Keywords and identifiers index

--instance-id, 119
--main-composite, 92
--standalone-application, 118
--template shared, 253
.splmm, 451
.spl, 451
Streams/key, 251
AdminGroup, 253
Aggregate, 93, 94, 130, 132, 135, 138–141, 154, 289, 298, 323, 452, 458
Average, 94
Sum, 136
partitionBy, 289
operator, see aggregate operator
ApplicationLiveness, 237
Associations, 391–393, 424
AuditLogFile, 255
Average, 294
Barrier, 97, 131, 289, 316, 458
directory, 289
operator, see barrier operator
Beacon, 89, 163, 295, 308
iterations, 89
period, 89
Canceling, 233
Classification, 391, 393
model, 391
Classifier, 391, 392
Clustering, 391–393
Custom, 156, 162, 163, 193, 294, 295, 298, 305, 308, 311, 316, 331, 333, 366, 379, 458, 476, 478
as a debugging tool, 193
DCT, 459, 460
size, 460
debug, 250
DataReceiver, 242
dataSender, 242
DeDuplicate, 129, 133
timeOut, 133
DirectoryScanner, 282, 311
directory, 282
sortBy, 282
date, 282
DirectoryScan, 98
DynamicFilter, 305, 306
addKey, 306
removeKey, 306
ERROR, 250
Export, 115, 116, 448–451
FPT, 459, 460
resolution, 460
Failed, 233, 234
FileSink, 90, 91, 193, 248, 280, 282
closeMode, 282
file, 91, 282
flush, 91, 282
format, 91
moveFileToDirectory, 282
tuplesPerFile, 282
as a debugging tool, 193
FileSource, 98, 133, 276, 282, 298, 305, 306, 311, 458
compression, 99
deleteFile, 99, 282
format, 98
csv, 282
gzip, 99
Filter, 93, 137, 160, 286, 287, 355
operator, see filter operator
First, 294, 300
Funct0r, 90, 93, 128, 131–134, 151, 158, 323, 366, 458, 460, 463
operator, see functor operator
INFO, 250
Import, 115, 116, 450–452
subscription, 452
InetSource, 280, 285
URIList, 285
doNotStreamInitialFetch, 285
fetchIntervalSeconds, 285
fetchInterval, 280
incrementalFetch, 285
Join, 95, 96, 131, 133, 134, 157, 151, 160, 289, 478
matchLHS, 96
matchRHS, 96
match, 96
outer, 96
partitionBy, 289
operator, see join operator
LLM_RUN_ID, 172
L_DEBUG, 251
L_ERROR, 251
L_INFO, 251
L_TRACE, 251
Last, 294, 300
LogFileMaxFiles, 249
LogFileMaxSize, 249
MatchRegex, 300, 301
First, 300
Last, 300
partitionBy, 300
pattern, 300
predicates, 300
Max, 130, 135, 294, 296
Merge, 318
Min, 130, 294, 296
ODBCAppend, 278, 279
ODBCSource, 278, 279
Operator.h, 261
OperatorContext.h, 262
OperatorContext, 259, 261, 262, 264
OperatorInputPort.h, 263
OperatorInputPort, 263
OperatorOutputPort.h, 264
OperatorOutputPort, 264
Operator, 154
PerfCounterService, 239
PressureValve, 448, 449
ProcessingElement.h, 260
ProcessingElement, 260, 261, 264
Punctuer, 458
position, 458
punctuate, 458
Regression, 391, 393
Restarting, 234
Running, 224, 225, 233, 234
SIGKILL, 236
Sort, 94, 133, 138, 289
partitionBy, 289
operator, see sort operator
Split, 96, 154, 173, 175, 318, 323
operator, see split operator
SrmNotificationService, 238
Starting, 234
Stopped, 234
Stopping, 234
Sum, 130, 136, 294, 296, 298
Switch, 303–305
Sys.DropFirst, 286
Sys.DropLast, 286
Sys.Exclusive, 172
Sys.Wait, 333
TCPServer, 448, 449
connectionCap, 449
numThreads, 448
port, 448
TCPSink, 99, 229, 277
address, 99
format, 99
port, 99
role, 99, 277
TCPSource, 99, 229, 283, 285, 472, 476
format, 99, 476
line, 283
port, 99
role, 99, 476
TRACE, 250
ThreadedSplit, 332
 Tuple, 155
Union, 323
UserService, 253
XMLParse, 283–285
XPathList, 284
XPath, 284
trigger, 283
XPathList, 284
XPath, 284
addConnection, 242
addKey, 306
addObject, 256
addStreamProperty, 235, 263
address, 99
allPortsReady, 156
applicationScope, 232
blob, 135, 151, 449, 451
boolean, 126, 132, 160
bufferSize, 332
cancelJob, 232, 233, 236
checkPermissionSession, 233, 256
checkPermissionUser, 256
clearM, 331
closeMode, 282
completeChallenge, 256
compression, 99
computeJobPlacement, 219, 234, 241
config, 168, 187, 248, 249, 296, 267, 447, 449
cleanup, 248
wrapper, 266
gdb, 267
valgrind, 267
connect, 242
Keywords and identifiers index

count, 282
createSessionByKey, 233, 256
createSessionByPwd, 233, 256
createThreads, 259
csv, 282
date, 282
decimal64, 149
deleteFile, 99, 282
deleteObject, 256
delta, 298
directory, 282
dnameserver, 220
false, 304
fetchInterval, 280
file, 91, 4, 156, 282
float64, 149, 392
flush, 91, 280
format, 98, 99, 282, 283
gdb, 194, 265–267
getApplicationState, 232, 233, 267, 268
getCheckpointDirectory, 259
getCheckpointingInterval, 259
getCheckpoint, 259, 262
gerectext, 259
getInfrastructureState, 238
getInputPortMetricValue, 333
getInstanceState, 238
gerelaunchCount, 259
geretMetricsData, 236, 239, 267
geretMetrics, 239, 259, 264
geretName, 259
geretPE, 157, 299
geretSAMPConnectionState, 267
geretSCHMetricsData, 236, 241
getShutdownRequested, 157
gerestreamProperty, 263
geretSubmissionTimeValue, 117
getSubscriptionExpression, 263
geretSystemState, 232, 233, 239
geretTimestamp, 295
graph, 89
hostChangeNotification, 236, 238
hostColocation, 169, 239, 240
hostExlocation, 169, 239, 240
hostIsolation, 169, 239, 240, 337
initialStatus, 304
insertR, 331
instanceChangeNotification, 238
int32, 284
int64, 48, 149
iterations, 89
key, 306
line, 283
list, 48, 150, 246, 284, 300, 328, 331
logActivity, 256
logLevel, 249
lookupEntry, 244
map, 150, 246
matchLHS, 96
matchRHS, 96
match, 96
moveFileToDirectory, 283
onMessage, 242
onPunct, 305
onTuple, 163, 331
on, 305
operatorDefinitions, 269
operatorInstances, 269
operatorDriven, 248
outer, 96
output, 89, 90, 126, 300
param, 89, 450
partitionBy, 131, 289, 298, 300
partitionColocation, 168, 239, 240, 322, 324, 447, 449
partitionExlocation, 168, 239, 240, 322
partitionIsolation, 168, 239, 240
pattern, 300
performSynchronousCheckpoint, 259
periodic, 248
period, 89
placement, 168
port, 99
position, 458
predicates, 300
processes, 154–157, 243, 259, 262
punctuate, 458
queue, 333
quiesceHost, 239
registerEntry, 244
registerJobWithAAS, 255
relocateable, 239, 248, 335, 447, 450
removeConnection, 243
removeKey, 306
removeStreamProperty, 263
resolution, 460
restartPE, 233
restartService, 239
restartable, 239, 248, 335, 447, 450
restoreCheckpoint, 259, 262
role, 99, 277
routingInfoNotification, 234, 262, 263
rpcTid, 247
rstring, 90, 161, 246, 392
run, 242
sbd, 63
sc, 91, 219, 264
--main-composite, 92
--standalone-application, 92
-g, 264
sdb, 192, 194, 230, 256, 264–267
setInputPortImportSubscription, 307
<table>
<thead>
<tr>
<th>Keywords and identifiers index</th>
<th>503</th>
</tr>
</thead>
<tbody>
<tr>
<td>setOutputPortExportProperties, 309</td>
<td></td>
</tr>
<tr>
<td>setPEState, 234, 236, 239</td>
<td></td>
</tr>
<tr>
<td>setStreamProperty, 263</td>
<td></td>
</tr>
<tr>
<td>setSubscriptionExpression, 263</td>
<td></td>
</tr>
<tr>
<td>setSubscription, 234, 236</td>
<td></td>
</tr>
<tr>
<td>set, 150, 246</td>
<td></td>
</tr>
<tr>
<td>shutdown, 242</td>
<td></td>
</tr>
<tr>
<td>size, 331, 460</td>
<td></td>
</tr>
<tr>
<td>sortBy, 282</td>
<td></td>
</tr>
<tr>
<td>startPE, 219, 234, 244, 258</td>
<td></td>
</tr>
<tr>
<td>startService, 236</td>
<td></td>
</tr>
<tr>
<td>stopPE, 233, 236</td>
<td></td>
</tr>
<tr>
<td>stopService, 239</td>
<td></td>
</tr>
<tr>
<td>streamViewability, 187</td>
<td></td>
</tr>
<tr>
<td>streams-aas, 220</td>
<td></td>
</tr>
<tr>
<td>streams-hc, 220</td>
<td></td>
</tr>
<tr>
<td>streams-pec, 220</td>
<td></td>
</tr>
<tr>
<td>streams-sam, 220</td>
<td></td>
</tr>
<tr>
<td>streams-sch, 220</td>
<td></td>
</tr>
<tr>
<td>streams-srm, 220</td>
<td></td>
</tr>
<tr>
<td>streams-srw, 220</td>
<td></td>
</tr>
<tr>
<td>streamtool, 118, 119, 219, 221, 224, 229, 230, 232, 236, 249, 251, 254–256, 267</td>
<td></td>
</tr>
<tr>
<td>-P, 119</td>
<td></td>
</tr>
<tr>
<td>-instance-id, 119</td>
<td></td>
</tr>
<tr>
<td>canceljob, 224</td>
<td></td>
</tr>
<tr>
<td>genkey, 251, 253</td>
<td></td>
</tr>
<tr>
<td>getacl, 254</td>
<td></td>
</tr>
<tr>
<td>getsentry, 229</td>
<td></td>
</tr>
<tr>
<td>getproperty, 221</td>
<td></td>
</tr>
<tr>
<td>help, 221</td>
<td></td>
</tr>
<tr>
<td>lshosts, 221</td>
<td></td>
</tr>
<tr>
<td>lsinstance, 221</td>
<td></td>
</tr>
<tr>
<td>lsjobs, 224</td>
<td></td>
</tr>
<tr>
<td>lspe, 224</td>
<td></td>
</tr>
<tr>
<td>man, 221</td>
<td></td>
</tr>
<tr>
<td>mkinstance, 221, 249</td>
<td></td>
</tr>
<tr>
<td>--template audit, 255</td>
<td></td>
</tr>
<tr>
<td>--template shared, 253</td>
<td></td>
</tr>
<tr>
<td>setacl, 254</td>
<td></td>
</tr>
<tr>
<td>setproperty, 249</td>
<td></td>
</tr>
<tr>
<td>startInstance, 221</td>
<td></td>
</tr>
<tr>
<td>submitjob, 224, 249</td>
<td></td>
</tr>
<tr>
<td>capturestate, 230</td>
<td></td>
</tr>
<tr>
<td>submitJob, 219, 232, 233, 247</td>
<td></td>
</tr>
<tr>
<td>submit, 155, 243, 259</td>
<td></td>
</tr>
<tr>
<td>subscription, 452</td>
<td></td>
</tr>
<tr>
<td>systemTopologyScoping, 268</td>
<td></td>
</tr>
<tr>
<td>tar, 249</td>
<td></td>
</tr>
<tr>
<td>telnet, 101</td>
<td></td>
</tr>
<tr>
<td>threadedPort, 259, 286</td>
<td></td>
</tr>
<tr>
<td>timestamp, 48</td>
<td></td>
</tr>
<tr>
<td>tracing, 249</td>
<td></td>
</tr>
<tr>
<td>trigger, 283</td>
<td></td>
</tr>
<tr>
<td>true, 305</td>
<td></td>
</tr>
<tr>
<td>tuplesPerFile, 282</td>
<td></td>
</tr>
<tr>
<td>tuple, 150, 284, 449</td>
<td></td>
</tr>
<tr>
<td>uint32, 89</td>
<td></td>
</tr>
<tr>
<td>uint8, 48</td>
<td></td>
</tr>
<tr>
<td>updateMetricsData, 239</td>
<td></td>
</tr>
<tr>
<td>ustring, 149, 246</td>
<td></td>
</tr>
<tr>
<td>valgrind, 267</td>
<td></td>
</tr>
<tr>
<td>wait, 242</td>
<td></td>
</tr>
<tr>
<td>wakeupHost, 239</td>
<td></td>
</tr>
<tr>
<td>wc, 101</td>
<td></td>
</tr>
<tr>
<td>wrapper, 266</td>
<td></td>
</tr>
<tr>
<td>write, 242</td>
<td></td>
</tr>
</tbody>
</table>
Index

A/D, 354
Analog-to-Digital, xxii
AAS, 220, 225, 227, 228, 232, 233, 236, 251–256
Authentication and Authorization Service, xxii
ACL, 227
audit logging, see audit logging
authentication, 253
LDAP, 253
PAM, 253
user, 253
authorization, 251
principal, 251
private security template, 253
session, 233, 251, 256
aberration, 427
absolute error, 364
Abstract Data Type, see ADT
access control, 251
Access Control List, see ACL
accuracy, 5, 12, 346, 395, 463, 472–474, 482
ACID, 494
Atomicity, Consistency, Isolation, and Durability, xxii
ACL, 204, 227, 251–253
Access Control List, xxii
action, 41
active database, 33, 40–42, 68
closed database, 41
open database, 41
Active MQ, 100
adapter operator, 92, 97
ADL, 118, 119, 219, 257, 258, 267
Application Description Language, xxii
, adl, 267
file, 257
Adobe Flash, 481
Adobe Flex, 13, 481
ADT, 46
Abstract Data Type, xxii
Advanced Encryption Standard, see AES
AES, 47
Advanced Encryption Standard, xxii
aggregate operator, 93
aggregation, 38, 49, 83, 109, 129, 288, 325
commutative operation, 319
function, 94, 154, 294, 296, 300
Average, 294
First, 294
Last, 294
Max, 294, 296
Min, 294, 296
Sum, 294, 296, 298
associative, 296
group, 94
operation, 131, 138
operator, 93
Agrawal, Imielinski, Swami, see AIS
AIS, 423
Agrawal, Imielinski, Swami, xxii
algorithm design, 21
algorithmic trading, 63
Alon, Matias, and Szegedy, see AMS
Amdahl’s law, 316
American National Standards Institute, see ANSI
American Standard Code for Information Interchange, see ASCII
AMS, 359, 360
Alon, Matias, and Szegedy, xxii
Analog-to-Digital, see A/D
analytical
application, 52
task, 50, 51, 204, 205
continuous, 50
exploration, see exploration task
single pass, see also single pass algorithm
analytics, 36, 37, 39, 107, 442, 467, 488, 495
continuous, 68
data
at rest, see data at rest
in motion, see data in motion
detection, see detection
exploration, see exploration
multi-modal, 34, 38, 292
multi-variate statistics, 38, 39
proactive, 79
reactive, 78
spatio-temporal, 37
stateful, 51
stream, 39
text, 324
streaming, 291
time-sensitive, 51
transformation, see transformation
unsupervised, 51
anomaly detection, 36, 373, 383, 388, 427, 433
classification-based, 429
clustering-based, 430
collective anomaly, see collective anomaly
contextual anomaly, see contextual anomaly
information-theoretic, 431
NN-based, 430
point anomaly, see point anomaly
spectral, 431
statistical, 430
ANSI, 61
American National Standards Institute, xxii
Apache, 65, 495
Apache Flex, see Adobe Flex
Apache Foundation, 100
Apache Maven, 490
Apache Thrift, 278
API, 14, 65, 154, 155, 338
Application Programming Interface, xxii
application, 56, 69, 116, 180, 204, 205, 207, 310
ADL file, see ADL file
analyst, 204
asset, 258
bottleneck, 317
checkpointing, 335, 336
incremental, 336
component, 205
composition, 275, 310
nested, 312
debugging, 230
decomposition, 21
degradation
graceful, 333
design, 310
developer, 171, 204
development, 82, 256
environment, 53, 80
life cycle, 256
entry point, 116, 122
fault tolerance, 247, 275
flow graph, 89, 111
health
monitoring, 239
high availability, 334
host, 220, 222, 226
instance, 113, 118
shutdown, 138
instantiation, 117
large scale, 310
latency
sensitive, 334
latency-sensitive, 170
life cycle, 207, 208
loss
intolerant, 334
tolerant, 334
maintenance, 114
model
instance, 257
monitoring, 230
namespace, see namespace
optimization, 275
parallelization, 275
parameter
submission-time, 313, 447
partition, 164, 166–169, 171, 179, 184, 185, 189, 195, 196, 198, 210, 239, 311, 321, 322, 324, 337
placement, 210
profiling, 199, 317, see also profiler, 320, 321, 325, 327–329
provisioning, 256, 258
recovery workflow, 236, 239
replica
active, 337
backup, 337
master, 337
replication, 335
restart
cold, 335
runtime, 87, 182, 184, 194, 195, 198, 199, 220, 247
instance, 249
scale up, 321
scope, 119, 232
default, 232
self-evolving, 303
solution, 310
speed up, 321
standalone, 92, 258
state
stationary, 333
transient, 333
throughput-sensitive, 171
topology, 106
Index
Index

composite operator, 85, 107, 116, 119, 122–126, 148, 150, 151, 162, 164, 166, 179, 312, 491
definition, 122
high-order, 127, 313
main, 122, 125, 232
reuse, 312
terminal, 310
use, 122
type, 149, 150
composition language, 87
compression, 100
computational environment, 220
node, 56
concept drift detection, 495
Concept-adapting Very Fast Decision Tree learner, see CVFDT
condition, 41
confusion matrix, 395, 463
precision, see precision recall, see recall
congestion policy, 286
conjugate gradient, 19
consistency, see ACID
contaminant, 427
content filter, 120, 449, 451
textual anomaly, 429
Continuous Query, see CQ
continuous query, 33, 40, 58, 61, 84, 205
Continuous Query Language, see CQL
control port, 302, 303
convolution, 16
CORBA, 15, 228, 491
Common Object Request Broker Architecture, xxii
Cornell University, 44
correlation, 16, 18, 37, 81, 347, 371, 469
alignment, 293
barrier, 292
coefficient, 372
join, 293
merge, 292
union, 292
covariance, 406
matrix, 406
CPU, 56, 185, 195, 198, 199, 209, 240, 314, 315, 317, 320, 329, 446
Central Processing Unit, xxii
cycle, 56
instruction cache
locality, 328
utilization, 195, 198, 320
CQ, 33, 40, 42
Continuous Query, xxii
CQL, 54, 55, 57, 58, 61, 62, 84
Continuous Query Language, xxii
critical section, 129
cross covariance, 347
cross validation, 12, 463, 464
CSV, 99, 278
Comma-Separated Value, xxii
custom function, 152
CVFDT, 403
Concept-adapting Very Fast Decision Tree learner, xxii
cybersecurity, 34, 35, 68, 302
botnet, see botnet
malware, see malware
social engineering, see social engineering vulnerability, 34
zero-day, 35
DAG, 59
Directed Acyclic Graph, xxii
dashboard, 10
data acquisition, 495
alignment, 5
analysis, 496
at rest, 84
cleaning, 5, 51
compression, 369
quantization, see quantization
correlation, 287
cube, 10
dicing, 10
slicing, 10
de-duplication, 51
declustering, 495
decorrelation, 375
gress, 87, 107
enrichment, 287
cache-based, 290
development, 132, 198
exploration, 5, 496
filtering, 51
fitting, 7, 416
flow graph, see flow graph
programming, 106
subgraph, 126
format, 275
conversion, 51
gathering, 20
healthcare, 38
in motion, 51, 84
ingest, 87, 107
management, 3, 7, 17, 53, 487
mining, xv, 3, 17, 49, 98, 109, 290, 292, 370, 375, 487, 493, 495
change detection, see change detection
classification, see classification

© in this web service Cambridge University Press
Index

- clustering, see clustering
- dependency modeling, see dependency modeling
- evaluation, see also model scoring
- model, 51, 115, 389
- modeling, 389
- online learning, 394
- PMML, see PMML
- regression, see regression summarization, see summarization toolkit, 291
- model, 46
- normalization, 5
- parallelism, 49, 80, 124, 312, 314, 315, 317, 320, 328, 493
- partitioned-stateful, 318, 323
- shuffling, 318
- stateful, 319
- stateless, 317
- tuple ordering, 319
- partition
  - attribute, 289, 318
  - partitioning, 287, 444, 445, 495
  - persistence, 130
- physiological, 38
- pre-processing, 51, 495
- processing, 87, 107
- protocol, 275
- reduction, see reduction, 354, 375
- schema, see schema
- semi-structured, see semi-structured data
  - sink, 45, 99
  - source, 45, 99
- external, 275
- streaming, 36
- standardization, 51
- stream, see stream
- structured, see structured data
  - synopsis, 82
- transform, see transform transformation, 495
- transport, see transport layer
- tuple, see tuple
- unstructured, see unstructured data
- warehouse, 9, 45, 378
- warehousing, 7, 10, 375
- Data Base Management System, see DBMS
- Data Definition Language, see DDL
- Data Manipulation Language, see DML
- data mining, 342
- Data Mining Group, see DMG
- database
  - active, see active database
  - append-only, 42
  - column-oriented, see column-oriented database design, 85
- embedded, 167
- in-memory, see in-memory database
- read-only, see read-only database
- table, 41, 126, 138, 279
- DataCaptor, 456
- DataCutter, 218
- db toolkit, 100
- DBMS, 8, 10, 11, 34, 41, 100, 278, 279
- Data Base Management System, xxii
  - type system, 278
- DBSCAN, 407
- Density-Based Spatial Clustering of Applications with Noise, xxii
- DCOM, 15
- Distributed Component Object Model, xxii
- DCT, 371, 372, 378, 379, 459
- Discrete Cosine Transform, xxii
- DDL, 9, 53, 63
- Data Definition Language, xxiii
- DDoS, 35
- Distributed Denial of Service, xxiii
- de-blur, 16
- de-noise, 16, 37
- debugger, 188, 214
- instrumentation point, 214
- native, 194
- stream processing, 194
- debugging, 62, 63, 87, 92, 178, 203, 213, 264, 487, 491, 492, 496
- breakpoint, see breakpoint definition, 192
- deployment, 189
- gdb, 267
- impact analysis, 193
- lineage, see lineage debugging backward, 193
- forward, 193
- mode, 264
- performance, 189
- sbd, 266
- semantic, 189
- tool, 188
- trace point, 192, see trace point user-defined operator, 189
- valgrind, 267
- Decision Support System, see DSS
decision tree, 394, 397, 398, 400, 402, 416, 429, 461
declarative interface, 81, 128
- Deep Packet Inspection, see DPI
delta threshold, 142
dendrogram, 407
- Denial of Service, see DoS
density similarity, 347
- Density-Based Spatial Clustering of Applications with Noise, see DBSCAN
dependency
Index

check, 496
management, 82, 149
modeling, 17
tracking, 490
deployment debugging, 194
descriptive statistics, 7, 16, 18, 288, 292, 345
design
pattern, 275
detection, 51
task, 44, 51, 61, 63, see also rule
determinism, 58, 490
device driver, 87
DFT, 371, 372, 375, 378
Discrete Fourier Transform, xxiii
DHT, 494
Distributed Hash Table, xxiii
dimension, see also attribute
dimensionality reduction, 370, 375, 379
Direct Drain Quiescent Current, see IDDQ
Directed Acyclic Graph, see DAG
directory service, 15
Discrete Cosine Transform, see DCT
Discrete Fourier Transform, see DFT
distance function, 407
Euclidean distance, see Euclidean distance
Mahalanobis distance, see Mahalanobis distance
Manhattan distance, see Manhattan distance
distributed
computing, 203, 487, 488
processing, 38, 58, 61
programming, 164
system, 3, 7
Distributed Component Object Model, see DCOM
Distributed Denial of Service, see DDoS
Distributed Hash Table, see DHT
divide and conquer, 20
DMG, 17, 390
Data Mining Group, xxiii
DML, 9, 53, 63
Data Manipulation Language, xxii
DNS, 35
Domain Name System, xxiii
Document Object Model, see DOM
Document Type Definition, see DTD
DOM, 283
Document Object Model, xxiii
domain, 141, 288, 375
Domain Name System, see DNS
Domain-Specific Language, see DSL
DoS
Denial of Service, xxiii
DPI, 146, 147, 175, 176
Deep Packet Inspection, xxiii
DSL, 22, 86
Domain-Specific Language, xxiii
DSO, 159, 227, 257, 258
Dynamically Shared Object, xxiii
DSS, 5, 10
Decision Support System, xxiii
DTD, 283
Document Type Definition, xxiii
durability, 11, see ACID
dynamic
adaptation, 275, 496
composition, 81, 106, 107, 112, 120, 232, 234,
302, 303, 306, 313, 314
content filter, see content filter
connection, 113–115, 119, 150, 182, 211, 227,
232, 258, 301
matching, 115, 232, 234
name-based, 119
flow
composition, 144
programming, 20
Dynamically Shared Object, see DSO
e-commerce
fraud detection, see fraud detection
ECA, 41
Event-Condition-Action, xxiii
ECG, 371, 454–458, 466, 482
Electrocardiogram, xxiii
Eclat algorithm, 423
Eclipse, 64, 230
École Polytechnique Fédérale de Lausanne, see
EPFL
eptic heartbeat, 455
detection, 456, 460, 482
ground truth, 458
EDA, 85
Electronic Design Automation, xxiii
ege
adaptation, 47, 49, 52, 57, 60, 78, 275, 378
adapter, 48–50, 275–277, 280, 283, 285, 286
back pressure, see back pressure
built-in, 276
communication protocol, 278
formatting, 277
parsing, 277
source, 280
user-defined, 276
EEG, 454
Electroencephalogram, xxiii
EKG, 371
Elektrokardiogramm, xxiii
electrocardiogram, see ECG
electroencephalogram, see EEG
Electronic Design Automation, see EDA
Elektrokardiogramm, see EKG
EM, 406, 430
Expectation Maximization, xxiii
e-mail, 41
Emergency Medical Services, see EMS
EMS
   Emergency Medical Services, xxiii
encapsulation, 122, 166
encryption, 100
dimensionless, 245
ensemble
   learning, 495
   model, 476
ensemble method, 397
Entity Relationship, see ER
time-based, 347, 398, 431, 461
rate, 347
enumeration, 151
EPL
   École Polytechnique Fédérale de Lausanne, xxiii
   Event Processing Language, xxiii
EPN, 61, 62
   Event Processing Network, xxiii
equi-join, 95
ER, 9, 69, 70
   Entity Relationship, xxiii
   model, 9
error
   absolute, see absolute error
   reporting, 248, 251
   squared, see squared error
ESP, 60
   Event Stream Processor, xxiii
Esper, 44
EsperTech, 44
   Ethernet, 14, 171, 244
ETL, 44, 378
   Extract/Transform/Load, xxiii
Euclidean distance, 405, 407, 410
event, 41, 294, 300
   composite, 41
   correlation, 44
   primitive, 41
Event Processing Language, see EPL
Event Processing Network, see EPN
Event Stream Processor, see ESP
Event-Condition-Action, see ECA
EventFlow, 54, 63
eviction
event, 143
   policy, 95, 138, 141–143, 155
   count-based, 138–142
delta-based, 142, 298
sliding window, 139
time-based, 143, 259
evolution strategy, 20
evolutionary
   programming, 20
search, 19
exception, 427
execution model, 80
Expectation Maximization, see EM
expiration policy, 141
exploration, 51
analysis, 36
task, 44, 51, 79, 107, 112–114
exported stream, 116, 120, 232, 445, 463
expression language, 81
extension, 54, 61, 63, 67, 79, 86, 87, 148, 487
eXtensible HyperText Markup Language, see XHTML
eXtensible Markup Language, see XML
ExTensible Stylesheet Language Transformations, see XSLT
Extract/Transform/Load, see ETL
face detection, 291
algorithm, 317
false, 133
Fast Fourier Transform, see FFT
Fault Detection and Classification, see FDC
fault tolerance, 6, 11, 51, 56, 58, 61, 64, 65, 68, 83, 129, 131, 169, 203, 206, 212, 219, 220, 229, 247, 275, 310, 314, 333, 343, 447, 453, 482
application, 212, 247
hot standby replica, 62, 64
partial, 6
request
   at most once, 247
   idempotent, 247
   non-idempotent, 247
FDC, 468
Fault Detection and Classification, xxiii
feature, see also attribute
   extraction, 16, 370, 371, 373, 458, 469, 482
   ranking, 370, 375
   selection, 370
   space, 470
   subset
   selection, 370
   vector, 409, 473, 476
feedback loop, 53, 79, 112, 301, 302, 327
FFT, 459
Fast Fourier Transform, xxiii
fidelity, 363
Field-Programmable Gate Array, see FPGA
file, 138
   adapter, 97
descriptor, 167
hot, 306
system, 45
distributed, 495
File Transfer Protocol, see FTP
filter, 132
condition, 305
high pass, see high pass filter
low pass, see low pass filter
Index

<table>
<thead>
<tr>
<th>Operator</th>
<th>93, 160</th>
</tr>
</thead>
<tbody>
<tr>
<td>Particle</td>
<td>357</td>
</tr>
<tr>
<td>Predicate</td>
<td>125, 133</td>
</tr>
<tr>
<td>Stateful</td>
<td>287</td>
</tr>
<tr>
<td>Stateless</td>
<td>287</td>
</tr>
<tr>
<td>Filtering</td>
<td>16, 39, 109, 127, 132, 287, 288, 325, 354</td>
</tr>
<tr>
<td>Finance engineering</td>
<td>318</td>
</tr>
<tr>
<td>Financial market</td>
<td>33</td>
</tr>
<tr>
<td>Firewall</td>
<td>See also Chinese wall</td>
</tr>
<tr>
<td>Fisher’s Linear Discriminant</td>
<td>See also LDA</td>
</tr>
<tr>
<td>Floating point number</td>
<td>288</td>
</tr>
<tr>
<td>Flow</td>
<td>288</td>
</tr>
<tr>
<td>Composition</td>
<td>84, 106</td>
</tr>
<tr>
<td>Dynamic, see dynamic composition</td>
<td></td>
</tr>
<tr>
<td>Nested, see nested composition</td>
<td></td>
</tr>
<tr>
<td>Static, see static composition</td>
<td></td>
</tr>
<tr>
<td>Fan-in</td>
<td>50, 52, 53</td>
</tr>
<tr>
<td>Fan-out</td>
<td>50</td>
</tr>
<tr>
<td>Feedback loop, see feedback loop</td>
<td></td>
</tr>
<tr>
<td>Logical</td>
<td>53, 164, 179, 194, 219</td>
</tr>
<tr>
<td>Physical</td>
<td>53, 164, 179, 194, 219</td>
</tr>
<tr>
<td>Manipulation</td>
<td>106, 275, 291</td>
</tr>
<tr>
<td>Splitting</td>
<td>287</td>
</tr>
<tr>
<td>Subgraph</td>
<td>122</td>
</tr>
<tr>
<td>Forecasting</td>
<td>37, 40, 373, 414, 469, 472</td>
</tr>
<tr>
<td>Format conversion</td>
<td>378</td>
</tr>
<tr>
<td>Fortran</td>
<td>14, 83</td>
</tr>
<tr>
<td>FP-growth algorithm</td>
<td>423</td>
</tr>
<tr>
<td>FPGAs</td>
<td>14</td>
</tr>
<tr>
<td>Field-Programmable Gate Array</td>
<td>xxiii</td>
</tr>
<tr>
<td>Fraud detection</td>
<td>34</td>
</tr>
<tr>
<td>Free Software Foundation, see FSF</td>
<td></td>
</tr>
<tr>
<td>Frequent pattern</td>
<td></td>
</tr>
<tr>
<td>Mining</td>
<td>383, 388, 420, 433</td>
</tr>
<tr>
<td>FSF</td>
<td>194, 265</td>
</tr>
<tr>
<td>Free Software Foundation, xxiii</td>
<td></td>
</tr>
<tr>
<td>FTP</td>
<td>100, 285</td>
</tr>
<tr>
<td>File Transfer Protocol, xxiii</td>
<td></td>
</tr>
<tr>
<td>Function</td>
<td>148, 151</td>
</tr>
<tr>
<td>Call</td>
<td>171</td>
</tr>
<tr>
<td>Custom</td>
<td>151, 152</td>
</tr>
<tr>
<td>Model</td>
<td>152</td>
</tr>
<tr>
<td>Native</td>
<td>152</td>
</tr>
<tr>
<td>Registration</td>
<td>152</td>
</tr>
<tr>
<td>Signature</td>
<td>152</td>
</tr>
<tr>
<td>Functional</td>
<td></td>
</tr>
<tr>
<td>Correctness</td>
<td>188</td>
</tr>
<tr>
<td>Design</td>
<td>275</td>
</tr>
<tr>
<td>Functor operator</td>
<td>90, 93</td>
</tr>
<tr>
<td>G.729</td>
<td>47</td>
</tr>
<tr>
<td>Garbage collection</td>
<td>62</td>
</tr>
<tr>
<td>Gauss–Newton</td>
<td>378</td>
</tr>
<tr>
<td>Gaussian distribution</td>
<td>406, 431</td>
</tr>
<tr>
<td>Gaussian Mixture Model, see GMM</td>
<td></td>
</tr>
<tr>
<td>Gbps</td>
<td></td>
</tr>
<tr>
<td>Gigabits per second</td>
<td>xxiii</td>
</tr>
<tr>
<td>gdb</td>
<td>194, 265–267</td>
</tr>
<tr>
<td>Gene expression</td>
<td>413</td>
</tr>
<tr>
<td>Genetic algorithms</td>
<td>20, 371</td>
</tr>
<tr>
<td>Programming</td>
<td>20</td>
</tr>
<tr>
<td>Sequence alignment</td>
<td>20</td>
</tr>
<tr>
<td>Georgia Tech</td>
<td>42</td>
</tr>
<tr>
<td>Gigabits per second, see Gbps</td>
<td></td>
</tr>
<tr>
<td>Gigascope</td>
<td>60</td>
</tr>
<tr>
<td>Gigascope SQL, see GSQL</td>
<td></td>
</tr>
<tr>
<td>Gini index</td>
<td>398</td>
</tr>
<tr>
<td>Global Positioning System, see GPS</td>
<td></td>
</tr>
<tr>
<td>Global Sensor Networks, see GSN</td>
<td></td>
</tr>
<tr>
<td>Global system for Mobile Communications, see GSM</td>
<td></td>
</tr>
<tr>
<td>Globus</td>
<td>209</td>
</tr>
<tr>
<td>GRAM</td>
<td>209</td>
</tr>
<tr>
<td>Globus Resource Allocation Manager, see GRAM</td>
<td></td>
</tr>
<tr>
<td>GMM</td>
<td>397, 406, 461</td>
</tr>
<tr>
<td>Gaussian Mixture Model, xxiii</td>
<td></td>
</tr>
<tr>
<td>Google</td>
<td>7</td>
</tr>
<tr>
<td>Google Protocol Buffers</td>
<td>278</td>
</tr>
<tr>
<td>GPFS</td>
<td>259</td>
</tr>
<tr>
<td>GPS</td>
<td>36, 37</td>
</tr>
<tr>
<td>Global Positioning System, xxiii</td>
<td></td>
</tr>
<tr>
<td>GPU</td>
<td>493</td>
</tr>
<tr>
<td>Graphics Processing Unit, xxiii</td>
<td></td>
</tr>
<tr>
<td>Gradient</td>
<td>418</td>
</tr>
<tr>
<td>Descent</td>
<td>378, 405, 417</td>
</tr>
<tr>
<td>Search</td>
<td>19</td>
</tr>
<tr>
<td>GRAM</td>
<td>209</td>
</tr>
<tr>
<td>Globus Resource Allocation Manager, xxiii</td>
<td></td>
</tr>
<tr>
<td>Graphical User Interface, see GUI</td>
<td></td>
</tr>
<tr>
<td>Graphics Processing Unit, see GPU</td>
<td>ground truth, 405, 458</td>
</tr>
<tr>
<td>Grubbs’s test</td>
<td>431</td>
</tr>
<tr>
<td>Gryphon</td>
<td>43</td>
</tr>
<tr>
<td>GSM</td>
<td>47</td>
</tr>
<tr>
<td>Global System for Mobile Communications, xxiii</td>
<td></td>
</tr>
<tr>
<td>GSN</td>
<td>60</td>
</tr>
<tr>
<td>Global Sensor Networks, xxiii</td>
<td></td>
</tr>
<tr>
<td>GSQL</td>
<td>60</td>
</tr>
<tr>
<td>Gigascope SQL, xxiii</td>
<td></td>
</tr>
<tr>
<td>GUI</td>
<td>53, 58, 62, 85, 146, 147, 186, 467, 476, 479, 481, 482</td>
</tr>
<tr>
<td>Graphical User Interface, xxiii</td>
<td>design and implementation, 85</td>
</tr>
<tr>
<td>Grzip</td>
<td>99</td>
</tr>
<tr>
<td>HA</td>
<td>62, 64, 212</td>
</tr>
<tr>
<td>High Availability, xxiii</td>
<td></td>
</tr>
<tr>
<td>Haar transform</td>
<td>359, 379–382</td>
</tr>
<tr>
<td>Error tree</td>
<td>381</td>
</tr>
<tr>
<td>Hadoop</td>
<td>12, 495</td>
</tr>
<tr>
<td>Index</td>
<td>513</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Hadoop Distributed File System, see HDFS hash</td>
<td>see</td>
</tr>
<tr>
<td>map, 150</td>
<td>HyperText Transfer Protocol Secure, xxiv</td>
</tr>
<tr>
<td>set, 150</td>
<td>HyperText Markup Language, see HTML</td>
</tr>
<tr>
<td>table, 289</td>
<td>HyperText Transfer Protocol, see HTTP</td>
</tr>
<tr>
<td>HC, 219, 220, 225–227, 232, 234–236, 238–240, 244, 258</td>
<td>HyperText Transfer Protocol Secure, see HTTPS</td>
</tr>
<tr>
<td>Host Controller, xxiii</td>
<td>hypothesis testing, 7</td>
</tr>
<tr>
<td>monitoring</td>
<td>hypothesis-based analysis, 5</td>
</tr>
<tr>
<td>cycle, 236</td>
<td>I/O, 4, 21, 56, 174, 175, 276, 327</td>
</tr>
<tr>
<td>HDFS, 495</td>
<td>Input/Output, xxiv</td>
</tr>
<tr>
<td>Hadoop Distributed File System, xxiii</td>
<td>bandwidth, 56</td>
</tr>
<tr>
<td>healthcare, 33, 34, 38, 39, 68</td>
<td>IBM, xv, xvi, xxxi, 3, 4, 23, 64, 86, 100, 101, 187, 259, 288, 318, 454, 467</td>
</tr>
<tr>
<td>heart, 38</td>
<td>IBM Cognos Now, 10</td>
</tr>
<tr>
<td>heartbeat mechanism, 208</td>
<td>IBM DB2 Server, 247</td>
</tr>
<tr>
<td>Hello World application, 87, 107, 118</td>
<td>IBM InfoSphere Data Warehouse, 10</td>
</tr>
<tr>
<td>Hertz, xxiv</td>
<td>IBM InfoSphere Streams, see Streams</td>
</tr>
<tr>
<td>Hidden Markov Model, see HMM</td>
<td>IBM MQ Websphere Low Latency Messaging, 244</td>
</tr>
<tr>
<td>hierarchical clustering, 407, 408</td>
<td>IBM Research, 43, 218</td>
</tr>
<tr>
<td>hierarchical model, 8</td>
<td>IBM SPSS Modeler, 10</td>
</tr>
<tr>
<td>High Availability, see HA</td>
<td>IBM T. J. Watson Research Center, 86</td>
</tr>
<tr>
<td>high pass filter, 458</td>
<td>IBM Websphere Business Events, 44</td>
</tr>
<tr>
<td>hill climbing greedy, 371</td>
<td>IBM WebSphere MQ, 356</td>
</tr>
<tr>
<td>HiPac, 41</td>
<td>IBM WebSphere MQ, 400</td>
</tr>
<tr>
<td>histogram, 346</td>
<td>IBM WebSphere MQ, 410</td>
</tr>
<tr>
<td>equi-depth, see equi-depth histogram</td>
<td>IBM WebSphere MQ, 500</td>
</tr>
<tr>
<td>v-optimal, see v-optimal histogram</td>
<td>ICU, 38, 39, 431, 454–456</td>
</tr>
<tr>
<td>HMM, 397</td>
<td>Intensive Care Unit, xxiv</td>
</tr>
<tr>
<td>Hidden Markov Model, xxiv</td>
<td>IDDO, 470, 472, 473, 476</td>
</tr>
<tr>
<td>Host Controller, see HC</td>
<td>Direct Drain Quiescent Current, xxiv</td>
</tr>
<tr>
<td>build, 221</td>
<td>IDE, 55, 64, 85, 190, 191, 219</td>
</tr>
<tr>
<td>execution, 221</td>
<td>Integrated Development Environment, xxiv</td>
</tr>
<tr>
<td>failure, 236</td>
<td>IDL, 15</td>
</tr>
<tr>
<td>isolation</td>
<td>Interface Definition Language, xxiv</td>
</tr>
<tr>
<td>hostIsolation, 337</td>
<td>type system, 278</td>
</tr>
<tr>
<td>placement, 166, 167, 239</td>
<td>IDSS, 288</td>
</tr>
<tr>
<td>co-location, 168, 240</td>
<td>Intrusion Detection System, xxiv</td>
</tr>
<tr>
<td>constraint, 169, 170</td>
<td>IEEE, 472</td>
</tr>
<tr>
<td>ex-location, 168, 240</td>
<td>Institute of Electrical and Electronics Engineers, xxiv</td>
</tr>
<tr>
<td>isolation, 169, 240</td>
<td>image</td>
</tr>
<tr>
<td>pool, 169, 170, 179</td>
<td>analysis, 97</td>
</tr>
<tr>
<td>exclusive, 170</td>
<td>compression, 357</td>
</tr>
<tr>
<td>fixed-size, 170, 172</td>
<td>processing</td>
</tr>
<tr>
<td>shared, 170</td>
<td>toolkit, 150</td>
</tr>
<tr>
<td>protection, 241</td>
<td>in-memory database, 11, 13</td>
</tr>
<tr>
<td>tag, 169, 170</td>
<td>incremental</td>
</tr>
<tr>
<td>Host Controller, see HC</td>
<td>deployment, 114, 119</td>
</tr>
<tr>
<td>Hotelling $\chi^2$ test, 431</td>
<td>learning, 396, 398</td>
</tr>
<tr>
<td>HTML, 46</td>
<td>scoring, 398</td>
</tr>
<tr>
<td>HyperText Markup Language, xxiv</td>
<td>indexing, 495</td>
</tr>
<tr>
<td>HTTP, 62, 87, 100, 102, 285</td>
<td>inet toolkit, 100, 285</td>
</tr>
<tr>
<td>HyperText Transfer Protocol, xxiv</td>
<td>inference test, 430</td>
</tr>
<tr>
<td>HTTPS, 285</td>
<td>inferential statistics, 17</td>
</tr>
<tr>
<td>Hotelling</td>
<td>InfiniBand, 171, 172, 211, 244</td>
</tr>
<tr>
<td>Host Controller, see HC</td>
<td>information</td>
</tr>
<tr>
<td>Hotelling $\chi^2$ test, 431</td>
<td>bottleneck, 347</td>
</tr>
<tr>
<td>flow</td>
<td>processing, 40, 179</td>
</tr>
<tr>
<td>processing system, 195</td>
<td></td>
</tr>
</tbody>
</table>
Index

gain, 347, 398–402
theory, 347, 357, 398
entropy, see entropy
information bottleneck, see information bottleneck
information gain, see information gain
Kolmogorov complexity, see Kolmogorov complexity
Information Technology, see IT
inheritance, 11
input, see Input/Output
inspector-executor, 210
instance, 218, 220
backbone, 228
bootstrapping, 230
configuration, 230
creation, 230
Instance Graph, 267
Institute of Electrical and Electronics Engineers, see IEEE
integer, 149
Integrated Development Environment, see IDE
integration, 488
Intensive Care Unit, see ICU
inter-arrival time, 281
inter-component communication protocol, 229
Inter-Process Communication, see IPC
Interface Definition Language, see IDL
Internet, 34, 100, 285
Internet Protocol, see IP
Interoperable Object Reference, see IOR
interpretation, 205
introspection, 159, 203
Intrusion Detection System, see IDS
IOR, 228
Interoperable Object Reference, xxiv
IP, 102, 141, 146, 147, 196, 302, 428
Internet Protocol, xxiv
address, 141, 243, 277, 288, 302, 305, 306
traffic, 302
IPC, 171
Inter-Process Communication, xxiv
ischemia, 454, 455
ISO/IEC, 9
isolation, see ACID
IT, 38, 63, 68, 178, 442–444, 482, 491
Information Technology, xxiv
itemset, 420–427
Java, xv, 46, 48, 62, 63, 65–67, 83, 148, 149, 151, 154, 193, 312, 461, 462, 489
class
method, 46
Java DataBase Connectivity, xxiv
Java Message Service, see JMS
Java Native Interface, see JNI
Java Virtual Machine, see JVM
JavaScript Object Notation, see JSON
JDBC, 62
JFreeChart, 13
jitter, 52
JMS, 61–63
Java Message Service, xxiv
JNI, 489
Java Native Interface, xxiv
job, 113, 118, 180, 204–207, 219, 224, 446
cancelation, 118, 232, 236
workflow, 227
identifier, 116, 118
management, 207, 219, 224, 230, 232, 235, 236, 338
component, 207
workflow, 232
manager, 212
state, 233
canceling, 233
failed, 233, 234
running, 224, 233, 234
submission, 117, 118, 232
workflow, 226, 227, 243, 267
join, 293, 359
algorithm, 95, 96
match
condition, 293
operator, 95
algorithm, 95
equi-join, 95
matching condition, 95, 131
outer join, 95
sliding window, 95
window, 293
joint optimization, 21
JPEG2000, 379
JSON, 65, 79
JavaScript Object Notation, xxiv
JVM, 62, 65
Java Virtual Machine, xxiv
k-means algorithm, 365, 405, 408, 410
k-median algorithm, 406
k-medoid algorithm, 406
k-Nearest Neighbors, see KNN
Kalman filter, 353
Karhunen–Loève Transform, see KLT
kbps
kilobits per second, xxiv
Kernel
kernel method, 371
Kernel Smoothing method, 353
KHz
kilohertz, xxiv
kilobits per second, xxiv
kilohertz, see KHz
Kleene
closure, 44
star operator, see also pattern matching
KLT, see also PCA, 375
Karhunen–Loève Transform, xxiv
KNN, 397, 430, 474–476, 478, 480
k-Nearest Neighbors, xxiv
Kolmogorov complexity, 347, 431
Kullback–Leibler divergence, 347
label, see category
LAN, 14
Local Area Network, xxiv
Latency, 198
latency, 37, 52, 59, 68, 171, 198, 211, 279, 283, 328–330
LBG, 366
Linde–Buzo–Gray, xxiv
LDA, 397
Linear Discriminant Analysis, xxiv
LDAP, 253
Lightweight Directory Access Protocol, xxiv
Least Frequently Updated Partition, see LFUP
Least Recently Updated Partition, see LRUP
least squares
fit, 18
regression, 415
legacy system, 97, 489
Levenberg–Marquardt, 378
LF, 473–476, 478, 480
Line Fit, xxiv
LFUP, 144
Least Frequently Updated Partition, xxiv
library, 148
Lightweight Directory Access Protocol, see LDAP
likelihood
function, 461
score, 406
Linde–Buzo–Gray, see LBG
Line Fit, see LF
lineage debugging, 193
linear
interpolation, 355
programming, 7, 19
regression, 415–417, 474
sketches, 359
transform, 361, 371, 376, 379
Linear Discriminant Analysis, see LDA
Linear Predictive Coding, see LPC
Linear Road, 60
Linux, 153, 220, 222
list, 150
liveness service, 237
LLM, 172, 244
Low Latency Messaging, xxiv
Lloyd’s algorithm, 405
load
balancing, 56, 78, 166, 194, 209, 240, 329, 444, 446
shredding, 58, 59, 78, 281, 329, 358, 445, 466
Local Area Network, see LAN
location-based
service, 36, 37
logging, 62, 63, 203, 213, 248, 251
LogFileMaxFiles, 249
LogFileMaxSize, 249
data management, 213
data retrieval, 213
log curation, 213
logic
temporal, 44
logistic regression, 397
lookup table, 98
loop sensor, 38
traffic volumetrics, 36
loss function, 416, 417
lossy counting algorithm, 424–426
Low Latency Messaging, see LLM
low pass filter, 458
LPC, 371
Linear Predictive Coding, xxiv
LRUP, 144
Least Recently Updated Partition, xxiv
machine
learning, 49, 370, 375
black box, 371
incremental, 51
machine learning, 342
Mahalanobis distance, 407
main, 336
main composite, 249
maintainability, 82
maintenance, 496
malware, 34–36
management, 222
management host, 220, 222
Manhattan distance, 407
Manifold Learning Algorithm, see MLA
manufacturing, 33
process control, 34
map, 132
MapReduce, 11, 15, 64
market
making, 34
Massachusetts Institute of Technology, see MIT
matching condition, 95, 131
Mathematica, 489
MATLAB, 389
mean, 17, 288, 347, 359, 473
sample, 346, 367
squared error, 355
Mean Time to Detect Failures, see MTDF
median, 17, 288, 346, 347, 359
mediation, see also edge adaptation
application, 442
Medusa, 59
megapixel, see MP
memory
usage, 195
merge, 49, 292, 293
message
queue, 45
Message Passing Interface, see MPI
message passing, 61
interface, 62, 63
JMS, see JMS
Microsoft, 15
MicroStrategy, 10
middleware, 53, 203
service, 203, 204
millisecond, see ms
millivolts, see mV
MIMD, 13
Multiple Instruction Multiple Data, xxiv
mining toolkit, 391, 392, 424
MIT, 54, 58, 456, 457, 459, 460, 463
Massachusetts Institute of Technology, xxiv
Physiobank, 456, 463
Mitrion-C, 14
mixed-mode, 451
MLA, 371
Manifold Learning Algorithm, xxiv
MLE, 461
Maximum-Likelihood Estimation, xxiv
mode, 346, 347
model
building, 20, 495
evaluation, 394, 404, 414, 428
formulation, 20
learning, 395, 404, 414
online, 301
robustness, 476
scoring, 389, 394, 404, 414, 428
training, 475
validation, 475
modeling, 495
modularity, 81, 107, 148, 316, 320, 487
mechanism, 82
modularization, 54, 122
module, 122
moment, 345–347, 359
first order, 409
second order, 409
moment algorithm, 424
Moment Preserving Quantization, see MPQ
monitoring, 5, 63, 194, 203, 209, 210, 236
natural resources, 34
task, 54
wildlife, 34
mote, 45
moving average, 82
MP, 42
megapixel, xxiv
MPI, 14, 491
Message Passing Interface, xxiv
MPQ, 365–369
Moment Preserving Quantization, xxiv
ms millisecond, xxiv
MTDF, 469, 471
Mean Time to Detect Failures, xxiv
Multi-Sequential LEast Squares, see MUSCLES
multicast, 244
Multiple Instruction Multiple Data, see MIMD
MUSCLES, 375
MUSCLES, xxv
mutual authentication, 214
mV millivolt, xxv
Naïve Bayes, see NB
Name Service, 59, 228
namespace, 81, 88, 150, 152
com.ibm.streams.cep, 300
com.ibm.streams.db, 100
com.ibm.streams.inet, 100, 285
com.ibm.streams.messaging, 100
management, 149
National Oceanic and Atmospheric Administration,
see NOAA
native function, 152
registration, 152
NB, 397, 461, 462, 464, 465, 467
Naïve Bayes, xxv
Nearest Neighbors, see NN
Neonatal Intensive Care Unit, see NICU
nested composition, 107, 122, 180
network
bandwidth, 56
model, 8
packet, 47, 141
packets, 354
protocol, 171, 288

© in this web service Cambridge University Press
Index

Online Transaction Processing, xxv
OM, 441, 444, 445, 451, 454, 482
Operations Monitoring, xxv
OMG, 15
Object Management Group, xxv
online learning, 441, 463, 464, 477, 495
scoring, 389, 390, 393, 463, 495
Online Analytical Processing, see OLAP
Online Transaction Processing, see OLTP
OO
Object-Oriented, xxv
OP, 144
Oldest Partition, xxv
Open Data Base Connectivity, xxv
OpenCQ, 42
OpenMP, 14
Operating System, see OS
operational monitoring, 281
Operations Monitoring, see OM
config clause, 168, 187
placement directive, 168
graph clause, 89, 122
onTuple clause, 163
output clause, 89, 90
param clause, 89
adapter, 92, 97
aggregation, see aggregation operator
analytic, 109
arity, 132
associations, 391
barrier, 97
checkpoint, 248
checkpointing, 247, 248
creating, 259
operator driven, 248, 259
periodic, 248
restoring, 259
classification, 391
classifying, 391
code generation
template, 159, 160, 162
code generator, 159
composite, see composite operator
computational cost, 326
per tuple, 326
configuration, 282
custom, 162
declarative, 87
development API
checkpointing, 155
code generation, 159
code generation, 159
code generation, 159
compilation-time introspection, 159
dynamic connection, 156
error handling, 155

...
Index

exception handling, 155
classification, 155
output attribute handling, 155
parameter handling, 155
runtime introspection, 159
thread startup, 155
thread management, 155
windowing management, 155
execution context, 261, 262
export, 115
file adapter, see file adapter
filter, 93, 160
fusion, 166, 248, 311, 327, 460
generic, 159
import, 115
input port, 91
interface, 263
threaded, 259
windowing policy, 259
instance, 89, 91, 111, 168, 170, 195, 198, 199, 289
name, 89
startup workflow, 244
internal flow graph, 219
invocation, 89, 108
join, see join operator
load shedding, 281
log level, 251
many-to-many, 132
many-to-one, 132
metric, 239
model, 128, 160
one-to-at-most-one, 132
one-to-many, 132
one-to-one, 132
output assignment, 134
output function, 134
plain, 134
output port, 91
interface, 264
parameter, 89, 132
bounded cardinality, 133
expression mode, 133
mandatory, 132
optional, 132
type, 132
partitioned stateful, 128, 130, 131
performance metric, 232
placement host, 239
partition, 166, 239
port, 91
primitive, see primitive operator
projection, see projection operator
pseudo operator, 115
push-based protocol, 276
regression, 391
relocation, 239, 240, 247
restart, 235, 239, 247
restartable, 248
restarting, 240
reusable, 81
runtime, 155
sampling, 353
selection, see selection operator
selectivity, 132, 326
fixed, 131
variable, 131
semantics, 136
single input/single output, 132
sink, see sink operator
sort, see sort operator
strict weak ordering, 133
source, see source operator
split, see split operator
state, 144
stateful, 93–95, 97, 128, 129, 289, 317
critical section, 129
stateless, 93, 96, 128, 129
system callback, 262
TCP adapter, 99
sink, 99, 277
termination, 236
toolkit, 148, 150, 151
function, 151
image processing, see image processing toolkit
registration, 160
type, 111
user-defined, 118
debugging, 194
utility, 92
windowed, 93–95
optimization, 417
constrained, 19
maximization, 19
minimization, 19
model, 21
non-linear programming, 19
objective function, 19, 417
optimal solution, 20
quadratic programming, 19
type, xv, 3, 18, 49, 487
unconstrained, 19
option, 3
Oracle, 61
Oracle CEP, 44, 62, 66

© in this web service Cambridge University Press www.cambridge.org
Oracle Data Mining, 10
Oracle Hyperion, 10
Oracle CEP, 60
Operating System, xxv
process, 166, 205, 206, 321, 322
scheduling unit, 321
shell, 258
system call, see system call
outer join, 95
outlier, 427
detection, 7
stream, see stream
property, 116
oxygen saturation, 38
package, 148
packet, 245
Padres, 43, 55
page rank, 12
PAM, 253
Plugabble Authentication Module, xxv
parallel
I/O, 495
processing, 487
Parallel Haskell, 14
parallel processing, 291
Parallel Virtual Machine, see PVM
parallelization, 129, 310, 314, see task parallelism
Amdahl’s law, see Amdahl’s law
automatic, 493
compute-bound task, 320
data parallelism, see data parallelism
hyperthreading, 314
I/O-bound task, 320
pipeline parallelism, see pipeline parallelism
speed up, 316
synchronization point, 324
barrier, 316
thread, see thread
parsing, 37, 109, 445
partition, 130, 143, 155, 227
attribute, 298
colocation, 312
expiration policy, 143
partition-count-based, 143, 144
partition-time-based, 143
sliding, 143
time-based, 144
tumbling, 143
tuple-based, 143, 144
tuple-count-based, 143, 144
expiration selection policy, 144
LFUP, 144
LRUP, 144
OP, 144
placement
constraint, 169
partitioning, 49, 94, 445, 446
Patient Data Management System, see PDMS
Patient Monitoring, see PM
pattern, 44, 51, 54, 61, 389, 420
detection, 5, 51, 129, 373
extraction, 16
matching, 61, 299
Kleene star operator, 294
language, 63
operation, 294
predicate, 294
recognition, 16
temporal, see temporal pattern
PCA, 369, 371, 372, 374, 379, 431, 472, 473, 475
Principal Component Analysis, xxv
principal axis, 473
components, 374
directions, 374
PDF, 346, 406
Probability Density Function, xxv
PDMS, 454
Patient Data Management System, xxv
Processing Element, xxv
augmented model, 244
binary, 257
cancellation, 236
checkpoint directory, 259
container, see PEC
DSO, see DSO
evaluation context, 259
input port, 259
launch count, 248
dlifecycle, 233
performance metric, 232
relocatable, 248
restart, 235
threshold, 235
restartable, 235, 236, 248
shutdown workflow, 260
standard error, 250
standard out, 250
starting, 234
Index

provisioning, 490, 492, 496
PSRO, 470, 472, 473, 476
Performance Sort Ring Oscillator, xxv
pub–sub, see publish–subscribe
Public Infrastructure for Processing and Exploring
Streams, see PIPES
public key, 253
publish–subscribe, 62
system, 33, 40, 42, 45
broker, 42
content-based, 43
publication, 42
publisher, 42
subscription, 42
topic-based, 43
pull
method, 208
model, 42, 208, 214, 276
punctuation, 136, 143, 155, 160, 191, 266, 356, 458
debugging, 266
expecting, 136
final, 136, 138
mode, 144
window, 136, 137, 161
free, 137
push
method, 208
model, 42, 208, 214, 276
PVM, 14
Parallel Virtual Machine, xxv
Python, xv, 489
QoS, 56, 59, 453
Quality of Service, xxv
QRS complex, 455
detection algorithm, 458
quadratic programming, 7
Quality of Service, see QoS
quantile, 347
quantization, 288, 325, 363, 364, 375, 405
binary clipping, 366
compressed sensing, 369
moment preservation, 367
MPQ, see MPQ
quantizer function, see quantizer function
rounding, see rounding
scalar, see scalar quantization
shape-preserving, 366
vector, see vector quantization
quantizer function, 363, 365, 366
binary, 366
quantization error, 363
query, 42
continuous, see continuous query
language, 205
plan, 57, 58, 205, 206
planning, 495
processing
approximate, 357, 358
relational, 42
scheduling, 495
snapshot, 42
queueing system, 281
R, 389, 489, 496
RAD, 64, 85
Rapid Application Development, xxv
Radial Basis Function, see RBF
radio astronomy, 34
radio telescope, 45
Radio-Frequency IDentification, see RFID
random
number, 356
sampling, 17, 354, 357
uniformly distributed, 356
variable, 38, 398, 400
entropy, 398
Random Decision Forest, see RDF
Rapid Application Development, see RAD
RBF
Radial Basis Function, xxv
RFID
Random Decision Forest, xxv
RDMA, 171
Remote Direct Memory Access, xxv
read-only database, 11, 13
Real-time Transport Protocol, see RTP
Really Simple Syndication, see RSS
recall, 395
Receiver Operating Characteristic, see ROC
reduction, 127, 132, 198, 287, 329, 375, 460, 473
aggregation, 288
filtering, 287
operation
max, 130
min, 130
projection, 288
quantization, see quantization
sampling, see sampling
redundancy elimination, 326
reflection, 67
regression, 7, 17, 36, 378, 383, 388, 390, 414, 417, 433
function, 414, 415
kernel regression, see kernel regression
least squares regression, see least squares
regression
linear, see linear regression
loss function, see loss function
model, 392, 415
non-parametric, 415
parametric, 415
NPMR, see NPMR
Index

relational database, xv, 494
Remote Direct Memory Access, see RDMA
Remote Procedure Call, see RPC
database, 7, 8, 45, 54, 79, 92, 95, 279, 289, 326
operator, 57, 63, 83, 128, 294
Table, 279
relation, 8, 9, 86
relation, 8, 9, 59, 84, 92
data management, 86
model, 8, 46
operation, 49, 132, 378
data-driven, 491
decay function, 357
resource allocation, 206
management, 56, 78, 203, 209, 236
workload, 236
manager, 208, 209
matching, 169
respiratory rate, 38
reusability, 79
reuse, 82, 316
RFID, 4
Radio-Frequency Identification, xxv
road congestion, 36
loop sensor, see loop sensor
network, 36
segment, 37
robot programming, 85
ROC, 396
Receiver Operating Characteristic, xxv
rounding, 365
RPC, 15, 491
Remote Procedure Call, xxv
RS232, 456
RSS, 100, 283
Really Simple Syndication, xxv
RTP, 47, 48, 50, 52
Real-time Transport Protocol, xxvi
packet, 52
rule, 7, 44, 51, 54, 378
ECA, 41
event, 63
logical predicate, 51
transformation, 54
runtime instance, see application runtime environment
instance, 230, 238, 240, 241
rebalancing, 241
runtime instance, 113, 118, 221–224, 226–228, 230, 236, 237
bootstrapping, 228
runtime performance, 166
S4, 64, 65
Scheduler, xxvi
Streams Application Manager, xxvi
matching process, 232, 234
Samos, 41
sampling, 127, 287, 288, 325, 329, 353, 375
bias, 355
selection, 355
compensative, 357
conceal, 358
data-driven, 354
ear, 355
residue, 355
hierarchical, 354
linear interpolation, 355
non-linear interpolation
cubic, 355
spline, 355
quota, 354, 356
random, see random sampling
random seed, 354
recent event, 355
reservoir, see reservoir sampling
systematic, 354
time-based, 294
fixed-interval, 295
inline, 294
uniform, 354
SAN, 64
Storage Area Network, xxvi
SAP, 60
SAS, 389, 496
SASE, 44
satellite imagery, 106
SAX, 276, 283
Simple API for XML, xxvi
sc, 219, 256, 264
SCADA, 22, 44
Supervisory Control and Data Acquisition, xxvi
scalability, 45, 62, 65, 68, 275, 441, 442, 445, 446, 482, 488
horizontal, 80
vertical, 80
scalar quantization, 364, 366
non-uniform, 365
uniform, 364
scale-out, 80
scale-up, 80

Cambridge University Press
978-1-107-01554-8 - Fundamentals of Stream Processing: Application Design, Systems, and Analytics
Henrique C. M. Andrade, Buğra Gedik and Deepak S. Turaga

© in this web service Cambridge University Press
www.cambridge.org
<table>
<thead>
<tr>
<th>Index</th>
<th>523</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCH, 219, 220, 225, 226, 232, 234, 240, 241</td>
<td></td>
</tr>
<tr>
<td>Scheduler, see SCH</td>
<td></td>
</tr>
<tr>
<td>scheduler, 211, 226</td>
<td></td>
</tr>
<tr>
<td>scheduling, 21, 56, 164, 206, 208, 209, 219, 240</td>
<td></td>
</tr>
<tr>
<td>best effort, 210</td>
<td></td>
</tr>
<tr>
<td>bipartite matching problem, 240</td>
<td></td>
</tr>
<tr>
<td>certificate of infeasibility, 240</td>
<td></td>
</tr>
<tr>
<td>constraint</td>
<td></td>
</tr>
<tr>
<td>checker, 239</td>
<td></td>
</tr>
<tr>
<td>workflow, 236, 239</td>
<td></td>
</tr>
<tr>
<td>online, 240</td>
<td></td>
</tr>
<tr>
<td>schema, 8, 48, 49, 81, 89, 90, 108, 109, 111, 122, 126, 158, 183, 446</td>
<td></td>
</tr>
<tr>
<td>scoring, 290, 461</td>
<td></td>
</tr>
<tr>
<td>offline, see offline scoring</td>
<td></td>
</tr>
<tr>
<td>online, see online scoring</td>
<td></td>
</tr>
<tr>
<td>scripting, 85</td>
<td></td>
</tr>
<tr>
<td>sdb, 192, 194, 230, 256, 264–267</td>
<td></td>
</tr>
<tr>
<td>breakpoint, 266</td>
<td></td>
</tr>
<tr>
<td>debugging</td>
<td></td>
</tr>
<tr>
<td>execution flow, 266</td>
<td></td>
</tr>
<tr>
<td>punctuation flow, see punctuation debugging</td>
<td></td>
</tr>
<tr>
<td>tuple flow, 266</td>
<td></td>
</tr>
<tr>
<td>injection point, 266</td>
<td></td>
</tr>
<tr>
<td>probe point, 265, 266</td>
<td></td>
</tr>
<tr>
<td>trace point, 266</td>
<td></td>
</tr>
<tr>
<td>SDE, 371</td>
<td></td>
</tr>
<tr>
<td>Semi-Definite Embedding, xxvi</td>
<td></td>
</tr>
<tr>
<td>Secure SHell, see SSH</td>
<td></td>
</tr>
<tr>
<td>Secure Sockets Layer, see SSL</td>
<td></td>
</tr>
<tr>
<td>security, 56, 62, 203, 205, 251</td>
<td></td>
</tr>
<tr>
<td>selection, 9, 132</td>
<td></td>
</tr>
<tr>
<td>operator, 93</td>
<td></td>
</tr>
<tr>
<td>selectivity, 131, 132, 144</td>
<td></td>
</tr>
<tr>
<td>fixed, 131</td>
<td></td>
</tr>
<tr>
<td>variable, 131</td>
<td></td>
</tr>
<tr>
<td>Self-Organizing Map, see SOM</td>
<td></td>
</tr>
<tr>
<td>semantic impedance, 53</td>
<td></td>
</tr>
<tr>
<td>search, 112</td>
<td></td>
</tr>
<tr>
<td>Semi-Definite Embedding, see SDE</td>
<td></td>
</tr>
<tr>
<td>semi-structured information, 79</td>
<td></td>
</tr>
<tr>
<td>semiconductor, 467</td>
<td></td>
</tr>
<tr>
<td>manufacturing, 467</td>
<td></td>
</tr>
<tr>
<td>deposition, 468</td>
<td></td>
</tr>
<tr>
<td>etching, 468</td>
<td></td>
</tr>
<tr>
<td>oxidation, 468</td>
<td></td>
</tr>
<tr>
<td>photo resist, 468</td>
<td></td>
</tr>
<tr>
<td>photolithography, 468</td>
<td></td>
</tr>
<tr>
<td>polish, 468</td>
<td></td>
</tr>
<tr>
<td>Semiconductor Process Control, see SPC</td>
<td></td>
</tr>
<tr>
<td>sensitivity analysis, 20</td>
<td></td>
</tr>
<tr>
<td>sensor, 45, see also mote</td>
<td></td>
</tr>
<tr>
<td>network, 45, 344, 357, 369, 413</td>
<td></td>
</tr>
<tr>
<td>sentiment analysis, 106, 291, 292</td>
<td></td>
</tr>
<tr>
<td>sepsis, 454</td>
<td></td>
</tr>
<tr>
<td>sequence, 420</td>
<td></td>
</tr>
<tr>
<td>serialization, 199</td>
<td></td>
</tr>
<tr>
<td>serviceability, 82</td>
<td></td>
</tr>
<tr>
<td>set difference, 9</td>
<td></td>
</tr>
<tr>
<td>SGD, 417–420, 474, 476, 479</td>
<td></td>
</tr>
<tr>
<td>Stochastic Gradient Descent, xxvi</td>
<td></td>
</tr>
<tr>
<td>shared memory, 244</td>
<td></td>
</tr>
<tr>
<td>segment, 171</td>
<td></td>
</tr>
<tr>
<td>Short Message System, see SMS</td>
<td></td>
</tr>
<tr>
<td>shortest path, 20</td>
<td></td>
</tr>
<tr>
<td>algorithm, 37</td>
<td></td>
</tr>
<tr>
<td>SickKids Hospital, 40</td>
<td></td>
</tr>
<tr>
<td>Siena, 43</td>
<td></td>
</tr>
<tr>
<td>sigil, 126</td>
<td></td>
</tr>
<tr>
<td>signal, 39</td>
<td></td>
</tr>
<tr>
<td>continuous, 354</td>
<td></td>
</tr>
<tr>
<td>discrete, 354</td>
<td></td>
</tr>
<tr>
<td>processing, 49, 354, 355, 375</td>
<td></td>
</tr>
<tr>
<td>quantization, see quantization</td>
<td></td>
</tr>
<tr>
<td>signal processing, xxv, 3, 7, 16, 292, 487</td>
<td></td>
</tr>
<tr>
<td>signature detection, 36</td>
<td></td>
</tr>
<tr>
<td>SIMD, 13, 314</td>
<td></td>
</tr>
<tr>
<td>Single Instruction Multiple Data, xxvi</td>
<td></td>
</tr>
<tr>
<td>Simple API for XML, see SAX</td>
<td></td>
</tr>
<tr>
<td>Simple Mail Transfer Protocol, see SMTP</td>
<td></td>
</tr>
<tr>
<td>Simple Object Access Protocol, see SOAP</td>
<td></td>
</tr>
<tr>
<td>simplex method, 19</td>
<td></td>
</tr>
<tr>
<td>simulated annealing, 20</td>
<td></td>
</tr>
<tr>
<td>simulation</td>
<td></td>
</tr>
<tr>
<td>model-based simulation, 85</td>
<td></td>
</tr>
<tr>
<td>Single Instruction Multiple Data, see SIMD</td>
<td></td>
</tr>
<tr>
<td>Single Logistic Regression, see SLR</td>
<td></td>
</tr>
<tr>
<td>single pass</td>
<td></td>
</tr>
<tr>
<td>algorithm, 51, 403, 408, 423–425, 495</td>
<td></td>
</tr>
<tr>
<td>transform, 383</td>
<td></td>
</tr>
<tr>
<td>Singular Value Decomposition, see SVD</td>
<td></td>
</tr>
<tr>
<td>sink operator, 50, 53, 88, 109</td>
<td></td>
</tr>
<tr>
<td>as debugging tool, 193</td>
<td></td>
</tr>
<tr>
<td>multi-input, 109</td>
<td></td>
</tr>
<tr>
<td>UDP, 99</td>
<td></td>
</tr>
<tr>
<td>SISAL, 14</td>
<td></td>
</tr>
<tr>
<td>sketch, 358</td>
<td></td>
</tr>
<tr>
<td>BJKST, 360</td>
<td></td>
</tr>
<tr>
<td>Count, 359</td>
<td></td>
</tr>
<tr>
<td>Count-Min, 359, 360</td>
<td></td>
</tr>
<tr>
<td>distinct value-based query, 359</td>
<td></td>
</tr>
<tr>
<td>Flajolet-Martin, 360</td>
<td></td>
</tr>
<tr>
<td>frequency-based query, 359</td>
<td></td>
</tr>
<tr>
<td>Gibbons-Tirthapura, 360</td>
<td></td>
</tr>
<tr>
<td>linear, see linear sketch</td>
<td></td>
</tr>
<tr>
<td>sketches, 375</td>
<td></td>
</tr>
<tr>
<td>skew, 346, 367</td>
<td></td>
</tr>
<tr>
<td>slicing, 10</td>
<td></td>
</tr>
<tr>
<td>sliding window, 95, 136, 138–143, 348, 357, 403, 424, 474</td>
<td></td>
</tr>
<tr>
<td>SLR</td>
<td></td>
</tr>
</tbody>
</table>
Index

Single Logistic Regression, xxvi
smoothing, 16
SMS, 288, 334
Short Message System, xxvi
SMTP, 100
Simple Mail Transfer Protocol, xxvi
Snmp, 41
SOAP, 101, 228
Simple Object Access Protocol, xxvi
social
engineering, 35
network, 52, 53, 414, 497
socket, 243, 277
software development
life cycle
software engineering, 79, 178, 487, 488, 490
modular design, 148
principle, 82
solution, 67, 310
SOM, 430
Self-Organizing Map, xxvi
sort operator, 94
key, 94
order
ascending, 94
descending, 94
strict weak ordering, 133
sorting, 109
sorting key, 94
source operator, 49, 50, 52, 87, 89, 97, 106, 108,
135, 195, 259
file, 276
multi-output, 108
TCP, 99
UDP, 99
XML
data, 276
SPA, 6, 12, 18, 39, 45, 47, 50, 51, 53–56, 60, 64–66,
69, 77–80, 82–86, 88, 95–98, 100–102,
112–114, 128, 129, 132, 137, 145–148, 164,
166, 167, 170–172, 178, 179, 181–184,
186–191, 194, 195, 199, 200, 203, 204, 209,
218, 247, 277–279, 281, 285, 289, 290, 299,
303, 315, 316, 321, 325, 326, 328, 330, 333,
334, 337, 339, 342, 347, 348, 358, 359, 374,
378, 383, 388, 389, 397, 398, 408, 414, 417,
419, 424, 428, 431, 433, 456, 476, 481, 482,
487–491, 493–496
Stream Processing Application, xxvi
SPADE, 86, 188
Stream Processing Application Declarative
Engine, xxvi
SPAs, 5
SPC, 442, 469–472, 474–479, 481, 482
Semiconductor Process Control, xxvi
spectral
analysis, 16
clustering, 414
SPIRIT, 373–375
Streaming Pattern Discovery in multiple
Time series, xxvi
SPL, xiv–xvi, 23, 24, 48, 54, 55, 65–67, 77, 80, 81,
86–101, 107–111, 113, 115–118, 122–124,
126–131, 133, 135–138, 140, 148–152, 154,
156, 157, 159, 160, 162, 168–170, 172, 180,
183, 187, 195, 219, 227, 239, 243–245,
247–249, 251, 256–258, 278, 279, 281, 283,
285, 289, 294, 296, 300, 303, 311–313, 321,
323, 324, 327, 330–332, 336, 339, 355, 357,
360, 366, 372, 379, 391, 403, 424, 446, 448,
449, 451, 459, 462, 477, 487, 496
acc compiler, 256
debugging mode, 264
options, 92
buildin function, 117
compiler, see acc
mixed-mode, see mixed-mode
standard toolkit, see standard toolkit
time series toolkit, see time series toolkit
type system, 135
split, 49
operator, 96
SpotFire, 13, 187
SPS, 21, 33, 40, 42–45, 48, 53, 56–61, 63–69,
78–80, 92, 94, 97, 100, 148, 149, 151, 153,
159, 164, 166, 171, 179, 182, 184, 185, 187,
195, 196, 198–200, 203–215, 218, 220, 227,
228, 232, 276, 278, 279, 283, 335, 487, 488,
490–492, 494–496
Stream Processing System, xxvi
application, 204, 205
life cycle, 207
application development environment, see
application development environment
application runtime environment, see application
runtime environment
job, see job, 205, 207
middleware, 53
processing element, see PE
service, 204, 206
user, 204, see user
ACL, see ACL
credential, 204
SPSS, 389, 496
SQL, 9, 41, 54, 57, 58, 60, 61, 63, 66, 67, 83, 84,
86, 93, 100, 279
Structured Query Language, xxvi
SQL9 standard, 61
trigger, 41
SQLAI, 54, 55, 59
Stream Query Algebra, xxvi
squared error, 364, 365
Index 525

SRM, 220, 225, 226, 228, 229, 232, 236–241, 250, 253, 267
Streams Resource Manager, xxvi
monitoring cycle, 236
SSH, 251
Secure SHell, xxvi
SSL, 251
Secure Sockets Layer, xxvi
stack space, 167
standalone application, 92, 258
standard, see standard toolkit
development, 17, 346, 347
output, 88, 91, 267
toolkit, 87, 89, 90, 98, 109, 129, 131, 133, 136, 154, 162, 287, 283, 289, 294, 298, 303, 305, 311, 332, 448
standardization, 378
Stanford University, 57
state machine, 299
management, 51, 128
variable, 298
stateful operator, 93–95, 97, 128, 129
stateless operator, 93, 96, 128, 129
statement language, 87
static composition, 106, 108, 314
statistical, 429
analysis, 469
inference, 18
statistics, 3, 16, 17, 49, 342, 354, 487
steepest descent method, 7
stochastic search, 19
Stochastic Gradient Descent, see SGD
stock exchange, 130, 280
market feed, 292
market, 293
ticker symbol, see ticker symbol, 293, 300
Storage Area Network, see SAN
Storm, 64, 65, 67
STREAM, 22, 54, 57, 58, 218
stream, 46, 49, 69, 80, 89, 108, 126, 136, 207, 344, 363, 370
connection, 50, 78, 80, 81, 91, 107, 148, 164, 170, 171, 183, 184, 267, 193, 195, 196, 219, 226, 228, 232, 234, 241, 244, 246, 302, 314, 327, 491
dynamic, see dynamic connection
fan-in, 110, 111
fan-out, 110, 112
static, see static connection
consumer, 246
data mining, see data mining
rate, 195
view, 187
definition, 108
demultiplexing, 108
dimensionality reduction, see dimensionality reduction
content filter, 120
importing, 115, 120, 306, 309
merging, 317
mining, 482
process, 345
processing, 3, 39, 45, 203, 375, 471, 487, 495, 497
language, 80, 159, 162, 312
personal application, 37, 497
punctuation, see punctuation
relational operator, 87, 92
schema, see schema, 149
splitting, 317
predicate, 212
Stream Mill, 60
Stream Processing Application, see SPA
Stream Processing Application Declarative Engine, see SPADE
Stream Processing Applications (SPAs), 4
Stream Processing System, see SPS
Stream Query Algebra, see SQAl
StreamBase, 22, 54, 59, 60, 63, 64, 66, 67
LiveView, 64
Studio, 64
streaming analytics, 129
streaming Pattern discovery in multiple Timeseries, see SPIRIT
StreamIt, 22, 54, 210, 218
StreamMon, 58
Instance Graph, 219, 230
Streams Application Manager, see SAM
Streams Console, 187, 225, 226, 228
Streams Exchange, 448, 449

© in this web service Cambridge University Press www.cambridge.org
Streams Resource Manager, see SRM
Streams Studio, 179–185, 190, 191, 195, 197, 219, 230–232, 236, 238, 250, 256, 258, 267
Streams Web Server, see SWS
StreamsQL, 54, 63
streamtool, 118, 119, 219, 221, 224, 229, 230, 232, 236, 249, 251, 254–256, 267
StreaQuel, 57
string, 149
structure, 150
structured data, 48, 68, 79, 276, 496
Structured Query Language, see SQL
Student’s t-test, 431
subscription expression, 116
summarization, 17
summation, 319
supervised
analysis, 5, 36, see also unsupervised analysis
learning, 395
Supervisory Control and Data Acquisition, see SCADA
Support Vector Machine, see SVM
surprise, 427
surveillance, 34, 47
SVD, 375
Singular Value Decomposition, xxvi
SVM, 397, 429
Support Vector Machine, xxvi
swarm intelligence, 20
Swig, 489
SWS, 219, 220, 225, 228, 251, 256
Streams Web Server, xxvi
Sybase, 60
Sybase ESP, 60
synchronization, 129
synopsis, 129, 379
syntax
highlighting, 230
syntax-aware
completion, 230
syslog, 249
system
administration, 213
administrator, 169, 204
call, 153, 215, 289
callback, 262
health, 210
management, 203
System S, 60, 64, 65, 86, 209, 218
System R, 8
T-Mobile, 4
table, 8
task parallelism, 49, 80, 291, 312, 314, 320, 324, 326, 328, 493
TB, 4
Transmission Control Protocol, xxvi
adapter, 99, see also source operator
client, 99, 246, 277
connection, 99
port, 244
server, 99, 243, 246, 277
socket, 278
port, 228
TEDS, 472, 476
Transducer Electrical Data Sheet, xxvi
Telegraph, 57
Telegraph Continuous Queries, see TelegraphCQ
TelegraphCQ, 57, 218
Telegraph Continuous Queries, xxvi
temperature
body, 38
templatization, 122
temporal
analysis, 58
pattern, 294, 300
event, see also event
matching, see also pattern matching, 294, 299
transform, 376, 377, 379
TeraBytes, see TB
Teradata Warehouse Miner, 10
testing, 92, 489
text, 46, 79
file, 45
format, 91
thread, 321
throughput, 68, 171, 195, 211, 279, 325, 328, 330
thunk, 67
TIBCO, 61
TIBCO BusinessEvents, 44, 60, 61, 66
TIBCO BusinessEvents Studio, 61
ticker symbol, 288
time series, 16
analysis, 129
discords, 433
forecasting, 353
Hoits–Winters model, 353
Kalman filter, 353
Kernel Smoothing method, 353
toolkit, 459, 460
Wiener filter, 353
toolkit, 148, 150, 151, 256
big data, 101
data mining, see mining toolkit
database, see db toolkit
financial services, 101
function, 151
custom, 151, 152
native, 152
Internet, see inet toolkit
messaging, 100
namespace, see namespace
text, 101
Toronto, 40
trace point, 191, 192
tracing, 63, 248
traffic
system, 33
transactional log, 130
Transducer Electrical Data Sheet, see TEDS
transform, 375
format conversion, see format conversion
linear, see linear transform
lossless, 376
lossy, 376
mediation, see mediation
non-linear, 371, see non-linear transform
normalization, see normalization
standardization, see standardization
temporal, see temporal transform
transformation, 51
task, 44, 51, 58, 61, 63
cleaning, see data cleaning
de-duplication, see data de-duplication
filtering, see data filtering
format conversion, see data format conversion
pre-processing, see data pre-processing
standardization, see data standardization
transient failure, 229
Transmission Control Protocol, see TCP
transport layer, 50, 59, 164, 170, 171, 173, 196, 211, 227, 234, 241, 243, 259, 327
connection, 260
handshake, 246
workload, 246
direct call, 243, 245, 246, 327
endpoint, 228, 243, 244
label, 243, 244
in-order delivery, 171
LLM, 244
multi-cast, 244
reconnection workflow, 246
reliable delivery, 171
TCP, 243
unicast, 244
transportation, 34, 68
traveling salesman, 20
trigger policy, 93, 94, 138, 139, 141, 142, 155, 259
count-based, 139–142
delta-based, 142
time-based, 143, 259
true, 133
Truviso, 57
tumbling window, 138–143, 356, 368, 377, 381
tuple, 8, 9, 45–52, 63, 81, 86, 89, 131, 134, 136, 151, 191, 266, 283, 293, 344, 363, 370, 472
attribute, see attribute
batching, 328, 330, 331
by count, 330
by time, 330
class, 245
complex, 46
creation, 445
decoding, 245
declaration, 48
deserialization, 245
encoding, 245
ingestion, 154
literal, 116
marshalling, 245, 283
process operation, 246
serialization, 245
submission, 110, 155, 246
transform, see transform
type, 81, 245
unmarshalling, 245
tweet, 65, 324
Twitter, 65, 106, 324
Storm, see Storm
type, 142, 148–150
blob, 151
collection
hash map, see hash map
hash set, see hash set
list, 150
composite, see composite type
collection, 150
structure, 150
enumeration, 151
floating point number
decimal64, 149
float64, 149
binary, 149
decimal, 149
hierarchy, 11
integer, 149
int64, 149
primitive, see primitive type
contract, 82
string, 149
wstring, 149
system, 53, 67, 79, 81, 278
tuple, 151
UCLA, 60
University of California, Los Angeles, xxvi
UDP, 99, 100, 492
User Datagram Protocol, xxvi
adapter, 99
datagram, 280
Index

protocol, 99
socket, 278
UML, 61
Unified Modeling Language, xxvi
uncertainty, 398
model, 44
unconstrained optimization, 19
unicast, 244
Unified Modeling Language, see UML
Uniform Resource Identifier, see URI
union, 9, 292
University of California, Berkeley, 57
University of California, Los Angeles, see UCLA
University of Colorado, 43
University of Massachusetts, 44
University of Ontario Institute of Technology, 454
University of Toronto, 43
University of Virginia, 454
University of Wisconsin, 42
Unix, 88, 91, 101, 146
shell, 91, 92
unsigned integer literal, 89
unstructured data, 38, 48, 68, 79, 276, 496
unsupervised
analysis, 5, 36, see also supervised analysis
learning, 404
URI, 285
Uniform Resource Identifier, xxvi
user, 204
credential, 204, 207
right, 207
User Datagram Protocol, see UDP
User Experience, see UX
utility operator, 92, 96, 294
UX, 187, 481, 497
User Experience, xxvi
variance, 17, 346, 365, 367
vector
quantization, 364
versioning, 82, 149, 490, 496
Very Fast Decision Tree, see VFDT
Very Fast Machine Learning, see VFML
Very Large Scale Integration, see VLSI
VFDT, 398, 400, 402, 403
Very Fast Decision Tree, xxvi
VFML
Very Fast Machine Learning, xxvi
video, 46, 79
analysis, 106, 107
compression, 357
processing, 288
virtualization, 64, 209
visual metaphor, 85
visualization, 17, 52, 53, 87, 178, 203, 267, 487, 491, 492
data, 186
perspective, 183
development-time, 183
runtime, 183
service, 214
status, 185
topology
hierarchy, 179
VLSI, 468
Very Large Scale Integration, xxvi
VMware ESX Server, 209
voice codec, see codec
Voice over IP, see VoIP
VoIP, 47, 48, 52, 53
Voice over IP, xxvi
Volume Weighted Average Price, see VWAP
Voronoi
cell, 365
region, 408
VWAP, 130, 187, 289, 318, 323
Volume Weighted Average Price, xxvi
wafer, 468
wall-clock time, 143, 144
Walmart, 3
WAN, 14, 42, 43
Wide Area Network, xxvi
water management, 34
wavelet transform, 371, 372, 375, 379
Haar transform, see Haar transform
web
browser, 226
server, 280
service, 15, 100, 101
web server, 34
Web Services Description Language, see WSDL
WebSphere MQ, 100
Weka, 389, 442, 461, 462, 496
Who is Talking To Whom, see WTTW
Wide Area Network, see WAN
Wiener filter, 353
window, 18, 44, 57, 58, 78, 82, 94, 129, 138, 143, 144, 293, 347
count-based, 296
eviction policy, see eviction policy
flush, 138, 139, 141, 143
partition, 94, 95, 141, 298
policy
trigger, 138
punctuation, see punctuation
sliding, 136, 138–140, 142, 143
partition, 140, 141, 143
time-based, 296, 478
toggle policy, see trigger policy
sliding, 139
tumbling, 138–143, 296
batch operation, 138
flush, 138, 139
partition, 140, 143

WSDL, 101, 228, 490

Web Services Description Language, xxvi

WTTW, 52, 53

Who is Talking To Whom, xxvi

XHTML, 283

eXtensible HyperText Markup Language, xxvii

XML, 17, 46, 60, 79, 230, 257, 276–278, 283, 284, 390, 391, 481

eXtensible Markup Language, xxvii

XML Path Language, see XPATH

XML Schema Definition, see XSD

XPATH, 283

XML Path Language, xxvii

XSD, 257, 278, 283

XML Schema Definition, xxvii

XSLT, 378

EXTensible Stylesheet Language Transformations, xxvii

Yahoo, 7, 64

S4, see S4

Yahoo S4, 64, 67

ZIP, 102

Zone Improvement Plan, xxvii

Zone Improvement Plan, see ZIP