Contents

1 Workflow Systems in the Cloud ........................................... 1
  1.1 Background: Cloud Computing ................................. 1
  1.2 Background: Workflow Systems .............................. 3
  1.3 Cloud Workflow Systems .......................................... 4
  1.4 Motivating Examples ............................................... 6
  1.5 Key Issues in the Design of Cloud Workflow Systems .... 11

2 Cloud Workflow System Architecture .............................. 13
  2.1 General Cloud Software Architecture ......................... 13
    2.1.1 Cloud Architecture ......................................... 13
    2.1.2 Example: Aneka Cloud Architecture ...................... 14
  2.2 General Architecture of Cloud Workflow Systems .......... 16
    2.2.1 Cloud Workflow System Architecture ..................... 16
    2.2.2 Example: Window Workflow Foundation Architecture .... 17

3 Cloud Workflow System Functionality ............................. 19
  3.1 Classical Workflow Reference Model .......................... 19
  3.2 Basic Functionalities of Cloud Workflow Systems ........... 22
    3.2.1 Cloud Workflow System Functionality .................... 22
    3.2.2 Example: Kepler Web/Grid Service Management .......... 24
    3.2.3 Example: CloudBus Cloud Resource Management ........... 25

4 Cloud Workflow System Quality of Service ....................... 27
  4.1 QoS of Cloud Services and Web Services ..................... 27
    4.1.1 General QoS ................................................ 27
    4.1.2 SLA Management ............................................ 29
  4.2 QoS of Cloud/Grid Workflows .................................. 31
  4.3 A Generic QoS Framework ...................................... 33
  4.4 Example 1: Time Management (on Temporal Constraints) .... 37
4.5 Example 2: Cost Management (on Data Storage) .......... 39
    4.5.1 Cost Model of Datasets Storage in the Cloud .......... 40
    4.5.2 Minimum Cost Benchmarking of Datasets Storage in the Cloud .......... 42
    4.5.3 Cost-Effective Datasets Storage Strategies .......... 42
4.6 Example 3: Reliability Management (on Data Replication) .... 44
    4.6.1 Data Replication .................................. 45
    4.6.2 Data Storage Reliability Model ..................... 46
    4.6.3 Cost-Effective Incremental Replication Strategy ....... 47
4.7 Example 4: Security Management (on Privacy) ............. 47
    4.7.1 Privacy Protection in Cloud ....................... 48
    4.7.2 Trust Based Privacy Protection ...................... 49

5 Case Study: SwinDeW-C Cloud Workflow System ............. 51
    5.1 Overview of SwinDeW-G Environment ................... 51
    5.2 SwinDeW-C System Architecture ....................... 53
        5.2.1 SwinCloud Infrastructure ....................... 54
        5.2.2 Architecture of SwinDeW-C ..................... 54
        5.2.3 Functionalities of SwinDeW-C Peers ............... 57
    5.3 QoS Management Components in SwinDeW-C ............... 58
        5.3.1 Performance Management in SwinDeW-C ............ 58
        5.3.2 Data Management (Storage and Replication) in SwinDeW-C ........ 60
        5.3.3 Security Management in SwinDeW-C ............... 61
    5.4 SwinDeW-C System Prototype ......................... 63
    5.5 Experiments ......................................... 64
        5.5.1 Evaluation on Performance Management ............ 64
        5.5.2 Evaluation on Data Storage Management .......... 66

Appendix A: Performance Management Strategies ............... 69
Appendix B: Data Storage Management Strategies .............. 77
Appendix C: Replication Management Strategies ............... 83
Appendix D: Trust-Based Noise Injection Strategy .......... 85
Appendix E: Literature Review ................................ 89
Bibliography ............................................... 93