Fault Management for Realtime Networks

by

Anindo Banerjea

B.Tech. (Indian Institute of Technology, New Delhi) 1989

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Computer Science

in the

GRADUATE DIVISION

of the

UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor Domenico Ferrari, Chair
Professor Pravin P. Varaiya
Professor Richard E. Barlow

1994

The dissertation of Anindo Banerjea is approved:

Chair	Date
	Date
	Date

University of California at Berkeley

1994

Fault Management for Realtime Networks

Copyright © 1994

by

Anindo Banerjea

Abstract

Fault Management for Realtime Networks

by

Anindo Banerjea

Doctor of Philosophy in Computer Science

University of California at Berkeley

Professor Domenico Ferrari, Chair

Realtime networks provide guaranteed performance communication to applications that require such guarantees. The Tenet Group, under Professor Domenico Ferrari, has developed a scheme to provide such Quality of Service (QoS) guarantees in a packet-switched internetwork. The scheme is based on the concept of the *realtime channel*, which is a network connection with associated traffic specifications and performance guarantees. Other schemes to provide realtime services exist, and share some fundamental similarities. However, none of these schemes address the problem of how to restore (or continue to provide) realtime service in the presence of network faults. This dissertation addresses the problem of dealing with faults in the context of realtime networks, using two classes of mechanisms: *proactive* and *reactive*.

Reactive mechanisms can be used to reroute realtime channels to surviving links in the network. This approach does not use any extra resources in the absence of faults, but involves a disruption of the service while the recovery action is being performed. In addition, some channels may not be successfully rerouted if the realtime load on the network is very high. Proactive schemes may be used to reserve redundant resources (such as on multiple paths in the network) so that there is no disruption (or only disruption that can be bounded *a priori*), as long as the fault scenario is one that is *covered*. The proactive scheme may be designed to cover fault scenarios such as single faults, double faults, and so on.

The first part of the dissertation describes the rerouting schemes for fault recovery. We start with a very general fault recovery model, and systematically refine our design till we are left with a well-structured and limited problem domain, which we explore through simulation. The simulation experiments were used to identify the best scheme within the chosen solution space, and to show that its performance is reasonable on a number of metrics of performance, such as speed, amount of traffic rerouted, and efficiency of resource usage, as well as for large variations in the external factors such as network load, fault scenario, traffic mix, and network topology. The dissertation also contains a high-level design of a protocol for fault recovery of Tenet realtime channels, the Real-Time Control Message Protocol, based on the results of the experiments.

The dissertation also describes the use of dispersity routing and forward error correction to provide fault tolerance for realtime channels. These techniques are used to design a variety of schemes that deliver various levels of service in the presence of restricted network faults. Some schemes provide transparent tolerance to multiple faults in the network, at the cost of increased network resource requirements; some others use no extra resources, as compared to a simple realtime channel, but the service degrades gracefully when a network failure occurs; still others suffer a total disruption in the event of a failure, but the duration of the disruption is limited to the time needed to notify the source of the fault. These mechanisms are end-to-end in nature, permitting implementation on top of a basic realtime service, as long as appropriate support for routing is provided by the network. The services are validated and the cost to the network is evaluated through simulation. The techniques are also compared to existing mechanisms that provide fault tolerance to realtime networks.

Committee chair: _____

Professor Domenico Ferrari

To my parents, Maya and Bidyut.

Table of Contents

Table of contents iv
List of Tables vi
List of Figures vii
Acknowledgements ix
Chapter 1: Introduction 1
1.1 The realtime channel paradigm 1
1.2 General requirements from the control mechanisms 4
1.3 Two classes of control mechanisms 5
1.4 Previous work 10
1.5 Scope of the dissertation 20
Chapter 2: The Tenet Realtime Protocol Suite 1 22
2.1 Introduction 22
2.2 Network model and assumptions 22
2.3 Service model and assumptions 23
2.4 The Tenet Suite 1 26
2.5 Channel management 28
2.6 Other realtime schemes 35
2.7 Conclusions 36
Chapter 3: A reactive scheme for fault recovery 38
3.1 Introduction 38
3.2 Objectives of the fault recovery mechanisms 39
3.3 Network model and assumptions 40
3.4 A fault recovery framework 44
3.5 Design issues for a reactive scheme for fault recovery 45
3.6 Issues not to be investigated 53
3.7 Remaining issues 54
3.8 Conclusion 64
Chapter 4: Evaluation of Fault Recovery 65
4.1 Introduction 65
4.2 Simulation design 65
4.3 Results 88
4.4 Conclusion 104
Chapter 5: High level design of a fault recovery protocol 111
5.1 Introduction 111
5.2 Summary of design decisions 112
5.3 Description of the protocol 114

5.4 Changes to the RCAP+DCM software 122
5.5 Conclusions 123
Chapter 6: A proactive scheme for fault tolerance 125
6.1 Introduction 125
6.2 Background 127
6.3 Schemes for fault tolerance 135
6.4 Simulator design 151
6.5 Simulation results 155
6.6 Implementation issues 177
6.7 Conclusions 180
Chapter 7: Conclusions 186
7.1 Introduction 186
7.2 Summary of the dissertation 186
7.3 Requirements re-examined 188
7.4 Contributions 189
7.5 Future work 191
Bibliography 193

List of Tables

Table	Label	Page
2.1	Performance parameters of the Tenet Suite 1 interface	25
2.2	Traffic parameters of the Tenet Suite 1 interface	25
2.3	RCAP messages	30
4.1	Load index levels	85
4.2	Properties of the fault recovery scheme selected	109
6.1	Properties of various dispersity systems	181
6.2	Properties of some more dispersity systems	181

List of Figures

Figure	Label	Page
1.1	Link vs. Global rerouting	16
1.2	Dispersity routing - (3,2,2) system	18
2.1	Service model for realtime network	24
2.2	Protocols in the Tenet Suite 1	26
3.1	Avoiding collisions	51
3.2	The sector of the solution space to be considered	52
3.3	4-dimensional solution space to be explored	55
3.4	Local rerouting	56
3.5	Valid regions of the 4-dimensional solution space	63
4.1	Rate Control Static Priority scheduler	71
4.2	Properties of the load index.	76
4.3	Square mesh topology	81
4.4	Square mesh topology with trunks	82
4.5	Core topology	83
4.6	Effect of timing on (i) reroute success (ii) time to reroute	88
4.7	Effect of locus of reroute on (i) reroute success (ii) excess resources used	90
4.8	Effect of retry policy. (i) Immediate retries (ii)- (iv) Delayed retries	92
4.9	Effect of retry interval on the histogram of time to reroute (i) Exponential (ii) Fixed (100 ms)	94
4.10	Effect of state prediction on (i) time to reroute (ii) realtime load rerouted on first attempt	95
4.11	Effect of state prediction on histogram of time to reroute (i) Local prediction (ii) Global prediction	97
4.12	Effect of number of faults on (i) reroute success (ii) success of prediction (iii) time to reroute (iv) excess resources used	99
4.13	Effect of new load mix on (i) Success ratio (ii) Time to reroute (iii) Histogram of time to reroute	101
4.14	Effect of topology. (i)-(ii) Mesh with added trunks (iii)-(iv) Core topology	103
4.15	Effect of retry interval on core topology (i) 100 ms interval (ii) 60 ms interval	104
5.1	Interactions for RTCMP	112
5.2	Hysteresis loop to avoid oscillations	116
6.1	One set of isolated link failures which the IFI channel can survive	151
6.2	No IFI set of links exists in a square mesh topology	152
6.3	Structure of reassembly buffer for N=4	154
6.4	(3,2,1) dispersity system simulated	157
6.5	Packet loss vs. loss rate for dispersity systems with N-K=1, no network failure case	160
6.6	Packet loss vs. loss rate for dispersity systems with N-K=1, network failure case	164
6.7	Packet loss vs. loss rate for dispersity systems with N-K=2, network failure case	164
6.8	Number of connections established for dispersity systems without redundancy	167

Figure	Label	Page
6.9	Effect of allowing S = 2 on dispersity systems without redundancy	171
6.10	Number of connections established for dispersity systems with N-K=1	171
6.11	Effect of allowing S = 2 on dispersity systems with N-K=1	174
6.12	Number of connections established for dispersity systems with a N-K=2	174
6.13	Effect of allowing S = 2 on dispersity systems with N-K=2	175
6.14	Equivalent dispersity systems with no shared links and N-K=1	175

Acknowledgements

First of all I would like to acknowledge gratefully the invaluable guidance and support that I have received from my advisor, Professor Domenico Ferrari. He provided me with the freedom to explore my own interests, and at the same time with the guidance to ensure that I did not waste my energies on unproductive avenues of research. In addition, his detailed technical and stylistic criticisms of a draft of this dissertation caused me to rewrite major parts of it, and improved its coherence and readability immensely.

I would like to thank Professors Pravin Varaiya and Richard Barlow, for serving on my qualifying examination and dissertation committees, and for their helpful comments on the thesis. I also thank Professor Alan Smith for chairing my qualifying examination committee.

Part of the work presented in Chapter 4 of this thesis was performed in cooperation with Colin Parris. The simulator used in the experiments was developed initially by Hui Zhang, and modified by Colin Parris. Many invaluable discussions with them, as well as with other members of the Tenet Group, contributed to the ideas presented here. I would specifically like to thank Amit Gupta, Ed Knightly, Bruce Mah, Steve Mccanne and Mark Moran for allowing me to use them as sounding boards for many ideas, and Dr. Andres Albanese for many fruitful discussions about network reliability. I fondly remember many debates about *Life, the Universe, and Everything* with various members of the Tenet Group, and the Thursday night beer gatherings with the Dave, Krste and the Realization boys. I gratefully acknowledge Keshav Srinivasan and Dinesh Verma for influence and advice during my early days at Berkeley, especially for advising me to join the Tenet Group, and for continued interaction even after they left Berkeley.

Finally, I am indebted to my parents for the emotional support to persevere in my undertakings, to my brother for his ready ear and quick wit, and to Karen for her love.