Index

Abilene network 157, 159, 207, 216–17
Access Control Lists (ACLs) 248–50
access routers 80–1, 85–6
ACF *see* autocorrelation function
ACK *see* acknowledgements
acknowledgements, TCP 22
ACLs *see* Access Control Lists
ACM SIGCOMM 11
ACM SIGMETRICS 10–11
active measurement
    concepts 118–24, 252, 255–6, 412–13
active monitoring 252, 254–5
active queue management 109
addresses
    see also IP addresses
    dark aspects 388–90
    MAC 388
    registries 80, 94–7
    space hijacking 386
    translators 117
administrative issues 117, 426
Advanced Research Project Agency (ARPA) 14–15
aggregated variance method 227
algebra 7, 45–8, 67, 151
alias resolution, traceroute 120–2
aliasing, IP addresses 325–6
all–pairs paths, graphs 66–70
alpha traffic 231, 235
analytic background 45–77
    graphs 7, 64–70
    linear algebra 7, 45–8, 67
    metrics 70–3
    models 208–11
    probability 7, 46, 48–58
    statistics 7, 58–63
ANEMOS 396–7
anonymization 355–78
    see also sensitive information
    alternatives 377–8
    application–level data 362, 373–4
    attacks against 359–60, 374–6
    concepts 9, 397
    configuration data 369–70
    data categorizing 360–5
    data sharing 357–60, 365
    definitions 356–7
    motivation 357–8
    packet–level traces 366, 370–3
    processes 365–7
    router–level data 369–70
    success metrics 376–7
    techniques 367–9
anycast 250–1
Apache Web server 287, 298–9, 307
application protocol layer
    see also DNS; FTP; HTTP; P2P; SMTP
    concepts 19–24, 31, 37–42
    data gathering 103–4
    instrumentation 103
    measurement 37–42
applications 4, 6–7, 13, 19–22, 42–4, 241–352
    see also DNS; emails; file sharing; FTP; multimedia; online games; P2P; Web
    anonymizable data 362
    concepts 4, 6–7, 13, 19–22, 42–4, 242–4, 362, 421–2, 429
    dynamic characteristics 4
    emerging questions 429
    historical background 421–2
    measurement area 6–7, 42–4
    mix 242–4
    sockets 22
    types 6–7, 20–2, 42–4
approximation, independent models 219–20
architecture
    see also infrastructure; protocols

concepts 13–44, 319–20, 425–6
core simplicity 425
games 334–5, 341–3
Internet 4–5, 13–44
measurement difficulties 425–7
P2P tools 319–20
ARPA *see* Advanced Research Project Agency
ARPANET
concepts 14–16, 43, 217, 419–20
history 14–16
measurement capabilities 16
router graphs 157–8
arrival processes
packets 172–4
stochastic processes 55–6
AS graphs 412–13
ASes *see* autonomous systems
asymmetry, paths 119
Asynchronous Transfer Mode (ATM) 421
ATM *see* Asynchronous Transfer Mode
ATMEN 88, 398, 417
attacks
against anonymization 359–60, 374–6
applications 394–5
backbones 82–4
DoS 303–4, 379, 381, 383, 387–8, 391
gateways 384–6
honeypots 388–90
inter–domain attacks 386
intranet 382–3
Man–In–The–Middle 381, 386
security 380–2
wide area measurement 387–94
audio 349, 351–2
authorities *see* hubs
autocorrelation
concepts 54–8, 60–3
confusions 58
dependence structures 211–12
self–similarity 224–5
time scales 219–21
autocorrelation function (ACF), concepts 60–1
autocovariance, concepts 53–8
autonomous systems (ASes)
see also backbones; Border Gateway Protocol
anonymization 361, 364, 369, 377
CDN 295
concepts 15, 26, 28–31, 111–12, 160–1, 386, 412–13, 426
connection points (peering/exchange points) 26
graphs 111–12, 125–6, 154–61
honeypots 388–90
Internet growth 160–1
network entry points 84–5
numbers 364, 369, 377
passive measurement 124–6
routing 28–31, 95, 97, 156–7, 159–61, 412–13
security 386
static graphical properties 154–7
top–level systems 426
traceroute 122–3
availability
bandwidth 128, 131–2
DNS 246
peers 325

backbones
see also autonomous systems; Internet Service Providers
concepts 15–17, 25–6, 27, 36, 71–2
default backbone 15–16
historical background 15–17
measurement locations 80–4
networks 177, 297
routers 80–1, 153
backscatter technique 387, 391
bandwidth
bulk transfer capacity (BTC) 135–6
concepts 67–70, 83, 127–36, 159, 293, 322, 350, 396–7
ISP allocations 83
multimedia streaming 350
network/traffic interaction 169
P2P 322
packet–pair methods 129–32
router graphs 159
self–induced congestion 134–5
size–delay methods 132–4
Web performance 293
basic packet sampling 196–8
benchmarks, mobile clients 286
best–effort service, IP 21
best–landmark method 144
beta traffic 231, 235

betweenness measure, vertex 65–6
BGP *see* Border Gateway Protocol
bias, sampling 71–3, 158
BIND 98, 255–6, 262, 267
biscaling behaviour 232
BitTorrent 42, 310–12, 317, 328–30, 429
blackhats 379, 390
blackhole lists 268
blogistan 306–7
blogs 305–9, 365
Bloom filters 199–201, 215
Bluetooth 27, 110
Boolean algebra 48, 151
Border Gateway Protocol (BGP) 30–1, 35, 80, 85, 398–9, 400–2, 406–7
  anonymizable data 362–3
  Beacon System 407
  control traffic 238–9
  packet capture 190
  passive measurement 125–6
  security 386
  system instability 163–164
botnets 381
bottleneck links 128–30
Box, George 74
bprobe 415
Bro intrusion detection tool 385
broadband 352
browsers 6, 15–16, 43–4
BTC *see* Bulk Transfer Capacity
building blocks, Internet 3–4
Bulk Transfer Capacity (BTC) 70–1, 129, 135–6, 408–9
bundling 298–9
bytes 172–3, 184, 216–33

cable modem 16, 27
cached/non–cached responses, DNS 39
caching
  DNS 39, 98, 247, 266–8
  P2P 313, 319–20
  Web 272, 275, 280–1, 289
CAIDA (Cooperative Association for Internet Data Analysis) 404, 410–14
campus study, P2P 330–1
capabilities, measurement 16–17
capacity
  awareness 319
  bandwidth 128–32, 135
  BTC 129, 135–6
  planning 82
capturing data 99–102
  see also packet capturing
case studies 9, 395–417
  ANEMOS 396–7
  CAIDA 404, 409–14
  Click 397–400
  dss 398, 400–3
  High Energy Physics community 404, 407–9
  individual toolsets 397–404
  large–scale measurement projects 404–17
  low–level monitoring tools 395–7
  PlanetLab 404, 414–17
  RIPE NCC 405–7
  software tools 397–400
  Windmill 397–9
categorical datasets
  concepts 58–63
  Zipf's law 62–3
CDF *see* cumulative distribution function
CDMA cellular phone system 91–2
CDNs *see* Content Distribution Networks
cell phones 91–2, 284
cellular telephony 24–5
Central Limit Theorem, concepts 51–2, 62
central tendency, concepts 51–2, 59–60, 62
centralized games architecture 334, 341
changes
  Internet 7, 424
  technology 419
  Web 275, 301
characteristic path length, graphs 65–6
characterization
  DNS 252–6, 258–61
  games 333–4, 342–3
  P2P 311–12, 318–19, 321–5, 327–30
  Web 270–1, 278–81, 286–8
cheating, games 332, 341
chunk sizes, P2P 327, 328–30
circuit–switching networks 17–18, 24
Cisco 369–70, 393, 400–3
Click 399–402
client/server roles, endsystems 25–6, 39–40, 43–4
clients
  see also users

classification 297–9
connectivity 296–9
customer–provider relationships 156
flash crowds 304
mobile clients 240, 284–6, 292–3, 330–1, 336, 345–6
proximity assumption 263–4
representative populations 274–5
tailoring responses 297–9
Web servers 274–5, 297
clocks 90–4
clustering
coefficient 65
dimensionality reduction 204–5
graph neighbours 65–6, 69–70
network–aware clustering 283–5, 295, 297–8, 303–5, 307, 319–21
small world phenomenon 156
CNAME 262
Code Red worm 304, 391, 413
CoDeen 416–17
codes, anonymization 378
CoDNS 416
coefficient of variation, concepts 59–60
commercial issues 5–6, 15–17, 269, 276, 283–4, 289, 332, 364–5, 393–4
see also electronic commerce
anonymization 364–5
caching 289
games 332
measurement reasons 5–6
NSF networks 15
performance issues 5
security products 393–4
Web measurement services 269, 276, 283–4
Common Log Format 279
communication media 24, 27
complex networks 423, 429
components, vectors 45–8
Computer Communication Review journal 11
Computer Networks journal 11
conditional distribution, concepts 51–8
conditional expectation, concepts 51–8
conditional probability, concepts 50–8
confidence intervals 185–7
configuration
anonymization 369–70
DNS 247
P2P 312
Web 271–2, 288
congestion 113–14, 134–6, 150–1, 166–7
TCP properties 21–2
connected graphs, concepts 65–70
connection points (peering/exchange points), autonomous systems 26
connections
ASes 155
interdisciplinary connections 423–4
networks/traffic 114–15
P2P 312–13
TCP 21–2
Web 296–9
wireless networks 109–11
connectivity issues 5, 16, 22–4, 26
console games 337, 343
see also online games
constancy, traffic stability 185
constant rate sampling 196–7
constraint–based method, geolocation 144
constructive data models, concepts 73–5
Content Distribution Networks (CDNs) 42, 257–8, 263, 266, 287, 294–6, 348, 416
DNS 257–8, 263, 266
multimedia 348
Web 287–8, 294–6
content modification, Web 276, 286
continuous random variables, concepts 49–58
continuous–time stochastic processes, concepts 52–5
control traffic 176, 190, 238–9
cookies 374
core simplicity 115–16, 177, 425
Cossack system 387
count–min sketch 202
Counter–Strike 342–4
counting 172–3, 184, 193–4, 200–1, 203
counting Bloom filters 200–1
covariance, concepts 46, 50–8
crawling
P2P 318, 321–2
Web 279, 281–2, 291–2, 299–302, 306
credit card numbers 363
critical path analysis 290
cross traffic 131–2, 134–5
crossover statistics 182–3
cryptographic techniques 372

cumulative distribution function (CDF) 49–63
curse of dimensionality 188
customer–provider relationships 156

dark addresses 388–90, 393
data
    see also modeling
    anonymization 360–5
    capture 99–102
    DNS 250
    games 337–8
    gathering 103–4, 290–3
    management 191–2
    models 73–7
    P2P 315–17
    sensitivity 356
    sharing 178–9, 357–60, 365
    volume 178, 424
    Web 274–6
data reduction
    counters 193–4
    dimensionality 203–7
    flow capture 194–6
    probabilistic models 207–12
    sampling 196–9
    summarization 199–203
    traffic 192–212
databases 79–80, 94–9, 146–7, 255–6, 398, 403–4
datagrams, UDP transport protocol 19–23, 34–5, 41
datasets
    concepts 58–63, 74–7, 400–3
    modeling 74–7
DDoS *see* denial of service attacks
de–identification 357, 376
dead links 302–3
decentralization
    design principles 22–3
    games architecture 334, 341
    Internet 426
decoys 314, 316, 320
default backbone, concepts 15–16
degree, vertex 65–70
degree distributions 67–70, 154–5, 158–60
delay–based geolocation 144–6
delays 67–70, 94, 112–13, 129–36, 144–50, 153, 166–7, 282, 338–9, 343–4
    games 338–9, 343–4
    internal to networks 147–50
    multimedia 351
    one–way delays 118
    packets 94, 112–13, 129–36, 153, 166–7
    Web 270
demographic aspects
    concepts 5, 164–5, 274–5
    geographic location 164–5
    Web clients 274–5
denial–of–service attacks (DoS) 303–4, 379, 381, 383, 387–8, 391, 411, 413
dependence
    distributional models 211–12
    long range 187, 218–21, 224–6, 228–32
    short range 220, 224, 229
dependent variables, concepts 50–8
descriptions, sets of measurement 59–60
descriptive data models, concepts 73–5
descriptive statistics, concepts 58–63
design principles
    decentralization 22–3
    end–to–end argument 24–5
    inference measurement 150–2
    Internet 13, 17–25
    IP hourglass design principle 23–4
    stateless switching 24
detection methods, attacks 382, 384–6, 392–3
deterministic delays 166
deterministic sampling 196–7
DEW *see* DNS–Enhanced Web
DHT *see* distributed hash table
dial–up connectivity 5, 16
diameter, small world phenomenon 156
dig (Domain Internet Groper) 98, 252, 267
digital divide, High Energy Physics community 409
digital subscriber line (DSL) 16, 27
dimensionality, traffic statistics 188, 203–7
DirectConnect 42, 324, 327, 331
directed graphs, concepts 64–70
discrete random variables, concepts 49–58
discrete–time stochastic processes, concepts 53–5
dispersion
    see also variance
    concepts 59–63
distance 141, 237
    see also paths
distributed data gathering 103–4

distributed games architecture 341
distributed hash table (DHT) 41
distributed internetworking 177
distributional models 181, 208–11
distributions
  concepts 46, 51–8, 62, 208–11, 415–16
  skewness 62–3
  tails 56–8, 62–3, 179–83, 210–11, 225–31, 234
  types 52, 56–8
DLLs (dynamically loaded libraries) 401–3
DNS (Domain Name System) 6, 20–2, 37–9, 42, 241, 244–69, 360, 362–4, 397, 402–3, 405–7, 412–13, 416–17
  see also applications
  anonymization 360, 362–4
  cached/non–cached responses 39
  caching 39, 98, 247, 266–8
  CDN 257–8, 263, 266, 294–5
  challenges 248–51
  characterization 252–3, 258–61
  concepts 6, 20–2, 37–9, 42, 245–56, 258–69, 397, 402–3, 405–7, 412–13, 416–17
  databases 94–5, 97–8
  Internet growth 162–3
  King 139
  location databases 146
  measurement 37–9, 397, 402–3, 405–7, 412–13, 416–17
  other applications 256–8, 262–9
  properties 245–8
  resource records 98
  servers 38–9, 244–69, 294–6, 360
  tools 251–6
DNS–Enhanced Web (DEW) 265–6
dnscache 267
dnschecker 255–6
dnsstat 254
dnstop 254–5
domain squatters 267
dominant congested links 150–1
DoS *see* denial–of–service attacks
download times, Web pages 5
downloading 312–14, 316–17, 321, 327–30
drop events 113
drop–tail service discipline 109
dsc 255
DSHIELD 392
DSL *see* digital subscriber line
dss 398, 400–3
duplicate blogs 307
duplicate packets 168
duration, packet train rate 236
dynamic characteristics, Internet 4

ECHO packets 35, 118, 121, 135
ECHO REPLY message 35
edges, graphs 64–70
eDonkey 41–2, 310–11, 326–9
educational multimedia 351–2
eigenvalues 48
eigenvectors 48
eighty–twenty rule 182
electrical links 190
electronic commerce 3
  see also commercial issues
'elephants and mice' phenomenon 183, 234
  see also size–count disparity
emails 15, 17, 43, 268, 390, 394
  see also SMTP
embedded content, Web 295, 297–9
embedding–based methods, latency 137, 139–42
empire, evil 281, 326
empirical distributions, concepts 58–63
eMule 310–12, 317
encryption 366, 372
end–to–end argument 24–5, 116
end–to–end protocols
  see also application protocol layer; transport protocol layer
  concepts 20
endsystems
  client/server roles 25–6, 39–40, 43–4
  concepts 18–22, 24–7, 39–40, 161–3
enterprise networks
  see also autonomous systems
  concepts 15–16
entertainment multimedia 351–2
entities
  DNS 246, 249, 251, 259–61
  P2P 310, 314–15, 319, 324
  Web 271–2, 277–8, 288
entropy rate, concepts 54–8
entry points, networks 84–6
Erdös–Rényi random graph, concepts 67–70
Erlang, A.K. 219

errors
  inference 358
  metrics 70
estimation
  latency 136–42
  traffic matrices 212–15
Ethereal 349
Ethernet 19, 23–4, 27, 136, 172, 189, 414
Euclidean space 46–7, 140–1, 166
eValid Test Suite 283
event averages 71–3
EverQuest 331–2, 335
exchange points 26, 86–7
expectation paradox 181
expected value, concepts 49–58
exponential distribution 52, 56–8, 183
exponential growth
  Internet 163
  traffic 217–18
external time sources 91
extreme distribution, concepts 52

false edges, paths 119–20
FastTrack 323–4
fault detection 67–70
FIFO queuing 131
file sharing 5–6, 25, 40–2
file transfers 20, 23, 40, 42–4
  see also FTP
filtering
  anonymization 372–4
  AS connections 126
  Bloom filters 199–201, 215
  multistage filters 201
  packets 86
  security 394
fingerprinting 356–7
firewalls 116, 268–9, 381–2, 384–5, 392, 399, 413
First Person Shooter (FPS) games 332, 335–7, 340, 342, 344–5
first–order stationarity, stochastic processes 54–8
flash crowds 272–3, 303–5, 330
flooding protocols, routers 30, 41
flows
  see also netflow
  anonymization 375
  data reduction 194–6
  intranet security 383
  IPFIX 195–6
  packets 101–2
  RTFM 196, 254
  sampling 283
  traffic 175–7, 194–6, 214–15, 236–8
Flowscan 403
Flowtools 195–6
forward secrecy 356
forwarding decisions, routers 28
forwarding engines 108–9
fpdns 255
FPS *see* First Person Shooter games
freeriding 313, 320, 322
FTP (File Transfer Protocol) application protocol 23, 40, 42–3, 242–3
  concepts 42–3
  TCP usage 43
fused measurements 127

games *see* online games
Gamespy 340–1
gamma distribution, concepts 52
Ganglia 417
gateways
  see also Border Gateway Protocol
  concepts 14–15, 26, 30
  protocols 85
  routers 80–1, 84
  security 384–6
Gaussian distributions *see* Normal
GAV (GPL Arcade Volleyball) 336, 345
generalized random graph, concepts 67–70
generic top–level domains, DNS 253, 262
Genii 388
geographic location 164–5
geolocation 142–7, 412–13
GIF images, CND 295
Gigascope 191–2, 398, 403–4
Gilbert model 167
Global Positioning System (GPS) 91–2
GNP embedding approach 139–41
GNU Zebra 190
Gnutella 41, 310–12, 321–5
Gomez 283
goodness–of–fit tests 208–11
goodput 115
GPS *see* Global Positioning System
graphs

see also incidence matrix
ASes 111–12, 125–6, 154–7, 159–61
cliques 65
commonly encountered models 67–70
concepts 7, 64–70
interface graphs 111–12
measurement issues 66–70
netflow data 259–61
P2P traffic 319
router graphs 111–12, 157–61
small world graphs 69–70, 156
structure characterization 65–6
Web graph modeling 280, 300
'ground–zero' hosts 413–14
group–based characterization, DNS 259–61
growing network model 68–70
growth
Internet 155, 160–3
traffic 217–18
Web sites 288
GSM 27
GSQL 191–2
GTrace 411

H.323 standard 347, 351
Half–Life 340, 342–5
hardware 79, 82, 92
hashing strings 369–70
headers
concepts 20–2, 32–4, 123, 370–2, 396–7
IP (Internet Protocol) 32–3
nested headers 20–1
TCP (Transmission Control Protocol) 34
UDP 35
Web example 21
heavy–tails 57–8, 62, 179–83, 225–31, 234
heuristic notions, sampling 71
hidden factors
DNS 250–1
games 337–9
P2P 315–17
Web 274–8
hierarchical routing
AS graphs 156–7
concepts 28–31, 37–8
P2P nodes 319, 322–3
high dimensionality, time series 188
High Energy Physics (HEP) community 404, 407–9
high variability
infrastructure 154–6, 158–60
Internet measurements 61–2
statistical issues 425
traffic 179–83, 237
high–speed connectivity 5
histograms, concepts 59–63
historical background, Internet 3–4, 14–17, 419–24
hit rates, Web caches 289
honeypots 387–90
hop–by–hop protocols
see also link protocol layer; network protocol layer
concepts 20–2, 25–6
host ID information, IP addresses 21–2, 37–8
Hostcount project, RIPE NCC 405
hosts 21–2, 25, 37–8, 161–3, 259, 263, 268
Host.txt file 37–8
hourglass design principle, IP 23–4, 116, 178, 425
HTTP (HyperText Transfer Protocol) application protocol
see also Web
CDN 294–5
concepts 20–2, 39–40, 42–4, 270–1, 276, 282, 398–403, 417
crawlers 282
DNS 257, 264–6
measurement 39–40, 398–403, 417
popularity 42–3
request–response pair 39–40
streaming multimedia 346
TCP usage 40, 42–4, 257
transport layer 320–1
HTTP/1.0 6, 39–40, 43
HTTP/1.1 6, 39–40, 43
httperf 291, 298–9
hubs 280, 300–1, 327
Human Pacman 334–5, 345
Hurst parameter (H) 221–4, 227–30
hybrid games architecture 334–5, 341

IANA *see* Internet Assigned Numbers Authority
ICFA *see* International Committee for Future Accelerators
ICMP (Internet Control Message Protocol) active measurement 118–19, 121

concepts 16, 35–6
ECHO packets 118, 121, 135
measurement 35–6
TIME EXCEEDED packets 118–21, 133
identity aspects 356, 361–3, 375–6
IDMaps 139
IDS *see* intrusion detection systems
IEEE *see* Institute of Electrical and Electronics Engineers
IEEE/ACM Transactions on Networking journal 11
IEPM–BW project, High Energy Physics community 409
IETF *see* Internet Engineering Task Force
ifEntry 193–4
Iffinder 411
IGPs 239
see also IS–IS; OSPF
ill–posed problems, traffic 212–13
IMC *see* Internet Measurement Conference
IMPs *see* Interface Message Processors
incidence matrix
see also graphs
concepts 65–6
indegree, vertex 65–6
independent models 213–14, 218–21
independent variables, concepts 50–8
indexers 281
inference 147–52
anonymization 358, 360
attacks 360
errors 358
searching 302–3
statistical 423
traffic tools 212–15
Infocom conference 10–11
information resources 10–11
information sharing 392–3, 426–8
infrared 25
infrastructure 3–4, 6–10, 17–22, 25–6, 107–69
see also architecture; links; routers
challenges 115–17
changes 102–3
concepts 3–4, 6–10, 17–22, 25–6, 107–69
measurement area 6–7
properties 107–15
tools 117–52
instability, routing 238–9
instantaneous RTT 137–8
see also round–trip time
Institute of Electrical and Electronics Engineers (IEEE) 110
instrumentation changes 102–3
inter–AS (domain) routing, concepts 28–31, 35
interarrival process, stochastic processes 55–6
interception proxies 277–8
interdisciplinary connections 423–4
interface graphs 111–12
Interface Message Processors (IMPs) 420
interfaces
concepts 21–2
traceroute 120–1
International Committee for Future Accelerators (ICFA) 407–9
Internet
see also applications; infrastructure; traffic
architecture 4–5, 13–44
building blocks 3–4
builders 3–4
changes 7
commercial issues 5–6, 15–17
concepts 3–10, 14–17
databases 94–9
definition 3, 15
design principles 13, 17–25
directories 94–9
dynamic characteristics 4
growth 155, 160–3
historical background 3–4, 14–17, 419–24
measurement concepts 3–10, 16–17
operation details 25–30
optimization issues 413
organization 14–22, 25–30
practical issues 79–104
security issues 4, 9, 380–2, 397, 411–12
uses 3, 16
versatility 3
Internet Assigned Numbers Authority (IANA) 95–6
Internet Engineering Task Force (IETF) 17, 36, 42
Internet Exchange Points 86–7
Internet Measurement Conference (IMC) 10
Internet Motion Sensor 393

Internet Official Protocol Standards 17
Internet Routing Registry (IRR) 95–7
Internet Service Providers (ISPs)
address registries 95–6
administrative barriers 117
anonymization 361
concepts 15–17, 26, 30–1, 117, 361, 414, 422–3
historical background 15–16
IDS 384
information sharing 426
measurement locations 80–3
network management 422–3
permissions 23
policy routing 30
routing 30–1
spam 394
top–level ISPs 17
types 16, 26
Internet Sink (iSink) project 393, 399–400
Internet2 414
interpretable parameters, modeling 75–6
intra–AS (domain) routing, concepts 28–31, 35
intranets 382–3
intrusion detection systems (IDS) 382, 384–6, 392–3
invariants, traffic 215–16
inverse matrices 46–7
InXs *see* Internet Exchange Points
IP addresses
aliasing 325–6
anonymization 361–3, 367, 369, 370–7
concepts 21–2, 26, 28, 37–9, 325–6, 361–3, 367, 369, 370–7, 411–13
DNS 37–9, 259–61, 263
network–aware clustering 283–4, 295, 308
P2P 319, 325–6
sockets 22
IP Flow Information Export effort (IPFIX) 195–6
IP Identifications (IP IDs) 121
IP (Internet Protocol)
see also network protocol layer
anonymization 371–2
concepts 14–15, 19–24, 26, 28, 31–3, 37–9, 325–6, 361–3, 367, 369–77, 398–9, 405–7, 411–13
firewalls 399
headers 32–3, 123, 371–2
hourglass design principle 23–4, 116, 178, 425
IP addresses 21–2, 26, 28, 37–9, 325–6, 361–3, 367, 369–77, 411–13
IPv4 32, 407, 412, 428–9
IPv6 32, 123, 407, 428
measurement 31–3, 37, 398–9, 405–7, 411–12
Record Route option 33
segments 22
traffic flows 175, 195–6
TTL values 32–3, 118
unreliability 21, 24–5
IP multicast, concepts 37
iperf 409
IPFIX *see* IP Flow Information Export effort
IRR *see* Internet Routing Registry
IS–IS intra–domain routing protocol 29, 35
ISDN 27
ISPs *see* Internet Service Providers

jitter
games 340, 343–4
multimedia 348, 350–2
packets 114, 168
video 348, 350
joint entropy, concepts 54–8
joint probability, concepts 50–8
journals 10–11

k–anonymity 376–7
Kademlia 309
KaZaa 6, 41–2, 310–11, 313, 316, 326, 429
kernel 396–7
Keynote 283
King 139, 263–4, 344

lag, concepts 54–8
lame delegations, DNS 256
landmarks
geolocation 144–6
latency 137, 139, 141
window model 203
LANS *see* local area networks
large ISPs 16
see also Internet Service Providers
large observations, statistics 179–82

large–scale measurement projects, case studies 404–17
'last mile' providers 16
latency 136–42, 412–13
  DNS 247, 250
  embedding–based methods 137, 139–42
  estimation 136–42
  games 335–6, 343–4
  King 263
  latency radius 340
  mirror sites 295–6
  proxy–based methods 137–9
  sensitivity 83, 86
  user–perceived latency 293, 295–7, 299
  Web 270, 272
layers
  hidden layers 116, 251, 276–7, 316–17, 338–9
  overview 18–22
  protocols 18–22, 31–42
leechers, BitTorrent 328–30
legal aspects
  BitTorrent 330
  data sharing 359
libpcap 188–9, 254, 371
  see also packet capturing
Limewire 42
linear algebra
  concepts 7, 45–8, 67
  notation list 46
link protocol layer
  see also Ethernet; MPLS; Sonet; WiFi
  concepts 18–22, 23–6, 29
link state advertisements (LSAs) 30–1, 126, 153, 403
  see also routing
link state protocol 29–30
  see also routers
link weights, paths 30
link/structure preserving transformation 368
links
  see also infrastructure; routers
  backbones 84
  bandwidth 129–32, 135–6
  common technologies 25–6, 27
  concepts 15, 18–26, 27, 108
  data sharing risks 360
  dead links 302–3
  diagrammatic overview 18
  dominant congested links 150–1
  inference 151
  packet capture 189–90
  speeds 25–6, 27
  traffic matrices 212
  Web 280–1
Linux Operating System 328, 329
Lipschitz embedding 141–2
live multimedia streaming 326, 350–2
LLCD *see* log–log complementary distribution
local area networks (LANs) 14–15, 21, 24, 27, 82, 110, 189
local clocks, routers 35
local data gathering 103–4
local scaling 232
locality, flows 237
location
  databases 146–7
  DNS servers 247
  geolocation 142–7
  measurements 80–9
  peers 325
  Web functions 271, 280, 288
log–log complementary distribution (LLCD) 209–10, 226
logging 102, 252, 394
lognormal distribution, concepts 52
logscale diagram method 227–8
long tails 56–8, 62–3, 179–83, 231
long time scales 216–18
long–range dependence 187, 218–21, 224–6, 228–32, 425
lookups 248, 266
loss
  see also packet loss
  inference 147–50
lossless transformation 367–8
lossy transformation 368
low–level monitoring tools, case studies 395–7
lower–level protocol data 99–100
LSAs *see* link state advertisements

MAC addresses 370, 388
malicious traffic 176, 388, 390–2
malware 390–2
Man–In–The–Middle attacks 381, 386
management, middleboxes 116
management information base (MIB) 36, 193–4

maps, routers 29–30, 399–400
MAPS (Mail Abuse Prevention System) project 268
marginal distribution, traffic 230–1
Markov models 167
matrices 46–7
matrix–vector multiplication 47–8
maximum likelihood methods 149, 215
MDS *see* multidimensional scaling
means
  concepts 46, 49–58, 61–2, 179–80
  sampling 179–80
measurement
  see also metrics; passive measurement
  absence 3–4
  active measurement 118–24, 252, 255–6
  application protocol layer 37–42
  areas 6–7, 16–17
  ARPANET 16
  building platforms 426–7
  capabilities 16–17
  case studies 9, 395–419
  collection difficulties 4
  concepts 3–10, 16–17, 23–4, 31–42, 73–6, 80–9, 118–24, 252, 255–6, 395–419
  descriptions 59–60
  difficulties 3–4, 424–6
  DNS 37–9, 397, 402–3, 405–7, 412–13, 416–17
  emerging questions 428–30
  games 332–6
  HTTP 39–40, 398–403, 416–17
  ICMP 35–6
  information sharing 427–8
  Internet concepts 3–10
  IP 31–3, 37, 398–9, 404–7, 411–12
  large–scale measurement projects 404–17
  load 122, 141
  location 80–9
  low–level monitoring tools 395–7
  metrics 70–3
  modeling 73–7
  network protocol layer 31–3, 37
  operational measurement 420–1
  P2P 40–2, 397
  practical issues 79–104
  probability issues 52–8, 76–7
  progress 4–5
  protocol layers 31–42, 99–104
  reasons for measuring 5–6
  routing 35, 397–400
  sampling uses 16–17, 70–3
  security 379–95
  SNMP 36
  statistical issues 58–63
  stochastic processes 52–5, 76–7
  TCP 31, 33–4, 71, 398–9, 415–17
  tools 4–5, 7, 9, 16–17, 395–417
  traffic 421–3
  transport protocol layer 31, 33–4
  trends 419–24
  UDP transport protocol 31, 34–5, 198–9, 244–5, 259–60, 413
memory 60–1, 185–7
Mercator crawler 291, 301
messages 153, 238–9
metric–induced topology 150–1
metrics
  see also measurement; sampling
  concepts 70–3, 150–1
  errors 70
  event averages 71–3
  instability 179–81
  success 376–7
  time averages 71–3
  uncertainty 70–1
MIB *see* management information base
Microsoft Media Server (MMS) 346–8
mid–size ISPs 16
  see also Internet Service Providers
middleboxes 116–17, 153
MINC inference method 148
minimum RTT (minRTT) 137–9, 166
  see also round–trip time
mirror sites 86–7, 293, 295–6, 301
mirrored server games architecture 334
mmdump, multimedia tool 349
MMORPG (Massive Multiplayer Online Role Playing Games) 332–4
MMS *see* Microsoft Media Server
mobile clients
  games 336, 345–6
  P2P 330–1
  traffic 240
  Web 284–6, 292–3
Mockapetris, Paul 37

modeling
concepts 73–7
distribution models 181, 208–11
graph modeling 280, 300
independent models 213–14, 218–21
multifractal processes 233
ON/OFF models 234–5
over–fitting problems 75
parameters 74–7
parsimonious models 74–5
probability models 76–7, 180–1, 207–12, 230–1
queuing models 180
self–similarity 227–8, 230
statistical difficulties 181
types 73–4
uses 74–7
validation 74–5
'what if' questions 76, 403
modular querying mechanisms 397–404
Mohonk project 389–90
monitoring 252, 254–5, 283, 392–3
monofractal processes 232
Morris worm 391
Mosaic 16
MPLS, link protocol layer 23
MRT format files 400–3
multi–session games 336
multi–site measurements 87–8
multicast 37, 123, 149–50, 152, 327
multidimensional scaling (MDS) 141
multidimensional time series 188
multifractal processes 232–3
multilateration 144–5
multimedia 44, 326, 346–52
see also applications
categories 351–2
challenges 348
properties 347–8
streaming 326, 346–52
tools 349
multiplayer games 331–4, 339–40, 342, 345
multipoint communication 321
multistage filters 201

name–based geolocation 143
NAPs *see* network access points
Napster 6, 40–1, 244, 309–11, 321–3, 326–7
see also P2P
narrow links 129–32
National Institute of Standards and Technology (NIST), US 91
National Science Foundation (NSF) 15
NATS *see* network address translators
neighbours, graphs 64–70
nested headers, concepts 20–1
netflow 84–85, 101, 400, 402–403
anonymization 361, 366, 370, 377
DNS 253–4, 259–61
P2P 318, 323
NETFLOW method, *dss* 402
NeTraMet (Network Traffic Flow Measurement Tool) 196, 254, 262, 411
nettimer 325
network access points (NAPs) 81, 85
network address translators (NATs) 117
network effects, games 343–4
network engineering 171, 184
network ID information, IP addresses 21–2, 30–1
network latency 137, 141
Network News Transport Protocol (NNTP) 243
network protocol layer
see also IP
concepts 19–24, 31–3, 37
measurement 31–3, 37
network proximity assumption 262–3, 295
network telescopes 388, 412–14
Network Time Protocol (NTP) 93–4
network tomography 66–70, 147–52, 213
network–aware clustering 283–5, 295, 297–8, 303–5, 307, 319–21
network–defined flows 175–6
network–internal delays 147–50
network–level support, P2P 320–1
network–oriented traffic measurement 422
networked games 337–40
networks
see also individual networks; network; SNMP
backbones 177, 297
complex networks 423, 429
concepts 21–2, 25–6, 28–33, 40–2, 147–52, 213, 320–1, 337–44, 422
core simplicity 115–16
entry points 84–6
exchange points 86–7

LANs 82, 110, 189
management 422–3
traffic 112–15, 165–9
wide area networks 87–9
ninety–ten rule 182
NIST *see* National Institute of Standards and Technology
NNTP *see* Network News Transport Protocol
nodes 64–70, 131–2, 155, 160, 214–15, 260–1, 316, 319–20, 324–6
ARPANET 420
availability 325–6
DNS 260–1
fanouts 214–15
infrastructure 131–2, 155, 160
P2P 316, 319–20, 324–6
noise 216–17
norm, vectors 47
Normal (Gaussian) distributions, concepts 46, 51–8, 62, 73–4, 230–1
NSF *see* National Science Foundation
NSFNET, concepts 15–16
NTP *see* Network Time Protocol
numerical datasets, concepts 58–63

objectivity 209
observability 115–16, 177–8
Occam's razor 75
OD *see* origin–destination flow
OFF times 234–5
off–the–shelf hardware 395–7, 403–4
offline processes 366, 376–7
offset, clocks 90–1, 93–4
ON/OFF processes 173–4, 234–5
one–way packet delays 94, 118
online games 6, 20–2, 44, 331–45
see also applications
anonymization 362–5
attacks 382
challenges 337–40
characterization 333–4, 342–3
commercial value 332
concepts 6, 20–2, 44, 331–45
future aspects 429
measurement 332–6
networked games 337–40
tools 339–40
types 332–3
online processes, anonymization 366–7, 376
operation details, Internet 25–30
operational measurement 420–1
optical fibers 24, 25, 27
optimization issues, Internet 413
optimizations, measurement load 122
organization, Internet 14–22, 25–30
organization specific information 363–4
origin servers, HTTP requests 40
origin–destination flow (OD) 66, 175–6
OSPF (Open Shortest Path First) intra–domain routing protocol 29, 31, 35, 126–7, 153, 403
outdegree, vertex 65–6
over–fitting problems, modeling 75
overhead, Web servers 276
overlays 80, 104
Overnet 41, 235, 309
overview 6–10
ownership 313–14

P2P (peer to peer) 6–7, 20–2, 40–3, 309–31, 397
see also application; Napster
anonymizable data 362
applications mix 244
BitTorrent 328–30
challenges 314–17
characterization 311–12, 318–19, 321–5, 327–30
concepts 6–7, 20–2, 40–3, 314–31, 397
eDonkey 327–9
games 341–2
legal considerations 40
measurement 40–2, 397
other applications 325–7
overlays 104
peer location 325–6
properties 310–14
protocols 40–2
tools 317–21
Web traffic study 292–3
wireless networks 330–1
packet capturing
see also libpcap; passive measurement
low–level monitoring tools 395–7
traffic 177–8, 188–90
packet delays 94, 112–13, 118, 129–36, 153, 166–7
packet drops 113

packet flows 101–2, 195–6
packet jitter 114, 168
packet loss
  games 335–6, 343
  infrastructure 113–14, 123, 167–8
  multimedia 348, 351
packet pairs 129–32, 134, 150
packet reordering 168
packet size distributions 233
packet switching 420
  concepts 14–22, 24, 26, 28
  diagrammatic overview 18
  stateless switching 24
packet tracing
  anonymization 366, 370–3
  backbones 82–3
  data capture 101–3
  data gathering 103
  DNS 252–3
  Web 276–8, 292, 301
packet trains
  autocorrelation 225
  flow capture 194–5
  heavy tails 225–7
  ON/OFF processes 174
  self–similarity 225–7
  sessions 174–5
  traffic structure 234–6
packets 172–3
  see also packet; traffic
  concepts 6, 14–22, 26, 28, 395–400
  duplication 168
  ECHO 118, 121, 135
  filtering 86
  headers 20–2, 32–4, 396–7
  paths 28–31
  routers 108–9
  segments 22
  TIME EXCEEDED 118–21, 133
  time scales 216–33
  window size 22
pages, Web 280–2, 300–3, 305–9
Palm calculus 73
parallel processes, anonymization 366–7
parameters, modeling 74–7
Pareto distribution, concepts 52, 57–8
Passive and Active Measurement workshop (PAM) 10
passive measurement 124–7, 188–90, 252–4, 256, 322–4, 411–14
  see also packet capturing
  BGP 124–6
  DNS 252–4, 256
  games 340
  OSPF 126–7
  P2P 322–4
  traffic 188–90
  Web 290, 293, 297
path vectors 30–1
pathchar 415
paths
  bandwidth 67–70, 133, 396–7
  concepts 28–31, 65–70
  graphs 65–70
  jitter 339
  length 237
  link weights 30
  matrices 176
  router graphs 160
  traceroute 119–20
patterns
  anonymization 365–6
  attack 359–60
PCA *see* principal component analysis
PCs
  endsystems 18, 25–6
  PC–based clocks 92
PDAs 25
PDF *see* probability density function
Peakflow 393
peer sets, BitTorrent 328–9
peering 26, 81, 84–5, 126, 156–7
peering/exchange points, autonomous systems 26
peers, concepts 25
percentiles, concepts 59–60
perception
  games players 338–9
  multimedia users 348
performance issues
  commercial issues 5
  DNS 256, 261–2
  measurement reasons 5–6, 36
  middleboxes 116
  traffic 171, 184
  Web 273, 276, 282–3, 293–9
permission, data sharing 358–9

personal–sensitive data 361, 363
personalized crawlers 291
physical devices, infrastructure 107–11
piggybacking 257, 264–6, 285
ping 118, 162, 325, 339, 415
PingER project, High Energy Physics community 408
pipelining, Web 43–4
PlanetLab 404, 414–17
PlanetSeer 416
players, games 338–9, 342
points of presence (PoPs) 82–3, 111–12, 237
Poisson distribution 56, 68, 72–3, 219
policy routing, ISPs 30
polling techniques 100
PoPs *see* points of presence
popularity
  DNS 266–8
  games 331–6, 341–2, 345
  P2P 312
  Web 273, 279–81, 290, 303, 376
ports
  concepts 22, 189
  mirroring 189
Postel, Jon 95
power laws 69–70, 155–6, 300–1
predictable variation, traffic 217–18
preferential attachment graph, concepts 67–70
prefixes, IP addresses 21–2, 371–2, 375, 377, 411–13
principal component analysis (PCA) 205–7
privacy 143, 179
  see also anonymization
PRO–COW compliance study 299
probabilistic counting 203
probability
  see also stochastic processes
  concepts 7, 46, 48–58, 76–7, 180–1, 207–12, 230–1
  measurement issues 52–8, 76–7
  notation list 46
probability density function (PDF), concepts 46, 49–58
probability models 76–7, 180–1, 207–12, 230–1
  concepts 76–7
  real data 75–7
  uses 76–7
probing 122–3, 131–2, 133
propagation delay 113, 132–3, 166
properties
  DNS 245–8
  games 332–7
  infrastructure 107–15
  measurement areas 6–7
  multimedia 348
  P2P 310–15
  traffic 172–6
  Web 270–3, 286–9
protocols 3–4, 6, 11, 18–22, 28–42, 80, 93–4
  see also architecture; individual protocols
  anonymizable data 362
  concepts 13–15, 18–22, 28–42
  gateways 85
  headers 20–2, 32–3
  hidden layers 116, 250–1, 276–7, 316–17, 339–40
  layers 18–22, 31–42
  measurement 31–42
  measurement across layers 99–104
  overview 18–22
  routing 28–31, 80
  standards 299
  time information 93–4
  types 14–15, 18–22, 28–42, 93–4
provisioning 83
  see also bandwidth
proxies
  anonymization 374
  games 338, 341
  interception proxies 277–8
  middleboxes 116
  mobile clients 285
  network–aware clustering 283–4
  P2P 313, 318
  Web 102, 277–8, 284–7, 374
proximity assumption 262–3, 295
proxy–based latency estimation 137–9

QQ *see* quantile–quantile plots
qstat 340
Quagga 190, 407
Quake games 342–5
quality, games 339–40
quantile–quantile (QQ) plots 209–10
quantiles, concepts 59–60, 209–10
queries

DNS 248–9, 254–6, 259, 261–5, 267
P2P 309, 310–12, 315–17, 323–4
queuing
delay 112, 129, 132, 153, 167–8, 186–7
FIFO queuing 131
management 108
models 180
self–similarity 229

RADbs *see* Routing Assets Databases
random additive sampling 72–3
random graphs, concepts 67–70
random sampling 196–7
random variables, concepts 49–58, 61–2
rates
clocks 90
packet trains 235–6
sampling 196–8
re–identification 357, 375–6
real data, probability models 75–7
Real Time Flow Metering (RTFM) 196
real–time requirements, games 335–6
Real–time Streaming Protocol (RTSP) 347–8, 350
Real–time Transport Protocol (RTP) 347
Realtime Traffic Flow Measurement (RTFM) 253–4
RealTracer tool 349–50
Record Route option 33, 121
records 98, 146
redirector devices 289
references, blogs 308–9
regional networks, concepts 15–16
registries 79–80, 94–7
regularization methods 213–15
relationships, ASes 156–7
relative variability, concepts 51–8
remote data gathering 103–4
reordering packets 168
representativeness
client populations 274–5
sampling 71–2
wide area networks 88–9
request counts 304–5
request–response pair, HTTP 39–40
Requests for Comments (RFCs) 10, 17, 31, 39–40
research 379, 421, 426–30
residual capacity 128
resolution, clock 90–1
resources
ownership 313–14
P2P 313–14, 315–16, 320
records 98, 146
segmentation 320
responses 264–5, 296–9, 346–7
reverse blacklisting 390
reverse engineering 349, 357, 359, 366, 375
reverse proxies 288
RFCs *see* Requests for Comments
'rich get richer' phenomenon 155, 302
RIPE NCC 405–7
risks, data sharing 358–60
Riverhead 393
RMX (reverse MX) 268
Rocketfuel project 369
root server study, DNS 261
root servers 261, 412–14
round–trip time (RTT) 118, 133, 137–9, 166–7, 232, 396–7, 415–16
latency estimation 137–8
packet delay 166–7
ping 118
size–delay methods 133
very short time scales 232
routers 6, 13, 15, 18–22, 24, 25–30, 71–2, 108–9, 152–3, 397–400
see also infrastructure; links
anonymization 369–70
BGP 30–1, 35, 398–402, 407
concepts 15, 18–22, 24, 25–30, 108–9, 152–3, 397–400
configuration 369–70
core simplicity 177
decision–making 26, 28, 29–31
diagrammatic overview 18
flooding protocols 30, 41
forwarding decisions 28
graphs 111–12, 157–60, 412–13
Internet growth 161–3
local clocks 35
maps 29–31, 399–400
measurement locations 80–1, 84–6
peering 80–1, 84–5
prime function 25
sampling uses 71–2
sizes 25
stateless switching 24

RouteViews project 124–6, 407
routing
  ASes 95, 97, 156–7, 159–61
  concepts 26, 28–31, 35, 397–400, 412–13
  control traffic 238–9
  delays 112, 166–7
  hierarchical routing 28–31, 37–8
  instability 163–164, 238–9
  levels 28–31
  loops 29, 33, 85
  measurement 35, 397–400
  overlays 104
  protocols 28–31, 35, 80
  registries 95–7
  tables 124
Routing Assets Databases (RADbs) 96
Routing Information Service (RIS) project, RIPE NCC 405, 407
routing matrix, graphs 66–70
RTCP (RTP Control Protocol) 347
RTFM *see* Real Time Flow Metering
RTP *see* Real–time Transport Protocol
rtpmon multimedia tool 349
RTSP *see* Real–time Streaming Protocol
RTT *see* round–trip time

S Statistical language 397
S–Net 397
sampled router graphs 412–13
sampling 16–17, 58, 70–3, 179–80, 196–9, 424–5
  see also metrics
  bias 71–3, 157–8
  concepts 16–17, 58, 71–3, 179–80, 196–9, 424–5
  data reduction 196–9
  event averages 72–3
  flow data 283
  intranet security 383
  packet traces/flows 101–2
  random additive sampling 72–3
  representative measurements 71–2
  sample means 179–80
  time averages 72–3
  unbiased 71–3
  uses 16–17, 71–3
satellite communications 24, 27
scalars 45–8
scale–free graphs 69
scaling issues 45–8, 69, 218–32, 335, 342, 419
SCIC *see* Standing Committee on International Connectivity
scriptroute 415
searching
  anonymization 363, 375
  search engines 281–2, 326
  Web 281–2, 300–3
second–order property, autocovariance 53–8
second–order stationarity, stochastic processes 53–8
security issues 4, 9, 110, 116, 246, 268–9, 379–95, 397, 411–14
  applications 394
  DNS 246
  DoS attacks 411, 413–14
  firewalls 116, 268–9, 381–2, 384–5, 392
  gateways 384–6
  inter–domain attacks 386
  Internet 380–2
  intranets 382–3
  measurement 379–95
  network telescopes 412, 413–14
  wireless LANs 110
  worms 411, 413
seeds, BitTorrent 328
segments, concepts 22
selection, peers 325
self–induced congestion 134–5
self–references 309
self–similarity 221–30, 232, 421
  asymptotic self–similarity 224–5, 230
semi–log plots 162–3
semi–lossy transformation 368
sensitive information 9
  see also anonymization
sensitivity, data 356, 361, 363
sensor networks 428–9
sequential inferencing 360
servers
  CDN 293–294
  DNS 38–9, 244–69, 294–6, 360
  eDonkey 327
  endsystems 18, 25–6, 39–40
  games 337–40, 343
  HTTP requests 39–40
  performance issues 6
  Web 274–7, 279, 288, 294–9, 360

service level agreements (SLAs) 86
service providers
see also Internet Service Providers
permissions 23
sessions
games 343
packet trains 174–5
traffic structure 236
settings, DNS parameters 248–9
sharing data 357–60
shingling 301
short tails, distributions 56–8, 62–3
short time scales 218–33
heavy–tails 225–30
long–range dependence 218–21, 224–6, 228–31
self–similarity 221–30
very short time scales 232–3
short–range dependence 220, 224, 229
signalling 216–17, 323
similarity
see also self–similarity
clustering 204
SimMud proposal 341
single–session games 336
size–count disparity 182–3
see also 'elephants and mice' phenomenon
size–delay methods 132–4, 136
sketches 201–2, 204
skew, clocks 90–1, 94
skewed distributions, concepts 62–3
Skitter project 411–13
Slammer worm 391, 413
SLAs *see* service level agreements
sliding window model 203
smacq 191
small world graphs, concepts 69–70, 156
SMTP (Simple Mail Transfer Protocol)
application protocol 20, 40, 43, 243
see also emails
SNMP (Simple Network Management Protocol) 16–17, 36
backbones 82
concepts 16–17, 36
data capture 100
historical background 420
instrumentation 102
intranet security 383
measurement 36
MIB 36, 193–4
spam 394
traffic counters 193–4
versions 36
Snort intrusion detection tool 385
sockets, TCP 22
software
case studies 397–404
clocks 92
DNS 247
malware 390–2
mobile clients 284
packet tracing 102
requirements 79
viruses 379, 381–3, 390–2
Web 290–2
worms 381–2, 388, 390–3
Sonet, link protocol layer 23
space–code Bloom filters 201, 215
spam 43, 268, 390, 394
spatial distribution, wireless traffic 240
special purpose systems, packet capture 189–90
speed
anonymization 377
link capacity 25–7, 136
packet capture 190
packet train rate 235–6
traffic 421
worms 391–3
spider programs 43, 279, 284–5
SQL 404
SRI *see* Stanford Research Institute
stability 61, 163–4, 183–5, 238–9, 425
standard deviation, concepts 50–63, 75
Standing Committee on International Connectivity (SCIC) 408–9
Stanford Research Institute (SRI) 14
statelessness 24, 115
static properties, AS graphs 154–60
stationarity condition 53–8, 183–5, 224–5
see also wide–sense stationarity
statistics
see also graphs; probability; sampling; tails; variability
autocorrelation 185–7, 211–12, 219–21, 224–5

concepts 7, 58–63, 179–88, 231, 423, 425, 429
difficulties 179–88, 425
dimensionality 188, 203–7
inference 423
measurement difficulties 425
measurement issues 58–63, 425
memory 185–7
stability 183–5
stationarity 183–5, 224–5
statistical physics 423, 429
traffic 179–88
Steam 340–1
stochastic delays 166–7
stochastic processes, concepts 52–8, 60, 76–7, 166–7, 217–18
stochastic variation, traffic 217–18
stored multimedia streaming 350–1
storms of instability 238–9
stratified sampling 196–7
stratum levels, synchronized clocks 93
streaming multimedia 20–2, 326, 346–52
STREAMS system 191
strictly–stationary condition, stochastic processes 53–8
strings 363, 369–70, 373, 375
strongly connected graphs, concepts 65–70
students, P2P 312, 331
stupid network 115
subexponential distribution, concepts 56–8, 62
subgraphs, concepts 65–70
success metrics 376–7
summarization 199–203
supercomputers 25
supernodes, P2P 41–2
surveys 349
switching
see also packet switching
ATM 421
Ethernet 189
infrastructure 115
Web 278
SYN flags 34, 411
synchronized time 93–4
synthetic traffic 228, 344–5
systems 73–4, 123–4, 185–7, 191

tails 56–8, 62–3, 179–83, 210–11, 225–31, 234
heavy tails 57–8, 62, 179–83, 225–31, 234
long tails 56–8, 62–3, 179–83, 231
TCP (Transmission Control Protocol)
see also protocols; transport protocol layer
acknowledgements 22
BTC 129, 135
concepts 1–5, 19–24, 31, 33–4, 38, 40–2, 71, 398–9, 415, 417
congestion 21–2
connections 21–2, 114–15, 312–13
distributional models 209–11
DNS 245, 254, 257, 259–61, 265
FTP 43
games 335, 338
header contents 34
HTTP 40, 42–4, 257
measurement 31, 33–4, 71, 398–9, 415, 417
P2P 312–13
packet reordering 168
packet trains 235–6
piggybacking 265
popularity 42
reliability function 21–2
segments 22
self–similarity 228
sockets 22
SYN flags 34
tcpdpriv 371–2, 375
tcpdump 254–5, 327–8, 339, 349, 416
Web 276–8
TCP/IP
see also protocols
concepts 14–15
historical background 14–15
tcpdpriv 371–2, 375
tcpdump 254–5, 327–8, 339, 349, 416
DNS 254–5
eDonkey 328
games 339
multimedia 349
technical factors 4–6, 16
telephone traffic analysis 219–20
telescopes, networks 388
Test Traffic Measurement (TTM) project, RIPE NCC 406–7
think time 234–5, 314, 339

three–way handshake 34
thresholds, interarrivals 174
throughput 114, 169
    see also bandwidth
tight links 129, 132
time averages 71–3
TIME EXCEEDED message 33, 35, 118–21, 133
time information 89–94
    background 90–1
    measurement 79, 89–94
    sources 91–2
    synchronized time 93–4
    terminology 90
time scales 216–33
    long time scales 216–18
    scaling behaviour 218–32
    short time scales 218–33
    traffic 216–33, 422
    very short time scales 232–3
time series 113, 172–3, 188
    concepts 55–6, 60–2
    memory 60–1
    stability 61
time series of counts, stochastic processes 55–6
Time Stamp Counter (TSC) register 92, 94
'time to glass' 273, 293
timers, segments 22
TIMESTAMP message 35
TIMESTAMP REPLY message 35
TiVo 373–4, 376
TLDs *see* top–level domains
tomography, networks 66–70, 147–52, 213
tools 117–52, 188–215
    see also individual tools
    active measurement 118–24
    anonymization 365
    bandwidth measurement 127–36
    case studies 9, 395–417
    data management 191–2
    data reduction 192–212
    DNS 251–6
    fused measurements 127
    games 339–40
    geolocation 142–7
    IDS 385
    inference 147–52, 212–15
    infrastructure 117–52
    large–scale measurement projects 404–17
    latency 136–42
    low–level monitoring tools 395–7
    measurement 4–5, 7, 9, 16–17, 118–42, 395–417
    multimedia 349
    P2P 317–21
    packet capture 188–90
    passive measurement 124–7
    security 385
    traffic 188–215
    Web 278–86
top–level domains (TLDs) 38, 253, 405–7
top–level ISPs 17
    see also Internet Service Providers
topology 111–12, 154–65
    AS graphs 154–7
    dynamic aspects 160–4
    geographic location 164–5
    inference methods 151–2
    router graphs 157–61
traceroute 16–17, 71, 410, 412, 416
    DNS 263–4
    historical background 420–1
    infrastructure 112, 118–23
    sampling bias 158
    Web 280
tracing *see* packet tracing
traffic 3–4, 6–7, 16–17, 67–70, 171–240, 333–4, 342–3, 395–417, 421–3
    backbones 83
    bytes 216–33
    challenges 176–88
    concepts 3–4, 6–7, 16–17, 67–70, 171–240, 395–417, 421–3
    data reduction 192–212
    DNS 245–6, 254, 258–9
    flows 254
    games 333–4, 342–3
    growth 217–18
    intranet security 382–3
    matrices 176, 212–15, 236–7
    networks 112–15, 165–9
    P2P 319
    packets 216–33
    properties 172–6
    RTFM 254
    self–similarity 221–30

shapers 116
statistical difficulties 179–88
synthetic traffic 228, 344–5
time scales 216–33
tools 188–215
volume 184
Web 270–3, 287, 289–93
trajectory sampling 198–9
transactions
DNS 249, 261
Web 270, 290
transfer size, packet trains 236
transformation techniques 367–9
transmission delay 113, 129, 132, 166
transport layer support, P2P 320–1
transport protocol layer
see also TCP; UDP
concepts 19–24, 31, 33–4, 40
measurement 31, 33–4
transposed matrices 46–7, 67–70
trees 149, 151–2, 327
trends 215–18, 419–24
tries 202
trust 365, 393
TSC *see* Time Stamp Counter register
TTL (time–to–live) field
DNS 247, 249–50, 262, 296–7
IP (Internet Protocol) 32–3, 118
traceroute 118, 121–2
tuples 376–7

UDP transport protocol
see also protocols; transport protocol layer
concepts 19–23, 31, 34–5, 38, 41–2, 119, 198–9, 244–5, 259–60, 413
header contents 35
measurement 31, 34–5, 198–9, 244–5, 259–60, 413
popularity 42
Slammer worm 413
unreliability 34–5
unbiased sampling 71–3
uncertainty, metrics 70–1
uncongested bandwidth 128
undirected graphs, concepts 64–70
unicast–based inference methods 149
Uniform Resource Identifier (URI) 39–40
University of California, Los Angeles (UCLA) 14
University of California, San Diego (UCSD) 412
University of California, Santa Barbara (UCSB) 14
UNIX 397–8, 400–3
Unreal Tournament 342, 344
unstable paths 119–20
unweighted graphs, concepts 64–70
uploading 313, 328–30
URI *see* Uniform Resource Identifier
URLs 38
user–perceived latency 293, 295–7, 299
users
see also clients
cancellation rates 293
mobility 240
multimedia 348–9
populations 89, 165
surveys 349
wireless access 239–40
workload models 271–2, 288–9
UTC 36

validation, modeling 74–5
variability 50–8, 61–2, 154–6, 158–60, 179–83, 237, 425
high variability 61–2, 154–6, 158–60, 179–83, 237, 425
random variables 50–8, 61–2
relative variability 51–8
variance
see also dispersion
concepts 46, 50–63
confusions 58
vectors 45–8
verification issues, anonymization 373
vertices
see also nodes
graphs 64–70
very short time scales, traffic 232–3
video 348–50, 351–2
viruses 379, 382–3, 390–2
volume
data 424
traffic 184
vulnerability testing 385

weak stationarity, concepts 53–4
Web 5, 6–7, 15–16, 20–2, 25, 38–40, 43–4, 243–4, 269–308, 397
see also applications; HTTP
anonymization 361–2, 374
applications 299–308
challenges 273–8
characterization 270–1, 278–81, 286–8
client/server roles 25, 39–40, 43–4
concepts 5, 6–7, 15–16, 20–2, 25, 38–40, 43–4, 270–308, 397
crawling 279, 281–2, 291–2, 299–302, 306
download times 5
heavy–tails 182
historical background 15–16, 43–4, 421–2
nested headers 20–1
pages 280–2, 300–3, 305–9
performance 273, 276, 283–4, 293–9
pipelining 43–4
ports 22
properties 270–3, 286–9
proxies 102, 278–9, 284–8, 374
research 429
search engines 326
servers 274–7, 279, 288, 294–9, 360
static structure 300
tools 278–86
URLs 38
Web sites
anonymization 363, 376
content/access dynamics 289
flash crowds 303–5
growth 288
latency 293
mirror sites 293, 296, 301
monitoring 283
popularity 273, 280–1, 289, 376
references 308
World Cup 1998 288
weblogs *see* blogs
Weibull distribution, concepts 52
weighted graphs, concepts 64–70
well–known ports 22
wget 291
'what if' questions, modeling 76, 403
whitehats 379, 390
whois database 146–7
wide area networks 87–9
Wide project 372
wide–area security measurement 386–94
wide–sense stationarity
see also stationarity
concepts 53–4
WiFi 19, 23–4, 27, 136
WiMAX 110
Windmill 397–9
window size, concepts 22
wireless networks 5, 25–6, 109–11, 239–40, 287, 330–1, 336, 345–6, 428
connectivity 109–11
games 336, 345–6
measurement 428
P2P 330–1
traffic 239–40
Web devices 287
Witty worm 414
workload models 271–2, 288
workshop papers 10
World Cup 1998 Web site 288
World Wide Web *see* Web
worms 304, 382, 388, 390–3, 411, 413–14

Zebra 407
zero skew, clocks 90
zing 118
Zipf's law 62–3, 289
zombies 260, 380–1