Economies of Scale in Information Dissemination
over the Internet

A dissertation submitted to the Graduate School
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Engineering and Public Policy

by

John Chung-I Chuang

Pittsburgh, Pennsylvania
November 1998
Acknowledgments

I would like to thank my committee members, Professors Marvin Sirbu, Granger Morgan, Alex Hills and Hui Zhang, for their guidance and support. Specifically, I am most indebted to my advisor and mentor, Marvin, for his generosity and kindness. My sincere gratitude also goes to the many colleagues who have provided invaluable feedback on this work.

Financial support from the National Science Foundation (Grant IRI-9411299), the Council on Library and Information Resources A. R. Zipf Fellowship in Information Management, and the Department of Engineering and Public Policy at Carnegie Mellon University, is gratefully acknowledged.

To the many wonderful people of Pittsburgh that I have come to know, especially those of EPP and of the Sacred Heart community, I thank them for their friendship and their prayers.

I thank my family, especially my mom and dad, for their love and encouragement.

I thank my wife, Melissa, for her love and her smile.

I thank God. Amen.
# Table of Contents

LIST OF FIGURES .................................................................................................................... IX

LIST OF TABLES ........................................................................................................................ XI

ABSTRACT ................................................................................................................................. XIII

1. INTRODUCTION ..................................................................................................................... 1

1.1 ECONOMICS OF INFORMATION AND INFORMATION DISSEMINATION ......................... 2

1.2 ECONOMIES OF SCALE: DIMENSIONS AND LEVELS .......................................................... 5

2. EOS IN OBJECTS - INFORMATION BUNDLING ................................................................. 11

2.1 BUNDLING AND UNBUNDLING OF INFORMATION GOODS ........................................... 12

2.2 ECONOMICS OF BUNDLING ............................................................................................. 15

2.3 N-GOOD BUNDLING MODEL ............................................................................................ 17

2.3.1 Modeling heterogeneity in consumer preferences ......................................................... 19

2.3.1.1 Consumer choice in pure bundling ........................................................................... 24

2.3.1.2 Consumer choice in pure unbundling ....................................................................... 25

2.3.1.3 Consumer choice in mixed bundling ........................................................................ 27

2.3.2 Production costs and economies of scale ....................................................................... 28

2.4 ANALYSIS AND EMPIRICAL RESULTS ............................................................................. 30

2.4.1 Optimal pricing and revenue mix .................................................................................. 33

2.4.2 Internet-based document delivery technology ............................................................... 35

2.5 CONCLUSION ....................................................................................................................... 39

3. EOS IN RECEIVERS - MULTICAST COMMUNICATION ...................................................... 43

3.1 PRICING MULTICAST COMMUNICATION: A COST-BASED APPROACH ....................... 44

3.2 COST QUANTIFICATION ..................................................................................................... 45

3.2.1 Quantifying Multicast Tree Cost ................................................................................... 47

3.2.2 Methodology .................................................................................................................. 50

3.2.3 Results ............................................................................................................................ 53

3.2.4 Tree Saturation .............................................................................................................. 55

3.3 MULTICAST PRICING ......................................................................................................... 57

3.3.1 Membership Accounting ............................................................................................... 59

3.3.2 Other Issues .................................................................................................................... 60
3.4 Dense vs. Sparse Mode Multicast ................................................................. 61
3.5 Conclusion ........................................................................................................ 67

4. EOS IN TIME - DISTRIBUTED NETWORK STORAGE ......................... 69

4.1 Distributed Network Storage Infrastructure with QoS Guarantees .......... 71
4.2 Related Work ................................................................................................... 75
  4.2.1 Network Caching ....................................................................................... 76
  4.2.2 Network Replication ................................................................................ 77
  4.2.3 Transmission-based QoS .......................................................................... 78
4.3 Service Specification ....................................................................................... 80
  4.3.1 Deterministic vs. Statistical Guarantees ..................................................... 84
  4.3.2 Performance-Oriented vs. Placement-Oriented Services ......................... 86
4.4 Service Provision ............................................................................................ 87
  4.4.1 Resource Reservation Protocol ............................................................... 87
  4.4.2 Resource Mapping .................................................................................... 88
  4.4.3 Admission Control ................................................................................... 90
  4.4.4 Service Provision Architecture ............................................................... 92
4.5 Real-Time Resource Management ............................................................... 96
  4.5.1 Local Storage Management ................................................................. 97
  4.5.2 Traffic Policing ....................................................................................... 98
  4.5.3 Hierarchical Resource Sharing .............................................................. 98
  4.5.4 Global Storage Management .............................................................. 101
4.6 Additional Mechanisms ............................................................................... 101
  4.6.1 Resource Discovery .............................................................................. 102
  4.6.2 Accounting, Billing and Payment ......................................................... 102
4.7 Economics ..................................................................................................... 103
  4.7.1 QoS Pricing ......................................................................................... 104
  4.7.2 Industrial Organization ......................................................................... 105
  4.7.2.1 Distributed Storage Economy ......................................................... 105
  4.7.2.2 Spot Market, Futures Market and Supplemental Insurance .............. 105
  4.7.2.3 Vertical Integration and Component-Based Competition ............... 106
4.8 Conclusion ..................................................................................................... 109

5. RESOURCE MAPPING FOR DISTRIBUTED NETWORK STORAGE SERVICES 111

5.1 Mathematical Model for Resource Mapping and Admission Control .... 111
  5.1.1 Traffic Profile .................................................................................... 112
5.1.2 Performance Requirements

5.1.3 Resource Mapping

5.1.3.1 Service with Worst Case Delay Bound

5.1.3.2 Service with Average Delay Bound

5.1.3.3 Service with Average and Maximum Delay Bounds

5.1.3.4 Service with Stochastic Guarantees

5.1.4 Admission Control

5.2 Resource Mapping for ARPANET

5.2.1 Base Case: Uniform Demand Distribution, Unconstrained Replica Locations

5.2.2 Non-Uniform Spatial Demand Distribution

5.2.3 Partial Replication of Multi-Object Collection

5.3 Network Storage Capacity Planning Problem

5.3.1 Resource Mapping with Constrained Replication Server Sites

5.4 Mapping into Storage and Transmission Resources

5.5 Conclusion

6. Conclusion

6.1 Policy Implications and Lessons

6.2 Contributions Made in this Dissertation

6.3 Future Work

APPENDIX 1. DERIVATION OF PRODUCER SURPLUS FOR ALTERNATIVE BUNDLING STRATEGIES

APPENDIX 2. SAMPLE LIST OF WEB-HOSTING SERVICE PROVIDERS

APPENDIX 3. SOURCE CODE FOR MULTICAST COST QUANTIFICATION

APPENDIX 4. TAXONOMY OF DATA DUPLICATION SCHEMES ACCORDING TO TRADITIONAL EX-POST VS. EX-ANTE DISTINCTION

APPENDIX 5. SOLUTION METHOD FOR RESOURCE MAPPING WITH PARTIAL REPLICATIONS OF MULTI-OBJECT COLLECTIONS

BIBLIOGRAPHY
List of Figures

Figure 1.1. Leveraging economies of scale along different dimensions (objects, receivers, time) at the information product level.................................................................7
Figure 1.2. Leveraging economies of scale along different dimensions (objects, receivers, time) at the bit transport level..................................................................8
Figure 2.1. Consumer choice regions for two-good bundling model. Alice and Bob will choose different product offerings under different bundling regimes. .................16
Figure 2.2. Total outlay vs. number of articles consumed. An *ex post* two-part tariff (in bold) offers a predictable price cap on consumer expenditure.........................18
Figure 2.3. Article valuation by an individual reader indexed by $\{w_o,k\}$ ...................21
Figure 2.4. This figure demonstrates the diversity of consumers that can be indexed by $\{w_o,k\}$. ........................................................................................22
Figure 2.5. Consumer choice in pure bundling scenario......................................25
Figure 2.6. Optimal article consumption level in pure unbundling scenario..........26
Figure 2.7. Consumer choice in mixed bundling scenario....................................28
Figure 2.8. Profit-maximizing bundling strategy: it is clear that mixed bundling is the dominant strategy across all marginal cost and economies of scale conditions......31
Figure 2.9. Optimal price ratio for mixed bundling strategy across various economies of scale and marginal cost conditions.................................................................34
Figure 2.10. Optimal revenue mix for mixed bundling strategy............................35
Figure 2.11. Effect of transmission cost on journal subscription pricing.................38
Figure 2.12. Effect of declining $k_d$ (transmission cost) on economies of scale and revenue mix..............................................................................................39
Figure 3.1. Transmission vs. computing cost trends............................................46
Figure 3.2. Example network shows that degree of link savings achievable is strongly dependent on spatial distribution of receivers. ..............................46
Figure 3.3. Quantifying economies of scale in multicast communication - a process overview............................................................................................................51
Figure 3.4. Normalized multicast tree length as a function of membership size - slope is constant (~0.8) across various network topological styles.................................54
Figure 3.5. Normalized multicast tree length as a function of membership size - slope is constant (~0.8) across various network sizes.................................................54
Figure 3.6. Normalized multicast tree length as a function of membership size - results confirmed with real networks.................................................................55
Figure 3.7. An illustration of the “tree saturation” effect: it takes just \(\sim 500\) randomly selected dial-in ports (or 0.5% of all ports) to subscribe to a multicast group before all 100 network nodes become part of the multicast tree. All subsequent subscribers can be served at no additional cost.

Figure 3.8. Comparing alternatives for sending one data packet to receivers in the MBone network.

Figure 3.9. Comparing alternatives for sending a 5kbps data stream to receivers in the MBone network - there appears to be no difference between sparse and dense mode multicast.

Figure 3.10. Comparing dense and sparse mode multicast for sending a 5kbps data stream to receivers in the MBone network - dense mode multicast clearly consumes more bandwidth when there are few receivers, but the two modes are comparable with subscription density as low as 4% (about 200 receivers).

Figure 4.1. From performance requirements to performance realization: the process flow of establishing a network storage service with QoS guarantees. The components in bold are the key components of the infrastructure.

Figure 4.2. Service provision architectural alternatives.

Figure 4.3. Hierarchical resource sharing example.

Figure 4.4. Example shows vertically integrated storage provider can internalize transmission cost savings not available to an independent storage provider.

Figure 5.1. Network topology of early ARPANET.

Figure 5.2. Resource Mapping for ARPANET.

Figure 5.3. Resource mapping for non-uniform spatial distribution.

Figure 5.4. Resource mapping for multi-object collections with non-uniform object distribution using full or partial replication.

Figure 5.5. Resource mapping for multi-object collections with non-uniform object distribution using partial replication.

Figure 5.6. Network capacity planning problem - different replication strategies may realize the lowest cost solution depending on the relative magnitudes of fixed \((c_0)\) and variable \((c_S)\) costs.

Figure 5.7. Resource cost relative to storage-only solution (4 replicas) at different storage to transmission cost ratios.

Figure 5.8. Optimal mapping decision for storage and transmission resources (ARPANET with \(\tau_{\text{max}} = 3\)).

Figure 5.9. Resource cost relative to storage-only solution (4 replicas).
List of Tables

Table 1.1. The various dimensions and levels of EoS explored in this work.................6
Table 2.1. Distribution of number of articles read in a journal. ...............................23
Table 2.2. Consumer choice in mixed bundling scenario.....................................27
Table 3.1. Networks used in this study.............................................................52
Table 3.2. Data and control/overhead for various options of sending data to multiple
destinations.........................................................................................................62
Table 4.1. Some examples of network storage services........................................81
Table 4.2. Entities involved in resource mapping and admission control functions in
different service provision architectural alternatives....................................95
Table 5.1. ARPANET Statistics..........................................................................120
Table 5.2. Resource Mapping for Service with Average Delay Bound....................122
Table 5.3. Comparing mapping efficiencies for constrained versus unconstrained
replication server sites.........................................................................................133
Abstract

This dissertation studies the different levels and dimensions along which economies of scale (EoS) savings may be realized when information is disseminated over the Internet. At the information product level, EoS savings may be realized along the object, consumer and temporal dimensions through strategies such as information bundling, site-licensing and subscriptions. At the information transport level, EoS savings may be realized along the same dimensions through just-in-time delivery, multicast, network caching and replication strategies. Each of these strategies is studied in this work.

Along the object dimension, a multi-product bundling model with multi-dimensional consumer taste characteristics is developed to study the optimal bundling and pricing strategy of information goods such as academic journals. Using empirical journal usage data and cost projections for information-delivery over the Internet, the model finds that metered usage (i.e., articles-on-demand) should account for a significant fraction of revenue when articles and subscriptions are optimally priced according to a mixed bundling strategy.

Along the receiver dimension, a communication cost model for multicast is developed. The model demonstrates that multicast group size can serve as an excellent proxy for multicast tree cost. Computer simulations show that, statistically, multicast tree length grows at the 0.8 power of the multicast group size until the point of tree saturation, beyond which additional receivers can be added to the group without further tree growth. In other words, the marginal cost of multicast declines according to an exponential decay function until it reaches zero at tree saturation. This result is validated
with both real and generated networks, and is robust across topological styles and network sizes. This suggests that a two-part tariff may be appropriate if providers choose to adopt a cost-based approach to multicast pricing.

Along the temporal dimension, economies of scale savings can be realized through network caching and replication. This work offers the vision of and motivation for a distributed network storage infrastructure with service guarantees. Caching and replication can be treated as different service classes within a unified Quality-of-Service (QoS) framework. Key components of the distributed network storage architecture include: service specification, resource reservation, resource mapping, admission control, real-time resource management and pricing. After establishing a research roadmap, this work focuses on the resource mapping problem and develops a formal mapping model, allowing services with different traffic profiles and performance specifications to be mapped into an optimal combination of storage and transmission resources. The model is also extended to tackle network storage capacity planning problems.

The work described in this dissertation promotes an understanding of how new network technologies have changed, and will continue to change, the economics of information dissemination. This understanding is essential to the design of engineering, economic and policy structures that will constitute the information infrastructure of the future.