

# Table of Contents

<b>Preface</b>	1
<hr/>	
<b>Section 1: Introduction to Elastic Stack and Elasticsearch</b>	
<hr/>	
<b>Chapter 1: Introducing Elastic Stack</b>	9
<b>What is Elasticsearch, and why use it?</b>	10
Schemaless and document-oriented	11
Searching capability	11
Analytics	12
Rich client library support and the REST API	12
Easy to operate and easy to scale	13
Near real-time capable	13
Lightning-fast	14
Fault-tolerant	14
<b>Exploring the components of the Elastic Stack</b>	14
Elasticsearch	15
Logstash	15
Beats	16
Kibana	17
X-Pack	17
Security	17
Monitoring	17
Reporting	18
Alerting	18
Graph	18
Machine learning	18
Elastic Cloud	19
<b>Use cases of Elastic Stack</b>	19
Log and security analytics	19
Product search	20
Metrics analytics	21
Web search and website search	22
<b>Downloading and installing</b>	22
Installing Elasticsearch	23
Installing Kibana	24
<b>Summary</b>	24
<b>Chapter 2: Getting Started with Elasticsearch</b>	25
<b>Using the Kibana Console UI</b>	26
<b>Core concepts of Elasticsearch</b>	29

Indexes	30
Types	31
Documents	32
Nodes	33
Clusters	33
Shards and replicas	34
Mappings and datatypes	36
Datatypes	36
Core datatypes	37
Complex datatypes	38
Other datatypes	38
Mappings	38
Creating an index with the name catalog	39
Defining the mappings for the type of product	39
Inverted indexes	42
<b>CRUD operations</b>	44
Index API	44
Indexing a document by providing an ID	44
Indexing a document without providing an ID	45
Get API	46
Update API	46
Delete API	48
<b>Creating indexes and taking control of mapping</b>	49
Creating an index	50
Creating type mapping in an existing index	50
Updating a mapping	52
<b>REST API overview</b>	54
Common API conventions	54
Formatting the JSON response	55
Dealing with multiple indexes	56
Searching all documents in one index	57
Searching all documents in multiple indexes	58
Searching all the documents of a particular type in all indexes	58
<b>Summary</b>	58
<hr/>	
<b>Section 2: Analytics and Visualizing Data</b>	
<hr/>	
<b>Chapter 3: Searching - What is Relevant</b>	61
<b>The basics of text analysis</b>	61
Understanding Elasticsearch analyzers	62
Character filters	63
Tokenizer	64
Standard tokenizer	65
Token filters	66
Using built-in analyzers	67
Standard analyzer	67
Implementing autocomplete with a custom analyzer	72
<b>Searching from structured data</b>	76

---

Range query	78
Range query on numeric types	79
Range query with score boosting	80
Range query on dates	81
Exists query	82
Term query	83
<b>Searching from the full text</b>	<b>84</b>
Match query	86
Operator	88
Minimum should match	88
Fuzziness	89
Match phrase query	90
Multi match query	92
Querying multiple fields with defaults	92
Boosting one or more fields	93
With types of multi match queries	93
<b>Writing compound queries</b>	<b>93</b>
Constant score query	94
Bool query	96
Combining OR conditions	97
Combining AND and OR conditions	98
Adding NOT conditions	99
<b>Modeling relationships</b>	<b>100</b>
has_child query	104
has_parent query	106
parent_id query	108
<b>Summary</b>	<b>109</b>
<b>Chapter 4: Analytics with Elasticsearch</b>	<b>111</b>
<b>The basics of aggregations</b>	<b>111</b>
Bucket aggregations	113
Metric aggregations	113
Matrix aggregations	114
Pipeline aggregations	114
<b>Preparing data for analysis</b>	<b>114</b>
Understanding the structure of the data	115
Loading the data using Logstash	118
<b>Metric aggregations</b>	<b>119</b>
Sum, average, min, and max aggregations	120
Sum aggregation	120
Average aggregation	122
Min aggregation	123
Max aggregation	123
Stats and extended stats aggregations	124
Stats aggregation	124
Extended stats aggregation	125
Cardinality aggregation	126

<b>Bucket aggregations</b>	127
Bucketing on string data	128
Terms aggregation	128
Bucketing on numerical data	133
Histogram aggregation	133
Range aggregation	134
Aggregations on filtered data	136
Nesting aggregations	138
Bucketing on custom conditions	141
Filter aggregation	142
Filters aggregation	143
Bucketing on date/time data	144
Date Histogram aggregation	144
Creating buckets across time periods	145
Using a different time zone	146
Computing other metrics within sliced time intervals	147
Focusing on a specific day and changing intervals	148
Bucketing on geospatial data	150
Geodistance aggregation	150
GeoHash grid aggregation	152
<b>Pipeline aggregations</b>	154
Calculating the cumulative sum of usage over time	154
<b>Summary</b>	156
<b>Chapter 5: Analyzing Log Data</b>	157
<b>Log analysis challenges</b>	157
<b>Using Logstash</b>	160
Installation and configuration	161
Prerequisites	161
Downloading and installing Logstash	162
Installing on Windows	163
Installing on Linux	164
Running Logstash	164
<b>The Logstash architecture</b>	165
<b>Overview of Logstash plugins</b>	168
Installing or updating plugins	169
Input plugins	169
Output plugins	170
Filter plugins	170
Codec plugins	171
Exploring plugins	171
Exploring input plugins	171
File	171
Beats	173
JDBC	176
IMAP	178
Output plugins	179
Elasticsearch	179
CSV	180

---

Kafka	181
PagerDuty	181
Codec plugins	182
JSON	182
Rubydebug	183
Multiline	183
Filter plugins	184
<b>Ingest node</b>	184
Defining a pipeline	185
Ingest APIs	185
Put pipeline API	185
Get pipeline API	187
Delete pipeline API	188
Simulate pipeline API	188
<b>Summary</b>	189
<b>Chapter 6: Building Data Pipelines with Logstash</b>	191
<b>Parsing and enriching logs using Logstash</b>	191
Filter plugins	192
CSV filter	193
Mutate filter	194
Grok filter	196
Date filter	198
Geoip filter	199
Useragent filter	200
<b>Introducing Beats</b>	201
Beats by Elastic.co	202
Filebeat	202
Metricbeat	202
Packetbeat	202
Heartbeat	203
Winlogbeat	203
Auditbeat	203
Journalbeat	203
Functionbeat	204
Community Beats	204
Logstash versus Beats	205
<b>Filebeat</b>	205
Downloading and installing Filebeat	206
Installing on Windows	206
Installing on Linux	207
Architecture	208
Configuring Filebeat	209
Filebeat inputs	213
Filebeat general/global options	216
Output configuration	217
Logging	219
Filebeat modules	220
<b>Summary</b>	223

---

<b>Chapter 7: Visualizing Data with Kibana</b>	225
<b>Downloading and installing Kibana</b>	226
Installing on Windows	227
Installing on Linux	227
Configuring Kibana	230
<b>Preparing data</b>	231
<b>Kibana UI</b>	236
User interaction	237
Configuring the index pattern	238
Discover	241
Elasticsearch query string/Lucene query	246
Elasticsearch DSL query	251
KQL	251
Visualize	261
Kibana aggregations	263
Bucket aggregations	263
Metric	265
Creating a visualization	265
Visualization types	267
Line, area, and bar charts	267
Data tables	267
Markdown widgets	267
Metrics	267
Goals	268
Gauges	268
Pie charts	268
Co-ordinate maps	268
Region maps	268
Tag clouds	269
Visualizations in action	269
Response codes over time	269
Top 10 requested URLs	271
Bandwidth usage of the top five countries over time	273
Web traffic originating from different countries	274
Most used user agent	276
Dashboards	278
Creating a dashboard	278
Saving the dashboard	280
Cloning the dashboard	281
Sharing the dashboard	282
<b>Timelion</b>	282
Timelion	283
Timelion expressions	283
<b>Using plugins</b>	288
Installing plugins	289
Removing plugins	289
<b>Summary</b>	290

---

## Section 3: Elastic Stack Extensions

---

<b>Chapter 8: Elastic X-Pack</b>	293
Installation	294
Activating X-Pack trial account	298
Generating passwords for default users	301
<b>Configuring X-Pack</b>	304
<b>Securing Elasticsearch and Kibana</b>	305
User authentication	305
User authorization	307
Security in action	309
Creating a new user	310
Deleting a user	312
Changing the password	313
Creating a new role	314
Deleting or editing a role	319
Document-level security or field-level security	321
X-Pack security APIs	326
User Management APIs	326
Role Management APIs	328
<b>Monitoring Elasticsearch</b>	330
Monitoring UI	332
Elasticsearch metrics	335
Overview tab	335
Nodes tab	336
The Indices tab	339
<b>Alerting</b>	341
Anatomy of a watch	342
Alerting in action	347
Creating a new alert	348
Threshold Alert	349
Advanced Watch	350
Deleting/deactivating/editing a watch	352
<b>Summary</b>	354

---

## Section 4: Production and Server Infrastructure

---

<b>Chapter 9: Running Elastic Stack in Production</b>	357
<b>Hosting Elastic Stack on a managed cloud</b>	358
Getting up and running on Elastic Cloud	358
Using Kibana	361
Overriding configuration	362
Recovering from a snapshot	362
<b>Hosting Elastic Stack on your own</b>	365
Selecting hardware	365
Selecting an operating system	366
Configuring Elasticsearch nodes	366
JVM heap size	367

Disable swapping	367
File descriptors	367
Thread pools and garbage collector	368
Managing and monitoring Elasticsearch	368
Running in Docker containers	368
Special considerations while deploying to a cloud	369
Choosing instance type	370
Changing default ports; do not expose ports!	370
Proxy requests	370
Binding HTTP to local addresses	370
Installing EC2 discovery plugin	371
Installing the S3 repository plugin	371
Setting up periodic snapshots	371
<b>Backing up and restoring</b>	372
Setting up a repository for snapshots	372
Shared filesystem	373
Cloud or distributed filesystems	374
Taking snapshots	375
Restoring a specific snapshot	375
<b>Setting up index aliases</b>	376
Understanding index aliases	376
How index aliases can help	377
<b>Setting up index templates</b>	378
Defining an index template	378
Creating indexes on the fly	379
<b>Modeling time series data</b>	380
Scaling the index with unpredictable volume over time	380
Unit of parallelism in Elasticsearch	380
The effect of the number of shards on the relevance score	381
The effect of the number of shards on the accuracy of aggregations	381
Changing the mapping over time	382
New fields get added	382
Existing fields get removed	382
Automatically deleting older documents	382
How index-per-timeframe solves these issues	383
Scaling with index-per-timeframe	383
Changing the mapping over time	384
Automatically deleting older documents	384
<b>Summary</b>	384
<b>Chapter 10: Building a Sensor Data Analytics Application</b>	385
<b>Introduction to the application</b>	385
Understanding the sensor-generated data	387
Understanding the sensor metadata	388
Understanding the final stored data	389
<b>Modeling data in Elasticsearch</b>	390
Defining an index template	390



---

Understanding the mapping	393
<b>Setting up the metadata database</b>	393
<b>Building the Logstash data pipeline</b>	394
Accepting JSON requests over the web	395
Enriching the JSON with the metadata we have in the MySQL database	396
The jdbc_streaming plugin	397
The mutate plugin	398
Moving the looked-up fields that are under lookupResult directly in JSON	399
Combining the latitude and longitude fields under lookupResult as a location field	399
Removing the unnecessary fields	400
Store the resulting documents in Elasticsearch	400
<b>Sending data to Logstash over HTTP</b>	401
<b>Visualizing the data in Kibana</b>	402
Setting up an index pattern in Kibana	402
Building visualizations	404
How does the average temperature change over time?	405
How does the average humidity change over time?	406
How do temperature and humidity change at each location over time?	407
Can I visualize temperature and humidity over a map?	409
How are the sensors distributed across departments?	410
Creating a dashboard	411
<b>Summary</b>	415
<b>Chapter 11: Monitoring Server Infrastructure</b>	417
<b>Metricbeat</b>	417
Downloading and installing Metricbeat	418
Installing on Windows	419
Installing on Linux	419
Architecture	420
Event structure	422
<b>Configuring Metricbeat</b>	424
Module configuration	424
Enabling module configs in the modules.d directory	425
Enabling module configs in the metricbeat.yml file	426
General settings	427
Output configuration	428
Logging	430
<b>Capturing system metrics</b>	431
Running Metricbeat with the system module	432
Specifying aliases	435
Visualizing system metrics using Kibana	437
<b>Deployment architecture</b>	440
<b>Summary</b>	441
<b>Other Books You May Enjoy</b>	443
<b>Index</b>	447

---