Index

A
abused user as interviewee, 101–102
access monitoring tables, 344
accumulating snapshot fact tables functions of, 274–276
table loader, 397–398
ad hoc access
access policy, 219
ad hoc queries, 160
BI applications, 475
meaning of, 475
performance testing, 552
security tables, 344
affinity grouping, 497–498
aggregates
aggregate builder, 403–404
aggregate table, 465–466
BI query and reporting tool for, 354
choosing aggregates, 354–355
definitions, 140
elements of, 136
maintenance of, 355–357
metadata, 140
navigation, 136–137, 146
OLAP engine for, 354
Oracle, use of, 353–354
and performance improvement, 353
physical database, 353–357
presentation server architecture, 134–137
problems, types of, 354–357
and query management, 146
sizing estimates, 341
usage based optimization, 136
usage statistics, 140
agile. See rapid application development
alerts
BI applications, 481–482, 491
allocation
fact tables, 277–278
analytical applications
BI applications, 142, 480, 488–490, 512,
527
developing, 527
pre-built applications, 489–490
read/write applications, 490
standard template for, 512
types of, 488–489
analytical cycle
BI applications, 479
analytical data
user requirements, 134
annual report
as research tool, 70
ANSI SQL 99 support
queries, 483
application interface
BI applications, 504
GUI, physical database, 358–359
application logic, 155
architecture
plan document, 187–189
See also enterprise data warehouse bus
architecture; parallel processing
hardware architectures; presentation
server architecture; technical architecture
archives
backup system function, 408–409
ETL system, 373, 408–409, 430
long-term archiving, 573
atomic data
atomic level models, 134
dimensional model, use of, 137, 299
fact tables, 240–241
user requirements, 133
attributes. See dimension attributes
audit(s)
process metadata, 131
audit columns, 376
audit dimension
ETL system, 127, 385–386, 416
keys, assigning, 453–454
audit tables
ETL system, 344, 385
authentication
defined, 145, 215
system set-up, 215
authorization
defined, 145
B
back room
activities of, 114
defined, 110
See also extract, transformation, and load (ETL) system
backdoors, 177
backup system
archive/retrieval, 408–409
basic requirements, 408
ETL system, 128, 407–409, 416
regular backups, 571–573
as security function, 220–221
balanced scorecard approach, 492
bandwidth
importance of, 169
base facts, 310
bastion server, 213
batch files
real time tradeoffs, 419
batch parameters
technical metadata, 132
bitmapmed index, 346
block size
physical database, 358
bottlenecks
avoiding. See performance improvement
disk controllers to avoid, 363
ETL system, common causes, 411–412
and hardware decisions, 168
symmetric multiprocessing (SMP), 160
bridge tables, 270–273
diagrams of, 269, 271–272
many-valued dimension table builder, 400
many-valued dimensions with, 270–273
navigation table, 269
organization bridge table, 269–270
pros/cons of, 270, 273
briefing letter
to interviewees, 76–77
browsers
metadata browser, 503
B-tree index, 345
bubble chart
dimensional modeling process, 303–304
bulk data loading, 444
bus matrix. See enterprise data warehouse bus matrix
business analysts
role of, 35
business drivers
communication with, 45
role of, 33
business group split, 139
business intelligence
architects, role of, 37
use of term, 10–11
business intelligence (BI) applications, 141–155
analytic cycle, steps in, 476–479
analytical applications, 480, 488–490
application interface options, 504
applications, types of, 142–143
architecture model, diagram, 142
BI tools, purchase of, 205–207
business intelligence (BI) portal, 499–504
customer modes related to, 475
dashboards, 480, 490–493
data mining, 494–499
deployment plan, 534
design/development of. See business intelligence (BI) applications development
desktop tool, 154
dimensional model, use of, 137–138
direct access query/reporting tools, 479–487
elements of, 7
importance of, 474–476
interface options, 143
maintenance of system, 534
metadata, 154–155
operational business intelligence, 493–494
portal developers, role of, 38
scorecards, 480, 490–493
standard reports, 476, 480, 487–488
types of, diagram of, 475
use of term, 10, 13
user requirements, 133–134
business intelligence (BI) applications development, 505–539
analytical applications, 512, 527
application content, 517–519
application developers, role of, 37–38, 506–507
blueprint for action, 535–538
common development issues, 533
dashboards, 512
data mining, 529–530
documentation, 534
formatting, 484, 526
install/test BI tools, 522–523
metadata, 522–524
naming project, 508–509
navigation framework, 519–520
operational business intelligence, 527–529
portals. See business intelligence (BI) portals
queries, creating, 526
review/validate system, 520–521
security, 487, 523
specifications, developing, 515–517
task list, 539
templates, creating, 509–512
user interactions, defining, 525
business intelligence (BI) data stores, 151–154
analytic application results, 152–153
analytical data, disposable, 152
application server caches, 151
databases, local, 151–152
downstream systems, 153
local stores, planning for, 152
reports, 151
security, 153–154
business intelligence (BI) management services, 143–151
enterprise reporting services, 147–149
metadata services, 144
operational write back, 150
portal services, 150
query management, 146–147
security services, 144–145
shared services, 143
usage monitoring, 145
vendor tools, 150
web access, 149–150
business intelligence (BI) portals, 499–504, 530–532
BI management function, 150
density factors, 500–502
developing, 530–532
effective criteria for, 499–500
functions of, 499, 502–503
home page, example of, 501
maintenance of, 565–566
navigation framework, 519–520
pages, levels of, 531
process metadata, 155
structural factors, 500–502
business metadata
BI applications, 155
elements of, 116
ETL system, 116, 132
presentation server architecture, 140
business needs
and ETL system design, 370–371
business requirements definition, 63–107
business users, role in, 64–65
dimensional modeling process, 292–293
elements of, 5
findings document, 88–91
interviews. See business requirements interviews; project requirements interviews
levels of analysis, 63
methods to avoid, 67–68
prioritization grid, 92–93
prioritization meeting, 91–93
program versus project requirements, 65
task list, 107
business requirements interviews, 66–91
compliance interviews, 83
conducting interview, 80–82
documentation of results, 87–91
interview team roles, 68–70
interviewees, types of, 71–73
IT/security interviews, 65–66, 83
knowledge of organization, 70–71
preparation of interviewees, 76–77
program business interviews, 71–72, 82–83
questions, developing, 73–74
results, reviewing, 85–87
scheduling, 73–76, 79
success criteria information from, 84–85
synthesizing information from, 86–87
wrapping-up, 84–85
business rules
business metadata, 132
business rule screens, 308
C
Caches
application server caches, 151
shared, 162
calculations
OLAP database, adding to, 352
Cap Gemini Ernst and Young Integrated Architecture Framework, 180
causal factors
BI applications, 477–478
central processing units (CPUs), 165, 203
change data capture (CDC)
ETL system, 127, 376–378
charting
BI applications, 484–485
checksums
data change comparisons, 458
ETL system, 131, 458
classifying, 496–497
clean functions
ETL system, 127
closed loop applications
BI applications, 480
clueless user as interviewee, 103
clusters/clustering, 161–162, 496
affinity grouping, 497–498
clustered index, 345–346
scale out, 161–162
columns
calculations on pivot results, 484
descriptions/documentation of, 559
enterprise data warehouse bus matrix, 252–253
comatose user as interviewee, 102
Common Warehouse Metamodel (CWM), 172
communication plan, 43–46
architecture plan as tool, 111
business user community, 45–46
for expansion of system, 584–585
IT professionals, 46
over-information, avoiding, 44
and project team, 44
competency centers, 57
compliance
ETL system, 128, 370–371, 415–416
professionals, interviews with, 83
and usage monitoring, 145
conforming dimensions, 244–246
benefits of, 245–246
bus matrix, 90
detailed versus shrunken dimension tables, 244–245
elements of, 13, 373
enterprise data warehouse bus matrix, 249–250
ETL system, 127, 373, 386–387
fact tables, 281–282
identifying, 309
political factors, 253
conforming facts, 239, 373
connectivity
bandwidth, 169
database connectivity, 170
directory services, 170
file transfer protocols, 169
infrastructure factors, 169–170
remote access, 169
security threats, 176–177
consolidated fact tables, 281–282
consultants
role of, 39–40
contract negotiation
with vendors, 201–202
Coordinated Universal Time (CTC), 256
costs/investments review, 27–31
costs, types of, 27–28
for expansion of system, 581–583
governance committee role, 55
and infrastructure decisions, 159
purchasing process, 193
return on investment (ROI), 30–31
tangible and intangible benefits, 28–30
cross-tabulation analysis
pivoting, 483–484

cubes
dimensional modeling, 128, 235
OLAP cube builder, 128, 404
currency conversion
fact tables, 278–279
customer relationship management (CRM)
data leverage for, 153

dashboards
alerts, 481–482, 491
BI applications, 142, 480, 490–493, 509–512
pros/cons of, 23–24
sales, example of, 491–492
standard template for, 512
user interface, 512
data
access, user requirements, 133
data track elements of, 6–7
delivery services, 129
information assets, threats to, 175
logical data maps, 132
service providers, 129
volumes, and infrastructure decisions, 157
data access policy, 215–220
ad hoc. See ad hoc access
e external reports, 219–220
filtered reports, 218
masking sensitive data, 216–217
open access policy, 215–216
restricted reports, 218
sensitive data matrix, 216
unrestricted reports, 218
data acquisition. See extract, transformation, and load (ETL) system
data architects
role of, 36
data audit
interviews, 66
data changes
change data capture (CDC) methods, 376–378
dimension updates, logic flow diagram, 460
incremental processing, 456–465
data cleansing
ETL system, 127, 381–382
data compression
ETL system, 379
data dictionary
functions of, 132
data extraction
source systems, 123, 436–437
See also extract, transformation, and load (ETL) system
data handler
late arriving data handler, 400–401
data handoff
ETL system, 373–374
data integration
ETL system, 372–373
data latency
ETL system, 373
data marts
defined, 134, 248
data mining, 494–499
affinity grouping, 497–498
anomaly detection, 498
BI applications, 142, 479–480, 529–530
classifying, 496–497
clustering, 496
complexity of, 498–499
defined, 494
elements of, 494–495
estimation/prediction, 497
importance of, 495
model, developing, 529–530
specialists, role of, 39
data modelers
role of, 36
data modeling tools
benefits of, 340
physical database design, 340
data parsing
historic data loading, 439–440
data profiling
business rule screens, 308
defined, 17, 297, 307
dimensional modeling process, 297, 307–308
ETL system, 127, 375–376
and feasibility of project, 17–19
junk dimensions, avoiding with, 265
outcomes of, 308
project requirements interviews, 99–100
steps in, 307–309, 376
tools for, 297
data propagation manager
role of, 128, 404–405
data quality
business metadata, 132
data cleansing, 381–382
errors, handling, 383–385
ETL system, 131, 371–372, 380–383, 431
improving, guidelines for, 380–381
live testing, 549
pre-deployment testing, 546–547
process metadata, 131
quality screens, 382–383
real time processing, 420
data sources
choosing, criteria for, 308–309
data profiling, 307–308
data source split, 139
for dimensional modeling, 296–297, 305–309
data staging
eliminating and real time processing, 420
data stewards
and dimensional modeling process, 291–292, 321–322
enterprise level role, 56–57
role of, 35–36, 321–322
data stores
ETL system, 130–131
physical database, 343–344
technical metadata, 132
See also business intelligence (BI) data stores
data transformations. See dimensional table transformations; fact table transformations
data warehouse
use of term, 10–11
data warehouse appliances
features of, 159, 162–163
data warehouse/business intelligence (DW/BI)
BI applications, 473–504, 473–539
build versus buy system, 120
business requirements definition, 63–107
deployment of system, 541–562
dimensional modeling, 233–324
elements of, 10, 112
evolution of system, 118–119
expansion of system, 579–591
extract, transformation, and load (ETL) system, 369–423
extract, transformation, and load (ETL) system development, 425–472
data warehouse/business intelligence (DW/BI)
(continued)
Kimball Lifecycle, 1–14
maintenance of system, 563–573
organizational factors, 139–140
physical database design, 327–367
technical architecture, 109–178
database(s)
administrator, role of, 36–37
connectivity, 170
design. See physical database design
infrastructure factors, 165–167
local, stored, 151–152
online analytical processing (OLAP). See
online analytical processing (OLAP) database
optimizers, 203
relational. See relational database
staging database, 342, 572
database management system platform
product selection criteria, 203–204
date/time dimension, 253–256
date/timestamps, 255
ETL system, 393
historic data loading, 446–447
outrigger dimensions, 268
surrogate date keys, 254–255, 331–332
time of day, 255
time zones, 256
decode tables
ETL system, 130–131
real time processing, 420
deduplication
ETL system, 386–387
degenerate dimensions, 256–257
delays
late arriving data handler, 400–401
delivery functions
ETL system, 127–128
denial of service attacks, 176, 177, 211–212
density
BI portals, 500–502
dependency analysis
ETL system, 128, 412–413, 416
deployment, 541–562
blueprint for action, 574–576
documentation, 558–560
elements of, 7–8
ETL system, 556
online analytical processing (OLAP) database,
556–557
pre-deployment testing. See deployment,
pre-deployment tests
of relational database, 554–556
standard BI reports, 557
task list, 577
training, 560–561
deployment, pre-deployment tests
data quality assurance testing, 546–547
desktop, 553–554
live testing, 548–549
operations process testing, 547–548
performance testing, 549–552
system testing, 543–546
usability testing, 552–553
desktop
cross-platform support, 168
infrastructure factors, 168
on infrastructure map, 224
and memory, 168
pre-deployment testing, 553–554
tool, BI applications, 154
development database
building, 343
development environment
security measures, 211
diff compare 377
dimension(s)
attributes. See dimension attributes
identifying, 247, 299–300
model, developing. See dimensional modeling;
dimensional modeling process
OLAP database, 351
dimension attributes 241–242
changes, approaches to, 431
dimension tables, adding to, 258–259
ETL system, 430
identifying, 304–305
dimension manager
ETL system, 128, 387–389, 402
role of, 292, 402
slowly changing dimension (SCD) manager,
387–389
dimension table(s), 241–246
attributes, adding, 258–259
date/time dimension, 253–256
design, documenting, 311, 313
dimension attributes 241–242
elements of, 13
ETL system, 393–395, 438–447
hybrid slowly changing dimension methods,
261–262
incremental processing, 456–459
indexing, 347
keys, 243–244, 332
loading, 443–445
mini-dimensions, 259–260
number in model, 243
organization of, 242–243
role-playing dimensions, 262–263
rows, adding, 258
transformations. See dimension table transformations
Type 2, loading history, 445–446
dimension table transformations, 439–443
data from multiple sources, 439–440
many-to-one/one-to-one relationships, validating, 441–443
production codes, decode, 440–441
simple data conversion, 439
surrogate key assignment, 443
Type 1, populating, 438–439
dimensional modeling, 233–286
for atomic level data, 137–138, 299
benefits of, 237–238
bridge tables, 270–273
changing dimensions, approaches to, 257–262
deformed dimensions, 244–246
degenerate dimensions, 256–257
design process, 246–248
dimension attributes, adding, 258–259
dimension tables, 241–246
dimensions in, 235
elements of, 6
extensibility of, 238
fact tables, 238–241, 273–282
facts in, 235
functions of, 234–235
hierarchies, 268–270
junk dimensions, 263–265
myths related to, 282–286
normalized data converted to, 236
outrigger dimensions, 267–268
process of. See dimensional modeling process review/validate, 316
rows, adding, 258
snowflake dimensions, 265–267
star schema, 13, 235, 347
compared to 3NF modeling, 235–236
dimensional modeling process, 287–324
attributes/metrics, identifying, 304–305
base/derived facts, identifying, 310
blueprint for action, 323–324
bubble chart, 303–304
bus matrix, updating, 314–315
and business requirements, 292–293
conformed dimensions, 309
data profiling, 297, 307–308
data sources, identifying/selecting, 305–309
and data stewards, 291–292, 321–322
detailed design worksheet, 311–314
dimensions, identifying, 299–301
facilities/supplies for, 297
finalization stage, 320–321
flow diagram of, 289
grain declarations, 298–300
initial design session, 302–303
issues list, 305, 317
IT data model review, 318–319
modeling tools, 293–295
naming conventions, 295–296
roles in, 289–292
source data in, 296–297
task list, 325
time requirements for, 302
users, presenting to, 319–320
direct access query/reporting tools, 479–487
director of project
role of, 33–34
directory servers, 213–214
functions of, 213
security features, 214
directory services
and infrastructure decisions, 170
disk controllers
to avoid bottlenecks, 363
disk drives
data warehouse requirements, 164
disk space monitoring, 570
fault tolerance, 164
on infrastructure map, 223
versus memory, 164
problems with, 163
subsystems, use of, 164
documentation. See project documentation
downstream systems
stored, 153
drill down
BI applications, 484
reports, 149
drill through
OLAP database, 350
dummy counter, 280
educators
role of, 39
encryption, 214–215
elements of, 215
ETL system, 379
terprise application integration (EAI) message bus, 126
terprise data warehouse
use of term, 10
terprise data warehouse bus architecture,
248–253
terprise data warehouse bus matrix,
249–250
terprise data warehouse opportunity
matrix, 90–91
planning issues, 248–249
enterprise data warehouse bus matrix, 249–250
conformed dimensions, 249–250
elements of, 6
example of, 89–90
importance of, 249
problems, types of, 252
supply chain management, 250–252
updating, 314–315
enterprise data warehouse opportunity matrix
example of, 90–91
elements of, 90
enterprise reporting services, 147–149
enterprise resource planning (ERP)
pros/cons of, 123
environmental factors
security, 174–175
ER models, 236
error event schema
ETL system, 127, 383–385
error rows
handling, 461–462
estimation, 497
evaluation matrix
product comparison, 193–194
event measurement
and process metadata, 117
exceptions
identifying, BI applications, 477
technical metadata, exception-handling logic, 132
expansion of system, 579–591
blueprint for action, 590–591
communication factors, 584–585
current system, assessment of, 585–587
data warehouse, marketing of, 584
growth opportunities, prioritizing, 587–588
iterative growth, 588–590
return on investment (ROI) evaluation, 581–583
service metrics, monitoring for, 583–584
steering committee for, 580–581
user input, 580
expectations about project
managing, 53, 80
external reports, 219–220
extract, transformation, and load (ETL) system
aggregate builder, 403–404
architects, role of, 37, 291
architecture model, diagram, 122
archiving and lineage, 373
auditing tables, 344, 385
backup system, 407–409
bottlenecks, common causes, 411–412
change data capture (CDC), 376–378
columns, defining, 252–253
dimension manager, 402
elements of, 7, 121, 375
encryption, 379
evaluation matrix
ETL system, 127, 383–385
error event schema, 127, 383–385
errors, handling, 383–385, 467–468
ETL tools, purchase criteria, 204
extract functions, 127
fact providers, 402–403
fact table builders, 395–398
failure, causes of, 409
file-based extraction, 379
functional service providers, 129
hierarchy manager, 393
importance of, 370
job schedulers, 406–407
late arriving data handler, 400–401
lineage and dependency analyzer, 412–413
lookup tables, 130–131
maintenance of system, 468, 568–569
management services, 128–129
master data management, 125–126
message queues/log files/redo files, 126
metadata, 116–117, 121, 131–132
metadata repository manager, 417
multi-valued bridge table builder, 400
OLAP cube builder, 404
operational data stores (ODS), 124–125
parallelizing, 414
performance testing, 550–551
pipelining, 414
problem escalation system, 413–414
proprietary source formats, 126–127
quality screens, 382–383
real time partition, 421–423
real time tradeoffs, 419–420
real time triage, 417–419
compliance manager, 415–416
columns, defining, 252–253
data cleansing, 381–382
data compression, 379
data delivery services, 129
data handoff, 373–374
data integration, 372–373
data latency, 373
data profiling, 375–376
data propagation manager, 404–405
data quality, 131, 380–383
data service providers, 129
data stores, 130–131
deduplicating, 386–387
delivery functions, 127–128
deployment of, 556
design/development of. See extract, transformation, and load (ETL) system
dimension manager, 402
elements of, 7, 121, 375
equality, 379
evaluation matrix
ETL system, 127, 383–385
error event schema, 127, 383–385
errors, handling, 383–385, 467–468
ETL tools, purchase criteria, 204
extract functions, 127
fact providers, 402–403
fact table builders, 395–398
failure, causes of, 409
file-based extraction, 379
functional service providers, 129
hierarchy manager, 393
importance of, 370
job schedulers, 406–407
late arriving data handler, 400–401
lineage and dependency analyzer, 412–413
lookup tables, 130–131
maintenance of system, 468, 568–569
management services, 128–129
master data management, 125–126
message queues/log files/redo files, 126
metadata, 116–117, 121, 131–132
metadata repository manager, 417
multi-valued bridge table builder, 400
OLAP cube builder, 404
operational data stores (ODS), 124–125
parallelizing, 414
performance testing, 550–551
pipelining, 414
problem escalation system, 413–414
proprietary source formats, 126–127
quality screens, 382–383
real time partition, 421–423
real time tradeoffs, 419–420
real time triage, 417–419

recovery system, 409–410
requirements for, 370–374
sandbox source system, 436–437
security, 372, 415–416
slowly changing dimension (SCD) manager, 387–392
sorting, 128, 412
special dimensions manager, 393–395
staging tables, 330–331, 343
stream extract, 379
surrogate key generator, 392
surrogate key pipeline, 398–400
use of term, 11–12
version control, 410
version migration system, 410
workflow monitor, 411–412
XML schema, 126
extract, transformation, and load (ETL) system development, 425–472
aggregate table, 465–466
blueprint for action, 469–471
default strategies, types of, 430–432
developers, role of, 37
development practices, 426
development time, estimate of, 121, 369
ETL tools, choosing, 429–430
hierarchies, cleaning, 432–434
high level data staging plan, diagram of, 428–429
incremental processing, dimension table, 456–459
incremental processing, fact table, 459–465
job schedulers, 466–467
one-time historic load processing. See historic data loading
process, steps in, 427
specification document, 435–436
table schematics, 434
task list, 472

F
facilitated sessions
versus interviews, 66–67
steps in, 94
fact(s)
base facts, 310
conformed facts, 239
derived facts, 310, 312
in dimensional modeling, 235, 238–239
identifying, 247–248, 301
fact providers
role of, 292, 402–403
fact table(s), 238–241, 273–282
accumulating snapshot tables, 274–276
allocation, 277–278
atomic, 240–241
audit statistics, 448
base and derived facts, 310, 312
builders. See fact table builders
consolidated tables, 281–282
degenerate dimensions, 256–257
design, documenting, 313
disk space monitoring, 570
ETL system, 447–455
factless tables, 280–281
granularity, 240–241, 246–247
granularity differences, 276–278
historic extracts, 447
incremental loading, 463–464
incremental processing, 459–465
indexes, 348
junk dimensions, 263–265
keys, 240, 333
load schematics, diagram of, 434
multiple currencies/units of measure, 278–280
partitioning, 360
periodic snapshot tables, 273–282
primary keys, 333
provider system, 128
transaction tables, 273, 463
transformations. See fact table transformations
fact table builders, 128, 395–398
accumulating snapshot fact table loader, 397–398
periodic snapshot fact table loader, 396–397
transaction grain fact table loader, 396
fact table transformations, 448–454
audit dimension keys, assigning, 453–454
content, improving, 449–450
data quality checkpoint, 460–461
error rows, handling, 461–462
fact data, restructuring, 450
null values, 448–449
surrogate key pipeline, 450–453, 461–463
factless fact tables, 280–281
fast data loading method, 444, 455
fault tolerance
disk drives, 164
physical database, 362
feasibility
date issues, 17–18
and prioritization grid, 92–93
federated data, 162
File Transfer Protocol (FTP), 169
filtered reports, 218
findings document
business requirements definition, 88–91
importance of, 89
firewalls
packet-filtering routers, 212–213
screened subnet configuration, 213
fixed hierarchies, 268
foreign keys
  disabling, issues related to, 334–335
  null values, avoiding, 449
  physical database design, 334–335
formatting
  BI applications, 484, 526
forums
  BI portals, 503
front room
  activities of, 114
  defined, 110
See also business intelligence (BI) applications
functional service providers, 129
fuzzy logic
  date transformation, 440
G
governance of project
  characteristics of, 54–55
  governance committee, role of, 55–56
grain
  declaring, 246–247, 298–300
  defined, 240
  elements of, 240–241
  granularity differences, 276–278
  and identifying dimensions, 299–300
  OLAP database, 350–351
graph(s)
  BI applications, 484–485
  graphical user interface (GUI)
  physical database, 358–359
H
handheld devices
  BI applications, 143
hardware
  costs/investments review, 27
  CPUs, 165
  disk drives 163–164
  on infrastructure map, 223
  installation of, 224–225
  memory, 164
  parallel processing hardware architectures, 159–163
  platform sharing, 167
  product selection criteria, 202–203
  secondary storage, 165
  security measures, 211
  security threats to, 174–175
  hash partitioning, 361
  hashing algorithms
  data change comparisons, 458
hierarchies
  bridge tables, use of, 269–270
  clean, construction of, 432–434
  dimensional modeling, 268–270
  ETL system, 432
  fixed, 268
  hierarchy manager, role of, 128, 393
  snowflake dimension, diagram of, 442
  variable depth, 268–269
hijacking, 177
historic data loading, 437–456
  date dimension, 446–447
  dimension table transformations, 439–443
  dimension tables, loading, 443–446
  dimension tables, Type 1, populating, 438–439
  fact table extracts 447–448
  fact table loading, 454–455
  fact table transformations, 448–454
  load cycle, speeding, 464–465
  performance improvement tips, 455
historical load
  sizing estimates, 342
horizontal partitioning
  presentation server architecture, 139
hybrid slowly changing dimension methods, 261–262, 391–392
impact analysis
  and metadata integration, 171
incremental processing, 456–465
  changed data, identifying, 456
  dimension tables, 456–459
  fact tables, 459–465
indexes, 344–350
  bitmapped index, 346
  B-tree index, 345
  clustered index, 345–346
  dimensional tables, 347
  fact tables, 348
  finalizing, 357
  initial plan, 344–345
  for loads, 349
  OLAP database, 349
  parallel indexing, 465
  sizing estimates, 341
infrastructure, 156–170
  and application server, 167–168
  choices, drivers of, 156–157
  connectivity factors, 169–170
  database factors, 165–167
  desktop factors, 168
  functions of, 156
  hardware performance, 163–165
parallel processing hardware architectures, 159–163
requirements, criteria for, 157–159
infrastructure map, 221–224
desktop on, 224
diagram of, 222
functions of, 221
network on, 223–224
security system on, 224
storage systems on, 223
instantaneous business intelligence, 143
intangible benefits
financial, 29–30
Internet
BI application web access, 149–150
product information on, 195
security threats, 211–212
TCP/IP protocol, 211–212
interview(s)
briefing letter to interviewees, 76–77
business requirements. See business requirements interviews
data audit interviews, 66
end, signs of, 80
versus facilitated sessions, 66–67
ground rules, 78–80
interviewees, challenging, 101–104
project requirements. See project requirements interviews
questions, developing, 73–74
sequencing, 73–74
time spent/group size, 75–76
write-ups, organization of, 88
interview team, 68–70
lead interviewer, 68
observers, 70
scribe, 68–70
issue tracking, 51–52
importance of, 51
log for, 51–52
issues list
dimensional modeling, 305, 317
IT professionals
business requirements interviews, 65–66, 83
communication with, 46
data model review, 318–319

junk dimensions
ETL system, 393–394
fact tables, 263–265

K
Kerberos authentication, 215
keys
audit dimension keys, 453–454
dimension tables, 243–244, 332
fact tables, 240, 333
performance indicators, 492
physical database design, 332–335
primary, 332–333
See also foreign keys; surrogate keys
kickoff meeting, 47–48
project requirements interviews, 95–96
Kimball Lifecycle, 1–14
business intelligence application track, 7
business requirements definition, 5
data track, 6–7
deployment, 7–8
diagram, use of, 8–9
growth stage, 8
maintenance, 8
project management, 4–5
project planning, 4
tasks, flowchart of, 3
technology track, 5–6
terminology related to, 10–14
See also specific stages of lifecycle
know-it-all user as interviewee, 103

lag facts, 276
late arriving data handler
ETL system, 128, 400–401
lead testers
role of, 39
leadership
governance committee, 55–56
legacy licenses, 374
Lightweight Directory Access Protocol (LDAP), 170
lineage analysis
ETL system, 128, 412–413, 416
functions of, 171, 412
and metadata integration, 171
live testing
pre-deployment, 548–549
log(s)
change request control log, 52
ETL system, 407
issue tracking, 51–52
log(s) (continued)
log files, 126, 444
log scraping, 377
usage monitoring, 567
logging database
sizing estimates, 342
lookup tables
ETL system, 130–131

M
maintenance of project. See project maintenance
maintenance of system, 563–573
archiving data, 573
backup system, 571–573
BI portals, 565–566
data reconciliation support, 568
disk space, 570
ETL system, 568–569
performance tuning, 570–571
periodic assessments, necessity of, 57–58
resources monitoring, 569–570
security, 566
usage monitoring, 566–567
usage reporting, 567
user support, 563–565
management of project. See project management
management services
ETL system, 128–129
many-to-one relationships
validating, 441–443
many-valued dimensions bridge table, 270–273
table builder, 400
market basket analysis, 497
market research, 194–196
marketing
and usage monitoring, 145
massively parallel processing (MPP)
pros/cons of, 161
master data management (MDM)
cleaning hierarchies, 433–434
functions of, 125–126
future view of, 126
MDX (Multi-Dimensional eXpressions), 352
meetings
kickoff meeting, 47–48
prioritization meeting, 91–93
status meetings, 48–49
vendors, product selection meeting, 197–198
memory
64-bit platforms, 164, 203
and desktop, 168
versus disk drive, 164
physical database, 358
message queues
ETL system, 126
monitoring, 377
real time processing, 419
metadata, 115–117, 170–173
BI applications, 144, 154–155, 485, 487, 522–524
browser, 503
business metadata, 116, 132, 155
Common Warehouse Metamodel (CWM), 172
ETL system, 116–117, 121, 131–132
and impact analysis, 171
importance of, 117
integration, importance of, 171–172
integration methods, 172–173
and lineage analysis, 171
management strategy, 208–210
managers, role of, 37, 207–210
metadata capture, 407
metadata catalog, 208
metadata management systems, 173
metadata-driven architecture, 117
presentation server architecture, 140
process metadata, 116–117, 131, 140, 154
quality, importance of, 172
regular backup/recovery, 571–572
repository manager, 129, 417
single repository for, 117
standard BI reports, 511–512, 524
tables, size of, 342
technical metadata, 116, 131–132, 155
MetaObject Facility (MOF), 172
Metaphor Computer Systems, 1–2
milestone counter, 276
mini-dimensions, 259–260
ETL system, 394
pros/cons of, 259–260
mock up
standard BI reports, 516–518
multipass queries, 481
multi-table views
avoiding, 332
multi-valued bridge table builder
ETL system, 128, 400

N
naming
BI applications, 508–509
dimensional modeling process, 295–296
physical database, 329–330
project, 31–32
navigation
aggregates, 136–137, 146
BI applications framework, 519–520
BI portals, 502, 519–520
bridge tables, 269
network(s)
on infrastructure map, 223–224
routers, function of, 212
network security, 211–215
directory servers, 213–214
encryption, 214–215
firewalls, 213
packet-filtering routers, 212–213
nonexistent user as interviewee, 103–104
non-uniform memory architecture (NUMA)
pros/cons of, 161
normalized modeling
converting to dimensional data, 236
myths related to, 285–286
3NF modeling, 235–236
notification
ETL system, 407
null values
avoiding, 241, 330
fact tables, 448–449
types of nulls, 449

O
observers
on interview team, 70
one-time historic load processing. See historic data loading
one-to-one relationships
validating, 441–443
online analytical processing (OLAP) database, 350–365
calculated facts/measures, adding, 352
data granularity, 350–351
deployment of, 356–357
dimensions, building, 351
disk space monitoring, 370
drill through, 350
indexing, 349
purchase decisions, 203–204
regular backup/recovery, 572–573
relational OLAP, 154
sizing estimates, 341–342
online analytical processing (OLAP) engine, 115, 128, 136
aggregates, implementing, 354
cube builder, 128, 404
dimensional modeling cubes, 128, 235
features of, 166–167
metadata, 140
pros/cons of, 166–167
open-ended comments, 265
operating systems
cross-platform support, 168
operational business intelligence
developing, 527–529
elements of, 143, 480, 493–494
operational write back, 150
operational data stores (ODS)
forms of, 124
reporting ODS, 125
use of term, 124
operations process testing, 547–548
opportunity matrix. See enterprise data warehouse opportunity matrix
Oracle
aggregation navigation, 353–354
materialized views, 353
organization bridge table, 269–270
outrigger dimensions, 267–268
See also snowflake dimensions
overbooked user as interviewee, 102
overwrite technique
slowly changing dimension (SCD) manager, 389–390
overzealous user as interviewee, 102–103
P
packet-filtering routers, 212–213
parallel processing hardware architectures, 159–163
clusters, 161–162
data warehouse appliances, 162–163
hardware partitioning, 163
massively parallel processing (MPP), 161
non-uniform memory architecture (NUMA), 161
options, diagram, 160
symmetric multiprocessing (SMP), 159–161
parallelizing
ETL system, 128, 414, 464–465
with multiple loads, 464
parallel execution, 465
parallel structures, 465
partitioning
fact tables, 360
hash partitioning, 361
parallel processing hardware, 163
partition settings, 140
and performance improvement, 361
physical database, 360–361
of presentation server architecture, 138–139
range partitioning, 361
real time partition, 421–423
passwords, 145
performance improvement
and aggregates, 353
and disk controllers, 363
ETL tools for, 430
performance improvement (continued)
of historic data loading, 455
load cycle, speeding, 464–465
and partitioning, 361
performance tuning, 570–571
ETL system, 550–551
query performance, 551–552
service level confirmation, 550
system tuning, 549–550
periodic snapshot fact tables
functions of, 274
table loader, 396–397
periodic snapshot real time partition, 422–423
phishing, 177
physical database design, 327–367
aggregations, 353–357
block size, 358
blueprint for action, 364–366
build scripts/parameter files, saving, 358–359
data modeling tools, 340
data stores, 343–344
data structure, 335–336
development database, building, 343
elements of, 6
fault tolerance, 362
file location, 331
graphical user interface (GUI), 358–359
indexes, 344–350, 357
keys, 332–335
memory, 358
naming project, 329–330
null values, avoiding, 330
online analytical processing (OLAP) database,
350–365
partitioning, 360–361
process, diagram of, 329
sizing estimates, 340–343
snowflake dimensions, 331–332, 339
source-to-target map, 336–338
staging tables, 330–331
standards for, 328–335
star schema, 338
storage area networks (SAN), 362
storage needs, 361–362
synonyms/views, use of, 331–332
task list, 367
test database, building, 352
pick lists
BI applications, 485–486
pipelining
ETL system, 128, 414
surrogate keys, 450–453, 461–463
pivoting
BI applications, 483–484
plans and planning. See project plan; project planning
political factors
conformed dimensions, 253
product selection decisions, 200
portals. See business intelligence (BI) portals
prediction, 497
presentation server architecture, 133–141
aggregates, 134–137
architecture model, diagram, 135
foundation of, 134
and infrastructure decisions, 157–159
metadata, 140
partitioning of, 138–139
real time partitioning, 421–423
server, contents of, 138
strain on server, causes of, 158
primary keys
physical database design, 333–334
prioritization grid
two-by-two, elements of, 92–93
prioritization meeting
business requirements definition, 91–93
private/public key encryption, 215
problems
ETL problem escalation system, 128, 413–414
issue tracking, 51–52
trouble signs, recognizing, 53–54
process metadata
BI applications, 154
elements of, 116–117
ETL system, 116–117, 131
presentation server architecture, 140
product selection, 191–207
BI tools, 205–207
comparison prototypes, 199–200
contract negotiation, 201–202
hands-on evaluation, 198–199
hardware considerations, 202–203
market research, 194–196
narrowing choices, 196–197
product evaluation matrix, 193–194
purchases, basic, 192
purchasing process, 193
shortcuts, 200
trial period, 201
vendors, meeting with, 197–198
production codes, decode, 440–441
productivity
architecture plan as tool, 111–112
ETL system, 120
program business interviews
  questions for, 82–83
project changes
  change request control log, 52
  controlling, 52
  project managers, handling of, 51
project charter
  as scope document, 25–26
project definition, 16–31
  costs/investments review, 27–31
  project scope, 22–26
  readiness for DW/BI factors, 16–18
  shortfalls, types of, 18–21
project documentation
  application architecture plan, 183–191
  architecture plan as, 112
  architecture plan document, 187–189
  BI applications, 534
  and business metadata, 116
  business process model descriptions, 558–559
  business requirements interview results, 87–91
  consolidating, 50
  deployment stage, 558–560
  dimensional model, 320–321
  ETL system, 120, 435–436
  project requirements interviews, 100–101
  scope document, 25–26
  self-documenting, ETL system, 413, 429
  table/column descriptions, 559
  and usability, 120
  user, business metadata, 155
  user support guide, 560
project identity, 31–32
project maintenance
  architecture plan as tool, 111–112
  cost of, 28
  elements of, 8
project management, 47–58
  blueprint for action, 58–60
  competency centers, 57
  data stewardship in, 56–57
  elements of, 4–5
  of expectations, 53
  governance roles, 54–55
  issue tracking, 51–52
  kickoff meeting, 47–48
  periodic assessments, 57–58
  project changes, 52
  project documentation, 50
  project plan evaluations, 50
  project status meetings
    risk reduction, 58–59
    scope of project, 51–52
    status meetings, 48–49
    status reports, 49
  task list, 61
  trouble signs, recognizing, 53–54
project managers
  challenges to, 248–249
  and kickoff meeting, 47–48
  role of, 34
  user requests, handling, 51
project plan, 40–43
  evaluating/updating, 50
  example of, 42
  implementing. See project planning
  information in, 41
  necessity of, 40–41
  roles in development of, 43
project planning, 31–47
  architecture plan as tool, 111
  communication plan, 43–46
  defining project. See project definition
  elements of, 4
  naming project, 31–32
  project identity, 31–32
  project plan, 40–43
  staffing project, 32–40
project requirements interviews, 94–101
  conducting interviews, 97–98
  data profiling, 99–100
  documentation of results of, 100–101
  facilitated sessions, 94
  interviewee preparation, 95–96
  interviewee selection, 95
  kickoff meeting, 95–96
  preparation for, 94
  questions, types of, 97–98
  results, review of, 100
  wrapping-up, 99
project scope, 22–26
  changes to project. See project changes
  driving forces for, 22–23
  managing, 51–52
  rapid application development, 24–25
  scope document, 25–26
  single business process, 22–23
project status
  status meetings, 48–49
  status reports, 49
project team
  communication plan, 44
  dimensional modeling team, 289–292
  expansion of system, committee for, 580–581
  interview team, 68–70
  kickoff meeting, 47–48
  sponsor/driver briefings, 45
  team building, 40
proof-of-concept, 20
proprietary source formats, 126–127
publish-and-subscribe capabilities, 148–149
quality assurance (QA)
  QA analysts, role of, 36
  See also data quality

query(ies)
  ANSI SQL 99 support, 483
  BI applications, creating, 526
  direct access, 142, 479, 480–481, 487
  direct query string, 483
  embedded, 486
  formulating, 481–483
  kill query, 486
  multipass/multiset queries, 481
  performance and dimensional modeling, 236
  performance testing, 551–552
  process metadata, 155
  query manager, role of, 147
  query optimizer, 349
  semi-additive summations, 482
  successive constraints, 482
  query management, 146–147
  aggregate navigation, 146
  query governing, 147
  query prioritization, 147
  query reformulation, 146
  query retargeting/drill across, 146

RAID technology, 164, 362, 363
range partitioning, 361
rapid application development, 24–25
  basic principles, 24
  situations for use, 25
readiness for DW/BI, 16–18
read/write analytical applications, 490
real time processing
  design, real time categorization, 417–418
  presentation server partitioning for, 421–423
  real time partition, 421–423
  real time tradeoffs, 419–420
  real time triage, 417–419
recovery system
  ETL system, 128, 409–410
  regular backup/recovery, 571–573
redo files, 126
redundant storage, 164, 362
referential integrity, 398
refresh
  ETL system, 408–409
  regression testing, 545
  relational database
  and atomic level data, 134
  B-tree index, 345
  deployment of, 554–556
  disk space monitoring, 570
features of, 165–166
loading data. See historic data loading
multiple indexes, 348
partitioning methods, 361
purchase decisions, 203
regular backup/recovery, 571–572
relational OLAP, 154
star schema optimization, 347
surrogate key assignment, 443
remote access, 169
replicated data, 162
reporting database
  sizing estimates, 342–343
reporting operational data stores, 125
reports
  administrative tools, 149
  BI applications. See standard BI reports
data access levels, 218–219
delivery methods, 148
document layout, 148
drilldown, 149
existing reports, handling of, 515
iterative execution, 148
mass distribution, 149
process metadata, 154
publish-and-subscribe capabilities, 148–149
report execution server, 148
report library, 149
scheduling, 148
security, 149
stored, 151
use reporting, 567
Request for Information (RFI)
to vendors, 197
requirements interviews
  business requirements interviews, 66–91
  project requirements interviews, 94–101
resources of system
  monitoring, 569–570
restart
  ETL system, 409–410
restore system
  ETL system, 409–410, 416
restricted reports, 218
results set
  basic calculations, 483
  retrieval of data
  ETL system, 408–409
return on investment (ROI)
  expansion of system criteria,
  581–583
  pre-project stage, 30–31
risk reduction
  basic principles, 58–59
  business requirements definition, 104
  project management, 58–59
role-playing dimensions, 262–263
rows
  enterprise data warehouse bus matrix, 252

S
sandbox source system
  ETL system, 436–437
scalability
  scale out, 161–162
  strain on server, causes of, 158
scheduling
  business requirements interviews, 73–76, 79
ETL system. See job schedulers
reports, 148
technical metadata, 132
scope of project. See project scope
corecards
  balanced scorecard approach, 492
  BI applications, 142, 480, 490–493
screened subnet firewall configuration, 213
scribe
  on interview team, 68–70
scripting
  BI applications requirement, 486
search
  BI portals, 502
Secured Sockets Layer (SSL), 169, 215
security, 173–177, 210–221
  authentication, 145, 215
  authorization, 145
backup/recovery plan, 220–221
BI applications, 144–145, 153–154, 487, 523
connectivity. See network security
do data of data. See data access policy
development environment, 211
encryption, 214–215
ETL system, 372, 415–416
firewalls, 213
  for hardware, 211
ID numbers, masking, 217
  on infrastructure map, 224
maintenance of, 566
process metadata, 154, 155
professionals, interviews with, 83
reports, 149
security manager, role of, 38–39, 174
  security tables, 344
usage monitoring, 220
security professionals
  business requirements interviews, 72–73
security threats
  backdoors/trapdoors, 177
  business interruption, causes of, 176
to connectivity/networks, 176–177
denial of service attacks, 176, 177, 211–212
hijacking, 177
to information assets, 175
phishing, 177
to physical assets, 174–175
to software assets, 175
spoofing, 177
terrorism, 176
viruses/worms/Trojan horses, 177
self-documenting, ETL system, 413, 429
semantic layer definition, 155
semi-additive summations
queries, 482
serialization of system, 551
server(s)
  data warehouse appliances, 159
directory servers, 213–214
infrastructure factors, 167–168
storage area networks (SAN), 362
virtual servers, 223
  See also presentation server architecture
service(s)
  defined, 117
  service layers, flexibility of, 117–118
service level agreements (SLAs), 550
service oriented architecture (SOA)
data service providers, 129
enterprise level components, 118
shared services
  BI applications, 143
shrunken dimensions
  ETL system, 394–395
signed software certificates, 215
single repository systems
  and metadata integration, 172
  sizing estimates
  for normalized ETL dimensions, 433–434
  physically changing dimension (SCD) manager, 128, 387–392
hybrid combinations, 391–392
new column, creating, 391
new row, creating, 390–391
overwrite technique, 389–390
small static dimensions
  ETL system, 395
snapshots
  ETL system, 410
snowflake dimensions, 265–267
  diagrams of, 266, 339
  hierarchical relationships, diagram of, 442
for normalized ETL dimensions, 433–434
outrigger dimensions, 267–268
physical database design, 331–332, 339
pros/cons of, 265–267
software costs/investments review, 27
development managers, role of, 544
and infrastructure decisions, 159
installation of, 224–225
security threats to, 175
sorting
BI applications, 484
ETL system, 128, 412
loading data, pre-sort, 444
sizing estimates, 341
source code
file location standards, 331
source systems
sandbox source system, 436–437
types of, 123
See also extract, transformation, and load (ETL) system
source-to-target map, 336–338
spark lines, 492
special dimensions manager
ETL system, 128, 393–395
specification document
ETL system, 435–436
speed
load cycle, methods for, 464–465
sponsors
business driver, 33
communication with, 45
importance of, 18–19
and kickoff meeting, 47
multiple, downside of, 20–21
role of, 16–17, 33
senior management as, 16, 33
weak, avoiding, 19–20
spoofing, 177
SQL Compare, 555
SQL scripts, 140
staffing project, 32–40
business analyst, 35
business drivers, 33
business intelligence application developers, 37–38
business intelligence architects, 37
business intelligence portal developers, 38
business project leads, 34
consultants, 39–40
cost of, 27
data architect, 36
data mining/statistical specialists, 39
data modeler, 36
data stewards, 35–36
database administrator (DBA), 36–37
director of project, 33–34
educators, 39
extract, transformation, and load (ETL) system architect, 37
extract, transformation, and load (ETL) system developer, 37
kickoff meeting, 47–48
lead testers, 39
metadata manager, 37
project managers, 34
quality assurance (QA) analysts, 36
security managers, 38–39
sponsors, 33
teams. See project team
technical architects, 38
technical support specialists, 38
staging database
regular backup/recovery, 572
sizing estimates, 342
staging tables
for ETL system, 330–331, 343
stand-alone systems
data marts, 248
standard(s)
for physical database design, 328–335
standard BI reports, 512–519, 532
in BI analytical cycle, 476
candidate report list, 512–514
components of, 510–511
cost, specifying, 517–519
deployment of, 557
descriptions/documentation of, 559
elements of, 142, 487–488
metadata, 511–512, 524
scheduling, 532
specifications, developing, 515–516
titles, examples of, 488, 508–509
user requirements, 148–149
star schema
benefits of, 347
denormalized, example of, 338
dimensional modeling, 13, 235, 347
physical database design, 338
relational database, 347
statistical specialists
role of, 39
status meetings, 48–49
issue tracking as topic, 52
steering committee
for expansion of system, 580–581
storage
disk space monitoring, 570
estimating needs, 362–363
fault tolerance, 362
physical database, 359–363
redundant storage, 164
secondary storage, 165
system on infrastructure map, 223
See also disk drives; memory; storage area networks (SAN)

storage area networks (SAN), 164, 223
  elements of, 362
  physical database, 362
  pros/cons of, 362
  regular backup/recovery, 572
  stovepipe systems, 249

stream extract, 379

substitute user as interviewee, 102

subsystems, design of, 187

successive constraints
  queries, 482

supply chain management
  bus matrix for, 250–252
  operational systems in, 250–251

surrogate keys
  date/time dimension, 254–255, 330–332
  defined, 332
  dimension table assignments, 443
  fact table transformations, 450–453
  physical database design, 332–334
  pros/cons of, 243–244, 333–334
  surrogate key generator, 128, 392
  surrogate key pipeline, 128, 398–400, 450–453, 461–463

survivorship, 386–387

symmetric multiprocessing (SMP), 159–161
  pros/cons of, 160

synonyms
  defined, 331
  for user accessible tables, 331

system testing, 543–546
  regression testing, 545
  test datasets, 545–546

tables
  access monitoring tables, 344
  aggregate tables, 465–466
  auditing tables, 344
  bridge tables, 270–273
  decode tables, 130–131
  descriptions/documentation of, 559
  dimension tables, 241–246
  ETL system schematics, 434
  fact tables, 238–241, 273–282
  lookup tables, 130–131
  partitioning, 360
  schematics, developing, 434
  security tables, 344
  sizing estimates, 341
  staging tables, 330–331, 343
  truncate table, use of, 444

tangible benefits
  financial, 28–30

tape recordings
  of interviewees, 69

teams. See project team

technical architecture, 109–178
  application architecture model, diagram, 190
  architecture model, creating, 185
  architecture plan document, 187–189
  architecture task force, 183–184
  BI applications, 141–155
  blueprint for action, 226–229
  data movement, scope of, 113–115
  development context, diagram, 182
  development estimated effort, diagram, 183
  development process, 180–183
  elements of, 5, 112
  enterprise IT approaches, 180
  extract, transformation, and load (ETL) system, 119–123
  hardware/software installation, 224–225
  high level model, diagram, 114
  implementation phases, determining, 186
  implications document, 184, 186
  importance of, 110–112
  infrastructure, 156–170
  infrastructure map, 221–224
  metadata, 115–117, 170–173, 207–210
  plan, evolution of, 118–119
  presentation server architecture, 133–141
  product selection, 191–207
  security, 173–177, 210–221
  service oriented architecture (SOA), 117–118
  subsystems, design of, 187
  task list, 230–232
  technical architect, role of, 38
  trends related to, 112–113

technical metadata
  BI applications, 155
  elements of, 116
  ETL system, 116, 131–132
  presentation server architecture, 140

technical support
  specialists, role of, 38
  from vendors, 196

technology track, 5–6

templates
  BI applications, 509–512
  terrorism
  as security threat, 176

3NF modeling
  compared to dimensional modeling, 235–236

time dimension. See date/time dimension

timed extracts, 376–377

trade publications, 195
training materials, 560–561
business metadata, 155
course outline, 561
designing materials, 560–561
training database, 562
training modules, developing, 561–562
from vendors, 195–196
transaction fact tables
functions of, 273
transaction grain fact table loader, 396
transaction grain real time partition, 421–422
transaction log files, 419
transformation logic
ETL tools for, 430
Transmission Control Protocol/Internet Protocol (TCP/IP), 211–212
trapdoors, 177
Trojan horses, 177
truncate table statement, 444
trusted certificate authorities, 215

U
units of measure
fact tables, 279–280
unrestricted reports, 218
usability
ETL system, 120
pre-deployment testing, 552–553
and project documentation, 120
usage monitoring
elements of, 566–567
events to monitor, 220
functions of, 145
user(s)
ad hoc access, 475
analytical data needs, 134
atomic data needs, 133
business intelligence needs, 133–134
and business requirements definition, 64–65
communication with, 45–46
data access needs, 133
data model review, 319–320
dimensional model, presenting to, 319–320
and expansion of system, 580
and infrastructure decisions, 157–158
interactions, defining for BI applications, 525
interviewees. See business requirements interviews
interviews
training. See training materials
user documentation and metadata, 155
user requests, project manager control of, 51
user requirements. See user(s)
user support, 563–565
three-tiered approach, 564–565
user support guide, 560
user-changeable variables
BI applications, 485

V
variable depth hierarchies, 268–269
vendors
cautions about, 205
contract negotiation, 201–202
product information from, 195–196
product selection meeting, 197–198
references, checking, 198, 199
relationship factors, 196
version control
ETL system, 128, 410, 416
technical metadata, 131
version migration
ETL system, 410
vertical partitioning
presentation server architecture, 138
views
defined, 331
materialized views, 353
for user accessible tables, 331–332
virtual private networks, 215
virtual servers, 223
viruses, 177
volatility
and infrastructure decisions, 157
volumes
physical database, 362–363
See also partitioning

W
workflow monitor
ETL system, 128, 411–412
worms, 177
write-ups
for interviews, 88

X
XML files
avoiding, 379
XML schema
standard schemas, 126

Z
Zachman Framework for Enterprise Architecture, 180