{intra}}^{\text{obs}}$ and $N_{\text{intra}}^{\text{gen}}$	137
8.16. Comparison between $B_{\text{intra}}^{\text{obs}}$ and $B_{\text{intra}}^{\text{gen}}$	138
8.17. QQ-plot of $B_{\text{intra}}^{\text{obs}}$ and $B_{\text{intra}}^{\text{gen}}$	138
8.18. Comparison between $B_{\text{inter}}^{\text{obs}}$ and $B_{\text{inter}}^{\text{gen}}$	139
8.19. QQ-plot of $B_{\text{inter}}^{\text{obs}}$ and $B_{\text{inter}}^{\text{gen}}$	139
8.20. Comparison between IAT^{obs} and IAT^{gen}	140