Replica Placement Algorithms for Efficient Internet Content Delivery

by

Shihong Xu

Thesis submitted in fulfillment of the
requirements for the degree of

Doctor of Philosophy

in

Computer Science

School of Computer Science
The University of Adelaide, Australia
November 26th, 2009

To my parents and my wife Jie,
who made all of this possible
with their endless encouragement and patience.

Contents

Abstract vi

Statement of Originality viii

Publications ix

Acknowledgments xi

1 Introduction **1**

1.1 Content Delivery in the Internet 2

1.1.1 Content Delivery Systems 2

1.1.2 Evolution of Techniques 4

1.1.2.1 Single proxy caching 4

1.1.2.2 Cooperative web caching 5

1.1.2.3 Content Distribution Networks 6

1.2 Issues and Challenges 6

1.2.1 Research Issues in Content Delivery 7

1.2.2 Challenges 9

1.3 Our Work 11

1.3.1 Issues Addressed in This Thesis 11

1.3.2 Overview of Our Work 13

1.3.3 Summary of Contributions 15

1.4 Thesis Outline 16

2 Replication for Content Delivery — An Overview **18**

2.1 Content Distribution Networks 18

2.1.1 Replica Server Placement 19

2.1.2 Content Replication 20
2.1.3 Other Topics 22
2.2 Web Caching 22
2.2.1 Caching Architecture 22
2.2.2 Object Replacement Algorithms 24
2.2.3 Coordinated En-Route Web Caching 25
2.3 Approaches and Methods 26
2.3.1 Linear Programming and Duality Theory 26
2.3.2 Approximation Algorithms 30
2.3.3 Dynamic Programming 31

3 Fault Tolerant Facility Allocation **33**
3.1 Introduction 33
3.1.1 Related Work 36
3.1.2 Our Technique 37
3.1.3 Our Results 38
3.2 Single-Factor Approximation 39
3.2.1 The Algorithm 39
3.2.2 Inverse Dual Fitting Based Analysis 44
3.2.2.1 High level analysis 44
3.2.2.2 The *p-th* phase 47
3.2.2.3 Performance ratio 49
3.3 Bi-Factor Approximation and Extension 50
3.3.1 The Algorithm 50
3.3.2 Inverse Dual Fitting Based Analysis 52
3.3.3 Scaling and Greedy Augmentation 57
3.4 Fault-Tolerant k-Facility Allocation 58
3.4.1 Bisection Search and Combination 60
3.4.2 Randomized Procedure for Rounding 63
3.4.2.1 Facility opening 63

3.4.2.2 Connection establishment 63
3.4.2.3 Approximation factor 66
3.4.3 Derandomization 66
3.5 Discussion 67
3.5.1 Dealing with Demand 67
3.5.2 Fault-Tolerant Network Design 67

4 QoS-Aware Content Replication for Parallel Access 69
4.1 Introduction 69
4.2 Problem Formulation 72
4.2.1 Object Placement for Parallel Access 72
4.2.2 An Alternative Integer Programming Formulation 74
4.2.3 Issues in the FTFL Problem 76
4.2.3.1 Triangle inequality 76
4.2.3.2 Access frequency 76
4.3 A Distributed ($|\mathcal{R}|$,2)-Approximation Algorithm 77
4.3.1 Facility/Client Side Pseudocodes 77
4.3.2 An alternative centralized algorithm 79
4.3.3 An upper bound of performance ratio 81
4.4 Numerical Results 85
4.4.1 Results on Grids 85
4.4.1.1 Impact of connectivity 85
4.4.1.2 Impact of facility costs 86
4.4.2 Results on Network Models 87
4.4.2.1 Random graph 87
4.4.2.2 GT-ITM model 88
4.4.3 Factors of Performance Ratio 88
4.5 Related Work 89
4.6 Conclusion and Discussion 91

5 Coordinated En-Route Web Caching in Multi-Server Networks 93

5.1 Introduction . . . 93

5.2 Related Work . . . 95

5.3 System Model . . . 98

5.3.1 Calculation of Caching Gain . . . 99

5.3.2 Problem Formulation . . . 102

5.4 Unconstrained *p*-Server CERC . . . 103

5.4.1 Exhaustive Algorithm . . . 104

5.4.2 Dynamic Programming Solution . . . 104

5.4.3 Caching Scheme . . . 109

5.4.4 Discussion . . . 110

5.4.4.1 Optimization . . . 110

5.4.4.2 Impact of Routing asymmetry . . . 111

5.4.4.3 Server selection and cache consistency . . . 111

5.5 QoS-Aware *p*-Server CERC . . . 111

5.5.1 QoS Constraints . . . 112

5.5.1.1 Individual Latency Constrained *p*-Server CERC . . . 113

5.5.1.2 Average Latency Constrained *p*-Server CERC . . . 115

5.5.2 Extended Constraints on Copy Number . . . 118

5.5.2.1 At most k copies . . . 118

5.5.2.2 At least k copies . . . 119

5.5.3 QoS-Aware Caching Schemes . . . 119

5.6 Simulation Experiments . . . 120

5.6.1 Simulation Model . . . 120

5.6.1.1 Parameter settings . . . 120

5.6.1.2 Caching schemes evaluated . . . 121

5.6.2 Result of Experiments . . . 122

5.6.2.1 Impact of cache size . . . 122

5.6.2.2 Impact of access pattern . . . 124

5.6.2.3 Impact of the number of cached copies . . . 125

5.6.2.4 Impact of the number of servers 127
5.7 Concluding Remarks 128

6 Conclusion and Future Work 130
6.1 Conclusion 130
6.2 Future Work – Towards Smart Content Delivery 132
6.2.1 Novel Architectures 132
6.2.2 Security Considerations 133
6.2.3 Employment of Data Mining Techniques 133

Bibliography 135

Appendix A: List of Symbols 152

Appendix B: Curriculum Vitae 156

Abstract

This thesis covers three main issues in content delivery with a focus on placement algorithms of replica servers and replica contents. In a content delivery system, the location of replicas is very important as perceived by a quotation: *Closer is better*. However, considering the costs incurred by replication, it is a challenge to deploy replicas in a cost-effective manner. The objective of our work is to optimally select the location of replicas which includes sites for replica server deployment, servers for replica contents hosting, and en-route caches for object caching. Our solutions for corresponding applications are presented in three parts of the work, which makes significant contributions for designing scalable, reliable, and efficient systems for Internet content delivery.

In the first part, we define the *Fault-Tolerant Facility Allocation (FTFA)* problem for the placement of replica servers, which relaxes the well known *Fault-Tolerant Facility Location (FTFL)* problem by allowing an integer (instead of binary) number of facilities per site. We show that the problem is NP-hard even for the metric version, where connection costs satisfy the triangle inequality. We propose two efficient algorithms for the metric *FTFA* problem with approximation factors 1.81 and 1.61 respectively, where the second algorithm is also shown to be (1.11,1.78)- and (1,2)-approximation through the proposed inverse dual fitting technique. The first bi-factor approximation result is further used to achieve a 1.52-approximation algorithm and the second one a 4-approximation algorithm for the metric *Fault-Tolerant k-Facility Allocation* problem, where an upper bound of facility number (i. e. k) applies.

In the second part, we formulate the problem of QoS-aware content replication for parallel access in terms of combined download speed maximization, where each client has a given degree of parallel connections determined by its QoS requirement. The

problem is further converted into the metric *FTFL* problem and we propose an approximation algorithm which is implemented in a distributed and asynchronous manner of communication. We show theoretically that the cost of our solution is no more than $2F^* + RC^*$, where F^* and C^* are two components of any optimal solution while R is the maximum number of parallel connections. Numerical experiments show that the cost of our solutions is comparable (within 4% error) to the optimal solutions.

In the third part, we establish mathematical formulation for the en-route web caching problem in a multi-server network that takes into account all requests (to any server) passing through the intermediate nodes on a request/response path. The problem is to cache the requested object optimally on the path so that the total system gain is maximized. We consider the unconstrained case and two QoS-constrained cases respectively, using efficient dynamic programming based methods. Simulation experiments show that our methods either yield a steady performance improvement (in the unconstrained case) or provide required QoS guarantees.

Statement of Originality

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution to Shihong Xu and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis (as listed in the next two pages) resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue, the Australasian Digital Theses Program (ADTP) and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signature:

Shihong Xu

Nov. 26th, 2009

Papers Published

[1] Hong Shen and Shihong Xu. Coordinated En-Route Web Caching in Multi-Server Networks. *IEEE Transactions on Computers*, vol.58, no.5, pp.605-619, May 2009.

[2] Shihong Xu and Hong Shen. A Distributed (R, 2)-Approximation Algorithm for Fault Tolerant Facility Location. In *The 10th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT 2009)*, Hiroshima, Japan, Dec. 8-11, 2009.

[3] Shihong Xu and Hong Shen. The Fault Tolerant Facility Allocation Problem. In *The 20th International Symposium on Algorithms and Computation (ISAAC 2009)*, Hawaii, USA, Dec. 16-18, 2009.

[4] Shihong Xu and Hong Shen. QoS-Oriented Content Delivery in E-Learning Systems. In *The 2nd International Symposium on Information Technology in Medicine and Education (ITME 2009)*, pages 665-670, Jinan, China, Aug. 2009, invited paper.

[5] Shihong Xu and Hong Shen. An Efficient Method for p-Server Coordinated En-Route Web Caching. In *The 8th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT 2007)*, pages113-117, Adelaide, Australia, Dec. 2007.

[6] Shihong Xu and Hong Shen. An O(nh) Algorithm for Dual-Server Coordinated En-Route Web Caching. In *The 7th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT 2006)*, pages

399-404, Taipei, Taiwan, Dec. 2006.

Papers under Review

[1] Shihong Xu and Hong Shen. Fault Tolerant Facility Allocation. Submitted to *SIAM Journal on Computing*, Aug. 2009.

[2] Shihong Xu and Hong Shen. QoS-Aware Content Replication for Parallel Access in the Internet. Submitted to *IEEE Transactions on Parallel and Distributed Systems*, Aug. 2009.

Acknowledgments

I am deeply grateful to my principal supervisor, Prof. *Hong Shen*, for guiding my research during the PhD candidateship which started at the Japan Advanced Institute of Science of Technology (JAIST) since year 2005. From the discussion of my proposal to the last corrections of this document, he has been part of my research in all the stages. And it is from him that I learned critical thinking and skills in technical writing, which laid solid foundation for my future research. Furthermore, I would like to thank my co-supervisor, Dr. *Michael Sheng*, for his support on the two and a half years study at the University of Adelaide.

I owe great thanks to the School of Computer Science at the University of Adelaide and the JAIST Foundation for granting me scholarships, which make it possible for me to pursue the degree. Thank the staff in the two institutions for providing their selfless help and professional assistance in my work and life.

I would like to thank all the people who either directly or indirectly provide me knowledge, experience and support. I would like to thank my friends *Zonghua Zhang*, *Yingpeng Sang* and *Hui Tian* for sharing their experience in research. Especially, I would like to thank my colleagues: *Haibo Zhang* and *Yawen Chen*, with whom I traveled around Japan and Australia; *Yidong Li*, who shared with me the office room and those stressful times when struggling on a research problem; and *Donglai Zhang*, who initiated the regular tea time in our group and shared his precious insights into technologies. I will always remember that it is these brilliant people that accompanied me intellectually and personally in these years.

I would like to thank my family for their support and love which make my life enjoyable even in those stressful times. I thank my parents for making it possible for

me to receive high-level education, and my wife Jie Feng for supporting my research faithfully all the time. They have been and will always be the reasons that I strive for excellence.

Last but not the least, I would like to express my gratitude to the anonymous examiners of this thesis. Review of a PhD thesis contains considerable amount of work which requires reading the full document and looking for problems and potential enhancements with a critical mind. I would like to thank them for their precious time and constructive comments on this thesis.

List of Figures

3.1 Credit Offers for Opening a Facility 43
3.2 Ranking of Contributions 48
3.3 Randomized Procedure 63

4.1 Factors of Performance Ratio 90

5.1 System Model in a 3-Server Network 98
5.2 An Example of Hardness for Problem Decomposition 103
5.3 Decomposition of Caching Gain 106
5.4 Reuse of Intermediate Results 108
5.5 Impact of Cache Size 123
5.6 Impact of Access Pattern 125
5.7 Impact of the Number of Copies 126
5.8 Impact of the Number of Servers 128

List of Tables

4.1 Performance Ratios Using Random Points on a Grid (with various maximum connectivity) 86

4.2 Performance Ratios Using Random Points on a Grid (with various maximum facility cost) 86

4.3 Performance Ratios on Network Models 88

5.1 Process of Algorithm (in terms of $G_x(x_r)$ and $P_x(x_r)$) 108

5.2 Parameter Settings 121

5.3 Groups of Experiments 122

List of Algorithms

3.1 1.861-Approximation Algorithm 41
3.2 Algorithm for the *p-th* Phase 42
3.3 Restatement of the *p-th* Phase Algorithm 44
3.4 Multi-Factor Approximation Algorithm 51
3.5 Improved Algorithm for the *p-th* Phase 51
3.6 1.52-Approximation Algorithm 57
4.1 Facility-Side Pseudocode 78
4.2 Client-Side Pseudocode 78
4.3 Extended JMS Algorithm [47, 52] 80
5.1 Unconstrained p-Server CERC 107
5.2 Individual Latency Constrained p-Server CERC 114
5.3 Average Latency Constrained p-Server CERC 117