

Contents

| | | |
|----------|--------------------------------------------------------------------------|-----------|
| 1 | Introduction..... | 1 |
| 1.1 | Loss is Part of the Internet | 3 |
| 1.2 | The need for multiple classes of Error Control | 5 |
| 1.3 | Error Control for Interactive, Continuous Media Applications | 6 |
| 1.3.1 | Our Solution | 8 |
| 1.3.2 | Contributions | 9 |
| 1.4 | Error Control for Large-Scale Multicast Applications | 10 |
| 1.4.1 | Our Solution: Light-weight Multicast Services (LMS) | 11 |
| 1.4.2 | Contributions | 12 |
| 1.5 | Dissertation Overview | 14 |
| 2 | Error Control for Interactive Continuous Media Applications | 15 |
| 2.1 | Requirements of Continuous Media Applications | 16 |
| 2.2 | Avoiding Losses | 17 |
| 2.2.1 | Forward Error Correction (FEC) | 17 |
| 2.2.2 | Concealment | 18 |
| 2.2.3 | Retransmission | 18 |
| 2.3 | Making Retransmission Work | 21 |
| 2.3.1 | Playout buffering | 21 |
| 2.3.2 | Gap-based loss detection | 22 |
| 2.3.3 | Implicit expiration of sender retransmission buffers | 23 |
| 2.3.4 | Conditional retransmission requests | 23 |
| 2.3.5 | Data integrity information delivery to the application | 24 |
| 2.4 | Design and Implementation | 24 |
| 2.4.1 | Protocol Operation | 25 |

| | | |
|----------|-------------------------------------------------------------------|-----------|
| 2.4.2 | Performance | 26 |
| 2.4.3 | Processing Overhead | 27 |
| 2.5 | Experiments | 27 |
| 2.5.1 | A Network Loss and Delay Emulation Tool | 28 |
| 2.5.2 | Experiment 1: Random Bidirectional Loss | 29 |
| 2.5.3 | Experiment 2: Bidirectional Bursty Loss | 30 |
| 2.5.4 | Experiment 3: Bidirectional Bursty Loss with MPEG Load ... | 31 |
| 2.5.5 | Experiment 4: Qualitative evaluation with raw video | 33 |
| 2.6 | Future Work: Extending Protocol to Multiple Retransmissions | 33 |
| 2.7 | Conclusions | 35 |
| 3 | Background and Related Work..... | 37 |
| 3.1 | Background: Positive vs. Negative Acknowledgments | 37 |
| 3.2 | Problems with Reliable Multicast | 39 |
| 3.2.1 | The Error Model | 40 |
| 3.2.2 | Implosion | 40 |
| 3.2.3 | Exposure | 41 |
| 3.3 | Overview of Related Work | 43 |
| 3.3.1 | Non-Assisted Schemes | 43 |
| 3.3.2 | Schemes Using FEC | 46 |
| 3.3.3 | Assisted Schemes | 46 |
| 3.4 | SRM | 48 |
| 3.5 | PGM | 50 |
| 4 | LMS: Light-weight Multicast Services | 53 |
| 4.1 | Defining an Efficient Scheme | 54 |
| 4.1.1 | A Near-Optimal Solution | 54 |
| 4.2 | Implementing Reliable Multicast Using a Hierarchy | 58 |
| 4.2.1 | A Logical Receiver Hierarchy | 58 |
| 4.2.2 | A Router Hierarchy | 59 |

| | | |
|--------|--------------------------------------------------|----|
| 4.3 | The LMS Concept | 62 |
| 4.4 | LMS Core Ideas | 63 |
| 4.4.1 | Selecting a Replier (surrogate) | 64 |
| 4.4.2 | Steering Messages to Repliers | 66 |
| 4.4.3 | Directed Multicast (DMCAST) | 69 |
| 4.4.4 | LMS Summary | 70 |
| 4.5 | Preventing Duplicate Retransmissions | 72 |
| 4.6 | LMS Specification: Forwarding Services | 73 |
| 4.6.1 | Replier State | 73 |
| 4.6.2 | Handling Requests at the Router | 74 |
| 4.6.3 | Handling Directed Multicasts at the Router | 74 |
| 4.7 | LMS Specification: Error Control | 75 |
| 4.7.1 | Sending Requests | 75 |
| 4.7.2 | Receiving Requests | 75 |
| 4.7.3 | Receiving Retransmissions | 76 |
| 4.8 | Problem: Duplicate Data Packets | 76 |
| 4.9 | Source Spoofing | 77 |
| 4.10 | Dealing with Shared Trees | 77 |
| 4.11 | Replier Failure | 78 |
| 4.12 | Selecting Repliers in a LAN | 80 |
| 4.12.1 | Coordinating Routers Interconnecting LANS | 83 |
| 4.13 | Routers with a Large Fan-out | 83 |
| 4.14 | Improvements | 84 |
| 4.14.1 | Proxy Directed Multicast | 84 |
| 4.14.2 | A DMCAST that Eliminates Exposure | 85 |
| 4.15 | Some Pathological Topologies | 86 |
| 4.15.1 | Long and Skinny Branches | 86 |
| 4.15.2 | Topologies that cause Request Implosion | 87 |
| 4.16 | Other LMS Applications | 88 |
| 4.16.1 | Simple ANYCAST | 88 |

| | | |
|----------|--------------------------------------------------------|-----------|
| 4.16.2 | Positive Acknowledgment-Based Reliable Multicast | 88 |
| 4.17 | Security Issues | 90 |
| 4.18 | Incremental Deployment | 90 |
| 4.19 | Summary | 92 |
| 5 | Simulation Results | 94 |
| 5.1 | Why simulation? | 95 |
| 5.2 | Why <i>ns</i> ? | 97 |
| 5.3 | Simulation Methodology | 98 |
| 5.4 | Evaluation Metrics | 100 |
| 5.4.1 | Normalized Recovery Latency (LMS, SRM, PGM) | 101 |
| 5.4.2 | Exposure (LMS) | 102 |
| 5.4.3 | Duplicate Requests/Repairs (SRM) | 103 |
| 5.4.4 | Repeated Retransmissions (PGM) | 103 |
| 5.5 | Topology Generation | 104 |
| 5.5.1 | Topology Generator: GT-ITM | 105 |
| 5.5.2 | Selected Topologies | 105 |
| 5.6 | Simulation Parameters | 107 |
| 5.7 | Simulation Verification | 109 |
| 5.7.1 | Verification using traces | 109 |
| 5.7.2 | Verification using Nam | 109 |
| 5.8 | LMS Experiments | 109 |
| 5.8.1 | Binary Trees | 110 |
| 5.8.2 | Random Topologies | 113 |
| 5.8.3 | Transit-Stub Topologies | 118 |
| 5.8.4 | LMS Trade-offs | 122 |
| 5.9 | SRM Experiments | 122 |
| 5.9.1 | SRM with Adaptive Timers | 126 |
| 5.9.2 | SRM Trade-offs | 128 |
| 5.10 | PGM Experiments | 128 |

| | | |
|----------|-----------------------------------------------------------|------------|
| 5.10.1 | PGM Trade-offs | 132 |
| 5.11 | Summary and Discussion | 132 |
| 6 | LMS IMPLEMENTATION..... | 135 |
| 6.1 | Background | 139 |
| 6.1.1 | Control Information Exchange: Ancillary Data | 140 |
| 6.1.2 | NetBSD Multicast Forwarding Architecture | 142 |
| 6.2 | LMS in NetBSD | 144 |
| 6.2.1 | IP Options for LMS | 145 |
| 6.3 | LMS Implementation at Endpoints | 146 |
| 6.3.1 | Sending/Receiving Requests | 147 |
| 6.3.2 | Sending/Receiving DMCASTS | 149 |
| 6.4 | Handling LMS Messages at Routers | 151 |
| 6.4.1 | New Router State | 151 |
| 6.4.2 | Handling LMS Packets at a Router | 151 |
| 6.4.3 | Handling Requests at a Router | 153 |
| 6.4.4 | Handling DMCASTs at a Router | 153 |
| 6.4.5 | A Conflict with NetBSD's Handling of Encapsulated Packets | 153 |
| 6.5 | Evaluation | 154 |
| 6.5.1 | Request Overhead at Endpoints | 154 |
| 6.5.2 | Dmcast Overhead at Endpoints | 155 |
| 6.5.3 | Processing Overhead | 155 |
| 6.6 | Summary | 159 |
| 6.7 | Future Work | 160 |
| 6.7.1 | The LMS-HIER Component | 161 |
| 6.7.2 | LMS in the Fast Path | 163 |
| 7 | Conclusions and Future Work..... | 165 |
| 7.1 | Interactive Continuous media Applications | 165 |
| 7.1.1 | Contributions | 165 |

| | | |
|-------|------------------------------------------|-----|
| 7.1.2 | Future Work | 166 |
| 7.2 | Large-Scale Multicast Applications | 167 |
| 7.2.1 | Contributions | 167 |
| 7.2.2 | Future Work | 169 |
| 7.3 | Source Code Availability | 170 |
| 7.4 | Closing Remarks | 170 |