Contents

Agraïments iii

Abstract v

Resum vii

Resumen ix

Contents xi

1 Introduction 1

1.1 Problem Description 1
1.2 Objectives. 9
1.3 Contributions of the Thesis 10
1.4 Thesis Outline 11

2 Related Work 13

2.1 Prefetching 13

2.2 Cache Interference Analysis and Fairness 15
2.3 Cache Partitioning 16

3 Experimental Framework 19
3.1 Simulation Framework 19
3.2 Real Machine: The Intel Xeon E5 2658A v3 22
3.3 Benchmark Suites 26

4 Selective Prefetching under Limited Memory Bandwidth 31
4.1 Introduction 31
4.2 Characterization Study 33
4.3 ADP Mechanism 37
4.4 Experimental Setup 42
4.5 Evaluation 44
4.6 Summary 51

5 Improving Fairness with LLC Partitioning 53
5.1 Introduction 53
5.2 Analysis of the Inter-Application Cache Interference 57
5.3 Analysis of Progress Estimation Approaches 58
5.4 FPCP Partitioning Approach 61
5.5 Experimental Setup 67
5.6 Evaluation 68
5.7 Summary 74

6 Improving Fairness with LLC Partitioning using Intel CAT 77
6.1 Introduction 77
6.2 Progress Characterization and Estimation 79
6.3 To Overlap or Not To Overlap Cache Ways 83
6.4 Cluster-Based Partitioning Policies 85
6.5 Experimental Setup 87

6.6 Evaluation 89
6.7 Summary 96

7 Conclusions 99

7.1 Contributions 99
7.2 Future Directions 102
7.3 Publications 102

Bibliography 107