

# Index

- A*, *see* Arrivals
- Abstraction in modelling, 20-22, 217
- Accumulated time in system ( $\bar{W}$ ), 42
- Accuracy of models, 14, 16, *see also* Separable models, robustness
- Active customer (memory constrained system), 185
- ADEPT (tool for modelling proposed systems), 329-32
- Aggregate (hierarchical modelling), 152-58, *see also* Flow equivalent service center
- Aggregate performance measures for multiple class models, 62-63
- Algorithms
  - asymptotic bounds, 79
  - balanced system bounds, 93
  - disk I/O
    - non-RPS, 228
    - RPS, 232
    - multipathing, 239
  - hierarchical decomposition, 161
  - mean value analysis
    - single class exact, 115
    - single class approximate, 118
    - multiple class exact, 141
    - multiple class approximate, 143
    - multiple class mixed exact, 146
  - memory constraints
    - single class, 188
    - multiple class, independent, 193-94
    - multiple class, shared, 195-96
  - open models
    - single class, 111
    - multiple class, 137
  - priority scheduling, 259
  - swapping
    - dedicated device, 199
    - shared device, 200
  - variable overhead, 307
- Amdahl (case studies) *see* IBM
  - IMS capacity, 106-9
  - using modelling software, 360-69
  - using RMF, 376-94
- Appropriateness of queueing networks, 14
- Approximate solution techniques
  - for separable models, *see* Mean value analysis, approximate
  - for non-separable models, *see* Disk I/O; Memory; Processors
- Approximation transformations (modelling software), 356
- Arrival instant queue length in separable models, 112, 114, 139-40
- Arrival instant theorem for separable models, 114, 140
  - and balanced system bounds, 126
- Arrival rate ( $\lambda$ ), 41
- Arrivals ( $\mathcal{A}$ ), 40
- Assumptions, testing of, 33-36, *see also* Validation
- Assumptions of separable models, *see* Separable models, assumptions
- Asymptotic bounds, 72-87, *see also* Bounds on performance
  - Algorithm 5.1, 79
- B*, *see* Busy time
- Balanced system bounds, 86-93, *see also* Bounds on performance
  - Algorithm 5.2, 93
- Balanced systems, utilizations in, 90
- Baseline model, *see* Existing systems
- Basic observable quantities, 40-42
- Batch workload, 58, *see also* Closed models
- Benchmarking, 3, 30-31
- Block size (I/O), 290-91
- Books on performance, 18-19
- Bottleneck, 72-73
  - effect of removing (example using asymptotic bounds), 82-84
  - secondary, 82-83
- Bounds on performance (Chapter 5), 70-97
  - advantages, 70-71, 94
  - asymptotic bounds, 72-87
    - Algorithm 5.1, 79
  - examples and case studies, 77-87

- Bounds on performance (*cont'd.*)  
 balanced system bounds, 86-93  
   Algorithm 5.2, 93  
   and arrival instant theorem, 126  
   optimistic and pessimistic bounds, 71  
 Busy time (*B*), 41
- C*, *see* Completions; Customer classes  
 Cached I/O devices, 244-45  
 CAD/CAM system (performance projection of proposed system), 329-32  
 Capacity planning, *see* Evolving systems  
 Capture ratio (measurement), 285  
 Carrier sense multiple access with collision detection (CSMA-CD), 340
- Case studies and examples  
 Amdahl  
   IMS capacity, 106-9  
   using modelling software, 360-69  
   using RMF, 376-92  
 asymptotic bounds, 78-87  
 bottleneck removal (asymptotic bounds), 82-84  
 CAD/CAM system (proposed), 329-32  
 capacity planning  
   Amdahl 470, IMS, 106-9  
   PDP-10, 132-33  
 communication network (proposed), 324-27  
 concurrency control (database), 343-47  
 CPU replacement, 27-29, 33-36  
 CTSS, 102-3  
 Cyber 173, 133-34  
 database concurrency control, 343-47  
 disk I/O  
   non-RPS, 227-29  
   RPS, 232-33  
 disk load balancing  
   asymptotic bounds, 85-87  
   Univac 1100, 312-13  
 Ethernet, 339-41  
 evolving systems, 309-15  
 FCFS scheduling  
   class-dependent service times, 263-64  
   highly variable service times, 266-67  
 FESC for memory constrained system, 189-90  
 forced flow law, 48-50  
 global balance, 163-69  
 hierarchical modelling, 163-66  
 hierarchical workload characterization (instructional computing acquisition, Prime, VAX), 30-34, 51-52
- Case studies and examples (*cont'd.*)  
 IBM  
   early virtual memory system, 206-9  
   processing complex (workload balancing), 24-27  
   360, 104-6  
   3790, 8130, 8140 (insurance company), 35-37, 78-82  
 IMS (Amdahl 470, capacity planning), 106-9  
 instructional computing acquisition (Prime, VAX), 30-34, 51-52  
 insurance company (IBM 3790, 8130, 8140), 35-37, 78-82  
 Little's law  
   and forced flow law, 48-50  
   and memory constrained systems, 49-50, 55-56  
   at various levels, 44-46  
 loosely-coupled multiprocessor (Cyber 173), 133-34  
 memory, 206-17  
 memory constrained system  
   FESC, 189-90  
   Little's law, 49-50  
 modelling software, use of (Amdahl 470), 360-69  
 modification analysis, 309-15  
   asymptotic bounds, 85-87  
   IBM processing complex, 24-27  
   IBM 360, 104-6  
 multiple class models, 129-34  
 multiple class solution techniques, 137-38, 142, 144-45, 147  
 multiprogramming level variability, 182-84
- MVS  
 Amdahl 470 (using RMF), 376-92  
 IBM processing complex, 24-27  
 SRM, 347-50  
 non-RPS disks, 227-29  
 paging, 202-5  
 PDP-10, 132-33  
 Prime (instructional computing acquisition), 30-34  
 priority scheduling, 259-61  
 proposed systems, 324-27, 329-32  
 RMF usage (Amdahl 470), 376-92  
 RPS disks, 232-33  
 single class models, 102-9  
   of heterogeneous workloads, 130-31  
 single class solution techniques, 112-13, 116-17, 119  
 single service center, 5-8

- Case studies and examples (*cont'd.*)
    - SNA flow control, 336-39
    - software resources, 342-43
    - swapping
      - moving from drum to disk (Univac 1100), 314-15
      - to a shared device, 200-201
    - tightly-coupled multiprocessor (Univac 1100), 310-11
    - Univac 1100, 309-15
    - upgrade from single to dual processor (Univac 1100), 310-11
    - variation in multiprogramming level, 182-84
  - VAX
    - instructional computing acquisition, 30-34, 51-52
    - memory modelling, 209-17
    - virtual memory, 202-5
    - workload balancing (IBM processing complex), 24-27
  - Center description, 59
    - for models of existing systems, 283
    - from RMF, 385-86
  - Central subsystem, 45, 58, 185-86
  - Channel Activity Report (RMF), 379-80
  - Channel contention (I/O)
    - non-RPS, 225-30
    - RPS, 230-33
  - Channel holding time (I/O), 227
  - Classes, *see* Customer classes; Customer description
  - Closed classes in mixed models, 144-48
  - Closed models, 58, 62, 112-19, 139-45
  - Communication networks, *see* Computer communication networks
  - Complementary network (hierarchical modelling), 152-55
  - Completions (*C*), 40, 47
  - Components of modelling software, 354-60
  - Composite queueing (FESCs), 157
  - Computational algorithms, *see* Algorithms
  - Computer communication networks (case studies)
    - Ethernet, 339-41
    - projecting performance of proposed system, 324-27
    - SNA, 336-39
  - Concurrency control (database) (example), 343-47
  - Conducting a modelling study (Chapter 2), 20-39
    - list of specific points, 38-39
  - Conferences on performance, 17-18
  - Contention time (I/O), 224
  - Controller (I/O), 234-35
  - Core computational routine for separable models, *see* Mean value analysis
    - evolution, 369-70
    - in modelling software, 354
  - CPU, *see* Processors
  - CPU Activity Report (RMF), 378-79
  - CPU replacement, parameterization to reflect, 301
    - case study, 27-29, 33-36
  - CRYSTAL (tool for modelling proposed systems), 327-28
  - CSMA-CD, 340
  - CTSS (case study), 102-3
  - Customer classes (*C*), 62
  - Customer description, 57-58
    - for models of existing systems, 279-82
    - from RMF, 383-85
  - Cyber 173 (case study), 133-34
- D, see* Service demand
- DASD, *see* Disk I/O
  - Database modelling
    - approaches, 343-47
    - high-level front ends, 359-60
  - Decomposition, *see* Flow equivalent service center; Hierarchical modelling
  - Definition of models, 9-12
  - Delay service center, 59
  - Derived observable quantities, 40-42
  - Device homogeneity, 120, 147
  - Direct Access Device Activity Report (RMF), 379, 381
  - Disk I/O (Chapter 10), 222-52, *see also* Algorithms, disk I/O
    - alternate modelling approaches, 248-49
    - inferring parameter values from measurement data, 245-47, 249
    - load balancing
      - asymptotic bounds example, 85-87
      - Univac 1100 case study, 312-13
    - non-RPS example, 227-29
    - representation using modelling software, 247, 367-68
    - RPS example, 232-33
    - service time, components of, 224
  - Domains, 191-96
    - in MVS, 378
    - representation using modelling software, 363
  - Dynamic reconnection (I/O), 237

- Effective disk service time, 224
- End effects in measurement, 43-44, 277
- Ethernet, 339-41
- Evaluation of models, 13-14
  - closed and mixed, *see* Mean value analysis
  - open, *see* Open models
- Event monitor, 276
- Evolution of core computational algorithms for separable models, 369-70
- Evolving systems (Chapter 13), 296-319, *see also* Parameterization
  - case studies and examples
    - simple example, 24-27
    - using asymptotic bounds, 85-87
    - IBM 360, 104-6
    - Univac 1100, 309-15
      - using modelling software, 365-69
    - primary and secondary effects, 296-97
  - Examples, *see* Case studies and examples
  - Execution graphs (for workload characterization), 329-30
  - Existing systems (Chapter 12), 274-95, *see also* Parameterization
    - RMF example, 376-94
  - Experimentation, 3
  - External arrival homogeneity, 120, 149
- FCFS, *see* First-come-first-served
- FESC, *see* Flow equivalent service center
- File placement
  - techniques for modelling, 303-4
  - Univac 1100 case study, 312-13
- First-come-first-served scheduling, 128
  - class-dependent service times, 262-64
  - highly variable service times, 263-66
- Flow balance assumption, 51-52
- Flow control (SNA case study), 336-39
- Flow equivalent service center (FESC), 152-75, *see also* Load dependent service center
  - and separability, 159-60, 191-92
  - and transaction workloads, 157
  - applications
    - Ethernet, 339-41
    - memory constrained system, 185-96
    - SNA flow control, 336-39
    - tightly-coupled multiprocessor, 254-56
  - cost of evaluating a model containing, 157, 159, 191-92
  - desired characteristics, 155-58
  - evaluating high-level models, 159-60
  - obtaining parameters, 158-59
- Flow equivalence and hierarchical modelling (Chapter 8), 152-175, *see also*
  - Flow equivalent service center;
  - Hierarchical modelling;
  - Load dependent service center
- Forced flow law, 47-50
- Front ends for modelling software, 358-60
- Fundamental laws (Chapter 3), 40-56
  - table of, 53
- Global balance, 162-69
  - cost, 162-63
  - details, 166-69
  - to evaluate model of SNA, 336-39
  - to model priority scheduling, 163-66
- Goal-oriented scheduling, 262
- Hardware modifications, *see* Evolving systems
  - parameterization to reflect, 300-303
  - representation using modelling software, 366-68
- Heads of string (I/O), 235-36
- Hierarchical modelling, 152-75, *see also*
  - Flow equivalent service center;
  - Hybrid modelling
    - Algorithm 8.1, 161
  - applications
    - Ethernet, 339-41
    - memory constrained systems, 185-96
    - SNA flow control, 336-39
    - detailed example, 163-66
- Hierarchical workload characterization, 30-34
- Homogeneity assumptions, 120, 147-49
  - and memory modelling, 179-80, 191-92
- Hybrid modelling, 16, 170-73
  - of MVS SRM, 347-49
- IBM (case studies), *see* Amdahl
  - early virtual memory system, 206-9
  - processing complex (workload balancing), 24-27
  - 360, 104-6
  - 3790, 8130, 8140 (insurance company), 35-37, 78-82
- IMS (Amdahl 470 capacity planning case study), 106-9
- Inflation of service demands, *see* Load concealment

- Inputs and outputs of models (Chapter 4), 4-14, 49, 57-68, 99-101, *see also* Parameterization
- Insight, sources of, 35-37
- Installation Performance Specification (IPS) (MVS), 377-78
- Instructional computing acquisition (case study, Prime, VAX), 30-34
- Insurance company case study (IBM 3790, 8130, 8140), 35-37, 78-82
- I/O subsystem, *see* Disk I/O
- I/O subsystem modifications, parameterization to reflect, 302-3
- IPS (IBM MVS Installation Performance Specification), 377-78
- Iterative solution techniques
  - for separable models, *see* Mean value analysis, approximate
  - for non-separable models, *see* Disk I/O; Memory; Processors
- Journals on performance, 17-18
- K*, *see* Number of centers
- Last-come-first-served (LCFS) scheduling, 129
- Latency time (I/O), 224
- LCFS, 129
- Little's law, 40, 42-46
  - and flow balance assumption, 51-52
  - and forced flow law, 48-50
  - and mean value analysis, 112, 139
  - and memory constrained systems, 49-50, 55-56
  - application at various levels, 44-46
  - used to derive asymptotic bounds on performance, 76-77
- Load concealment
  - in evaluating mixed models, 145
  - in modelling shared disks, 243
- Load dependent service center (Chapter 20), 120-21, 149, 156-58, 403-8 *see also* Flow equivalent service center
  - evaluating models containing, 159-60, 403-8
  - MVA algorithms for, 403-8
  - use in modelling tightly-coupled multiprocessor, 254-56, 310-11
- Load independent service center, 120-21, 149, *see also* Queueing service center
- Local area networks, 339-41
- Logical I/O operation, 287-91
- Loosely-coupled multiprocessor, 237, 242-43, 253
  - Cyber 173 case study, 133-34
- M*, *see* Memory constraint
- MAP (queueing network modelling software package), 360-69
- Mean value analysis (MVA), 108-19, 134-48, *see also* Algorithms, mean value analysis; Evaluation of models; Open models
  - approximate, 117-19, 142-45
  - calculating outputs from, 116, 142
  - exact, 112-17, 139-42
  - implementations, 395-402
  - key equations, 112, 139
  - load dependent centers, 403-8
  - mixed models, 144-48
  - population precedence, 115-16, 140-41
  - residence time equation, 112, 139
- Measurement, 21
  - choice of interval, 43-44
  - event recording, 276
  - sampling, 276
  - to parameterize FESCs, 158
  - useful data items, 275-79
  - using RMF, 376-94
- Measurement data, inadequacy of, 34-36
- Measurement interval duration ( $\bar{T}$ ), 40
- Memory (Chapter 9), 179-221, *see also* Memory constraint; Multiprogramming level; Paging
  - admission policies, hybrid modelling of, 172-73
  - assumptions made in separable models, 179-80
  - case studies
    - early IBM virtual memory system, 206-9
    - VAX/VMS, 209-17
  - robustness of separable models, 180
- Memory constraint ( $\bar{M}$ ), 184-96, *see also* Algorithms, memory constraints
  - Little's law used to evaluate, 49-50, 55-56
  - representation using modelling software, 363
- Memory expansion, parameterization to reflect, 209-17, 301-2
- Method of stages (global balance), 268-69

need for, 65-67  
Notation, table of, 53  
Number in system ( $N$  as input,  $Q$  as output), 60-61, *see also* Little's law  
Number of centers ( $K$ ), 58, 63  
  
Objectives of modelling study, importance of understanding, 27-30  
Observation interval, 43-44  
One step behavior, 120, 147  
Open classes in mixed models, 144-48

performance-oriented design, *see* proposed systems  
Performance period (MVS), 377  
Performance projection, *see* Evolving systems; Proposed systems  
    simple example, 24-27  
Pessimistic bounds on performance, 71  
Physical I/O operations, 287-91  
Population, *see* Number in system; Little's law  
Population precedence (MVA), 115-16, 140-41

- Primary and secondary effects of modifications, 12-13, 20-21, 248, 296-97
- Prime (case study, instructional computing acquisition), 30-34
- Priority scheduling, 256-62  
 Algorithm 11.1, 259  
 representation using modelling software, 362-63  
 an approach using global balance, 163-66
- Processing capacity, *see* Bounds on performance  
 of open models, 109, 134-35
- Processor sharing scheduling, 129
- Processors (Chapter 11), 253-71, *see also* CPU replacement; Multiprocessor; Scheduling discipline
- Program listings  
 single class exact MVA, 395-97  
 multiple class exact MVA, 398-402
- Projecting performance, *see* Evolving systems; Proposed systems
- Projection phase (modelling cycle), 22-27
- Proposed systems (Chapter 14), 320-34  
 case studies, 324-27, 329-32
- $Q$ , *see* Number in system; Queue length
- Queue length ( $Q$ ), 60-61, *see also* Little's law  
 at channel (I/O), 227  
 in open models, 111, 136
- Queue length distribution, 62  
 for load dependent centers, 405-7
- Queueing network modelling software (Chapter 16), 354-73  
 approximation transformations, 356  
 core computational routine, 354  
 example use, 360-69  
 high-level front ends, 358-60  
 user interface, 356-58
- Queueing service center, 59
- Queueing theory, 16
- $R$ , *see* Residence time; Response time
- Ready customer (memory constrained system), 185
- Reconnect (I/O), 230  
 algorithms, 237  
 failure, estimating probability of, 241
- Relationships among observed quantities, table of, 54
- Remote terminal emulation, 27, *see also* Benchmarking
- Residence time ( $R$ ), 42, 61, *see also* Mean value analysis; Open models; Response time
- Residence time equation (MVA), 112, 139, *see also* Mean value analysis  
 for load dependent centers, 405-7  
 modifications  
 for biased processor sharing, 261-62  
 for FCFS with class-dependent service times, 262-63  
 for FCFS with highly variable service times, 265  
 for priority scheduling, 258
- Residual service time, 265
- Resource Measurement Facility (RMF) (IBM), use of, 376-94
- Response time ( $R$ ), 42, 61, *see also* Mean value analysis; Open models; Residence time
- Response time law, 46  
 to estimate think times, 282  
 to evaluate memory constrained system, 207
- Retry (I/O), 230
- RMF (IBM Resource Measurement Facility), use of, 376-94
- Rotation time (I/O), 230
- Rotational position sensing (RPS), 230-33, *see also* Disk I/O  
 Algorithm 10.2, 232
- Routing homogeneity, 120, 147
- RPS, *see* Rotational position sensing
- $S$ , *see* Service time per visit
- Sampling monitor, 276
- Saturated system, 72, 109, 134-35
- Scheduling discipline, *see* First-come-first-served; Last-come-first-served; Priority; Processors; Separable models, scheduling disciplines in
- Secondary bottleneck, 82-83
- Secondary effects of modifications, 12-13, 296-97, 306-9
- Seek time (I/O), 224
- Sensitivity analysis, 33-36
- Separable models, 15  
 advantages, 121-22  
 arrival instant theorem, 114  
 assumptions, 20-22, 62, 147, 149  
 inputs and outputs, 57-68  
 limitations, 64-66, 121  
 robustness, 121, 180, 222-25, 248, 261, 264

- Separable models (*cont'd.*)  
 scheduling disciplines in, 59, 62, 128-29, 253-54  
 theoretical foundations, 119-21, 147, 149
- Service center, 4-8, *see also* Center description
- Service center flow balance, 120, 147
- Service center types, 59
- Service demand ( $D$ ), 48-49, 59-60  
 estimation of  
   for models of existing systems, 283-91  
   CPU, 285-87  
   I/O, 287-91  
   difficulties, 284  
   using RMF, 386, 388-91  
 inflation, *see* Load concealment
- Service objective (MVS), 377
- Service rate ( $\mu$ ) (FESC), 156-58, *see also*  
 Flow equivalent service center;  
 Load dependent service center
- Service rate of tightly-coupled multiprocessor, 255-56
- Service time per visit ( $S$ ), 41
- Service time homogeneity, 120, 149
- Shadow CPU technique (priority scheduling), 258-61
- Shared disks, 242-43
- Simulation, 15-16  
 in hybrid modelling, 170-73  
 to parameterize FESCs, 158
- Simultaneous resource possession, 185
- Single class models (Chapter 6), 98-126, *see also* Algorithms  
 advantages, 98, 122, 127-28  
 case studies illustrating use, 102-9  
 examples illustrating evaluation, 112-13, 116-17, 119  
 inaccuracy when workload is heterogeneous, 130-31  
 inputs and outputs, 57-62  
 limitations, 98-99
- Single server queue, 4-8
- SNA (IBM System Network Architecture), 336-39
- Software modifications, parameterization to reflect, 303-6
- Software resources, 342-43
- Software specifications, refinement of (modelling proposed systems), 324-36
- Solution of models, *see* Evaluation of models
- SRM (IBM MVS System Resources Manager)  
 description, 377-78  
 hybrid model of, 346-49
- State, *see* Global balance, details of
- States of a customer (memory constrained system), 185
- Static reconnection (I/O), 237
- Stochastic analysis, xii-xiii, 123, 150
- Swapping, 196-201  
 moving from drum to disk (Univac 1100 case study), 314-15  
 to a dedicated device, 197-99  
   Algorithm 9.4, 199  
 to a shared device, 198-201  
   Algorithm 9.5, 200
- System Network Architecture (SNA) (IBM), 336-39
- System Resources Manager (IBM MVS)  
*see* SRM
- $T$ , *see* Measurement interval duration
- Tape I/O, 290-91, 386, 387, 391
- Terminal workload, 58
- Think time ( $Z$ ), 46  
 estimating, 282, 384
- Thrashing, 202-5
- Throughput ( $\bar{X}$ ), 41, 61  
 bounds on, 72-73, 76-77  
 in memory constrained systems, 187-88  
 in open models, 109, 134-35  
 versus multiprogramming level, 181-84
- Tightly-coupled multiprocessor, 253, 254-56  
 upgrading from single to dual processor (Univac 1100 case study), 310-11
- Tolerances, in validation, 292
- Transaction (database), 344-46
- Transaction workload, 58, *see also* Open models  
 calculating utilization, 51-52
- Transfer time (I/O), 224
- TSO (IBM Time Sharing Option) (case studies), 24-27, 360-69, 376-94
- $U$ , *see* Utilization
- Unattributed busy time, apportioning, 285-91
- Univac 1100 case studies, 309-15
- User interface for modelling software, 356-58

- Utilization ( $U$ ), 41, 61  
Utilization law, 41-42, 45  
    and flow balance assumption, 51-52
- $V$ , *see* Visit count  
Validation, 22-27, 291-92  
Validation phase (modelling cycle), 22-27  
Variability in multiprogramming level,  
    181-84  
Variance in service time, 265-66  
VAX (case studies)  
    instructional computing acquisition, 30-  
        34, 51-52  
    memory modelling, 209-17  
Verification phase (modelling cycle), 22-27  
Virtual memory, 201-5  
Visit count ( $V$ ), 47-50
- $W$ , *see* Accumulated time in system  
Waiting customer (memory constrained  
    system), 185  
Workload Activity Report (RMF), 379-80,  
    382  
Workload balancing (IBM processing com-  
    plex case study), 24-27  
Workload characterization  
    hierarchical, 30-33  
    in modelling proposed systems, 323-34  
    using RMF, 376-94
- Workload components  
    identification, 279-81  
    modifications, parameterization to  
        reflect, 399-400  
Workload intensity, 57-58, 62, 63, 100-101  
    determining for existing systems, 281-82  
    from RMF, 383-85  
Workload measures, 22  
Workload modifications  
    parameterization to reflect, 297-300  
    PDP-10 case study, 132-33  
    representation using modelling software,  
        365  
Workload representation in single class  
    models, 99-101  
Workload types, 57-58  
    selecting for models of existing systems,  
        279-81  
    from RMF, 383-85
- $X$ , *see* Throughput
- $Z$ , *see* Think time
- $\lambda$ , *see* Arrival rate
- $\mu$ , *see* Service rate