Index

Accessibility, 11, 13  
Agility, 118  
Analyze application tolerances, 126  
Application, 58, 60  
Application capacity management, 31, 33–35  
Application demand management, 124  
Application latency, 12  
Application performance management, 73  
Application service level, 68  
Application service provider strategy, 68–69  
Application service quality, 11–13  
Assurance, 23, 113  
Automation, 5, 71, 79, 170, 177  
Balance, 100–101  
Become a learning organization, 59  
Bottled capacity, 47, 83, 95  
Broad network access, 2, 3  
Buffers, 124, 129, 135, 196  
Business benefits of cloud computing, 4  
Cadence, 30, 42, 61–63, 137, 171–172, 180, 194  
Capability maturity model, 184  
Capacity, 19–20  
Capacity emergency, 21, 47, 85, 103, 129, 133–134, 144, 175, 182  
Capacity fulfillment, 26  
Capacity management, 22–24  
Capacity management metrics, 172–173  
Capacity management overhead, 48–49  
Capacity metrics, 168–169  
Capacity rating, 94–95  
Carbon footprint, 151, 180  
Carrying costs, 114–115  
Catastrophic failure, 139  
Changeover time, 63  
Cloud auditor, 7  
Cloud computing, 1–2, 6, 9, 18, 20, 24  
Cloud service broker, 7  
Cloud service customer, 6–7, 57, 79–80  
Cloud service developer, 7  
Cloud service provider, 6  
Cloud service user, 6  
Complexity overhead, 51, 77  
Component failure, 129, 159  
Consumption, 50, 72, 89, 90  
Continuous improvement, 61–62  
Cost of poor quality, 175, 177, 179, 181–182  
Curtailment, 76 93  
Dead on Arrival, 48, 116  
Decision and planning, 20, 24, 26, 39, 68  
Delivery relationships, 6, 8  
Demand, 13–16, 20–21  
Demand driven capacity management, 4, 6, 108–109  
Demand management, 21, 39, 55, 64, 71, 97–98, 102, 119, 122, 124, 196  
Demand variability, 16–18  
Develop exceptional people, 58–59  
Disaster recovery, 58, 95, 136, 142–143  
Dispatch curve, 85, 97  
Disruption of IT service management, 5  
DOA, 48, 52, 179, 182  
Drain time, 154, 159  
Economic dispatch, 84, 101–103, 115, 150, 152  
Economically perfect capacity, 169–170  
Efficiency, 43, 109
Elasticity, 8, 109
Element shutdown time, 157–158
Element startup time, 155–157
Emergency rating, 94
Emergency reserve(s), 139, 142–144
Energy balance, 100, 101
Essential characteristics of cloud computing, 108–109
eTOM capacity management, 23–24, 28–29
Exceeding forecast, 137, 145
Excess application capacity, 46, 108
Excess online infrastructure capacity, 46, 74
Excess physical infrastructure capacity, 46–47, 175
Explicit demand management actions, 125
Failure, 51, 58, 76–77, 136
Force majeure, 139, 142–143, 146–147, 197
Foundation, 62
Fulfillment, 26, 30
Full stream analysis methodology, 126–127
Full stream optimization, 67, 77–79
Functional component(s), 59, 70
Fungibility, 5, 13, 91–92
Georedundancy, 143
Go see, 55, 59
Goal, 41–43
Grow, learn and teach others, 58
Heat, 53–54, 65
Horizontal growth, 33–34
Ideal reserve, 144–145
Immaturity model, 185
Implement rapidly, 59
Inadequate capacity, 45, 47
Indicators, 180, 182, 184, 203
Infrastructure commitment, 173–174
Infrastructure commitment metrics, 173–174
Infrastructure demand management, 120–121
Infrastructure element failure, 136–137
Infrastructure overhead, 48
Infrastructure service provider, 86, 88, 92, 98, 102
Inventory, 57, 106
Inventory costs, 45–46
ITIL capacity management, 27–28
Kaizen, 61
Key performance indicators, 57, 167
Key principle indicators, 180–181
Key principles, 54–55
Latency, 12, 52, 72, 88
Lead time, 27, 30
Lead time demand, 137–139
Leaking and lost resources, 53
Lean application service provider, 68–69
Lean demand management, 119–120
Lean infrastructure commitment, 199–201
Lean infrastructure service provider, 73–74
Lean reserve, 45, 133
Lean thinking, 41–43
Learn-do-lean, 204
Learning organization, 59, 181
Level the workload, 55
Load, 125, 127
Load balancers, 124–125
Load balancing, 10, 35
Load placement, 71–72
Locational considerations, 103
Locational marginal price (LMP), 95
Long-term perspective, 58, 181
Maintenance actions, 76, 88, 121, 124
Mandatory demand shaping, 122
Master practices, 56–57
Mean opinion score, 12–13, 96
Measured service, 2, 4, 109
Metrics, 168, 172
MIPS, 90
Move towards flow, 168, 204
Multi-tenancy, 2, 70
Natural demand patterns, 127
Network provider, 7
Newsvendor problem, 105
NIIT, 1, 6
Normal rating, 94
Normal reserve, 139, 197
On-demand self-service, 2, 4, 53, 87, 107, 109
Operating reserve, 98
Order completeness, 115
Order lead time, 115
Order reliability, 115
Index

215

Outgrowth, 35
Overload, 113, 115, 124, 125
Partners help each other, 58–59
Pay-as-you-use licensing, 70
Perfect capacity metrics, 168–172
Perfect dispatch, 201
Perfection challenge, 61–62
Performance indicators, 182
Performance management, 21–23
Philosophy of lean cloud, 194
Physical resource capacity management, 30–32, 37, 38
Placement, 93
Planning cycles, 62–63
Power management overhead, 45, 50
Pricing, 21, 78
Pricing models, 78, 79
Private cloud, 8
Provisioning reliability, 52, 178
Public cloud, 8
Pull versus push, 55, 180
Queues, 134, 196
Random demand peaks, 135
Rapid elasticity, 2, 4, 20, 24, 29
Rating, 94
Recovery point objective, 142
Recovery time objective, 136
Reliability, 117
Reserve, 45–46
Reserve capacity, 45–46, 133–135
Resource, 129–130
Resource allocation failure, 51–52
Resource curtailment, 76, 93
Resource failure, 136
Resource overhead, 49
Resource pooling, 3, 150
Resource scheduling, 80, 130, 198
Respect, 59–61
Retail analog, 110–112
Retainability, 11
Roles in cloud computing, 1, 6–9
Sample application, 10, 20, 34, 179
Scheduling, 76, 80
Seeing waste, 42
Self service by end users, 71, 79
Service, 39, 83
Service curtailment, 99–100, 141
Service delivery chain, 41
Service level, 67, 69, 71, 80
Service quality, 11–13, 15–16, 18
Shutdown time, 157–160
Slashdot events, 17
Spot VM instances, 122
Stability limit, 95
Startup time, 155–156
Stop and fix problems, 55–56
Strategy for lean cloud, 42, 55, 59
Supply v. demand, 53, 60, 95
Take a long-term perspective, 181
Teach others, 58
Technically perfect capacity, 169
Testing, 116
Tolerances, 128–129
Traditional capacity management, 25–26, 61, 93
Traditional production chain, 62
Transport latency, 12, 96
Unit commitment, 74, 101
User workload placement, 71–72
Value, 36, 39, 41–42, 44, 47, 55
Value chain, 39, 58, 61, 67, 86
Variability, demand, 16–18
Vertical growth, 34
Virtual resource balance, 101–102
Virtual resource capacity management, 32–33, 39, 61, 73
Visual management, 57, 181
Voluntary demand shaping, 123, 142
Waste, 38–39
Waste heat, 179–180
Waste metrics, 174
Well-tested technology, 57–58
Whipsaw events, 18
Work toward flow, 61
Working stock, 112
Workload migration, 50–51, 177–178
Workload placement, 71–72
Yield management, 21