

Table of Contents

1. Introduction	1
1.1 Minimizing Response Time: Sources and Deterrents	1
1.1.1 Sources of Speedup	2
1.1.2 Deterrents to Speedup	3
1.2 Model for Parallel Query Optimization	4
1.2.1 Annotated Query Trees	5
1.2.2 Operator Trees	5
1.2.3 Parallel Machine Model	7
1.3 Organization of Thesis	8
1.4 Related Work	9
1.4.1 Query Optimization for Centralized Databases	9
1.4.2 Query Optimization for Distributed Databases	9
1.4.3 Query Optimization for Parallel Databases	10
2. Price of Parallelism	13
2.1 Introduction	13
2.2 Tandem Architecture: An Overview	14
2.2.1 Parallel and Fault-Tolerant Hardware	14
2.2.2 Message Based Software	16
2.2.3 Performance Characteristics	16
2.3 Parallelism in NonStop SQL/MP	17
2.3.1 Use of Intra-operator Parallelism	17
2.3.2 Process Structure	18
2.4 Startup Costs	20
2.5 Costs of Operators and Communication	20
2.5.1 Experimental Setup	22
2.5.2 Costs of Scans, Predicates and Aggregation	23
2.5.3 Costs of Local and Remote Communication	24
2.5.4 Cost of Repartitioned Communication	26
2.5.5 Costs of Join Operators	27
2.5.6 Costs of Grouping Operators	30
2.6 Parallel Versus Sequential Execution	31
2.6.1 Parallelism Can Reduce Work	31
2.6.2 Parallelism Can Increase Response Time	33

2.7	Summary of Findings	33
3.	JOQR Optimizations	35
3.1	A Model for Minimizing Communication	36
3.1.1	Partitioning	36
3.1.2	Repartitioning Cost	38
3.1.3	Optimization Problem	38
3.2	Algorithms for Query Tree Coloring	39
3.2.1	Problem Simplification	40
3.2.2	A Greedy Algorithm for Distinct Pre-Colorings	42
3.2.3	Algorithm for Repeated Colors	43
3.2.4	Extensions: Using Sets of Colors	46
3.3	Model for Methods and Physical Properties	48
3.3.1	Annotated Query Trees and Their Cost	50
3.4	Extension of ColorSplit for Methods and Physical Properties ..	52
3.5	Model with Join Ordering	53
3.5.1	Join Ordering Without Physical Properties	54
3.5.2	Join Ordering with Physical Properties	55
3.6	Usage of Algorithms	56
4.	Scheduling Pipelined Parallelism	59
4.1	Problem Definition	59
4.2	Identifying Worthless Parallelism	62
4.2.1	Worthless Edges and Monotone Trees	63
4.2.2	The GreedyChase Algorithm	65
4.2.3	Lower Bounds	65
4.3	The Modified LPT Algorithm	66
4.4	Connected Schedules	68
4.4.1	Connected Schedules When Communication is Free ...	68
4.4.2	BalancedCuts with Communication Costs	73
4.5	Connected Schedules as an Approximation	73
4.6	Heuristics for POT Scheduling	77
4.6.1	A Hybrid Algorithm	78
4.6.2	The Greedy Pairing Algorithm	78
4.7	Approximation Algorithms	79
4.7.1	A Two-Stage Approach	80
4.7.2	The LocalCuts Algorithm	82
4.7.3	The BoundedCuts Algorithm	84
4.8	Experimental Comparison	89
4.8.1	Experimental Setup	90
4.8.2	Experimental Comparison	90
4.8.3	Performance of Hybrid	91
4.8.4	Comparison of Hybrid, LocalCuts and BoundedCuts ..	91
4.8.5	Behavior of Lower Bound	92
4.9	Discussion	94

5. Scheduling Mixed Parallelism	95
5.1 Problem Definition	95
5.2 Balanced Schedules	99
5.3 Symmetric Schedules	102
5.4 Scheduling Trees with Two Nodes	111
5.5 Discussion	112
6. Summary and Future Work	115
6.1 Summary of Contributions	115
6.2 Future Work	118
References	121
Index	131