



Contents

I	CONCEPTS & TECHNOLOGIES	23
1	Introduction to Cloud Computing	25
1.1	Introduction	26
1.1.1	Definition of Cloud Computing	26
1.2	Characteristics of Cloud Computing	27
1.3	Cloud Models	29
1.3.1	Service Models	29
1.3.2	Deployment Models	29
1.4	Cloud Computing Concepts & Technologies	32
1.4.1	Virtualization	32
1.4.2	Load Balancing	32
1.4.3	Scalability & Elasticity	37
1.4.4	Deployment	38
1.4.5	Replication	39
1.4.6	Monitoring	40
1.4.7	Software Defined Networking	42
1.4.8	Network Function Virtualization	46
1.4.9	Identity and Access Management	47
1.4.10	Service Level Agreements	49
1.4.11	Billing	50
1.5	Cloud Computing Services & Platforms	51

1.6	Compute Services	51
1.6.1	Amazon Elastic Compute Cloud	53
1.7	Storage Services	53
1.7.1	Amazon Simple Storage Service	53
1.8	Database Services	54
1.8.1	Amazon Relational Data Store	54
1.8.2	Amazon DynamoDB	54
1.9	Application Services	55
1.9.1	Application Runtimes & Frameworks	55
1.9.2	Queuing Services	55
1.9.3	Email Services	55
1.9.4	Notification Services	56
1.9.5	Media Services	56
1.10	Content Delivery Services	56
1.10.1	Amazon CloudFront	57
1.11	Analytics Services	57
1.11.1	Amazon Elastic MapReduce	57
1.12	Deployment & Management Services	58
1.12.1	Amazon Elastic Beanstalk	58
1.12.2	Amazon CloudFormation	58
1.13	Identity & Access Management Services	58
1.13.1	Amazon Identity & Access Management	58
1.14	Open Source Private Cloud Software	59
1.14.1	CloudStack	59
1.14.2	Eucalyptus	59
1.14.3	OpenStack	60
1.15	Generic Case Studies	61
1.15.1	Cloud Computing for Healthcare	61
1.15.2	Cloud Computing for Energy Systems	62
1.15.3	Cloud Computing for Transportation Systems	63
1.15.4	Cloud Computing for Manufacturing Industry	64
1.15.5	Cloud Computing for Government	64
1.15.6	Cloud Computing for Education	65
1.15.7	Cloud Computing for Mobile Communication	66
2	Virtual Machines & Compute Services	69
2.1	Virtualization	70
2.2	Compute Services	72
2.3	Amazon EC2	73
2.3.1	EC2 Instance Types	73

2.3.2	Amazon Machine Image	74
2.3.3	Security Groups	74
2.3.4	Tenancy Options	74
2.3.5	Pricing Options	75
2.3.6	Placement Groups	75
2.4	AWS EC2 Auto Scaling	81
2.4.1	Launch Configuration	81
2.4.2	Auto Scaling Group	81
2.4.3	Scaling Policy	81
2.5	AWS Elastic Load Balancing	85
2.5.1	Internet Facing or Internal Load Balancers	85
2.5.2	Listeners	85
2.5.3	Health Checks	85
2.5.4	Sticky Sessions	85
2.5.5	Connection Draining	86
2.5.6	Using an Application Load Balancer	86
2.6	Amazon Virtual Private Cloud (VPC)	90
2.6.1	Subnets	90
2.6.2	Internet Gateways	93
2.6.3	Route Tables	93
2.6.4	NAT Instances and NAT Gateways	95
2.6.5	Network ACLs	96
2.6.6	Security Groups	98
2.6.7	Endpoints	100
2.6.8	DHCP Option Sets	102
2.6.9	Peering	102
2.6.10	Virtual Private Network (VPN)	102
3	Cloud Application Development	105
3.1	Design Considerations for Cloud Applications	106
3.1.1	Scalability	106
3.1.2	Reliability & Availability	107
3.1.3	Security	107
3.1.4	Maintenance & Upgradation	108
3.1.5	Performance	108
3.2	Cloud Application Design Methodologies	108
3.2.1	Service Oriented Architecture	108
3.2.2	Cloud Component Model	111
3.2.3	Model View Controller	116
3.2.4	RESTful Web Services	117

3.3	Reference Architectures for Cloud Applications	119
3.4	Python Web Application Framework - Django	122
3.4.1	Django Architecture	122
3.4.2	Starting Development with Django	123
3.5	Designing a RESTful Web API	137
3.6	Case Studies	144
3.6.1	Image Processing App	144
3.6.2	Document Storage App	152
3.6.3	Social Media Analytics App	165
4	Serverless Applications	183
4.1	Introduction to Serverless Computing	184
4.1.1	Pros and Cons of Serverless	186
4.2	Serverless Use Cases	187
4.2.1	Web Application	187
4.2.2	Real-time Stream Processing	188
4.2.3	Real-time File Processing	188
4.2.4	IoT Backends	189
4.2.5	Mobile Application Backends	189
4.2.6	Extract, Transform, Load (ETL)	190
4.2.7	Scheduled Tasks	190
4.2.8	Bots	190
4.3	Serverless Design Patterns	191
4.3.1	Asynchronous Processing with Messaging Queues	191
4.3.2	Load Balancing with Multiple Consumers	191
4.3.3	Priority Queues	192
4.3.4	Command Pattern	192
4.3.5	Fan-Out Pattern	192
4.3.6	Pipes and Filters	194
4.4	AWS Lambda	195
4.4.1	Triggers	196
4.4.2	Lambda Function Example	198
4.5	Serverless Concepts	204
4.5.1	Push and Pull Models of Invocation	204
4.5.2	Concurrent Execution	204
4.5.3	Execution Duration	204
4.5.4	Container Reuse	205
4.5.5	Cold and Warm Functions	205
4.6	Case Study: Serverless Photo Gallery Application	206

II	STORAGE, ANALYTICS & SECURITY	217
5	Cloud Storage	219
5.1	Amazon Simple Storage Service (S3)	220
5.1.1	Buckets	221
5.1.2	Objects	223
5.1.3	Managing Access with Bucket Policies and ACLs	225
5.1.4	Cross-Origin Resource Sharing (CORS)	227
5.1.5	Versioning	228
5.1.6	Encryption	228
5.1.7	Static Website Hosting	229
5.1.8	Transfer Acceleration	229
5.1.9	Durability and Availability	229
5.1.10	Consistency	230
5.1.11	Storage Classes	230
5.1.12	Object Lifecycle Management	231
5.1.13	Cross-Region Replication	231
5.2	Amazon Elastic File System (EFS)	231
5.2.1	EFS Performance Modes	235
5.2.2	EFS Throughput Modes	235
5.2.3	Encryption	236
5.3	Amazon Elastic Block Store (EBS)	236
5.3.1	Volumes	236
5.3.2	Snapshots	238
5.3.3	Encryption	239
5.4	AWS Storage Gateway	239
6	Relational (SQL) Databases	241
6.0.1	OLTP vs OLAP	243
6.1	MySQL	244
6.2	Amazon Relational Data Store	246
6.2.1	Setting up an RDS Database Instance	247
6.3	Case Study: Photo Gallery Application	255
6.4	Amazon Redshift	269
6.5	Amazon ElastiCache	279
7	NoSQL Databases	283
7.1	Consistency, Availability & Partition Tolerance (CAP)	285
7.2	Key-Value Databases	286
7.2.1	Amazon DynamoDB	287

7.3	Case Study: Photo Gallery Application	292
7.3.1	Cassandra	299
7.4	Document Databases	302
7.4.1	MongoDB	302
7.5	Column Family Databases	306
7.5.1	HBase	306
7.6	Graph Databases	314
7.6.1	Amazon Neptune	314
7.6.2	Neo4j	319
8	Big Data Analytics in the Cloud	327
8.1	What is Analytics?	328
8.1.1	Descriptive Analytics	328
8.1.2	Diagnostic Analytics	330
8.1.3	Predictive Analytics	330
8.1.4	Prescriptive Analytics	330
8.2	What is Big Data?	331
8.3	Characteristics of Big Data	332
8.3.1	Volume	332
8.3.2	Velocity	332
8.3.3	Variety	332
8.3.4	Veracity	333
8.3.5	Value	333
8.4	Domain Specific Examples of Big Data	333
8.4.1	Web	333
8.4.2	Financial	335
8.4.3	Healthcare	335
8.4.4	Internet of Things	336
8.4.5	Environment	337
8.4.6	Logistics & Transportation	338
8.4.7	Industry	340
8.4.8	Retail	341
8.5	Analytics Flow for Big Data	341
8.5.1	Data Collection	341
8.5.2	Data Preparation	342
8.5.3	Analysis Types	342
8.5.4	Analysis Modes	342
8.5.5	Visualizations	344
8.6	Big Data Stack	344
8.6.1	Raw Data Sources	344
8.6.2	Data Access Connectors	345
8.6.3	Data Storage	345

8.6.4	Batch Analytics	347
8.6.5	Real-time Analytics	347
8.6.6	Interactive Querying	348
8.6.7	Serving Databases, Web & Visualization Frameworks	348
8.7	Mapping Analytics Flow to Big Data Stack	349
8.8	Case Study: Genome Data Analysis	349
8.9	Case Study: Weather Data Analysis	358
8.10	Analytics Patterns	361
9	Queues & Connectors	369
9.1	Data Acquisition Considerations	370
9.1.1	Source Type	370
9.1.2	Velocity	370
9.1.3	Ingestion Mechanism	370
9.2	Publish - Subscribe Messaging Frameworks	371
9.2.1	Apache Kafka	372
9.2.2	Amazon Kinesis	377
9.2.3	AWS IoT	391
9.3	Big Data Collection Systems	400
9.3.1	Apache Flume	400
9.3.2	Apache Sqoop	413
9.3.3	Importing Data with Sqoop	413
9.3.4	Selecting Data to Import	414
9.3.5	Custom Connectors	415
9.3.6	Importing Data to Hive	415
9.3.7	Importing Data to HBase	415
9.3.8	Incremental Imports	415
9.3.9	Importing All Tables	416
9.3.10	Exporting Data with Sqoop	416
9.4	Messaging Queues	416
9.4.1	RabbitMQ	417
9.4.2	ZeroMQ	418
9.4.3	RestMQ	420
9.4.4	Amazon SQS	421
9.5	Custom Connectors	423
9.5.1	REST-based Connectors	423
9.5.2	WebSocket-based Connectors	426
9.5.3	MQTT-based Connectors	427
9.6	Case Study: Air Quality Monitoring Application	429
9.6.1	Variant-1 using AWS IoT, DynamoDB, Lambda, CloudWatch & EC2	431
9.6.2	Variant-2 using Kinesis Streams, DynamoDB, Lambda, CloudWatch & EC2	448
9.6.3	Variant-3 using Kinesis Streams, Kinesis Analytics, Lambda, & EC2	452

9.6.4	Variant-4 using Kinesis Streams, Kinesis Firehose, S3 & EMR	458
10	Batch Analytics	463
10.1	HDFS	466
10.1.1	HDFS Architecture	466
10.1.2	HDFS Usage Examples	470
10.2	Hadoop	472
10.2.1	Hadoop YARN	472
10.2.2	Hadoop Schedulers	476
10.3	MapReduce	477
10.3.1	Numerical Summarization	480
10.3.2	Top-N	488
10.3.3	Filter	490
10.3.4	Distinct	492
10.3.5	Binning	494
10.3.6	Inverted Index	496
10.3.7	Sorting	498
10.3.8	Joins	500
10.4	Hadoop - MapReduce Examples	505
10.4.1	Batch Analysis of Sensor Data	505
10.4.2	Batch Analysis of N-Gram Dataset	507
10.4.3	Find top-N words with MapReduce	509
10.4.4	Batch Analysis of MovieLens Dataset	510
10.5	Pig	520
10.5.1	Loading Data	521
10.5.2	Data Types in Pig	521
10.5.3	Data Filtering & Analysis	522
10.5.4	Storing Results	523
10.5.5	Debugging Operators	523
10.5.6	Pig Examples	525
10.6	Case Study: Batch Analysis of News Articles	525
10.7	Apache Oozie	531
10.7.1	Oozie Workflows for Data Analysis	531
10.8	Apache Spark	539
10.8.1	Spark Operations	540
10.9	Search	544
10.9.1	Apache Solr	544
10.9.2	Elasticsearch	554
10.9.3	Case Study: Book Search Application	568

11	Real-time Analytics	575
11.1	Apache Storm	576
11.1.1	Concepts	576
11.1.2	Stream Groupings	577
11.1.3	Architecture	578
11.1.4	Reliable Processing	579
11.1.5	Case Study: Real-time Twitter Sentiment Analysis	580
11.1.6	Case Study: Real-time Weather Data Analysis	592
11.2	Apache Spark	601
11.2.1	Case Study: Real-time Sensor Data Analysis	604
11.2.2	Case Study: Real-Time Parking Sensor Data Analysis	605
11.2.3	Case Study: Real-time Twitter Sentiment Analysis	611
11.2.4	Case Study: Windowed Analysis of Tweets	617
11.3	Apache Flink	618
11.3.1	Flink Sources	622
11.3.2	Flink Sinks	622
11.3.3	Flink Transformations	622
11.3.4	Checkpointing	626
11.3.5	Getting Started with Flink	627
11.3.6	Case Study: Real-time Air Quality Monitoring	631
11.3.7	Case Study: Real-time Twitter Sentiment Analysis	636
12	Cloud App Monitoring & Benchmarking	639
12.1	Introduction	640
12.1.1	Trace Collection/Generation	641
12.1.2	Workload Modeling	641
12.1.3	Workload Specification	641
12.1.4	Synthetic Workload Generation	641
12.1.5	User Emulation vs Aggregate Workloads	642
12.2	Workload Characteristics	642
12.3	Application Performance Metrics	643
12.4	Design Considerations for a Benchmarking Methodology	643
12.5	Benchmarking Tools	644
12.5.1	Types of Tests	649
12.6	Deployment Prototyping	649
12.7	Load Testing & Bottleneck Detection Case Study	650
12.8	Amazon CloudWatch	654
13	Cloud Security	659
13.1	Introduction	660

13.2	CSA Cloud Security Architecture	661
13.3	Authentication	663
13.3.1	Single Sign-on (SSO)	663
13.4	Authorization	666
13.5	Identity & Access Management	669
13.6	Data Security	670
13.6.1	Securing Data at Rest	670
13.6.2	Securing Data in Motion	674
13.7	Key Management	675
13.8	Auditing	677
13.9	AWS Identity and Access Management (IAM)	677
13.9.1	Users	677
13.9.2	Groups	681
13.9.3	Policies	682
13.9.4	Roles	682
13.10	AWS Key Management Service (KMS)	684
13.11	AWS CloudHSM	685
13.12	AWS Directory Service	685
14	Other Cloud Services	689
14.1	Amazon Route 53	690
14.2	Amazon CloudFront	691
14.3	AWS Elastic Beanstalk	692
14.4	AWS CloudFormation	693
14.5	AWS CloudTrail	693
14.6	AWS OpsWorks	693
14.7	AWS Config	694
14.8	AWS Trusted Advisor	695
14.9	AWS Data Pipeline	695
14.10	Amazon SNS	695
14.11	Amazon SWF	696
III	AWS WELL-ARCHITECTED FRAMEWORK	699
15	AWS Well Architected Framework	701

16	Applying the Operational Excellence Pillar	707
16.1	Best Practice Area: Prepare	708
16.2	Best Practice Area: Operate	710
16.3	Best Practice Area: Evolve	710
16.4	Applying the Operational Excellence Pillar	710
16.5	Recipe for Operational Excellence Pillar	712
17	Applying the Security Pillar	735
17.1	Best Practice Area: Identity and Access Management	736
17.2	Best Practice Area: Detective Controls	736
17.3	Best Practice Area: Infrastructure Protection	737
17.4	Best Practice Area: Data Protection	739
17.5	Best Practice Area: Incident Response	739
17.6	Applying Security Pillar	740
17.7	Recipe for Security Pillar	740
18	Applying the Reliability Pillar	765
18.1	Best Practice Area: Foundations	766
18.2	Best Practice Area: Change Management	766
18.3	Best Practice Area: Failure Management	769
18.4	Applying Reliability Pillar	769
18.5	Recipe for Reliability Pillar	769
19	Applying the Performance Efficiency Pillar	783
19.1	Best Practice Area: Selection	784
19.2	Best Practice Area: Review	786
19.3	Best Practice Area: Monitoring	786
19.4	Best Practice Area: Tradeoffs	787
19.5	Applying Performance Efficiency Pillar	787
19.6	Recipe for Performance Efficiency Pillar	787
20	Applying the Cost Optimization Pillar	799
20.1	Best Practice Area: Expenditure Awareness	800
20.2	Best Practice Area: Cost-Effective Resources	800

20.3	Best Practice Area: Matching supply and demand	801
20.4	Best Practice Area: Optimizing Over Time	801
20.5	Applying Cost Optimization Pillar	802
20.6	Recipe for Cost Optimization Pillar	803
	Bibliography	820
	Index	821