

INDEX

Symbols

3Com, 343

16-way RAID-0 setup,
414-415

8-way NUMA RAID setup,
412-414

A

Above Idle tool (Performance
Inspector), 102, 104-105

accept_redirects parameter
(kernel), 517

accept_source_route
parameter (kernel), 517

access bits, 203

ACEs (access control entries),
file systems, 236

ACLs (access control lists), 236

active balancing, 191

address spaces, 21

kernel address space, 21

Moore's Law, 200

user address space, 21

virtual address spaces,
196-197

area structures, 198-199

highmem interface, 201

kernel address

spaces, 199

kernel image

section, 200

kernel module
section, 199

paging, 201-203

swapping, 201-203

user address spaces,
197-198

virtual memory generic
view, 196

address translation process,
204-205

addressing, I/O devices,
229-230

affinity, IRQ, 462-464

AFP file system, 347

AFS/DFS, 347

agebuffer parameter
(bdflush), 249

AIM7 benchmarks, 142

AIM9 benchmarks, 143, 147

aio-max-nr parameter
(kernel), 514

aio-nr parameter (kernel), 514

algorithms, code tuning,
305-328

allocation, pages, 207

allocation groups

file systems, 236

XFS (Next-Generation File
System), 269

analysis methodologies

(benchmarks), 175-176

exit strategies, 179

measurements, 177-179

run rules, 176

target setting, 177

tuning, 177-179

anticipatory I/O elevators, 9

anticipatory I/O scheduler
(AS), 222-225

decayed frequency
tables, 224

lowest logical block
information, 226

positioning time, 224

read anticipation process,
227-228

tuning, 224-225

antic_expire parameter
(AS I/O scheduler), 225

anycast_delay parameter
(kernel), 516

Apache web servers, tuning,
338-339

APIs (application programming
interfaces), 21

NAPI (network device
driver API), 464-466

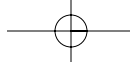
SendFile, 458-460

applets (Java), 376

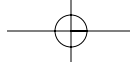
application benchmarks,
133, 161

database benchmarks,
164-166

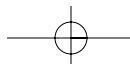
Java benchmarks, 161-163

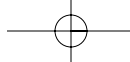


- Oracle Applications
 - Standard Application Benchmark, 168
 - PostMark, 164
 - SAP Standard Application Benchmark, 168
 - application programming interfaces (APIs). *See* APIs (application programming interface)
 - application servers, 375-376, 384-385
 - availability, 378
 - clustering, 391-394
 - commercial workload
 - J2EE model, 480-481
 - Trade3 benchmark, 481-504
 - tuning, 479-480
 - concurrency, 379
 - connection pools, 502-503
 - data integrity, 378
 - database servers,
 - interactivity, 496-499
 - databases, support, 378
 - distribution, 379
 - enterprise application, 383
 - failover, 378
 - fault tolerance, 378
 - file systems, 390
 - hardware, 387-388
 - heterogeneousness, 379
 - hyperthreading, 504
 - interactivity, 379
 - J2EE (Java 2 Enterprise Edition), 376-377
 - Java, 376
 - JVM (Java Virtual Machine), 386-389
 - load balancing, 397
 - memory usage, 389
 - multiple-applications-in-one, 380
 - multiple-services-in-one, 380
 - native code, 386
 - networking, 389
 - pass-by-reference, 502
 - performance
 - requirements, 377-390
 - requests, 380
 - scalability, 378
 - security, 378
 - SMP scaling, 390-391
 - software performance
 - stack, 382-383
 - string manipulation, 390
 - synchronization, 389
 - thread pools, 501-502
 - threads, 389
 - topologies, 394-398
 - transactions, 380
 - tuning, 398-402
 - vendor server
 - portability, 378
 - vendors, Linux support, 380
 - web servers, 390
 - applications
 - designing, 305-309
 - clients, 323-324
 - code, 309
 - file I/O, 322-323
 - servers, 309-310
 - sockets, 311, 313-314
 - synchronization, 318-321
 - threads, 315-317
 - timing routines, 310-311
 - profiling, 302-303
 - app_solicit parameter (kernel), 516
 - architecture models, web servers, 336-337
 - architectures
 - databases, 362
 - kernel, 20-23
 - multiprocessing
 - clusters, 38, 43-45
 - NUMA (Non-Uniform Memory Architecture), 38, 48-54
 - processor mixing, 38
 - SMP (Symmetrical Multiprocessing), 36-39
 - area structures, virtual memory, 198-199
 - AS (anticipatory I/O scheduler), 222-225
 - deadline behavior, 416-417
 - decayed frequency tables, 224
 - lowest logical block information, 226
 - positioning time, 224
 - read anticipation process, 227-228
 - sequential read
 - performance, 415-416
 - tuning, 224-225
 - asymmetrical multiprocessing systems, 28
 - asynchronous I/O, 246
 - databases, 371-372
 - file systems, 249-250
 - availability, application servers, 378
- ## B
- badblocks utility (Ext2/Ext3 file system), 258
 - banks, physical memory, 22
 - Baseline column (NetBench), 469
 - base_reachable_time parameter (kernel), 516
 - bdflush kernel daemon, 512
 - parameters, 248-249
 - BEA WebLogic, application servers, 377
 - benchmarks, 131, 173
 - analysis, 177-179



- application benchmarks, 133, 161
- database benchmarks, 164-166
- Java benchmarks, 161-163
- Oracle Applications Standard Benchmark, 168
- PostMark benchmarks, 164
- SAP Standard Application Benchmark, 168
- consistency, 132
- enterprise benchmarks, 133
- exit strategies, 179
- FFSB (flexible file system benchmark), 409
- file servers, 408
- I/O schedulers, 408-409
- Linux kernel performance benchmarks, 179-184
- mail servers, 409
- measurements, 177-179
- metadata, 409
- microbenchmarks, 133-134
 - disk benchmarks, 150-159
 - network benchmarks, 159-160
 - operating system benchmarks, 134-143, 147-150
- network performance, 456-457
- NetBench, 468-470
- Netperf3 benchmark, 471-472
- SPECWeb99 benchmark, 475-476
- VolcanoMark benchmark, 473-474
- operating system
 - benchmarks, 173
 - analysis methodology, 175-179
 - kernel performance benchmarks, 179-184
 - methodology case study, 175
 - numerical analysis, 174
 - simulators, 174
 - tracing, 173
 - workload characterization, 174
- repeatability, 132
- run rules, 176
- sequential read benchmarks, 409
- sequential write benchmarks, 409
- targets, setting, 177
- tuning, 177-179
- web server benchmarks, 166
 - ECPerf, 168
 - SPECjAppServer, 168
 - SPECweb, 166-168
 - SPECweb SSL, 166-168
 - TCP-W, 166-168
- web servers, 408
- workloads, 132, 134
- Berkeley sockets, 311, 313-314
- BI (business intelligence), 164
- binding, IRQ, 462-464
- BIO (block I/O) layer I/O schedulers, 217-219
 - AS (anticipatory I/O scheduler), 222-225
 - CFQ (Completely Fair Queuing) I/O scheduler, 225
 - deadline I/O scheduler, 219-220
 - noop I/O scheduler, 226
 - performance issues, 226
 - tuning, 221-222
- struct page pointers, 219
- blades, 48
- block allocation
- block groups, Ext2 file system, 253
- block size I/O, databases, 372
- blocks
 - allocation
 - Ext2 file system, 254
 - file systems, 235
 - increasing, 238
 - super blocks, 236
- Bonnie/Bonnie++ benchmarks, 150-152
- bonuses, interactive tasks, giving to, 194
- bootp_relay parameter (kernel), 517
- bottlenecks, system tracing, 87-88
 - OProfile command, 93-101
 - PI (Performance Inspector), 102-107, 110-112, 125-130
 - requirements, 88
 - strace command, 90-93
 - top utility, 88-90
- bottom half handlers, 25
- bounce buffers, databases, 370
- browsers, 331
- BSD (Berkeley Software Design), 12-16
 - initialization table
 - Slackware, 14-15
 - system initialization, 12-16
 - System V, compared, 12-16
- buffers, read/write buffers, TCP, 281-283
- burst transfer rates, data movement, 243
- business intelligence (BI), 164
- byte range requests, 333



**C**

C programming language,
301, 309

C++ programming language,
301, 309

caches, 39-40

- coherency, SMP
(Symmetrical
Multiprocessing)
system, 29
- locality of reference, 40

case sensitivity enforced
column (NetBench), 470

case studies

- benchmarking
methodology, 175
- commercial workload,
Trade3 benchmark,
481-504
- Ext2/Ext3 file system
layout analysis, 424-430
- file systems

 - I/O measurement,
446-452
 - tuning, 436-445

- I/O schedulers, 407

 - 16-way RAID-0 setup,
414-415
 - 8-way NUMA RAID
setup, 412-414
 - AS scheduler, 415-417
 - benchmarks, 408-409
 - CFQ I/O scheduler,
417-420
 - performance, 409
 - single-CPU single-disk
setup, 410-411

- JFS (Journaled File
System) layout
analysis, 430-431
- network performance,
455, 468

 - benchmarks, 456-457
 - kernel enhancements,
457-468

- NetBench, 468-470
- Netperf3 benchmark,
471-472
- SPECWeb99 benchmark,
475-476
- VolcanoMark bench-
mark, 473-474
- ReiserFS layout analysis,
432, 434
- XFS layout analysis, 435

CFQ (Completely Fair
Queuing) I/O
scheduler, 225

AS scheduler, compared,
415-417

benchmarks, 408

- 8-way NUMA RAID setup,
benchmark, 412-414
- performance, 409
- single-CPU single-disk
setup, benchmark,
410-411
- 16-way RAID-0 setup,
benchmark, 414-415
- tuning, 417-420

CGI (Common Gateway
Interface), 376

channel cards, 47

chattr utility (Ext2/Ext3 file
system), 258

children, sleep averages,
tuning, 192

CHILD_PENALTY parameter
(scheduler), tuning, 192

chunked encoding, 333

client server model, 331-333

clients

- designing, 323-324
- tuning, 350-354

 - file change notifications,
353
 - metadata, 353
 - opportunistic locking,
352-353

- protocol layering,
351-352
- read-only clients,
353-354

cluster file systems, 348

clustering

- application server, 391-394
- interconnect technology,
databases, 362

clusters, 38, 43

- HA (high availability)
clusters, 43-45
- HPCs (high-performance
clusters), 43

Coda, 347

code

- compilers, 326-327

 - as tuning tools, 303-304
 - gcc (GNU Compiler
Collection)
compiler, 301, 309

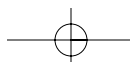
- native code, application
servers, 386
- programs, designing, 309
- tuning, 301-304

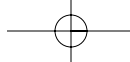
 - algorithms, 305-328

Comer, Douglas E., 505

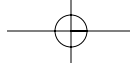
commands

- debugfs, 424
- dump2efs, 425
- gprof, 303
- gtop, 65-67
- hdparm, 243-246
- iostat, 75-75, 446-450, 452
- ipconfig, 82
- ipcs, 289-291
- jfs_mkfs, 266
- less, 26
- logrotate, 11
- mk fs, 234-237
- mke2fs, 255-256
- mkfs, 424
- netstat, 80-83
- nfsstat, 83

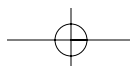


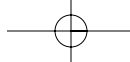


- OProfile
 - opcontrol, 94-95
 - oprofpp, 96
 - op_merge, 101
 - op_time, 97
 - op_to_source, 98
 - system tracing, 93-96, 98-99, 101
 - PI (Performance Inspector)
 - picntr, 129
 - post, 129
 - ptt, 130
 - pttstats, 130
 - swtrace, 128-129
 - prof, 302
 - ps, 71-72
 - pttstats, 105, 107
 - run.itrace, 110-111
 - sar, 67-69, 76, 78
 - sh, 339
 - strace, 90-93
 - sysctl, 278, 299
 - top, 65-67, 88, 90
 - tprof, 107, 110, 112
 - tune, 234
 - vmstat, 62-65, 72-73
 - commercial workload
 - J2EE model, 480-481
 - Trade3 benchmark, 481-492, 496-499
 - connection pools, 502-503
 - database I/O activity, 495-496
 - hyperthreading, 504
 - JVM (Java Virtual Machine), 499-501
 - pass-by-reference, 502
 - thread dumps, 492-495
 - thread pools, 501-502
 - web servers, 491
 - tuning, 479-480
 - Common Gateway Interface (CGI), 376
 - compilers
 - as tuning tools, 303-304
 - gcc (GNU Compiler Collection) compiler, 301, 309
 - compile_et utility (Ext2/Ext3 file system), 258
 - compiling code, 326-327
 - Completely Fair Queuing (CFQ) I/O scheduler, 225
 - component benchmarks. *See* microbenchmarks
 - concurrency
 - application servers, 379
 - web servers, 336
 - CONFIG kernel
 - parameters, 517
 - configuration
 - Ext3 file system, 256
 - hard drives, multiple hard drives, 4-5
 - IPC (interprocess communication), 298-299
 - JFS (Journaled File System), 266
 - kernel, configurable features, 9
 - RAID (Redundant Array of Inexpensive Disks), 7-8
 - ReiserFS, 260
 - XFS (Next-Generation File System), 269
 - connection management, TCP, kernel parameters, 285-286
 - connection pools, application servers, 502-503
 - contention statistics, generating, JLM (Java Lock Monitor), 125-128
 - context switching, 24-25
 - conversions, file systems, 6-7
 - counters
 - JLM (Java Lock Monitor), 125-128
 - OProfile, 94
 - CPU scheduler, virtual memory, 212-214
 - CPUs (central processing units), utilization, monitoring, 59-69
 - create benchmarks, 409
 - crediting interactive tasks, scheduler, 192
 - CRO (Continuous Reliable Operation), IBM zSeries mainframes, 47
- D**
- D-SACK, TCP, enabling, 285
 - data integrity
 - application servers, 378
 - assuring, 247
 - data movement
 - burst transfer rates, 243
 - disks, 241-242
 - internal transfer rates, 243
 - sustained transfer rates, 243
 - data=writeback column (NetBench), 469
 - database benchmarks, 164-166
 - results, 182
 - database servers, 361
 - application servers, interactivity, 496-499
 - databases
 - application servers, support for, 378
 - architectures, 362
 - clustering interconnect technology, 362
 - decision support system (DSS) workloads, 362



- I/O management, 370
 - asynchronous I/O, 371-372
 - block size I/O, 372
 - bounce buffers, 370
 - direct I/O, 372
 - raw I/O, 371
 - request locks, 373
 - vectored I/O, 371
- logging devices, 367-368
- memory management, 369-370
- OLTP (online transaction processing), 362
- process management, 368-369
- tuning, 362
 - global memory, 365-366
 - I/O, 363
 - load balancing, 363, 365
 - queue length, 363
 - response times, 363
 - workloads, 367-368
- dbench, 155-159
- dcache read copy update column (NetBench), 470
- deadline behavior, AS scheduler, 416-417
- deadline I/O elevators, 9
- deadline I/O schedulers, 219-220
- deadlocking, semaphores, 27
- debugfs utility, 259, 424
- debugging
 - Ext2 file system, 259
 - kernel parameters, 511
 - ReiserFS, 264
- decayed frequency tables, 224
- decision support (DS), 164
- decision support system (DSS) workloads, databases, 362
- dedicated network storage servers, 344
- default threads, virtual memory, 195
- defragmentation, file systems, 246
- delay_first_probe_time parameter (kernel), 516
- dentry-state parameter (kernel), 514
- designing programs, 305-309
 - clients, 323-324
 - code, 309
 - file I/O, 322-323
 - servers, 309-310
 - sockets, 311-314
 - synchronization, 318-321
 - threads, 315-317
 - timing routines, 310-311
- detailed symbol mode reports (OProfile), 96
- device special files, 32
- devices, I/O devices, addressing, 229-230
- dir-notify-enable parameter (kernel), 514
- direct I/O, databases, 372
- directories, 236
 - Ext2 file system, 253-254
 - JFS (Journaled File System), 265
- dirty bits, 203
- dirty_background_ratio parameter (virtual memory), 209
- dirty_expire_centisecs parameter (virtual memory), 209
- dirty_ratio parameter (virtual memory), 209
- dirty_writeback_sentisecs parameter (virtual memory), 209
- disk benchmarks, 150
 - Bonnie/Bonnie++ benchmarks, 150-152
 - dbench, 155-159
 - IOMeter benchmarks, 152
 - IOzone benchmarks, 152
 - Penguinmeter benchmarks, 153
 - tiobench, 153-155
- disks
 - file systems, 241
 - data movement, 241-242
 - hdparm command, 243-246
 - rotational speed, 242-243
 - track-to-track seek speed, 242-243
 - transfer rates, 243
 - multiple disks, configuring, 4-5
 - partitions, placement, 4-5
 - RAID (Redundant Array of Inexpensive Disks), configuring, 7-8
- distribution, application servers, 379
- DMZ topology, application servers, 395
- DNS servers, 35
- domainname parameter (kernel), 511
- dquot-max parameter (kernel), 514
- dquot-nr parameter (kernel), 514
- DS (decision support), 164
- dumpe2fs utility (Ext2/Ext3 file system), 259, 425
- dynamic content
 - SPECWeb99 benchmark, 475-476
 - web servers, 331
- dynamic kernel patches, PI (Performance Inspector), 103
- dynamic link libraries, 327-328



**E**

e2fsck utility (Ext2 file system), 259

e2image utility (Ext2 file system), 259

e2label utility (Ext2 file system), 259

EAs (extended attributes), file systems, 236

ecn parameter (kernel), 519

ECPerf benchmarks, 168

EJB (Enterprise Java Beans), 376

elevators (I/O), 9

enterprise applications

- clustering, 391-394
- J2EE, 383
- SMP scaling, 390-391
- topologies, 394-398
- tuning, 398-402

enterprise benchmarks, 133

Enterprise Java Beans (EJB), 376

enterprise servers, 42

Enterprise Storage Server (ESS), 48

error_burst parameter (kernel), 517

error_cost parameter (kernel), 517

ESCON channel cards, 47

ESS (Enterprise Storage Server), 48

etc/inittab initialization table, 12-16

Ethereal packet capture tool, 334-335

evaluation methodologies (performance), 173

- analysis methodology, 175-179
- benchmarking methodology case study, 175

kernel performance

- benchmarks, 179-184
- numerical analysis, 174
- simulators, 174
- tracing, 173
- workload characterization, 174

evenly affined IRQs column (NetBench), 469

event balancing, 191

event-driven web servers, tuning, 339

exit strategies, benchmarks, 179

EXIT_WEIGHT parameter (scheduler), tuning, 192

Ext2 file system, 5, 30, 253

- block allocation, 254
- block groups, 253
- creating, 255-256
- directories, 253-254
- Ext3 extensions, 256
- inodes, 253
- layout, analyzing, 424-430
- relationships, 254
- upgrading, 258
- utilities, 258-259

Ext3 file system, converting to, 6-7, 258

Ext3 file system, 5

Ext2 file system, converting from, 6-7, 258

extensions, 256

external journal devices, using on, 258

kernel configuration support, 256

layout, analyzing, 424-430

partitions, creating, 257

tuning, 256, 436-440

utilities, 258-259

extended attributes (EAs), file systems, 236

extensions, Ext3 file system, 256

extents

- allocation, file systems, 235
- file offset, 235
- file systems, 235
- JFS (Journaled File System), 265
- starting block numbers, 235

external fragmentation, file systems, 235

external journal devices, Ext3 file systems, using on, 258

external logs, 423

- JFS (Journaled File System), 267
- tuning, 442-443
- ReiserFS, 262
- tuning, 440-441
- XFS (Next-Generation File System), 271

F

FACK (Forward Acknowledgement), TCP, enabling, 285

failover, application servers, 378

fault tolerance, application servers, 378

fdatasync function, 247

Feigenbaum, Barry, 346

FFSB (flexible file system benchmark), 409

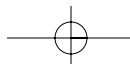
FHS (Filesystem Hierarchy Standard), 13

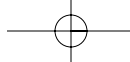
fibre channel, 41

FICON channel cards, 47

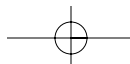
FIFOs, 289

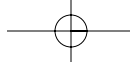
fifo_batch parameter (I/O scheduler), 221



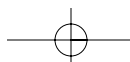


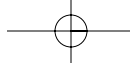
- file change notifications,
 - clients, 353
- file I/O, 322-323
- file integrity, assuring, 247
- file offset, 235
- file servers, 343
 - benchmarks, 408
 - NAS (network attached storage), 344
 - network file system protocols, 345-350
 - tuning, 354-355
 - load measurement, 358
 - NFS (Network File System), 355-356
 - performance measurement, 357
 - Samba, 356-357
- file status flags, setting, 247
- file synchronization, 246-248
- file systems, 29, 233
 - ACEs (access control entries), 236
 - ACLs (access control lists), 236
 - allocation groups, 236
 - application servers, 390
 - asynchronous I/O, 249-250
 - bdflush kernel daemon, parameters, 248-249
 - block allocation, 235
 - converting, 6-7
 - creating
 - mk fs command, 234-237
 - planning, 234
 - defragmentation, 246
 - device special files, 32
 - directories, 236
 - disks, 241
 - data movement, 241-242
 - hdparm command, 243-246
 - rotational speed, 242-243
 - track-to-track seek speed, 242-243
 - transfer rates, 243
- EAs (extended attributes), 236
- Ext2 file system, 5, 30, 253
 - block allocation, 254
 - block groups, 253
 - creating, 255-256
 - directories, 253-254
 - Ext3 extensions, 256
 - inodes, 253
 - layout analysis, 424-430
 - relationships, 254
 - upgrading, 258
 - utilities, 258-259
- Ext3 file system, 5
 - external journal devices, 258
 - kernel configuration support, 256
 - partitioning, 257
 - tuning, 256, 436-440
 - utilities, 258-259
- extents, 235
- external logs, 423
- FFSB (flexible file system benchmark), 409
- FHS (Filesystem Hierarchy Standard), 13
- file synchronization, 246-248
- fragmentation, 235, 246
- I/O, measuring, 446-452
- IBM JFS (Journaled File System), 5
 - implementation, 5, 233-241
 - inodes, 235
 - JFS (Journaled File System), 265-266
 - comparison list, 274
 - creating, 266-267
- directories, 265
- dynamic inode allocation, 265
- extent-based addressing structures, 265
- external logs, 267
- kernel configuration support, 266
- layout analysis, 430-431
- mounting, 240, 267
- transactions, 238-239
- tuning, 266, 268, 442-443
- utilities, 268
- kernel parameters, 513-514
- layouts, 423
- logical blocks, 234
- logical volumes, 235
- LVM (Linux Volume Manager), 30
- meta data, 235
- mk fs command, 234
- network file system protocols, 345-350
- proc file system, 21, 508-509
 - memory utilization, 70-71
- RAID, 30-31
- raw disk I/O, 251-252
- ReiserFS, 5, 259-260
 - creating, 261-262
 - external logs, 262
 - kernel configuration support, 260
 - layout analysis, 432-434
 - mounting, 262-263
 - tuning, 260, 264, 440-441
 - utilities, 264
- selecting, 5
- super blocks, 236
- sysfs, 509
- tuning, 234, 436



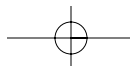


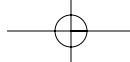
- VFS (virtual file system), 30
- VGs (volume groups), 31-32
- XFS (Next-Generation File System), 5, 268-269
 - allocation groups, 269
 - creating, 270-271
 - external logs, 271
 - fullness, 271
 - kernel configuration support, 269
 - layout analysis, 435
 - mounting, 271-272
 - partitions, 269
 - tuning, 270-272, 443-445
 - utilities, 272-273
- file-max parameter (kernel), 513
- file-nr parameter (kernel), 513
- file sizes, increasing, 238
- Filesystem Hierarchy Standard (FHS). *See* FHS (Filesystem Hierarchy Standard)
- Flash web servers, tuning, 339
- flexible file system benchmark (FFSB), 409
- forced file synchronization, 248
- FORTTRAN programming language, 309
- Forward Acknowledgement (FACK), TCP, enabling, 285
- forwarding parameter (kernel), 517
- fragmentation, file systems, 235, 246
- front_merges parameter (I/O scheduler), 221
- fsck.ext3 utility (Ext3 file system), 259
- functions
 - asynchronous I/O, 250
 - file synchronization, 247
- G**
- gcc (GNU Compiler Collection) compiler, 301, 309
 - tuning options, 303-304
- gc_stale_time parameter (kernel), 516
- GFS, 348
- Gigabit Ethernet, tuning, Netperf3, 471-472
- global memory, databases, 365-366
- GNU C Library (GLIBC) asynchronous I/O interface, 372
- GPL (GNU General Public License), Linux kernel, 19
- gprof command, 303
- grain locking, processors, 29
- gtop command, CPU utilization, 65-67
- Gunther, Harvey, 505
- H**
- HA (high availability) clusters, 43-45
- hard drives. *See* disks
- hardware
 - analysis methodologies, 175
 - application servers, 387-388
 - IBM zSeries mainframes, 46
 - NUMA servers, 51
- hdparm command, drive information, viewing, 243-246
- heterogeneousness, application servers, 379
- hierarchies
 - NUMA servers, 51
 - Processes, viewing, 88, 90
- high availability (HA) clusters, 43-45
- high bandwidth interconnect, NUMA servers, 49
- high-performance clusters (HPCs), 43-44
- highmem interface, 201
- hold time accounting, JLM (Java Lock Monitor), 125-128
- home partition, placement, 4-5
- horizontal clustering, application servers, 393
- hostname parameter (kernel), 511
- hot_list_length parameter (kernel), 515
- HPCs (high-performance clusters), 43-44
- HTML generators, 376
- HTTP (HyperText Transfer Protocol), 331
 - ISO layering model, 333
 - requests, 332-333
 - byte range requests, 333
 - responses, 331-333
 - chunked encoding, 333
 - WebDAV extensions, 347
- huge TLB page support, 9
- hyperthreading application servers, 504
- I**
- I/O
 - buses, 41
 - databases, 370
 - asynchronous I/O, 371-372
 - block size I/O, 372
 - bounce buffers, 370
 - direct I/O, 372
 - raw I/O, 371
 - request locks, 373
 - tuning, 363
 - vectored I/O, 371



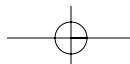


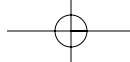
- devices, 35
 - addressing, 229-230
- elevators, 9
- file I/O, 322-323
- file systems
 - asynchronous I/O, 249-250
 - measuring, 446-452
- operations, 246
 - IBM zSeries mainframes, 47-48
 - raw disk I/O, file systems, 251-252
- request locks, databases, 373
- utilization, monitoring, 73-78
- I/O schedulers, 217-219, 407
 - AS (anticipatory I/O scheduler), 222-228, 415-416
 - deadline behavior, 416-417
 - decayed frequency tables, 224
 - lowest logical block information, 226
 - positioning time, 224
 - read anticipation process, 227-228
 - tuning, 224-225
 - benchmarks, 408-409
 - CFQ (Completely Fair Queuing) I/O scheduler, 225
 - tuning, 417-420
 - deadline I/O schedulers, 219-220
 - 8-way NUMA RAID setup, benchmark, 412-414
 - noop I/O scheduler, 226
 - performance issues, 226, 409
 - performance rankings, reshuffling, 420
 - performance of requests, 223
 - single-CPU single-disk setup, benchmark, 410-411
 - 16-way RAID-0 setup, benchmark, 414-415
 - tuning, 221-222
- I/O subsystem, 217
- I/O schedulers, 217-219
 - AS (anticipatory I/O scheduler), 222-228
 - CFQ (Completely Fair Queuing) I/O scheduler, 225
 - deadline I/O schedulers, 219-220
 - noop I/O scheduler, 226
 - performance issues, 226
 - tuning, 221-222
 - Trade3 benchmark, commercial workload, 481-504
- integration, lack of, 217
- MPIO (Multipath I/O), 41
- performance issues, 228-229
- servers, 40-41
- IBM, 343
 - application servers, 377
 - JFS (Journaled File System), 5, 239, 265
 - Netperf3, 456
 - WebSphere Application Server, 377, 481-504
 - zSeries mainframes, 45-48
- ICMP, kernel parameters, 515-516
- icmp_echo_ignore_all parameter (kernel), 516
- icmp_echo_ignore_broadcasts parameter (kernel), 516
- icmp_ignore_bogus_error_responses parameter (kernel), 516
- icmp_ratelimit parameter (kernel), 515
- icmp_ratemask parameter (kernel), 515
- IDE hard drives, 243-246
- identifiers, IPC (interprocess communication), 292
- identity mapped segment (virtual memory). See kernel image section (virtual memory), 195
- igmp_max_membership parameter (kernel), 516
- init process, initialization table, 12-16
- initialization
 - BSD, 12-16
 - init process, 12-16
 - System V, 12-16
- initialization table, 13-14
 - BSD initab, Slackware, 14-15
 - init process, 12-16
 - System V initab, Red Hat, 15-16
- inode-nr parameter (kernel), 514
- inode-state parameter (kernel), 514
- inodes
 - Ext2 file system, 253
 - file systems, 235
 - JFS (Journaled File System), 265
- installation, Linux, 3, pre-installation planning, 3-8
- instruction tracing, PI (Performance Inspector), 103, 110-111
- integration, I/O subsystems, lack of, 217
- interactive tasks, 192-194
- INTERACTIVE_DELTA parameter (scheduler), tuning, 193



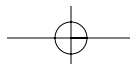


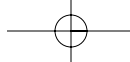
- interactivity, application servers, 379
- interconnects, NUMA servers, 50, 52
- interfaces
 - APIs (application programming interfaces), 21
 - highmem interface, 201
 - information, displaying, 82
 - kernel, 20
 - procfs, 508-509
 - sysctl, 507-508
 - sysfs, 509
- Intermezzo, 347
- internal fragmentation, file systems, 235
- internal transfer rates, data movement, 243
- Internet Explorer, 331
- interprocess communication (IPC). *See* IPC (inter-process communication)
- interrupt handlers, 25
- interrupts, monitoring
 - Above Idle tool (PI), 104-105
 - sar command, 68-69
- interval parameter (bdflush), 249
- IOmeter benchmarks, 152
- iostat command, I/O monitoring, 75-76, 446-452
- IOzone benchmarks, 152
- IP fragmentation, kernel parameters, 519
- IPC (interprocess communication), 25, 289
 - FIFOs, 289
 - identifiers, 292
 - message queues, 289
 - parameters, 295-296
 - parameters, modifying, 298-299
 - pipes, 26, 289, 299
 - semaphores, 27-28, 289, 292-294
 - shared memory segments, 289, 297-298
 - signals, 25
 - SysV (System V), 26, 289-292
- ipconfig command, 82
- ipcs command, 289-291
- ipfrag_high_thresh parameter (kernel), 520
- ipfrag_low_thresh parameter (kernel), 520
- ipfrag_time parameter (kernel), 519
- ip_autoconfig parameter (kernel), 519
- ip_default_ttl parameter (kernel), 519
- ip_dynaddr parameter (kernel), 520
- ip_forward parameter (kernel), 519
- ip_local_port_range parameter (kernel), 520
- ip_nonlocal_bind parameter (kernel), 519
- ip_no_pmtu_disc parameter (kernel), 519
- IRQ, affinity, 462-464
- ISO layering model, 333
- J**
- J2EE (Java 2 Enterprise Edition), 384-385
 - application servers, 376, 377, 384-385
 - enterprise application, 383
 - performance requirements, 377-390
 - software performance stack, 382-383
 - clustering, 391-394
 - commercial workload, 480-481
 - file systems, 390
 - hardware, 387-388
 - JVM (Java Virtual Machine), 386, 388-389
 - memory usage, 389
 - native code, 386
 - networking, 389
 - SMP scaling, 390-391
 - string manipulization, 390
 - synchronization, 389
 - threads, 389
 - topologies, 394-398
 - tuning, 398-402
 - web servers, 390
- Jamison, Wilfred C., 505
- Java
 - applets, 376
 - application servers, 376
 - benchmarks, 161-163
 - JLM (Java Lock Monitor), hold time accounting and contention statistics, 125-128
 - JPROF tool, JIT address-to-name resolution, 112
 - JVM (Java Virtual Machine), 376
 - servlets, 376
- Java Database Connections (JDBCs), 376
- Java Messaging Service (JMS), 377
- Java Transaction Architecture (JTA), 377
- JBoss application servers, 377
- JDBC (Java Database Connections), 376
- JFS (Journaled File System), 5, 237-241, 265-266 creating, 266-267



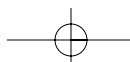


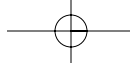
- directories, 265
 - dynamic inode
 - allocation, 265
 - extent-based addressing
 - structure, 265
 - external logs, 267
 - kernel configuration
 - support, 266
 - layout, analyzing, 430-431
 - mounting, 240, 267
 - tuning, 266-268, 442-443
 - transactions, 238-239
 - utilities, 268
 - jfs_mkfs command, 266
 - JIT (just-in-time) compilers,
 - JVM (Java Virtual Machine), 386
 - JLM (Java Lock Monitor) tool (Performance Inspector), 102, 125-128
 - JMS (Java Messaging Service), 377
 - journaling file systems
 - comparison list, 274
 - Ext2 file system, 253
 - block allocation, 254
 - creating, 255-256
 - directories, 254
 - Ext3 extensions, 256
 - inodes, 253
 - relationships, 254
 - tuning, 256
 - upgrading, 258
 - utilities, 258
 - Ext3 file system
 - external journal devices, 258
 - kernel configuration support, 256
 - partitioning, 257
 - utilities, 258
 - JFS (Journaled File System), 265-266
 - creating, 266-267
 - directories, 265
 - dynamic inode
 - allocation, 265
 - extent-based addressing
 - structures, 265
 - external logs, 267
 - kernel configuration
 - support, 266
 - layout, analyzing, 430-431
 - mounting, 267
 - tuning, 266-268, 442-443
 - utilities, 268
 - ReiserFS, 259-260
 - creating, 261-262
 - external logs, 262
 - kernel configuration support, 260
 - mounting, 262-263
 - tuning, 260, 264
 - utilities, 264
 - XFS (Next-Generation File System), 268-269
 - allocation groups, 269
 - creating, 270-271
 - external logs, 271
 - fullness, 271
 - kernel configuration support, 269
 - mounting, 271-272
 - partitions, 269
 - tuning, 270, 272
 - utilities, 272-273
 - JPROF tool (Performance Inspector), 102, 112
 - JTA (Java Transaction Architecture), 377
 - JVM (Java Virtual Machine), 376, 386
 - application servers, 388-389
 - commercial workload, 499-501
 - thread dumps, 492-495
- K**
- keep-alive management,
 - TCP, kernel parameters, 286-287
 - kernel
 - architecture, 20-23
 - configurable features, 9
 - file systems, 29-32
 - IPC (Interprocess Communication) mechanisms, 25-28
 - memory management, 21-23
 - microkernel operating systems, compared, 20
 - modular design, 35
 - modules, 20
 - monolithic aspects, 20
 - network performance, enhancements, 457-468
 - networks, parameter tuning mechanisms, 278-287
 - open source, 19
 - organization, 20
 - parameters, 509
 - CONFIG, 517
 - debugging, 511
 - file system, 513-514
 - ICMP, 515-516
 - IP fragmentation, 519
 - NEIGHBOR, 516
 - net core, 515
 - network, 515-519
 - processes, 510
 - profiling, 511
 - routing, 517
 - shared memory, 510
 - signals, 510
 - system, 511
 - TCP, 518-519
 - virtual memory, 511-513



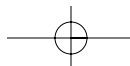


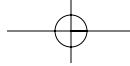
- performance benchmarks, 179-184
- proc file system, 21
 - parameter tuning, 508-509
- process management, 24
 - context switching, 24-25
 - kernel threads, 24
 - process scheduling, 24
 - processes, 24
 - tasks, 24
 - threads, 24
- responsibilities, 20
- services, 20
- SMP (Symmetrical Multiprocessing) system, 28-29
- sysfs interface, parameter tuning, 509
- sysctl interface, parameter tuning, 507-508
- system calls, 20
- variables, 277
- kernel address spaces, 21, 199
- kernel image section (virtual memory), 195, 200
- kernel module section (virtual memory), 199
- kernel threads, 24
- Krishnamurthy, Balachander, 333, 341
- kswapd parameter (kernel), 512
- L**
- lacency ratios, NUMA servers, 52
- large pages, support, 206
- layers, BIO (block I/O) layer, I/O schedulers, 217-226
- layouts
 - Ext2/Ext3 file systems, analyzing, 424-430
 - file systems, 423
 - JFS (Journaled File System), analyzing, 430-431
 - ReiserFS, analyzing, 432, 434
 - XFS, analyzing, 435
- less command, 26
- levels
 - interactive tasks, raising, 192
 - RAID, 31
- libraries, dynamic link libraries, 327-328
- Linux
 - evolution of, 19
 - installing, 3
 - pre-installation planning, 3-8
 - processor support, 19
 - UNIX, compared, 19
 - version 2.6
 - new features, 33
- Linux network stacks, 455
- list symbol mode reports (OProfile), 96
- listeners (thread), 315
- LMbench, 134-141
- load balancing
 - application servers, 397
 - databases, 363-365
 - schedulers, 191
- locality of reference, caches, 40
- locks
 - opportunistic locking, clients, 352-353
 - SMP (Symmetrical Multiprocessing) system, 28-29
- locktime parameter (kernel), 516
- log analysis tools, web servers, 340
- logger facility, 12
- logging devices, databases, 367-368
- logical blocks, file systems, 234
- logical volumes, file systems, 235
- logrotate command, 11
- logs
 - customized logging, 12
 - external logs, 423
 - finding, 11
 - logger facility, 12
 - rotating, 11
 - /var/log/messages logs, 11
 - /var/log/XFree86.0.logs, 11
 - viewing, 10
- loosely coupled multiprocessing systems, 28
- lower_zone_protection parameter (virtual memory), 210
- lowest logical block information, AS (anticipatory I/O scheduler), 226
- lsattr utility (Ext2 file system), 259
- Lustre, 348
- LVM (Linux Volume Manager), 30
- M**
- mail servers, benchmarks, 409
- management policies, schedulers, 187
- 2.6 Linux scheduler, 189-191
- load balancing, 191
- NUMA (non-uniform memory access), 188
- single processor systems, 187
- SMP (symmetric multiprocessing), 188



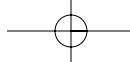


- SMT (symmetric multithreading), 189
- tuning, 192-193
- mapping
- memory, 22-23
- ObjRMAP (object-based reversed mapping), 206
- rmap (reversed mapping), 205
- max-readahead parameter (kernel), 511
- max_delay parameter (kernel), 517
- max_map_count parameter (kernel), 511
- max_size parameter (kernel), 517
- MAX_SLEEP_AVG parameter (scheduler), tuning, 193
- MAX_TIMESLICE parameter (scheduler), tuning, 193
- mcast_solicit parameter (kernel), 516
- mc_forwarding parameter (kernel), 517
- measurements, benchmarks, 177-179
- memory
 - caches, 39-40
 - locality of reference, 40
 - global memory, databases, 365-366
 - managing, databases, 369-370
 - memory management, 21-23, 283-284
 - nodes, 22
 - NUMA (non-uniform memory access), schedulers, 188
 - processes, tracing, 88, 90
 - servers, 39-40
- shared memory
 - kernel parameters, 510
 - semaphores, 27
- utilization, monitoring, 69-73
- virtual memory, 195
 - address spaces, 196-199
 - address translation process, 204-205
 - area structures, 198-199
 - CPU scheduler, 212-214
 - highmem interface, 201
 - kernel image section, 195, 200
 - kernel module section, 199
 - kernel parameters, 511-513
 - largepage support, 206
 - ObjRMAP (object-based reversed mapping), 206
 - page allocation, 207
 - page replacement, 207
 - page tables, 203, 205
 - paging, 201-203
 - processes, 195
 - rmap (reversed mapping), 205
 - slab allocators, 208
 - swapping, 201-203
 - tuning, 209-212
- memory management, 21-23, 283-284
- memory mapping, 22-23
- memory usage, application servers, 389
- message queues, 289
 - IPC (interprocess communication), parameters, 295-296
 - System V IPC, 27
- message_burst parameter (kernel), 515
- message_cost parameter (kernel), 515
- metadata
 - benchmarks, 409
 - clients, 353
 - file systems, 235
- microbenchmarks, 133-134
 - disk benchmarks, 150
 - Bonnie/Bonnie++ benchmarks, 150-152
 - dbench, 155-159
 - IOmeter benchmarks, 152
 - IOzone benchmarks, 152
 - Penguinmeter benchmarks, 153
 - tiobench, 153-155
 - network benchmarks
 - dbench, 159
 - Netperf, 159-160
 - SPEC SFS, 160
 - operating system
 - benchmarks, 134
 - AIM7, 142
 - AIM9, 143, 147
 - LMbench, 134-141
 - Reaim, 147
 - SPEC SDDET, 147-150
- microkernel operating systems, 20
- min-readahead parameter (kernel), 511
- min_delay parameter (kernel), 517
- min_free_kbytes parameter (virtual memory), 210
- MIN_TIMESLICE parameter (scheduler), tuning, 193
- mixing processors, 38
- mk fs command, file systems, creating, 234-237, 424
- mke2fs utility (Ext2 file system), 259

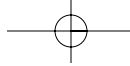




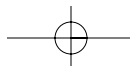
- mkesfs command,
 - Ext2/Ext3 file systems, creating, 255-256
 - mkfs.ext3 utility (Ext3 file system), 259
 - mklost+found utility (Ext2 file system), 259
 - mkreiserfs utility (ReiserFS), 264
 - mk_commands utility (Ext2/Ext3 file system), 259
 - modular design, kernel, 20, 35
 - monitoring performance, 58-59, 131
 - analysis methodology, 175-179
 - benchmarking
 - methodology case study, 175
 - CPU utilization, 59-69
 - I/O utilization, 73-78
 - kernel performance
 - benchmarks, 179-184
 - memory utilization, 69-73
 - network utilization, 78-83
 - numerical analysis, 174
 - simulators, 174
 - system tracing, 87-88
 - OProfile command, 93-101
 - PI (Performance Inspector), 102-112, 125-130
 - requirements, 88
 - strace command, 90-93
 - top utility, 88-90
 - tracing, 173
 - workload characterization, 174
 - Moore's Law, address spaces, 200
 - Morton, Andrew, 275
 - mounting
 - JFS (Journaled File System), 267
 - ReiserFS, 262-263
 - XFS (Next-Generation File System), 271-272
 - Mozilla, 331
 - MPIO (Multipath I/O), 41
 - NUMA servers, 53
 - msgmax parameter (kernel), 510
 - msgmnb parameter (kernel), 510
 - msgmni parameter (kernel), 510
 - multidomain topology, application servers, 398
 - Multipath I/O (MPIO). *See* MPIO (Multipath I/O)
 - multiprocessing
 - clusters, 38, 43-45
 - HA (high availability) clusters, 43-45
 - HPCs (high-performance clusters), 43
 - NUMA (Non-Uniform Memory Architecture), 38, 48-54
 - hardware, 51
 - hierarchy, 51
 - high bandwidth
 - interconnect, 49
 - interconnects, 50-52
 - latency ratios, 52
 - MPIO (Multipath I/O), 53
 - nodes, 50-51
 - timers, 54
 - processor mixing, 38
 - SMP (Symmetrical Multiprocessing), 36-39
 - mutual exclusion, 36
 - schedulers, 188
 - topologies, 37-38
- multiprocessor systems, 28-29
- mutual exclusion, shared system resources, 36
- ## N
- named pipes, 26, 289
 - Namesys, ReiserFS, 239
 - NAPI (network device driver API), 464-466
 - NAS (network attached storage), 344-345
 - native code, application servers, 386
 - ndirty parameter (bdflush), 249
 - NEIGHBOR kernel parameters, 516
 - net core, kernel parameters, 515
 - NetBench, 456, 468-470
 - netdev_max_backlog parameter (kernel), 280, 515
 - Netperf benchmarks, 159-160
 - Netperf3 microbenchmark, 456
 - Gigabit Ethernet tuning, 471-472
 - netstat command, network utilization, monitoring, 80-83
 - network attached storage (NAS). *See* NAS (network attached storage)
 - network benchmarks, 159-160
 - network file system protocols, 345-350
 - network initialization, Slackware, 15
 - Network Operating Systems (NOS), 343
 - network protocol stack, 277-278
 - network storage servers, 344

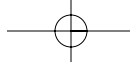


- networks
- application servers, 389
 - kernel parameters, 515-519
 - performance, 455, 468
 - benchmarks, 456-457
 - IRQ affinity, 462-464
 - kernel enhancements, 457-468
 - NAPI (network device driver API), 464-466
 - NetBench, 468-470
 - Netperf3 benchmark, 471-472
 - process affinity, 462
 - SendFile API, 458-460
 - SPECWeb99 benchmark, 475-476
 - TOE (TCP Offload Engine), 467-468
 - TSO (TCP Segmentation Offloading), 460-462
 - VolcanoMark benchmark, 473-474
 - Zerocopy, 458-459
 - tuning
 - kernel parameter
 - tuning mechanisms, 278-287
 - network protocol stack, 277-278
 - utilization, monitoring, 78-83
 - nfract parameter (bdflush), 249
 - nfract_stop_bdflush parameter (bdflush), 249
 - nfract_sync parameter (bdflush), 249
 - NFS (Network File System), 346-347
 - client/server implementations, 350
 - tuning, 355-356
 - nfsstat command, network utilization, monitoring, 83
 - NICs, Gigabit Ethernet NICs, tuning, 471-472
 - nodes
 - device nodes, 32
 - memory, 22
 - NUMA servers, 50-51
 - non-interactive tasks, starving, 194
 - Non-Uniform Memory Architecture (NUMA). *See* NUMA (Non-Uniform Memory Architecture)
 - noop I/O scheduler, 226
 - NOS (Network Operating Systems), 343
 - Novell, 343
 - nr_pdflush_threads
 - parameter (virtual memory), 210
 - NUMA (Non-Uniform Memory Architecture), 22, 38, 48-54
 - hardware, 51
 - hierarchy, 51
 - high bandwidth
 - interconnect, 49
 - interconnects, 50-52
 - latency ratios, 52
 - MPIO (Multipath I/O), 53
 - nodes, 50-51
 - schedulers, 188
 - timers, 54
 - numerical analysis, 174
- O**
- O(1) scheduler column (NetBench), 469
 - objects, reference counted objects, 197
 - ObjRMAP (object-based reversed mapping), 206
 - OLTP (online transaction processing), 164
 - databases, 362
 - logging devices, 367-368
 - opcontrol command (OProfile), 94-95
 - open source, kernel, 19
 - Open Source Development Lab (OSDL), Reaim benchmarks, 147
 - operating system
 - benchmarks, 134, 173
 - AIM7, 142
 - AIM9, 143, 147
 - analysis methodology, 175-179
 - Linux kernel performance benchmarks, 179-184
 - LMbench, 134-141
 - methodology case study, 175
 - numerical analysis, 174
 - Reaim, 147
 - simulators, 174
 - SPEC SDET, 147-150
 - tracing, 173
 - workload characterization, 174
 - operating systems. *See also* kernel
 - kernel performance benchmarks, 179-184
 - microkernel operating systems, 20
 - NOS (Network Operating Systems), 343
 - performance evaluation methodologies, 173
 - analysis methodology, 175-179
 - benchmarking
 - methodology case study, 175
 - numerical analysis, 174
 - simulators, 174

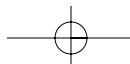


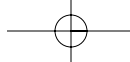
- tracing, 173
- workload characteriza-
tion, 174
- opportunistic locking, clients,
352-353
- OProfile tool
 - counter, 94
 - opcontrol, 94-95
 - oprofpp, 96
 - op_merge, 101
 - op_time, 97
 - op_to_source, 98
 - system tracing, 93-101
- optmem_max parameter
(kernel), 281, 515
- Oracle, application
servers, 377
- Oracle Applications Standard
Benchmark, 168
- Orion, application
servers, 377
- OSDL (Open Source
Development Lab),
Reiam benchmarks, 147
- osrelease parameter
(kernel), 511
- ostype parameter
(kernel), 511
- overcommit_memory
parameter (virtual
memory), 211
- overcommit_ratio parameter
(virtual memory), 211
- overhead (locking),
processors, 29
- P**
- packets, Ethernet packet
capture tool, 334-335
- PAE (Physical Address
Extension), 22
- page frame maps, 200
- page tables, 203, 205
- page-cluster parameter
(kernel), 512
- page-cluster parameter
(virtual memory), 211
- pagesets, 207
- pagetable_cache parameter
(kernel), 512
- paging
 - virtual address spaces,
201-203
 - virtual memory, 23
- parameters
 - IPC (interprocess
communication),
298-299
 - kernel, 509
 - CONFIG, 517
 - debugging, 511
 - file system, 513-514
 - ICMP, 515-516
 - IP fragmentation, 519
 - NEIGHBOR, 516
 - net core, 515
 - network, 515-519
 - network tuning, 278-287
 - processes, 510
 - profiling, 511
 - routing, 517
 - shared memory, 510
 - signals, 510
 - system, 511
 - TCP, 518-519
 - virtual memory, 511-513
 - message queues, 295-296
 - semaphores, 292-294
 - shared memory segments,
297-298
- parents, sleep averages,
tuning, 192
- PARENT_PENALTY
parameter (scheduler),
tuning, 193
- partitions
 - Ext3 partitions,
creating, 257
 - placement, 4-5
 - XFS (Next-Generation File
System), 269
- pass-by-reference, application
servers, 502
- PCI I/O buses, 41
- Penguinometer
benchmarks, 153
- per-thread time (PTT). *See*
PTT (per-thread time)
- performance
 - application servers,
requirements, 377-390
 - enterprise applications
clustering, 391-394
 - SMP scaling, 390-391
 - topologies, 394-398
 - tuning, 398-402
- evaluation methodologies,
173
 - analysis methodology,
175-179
 - benchmarking
methodology case
study, 175
 - numerical analysis, 174
 - simulators, 174
 - tracing, 173
 - workload characteriza-
tion, 174
- file systems, disks, 241-246
- I/O schedulers, 226, 409
 - 16-way RAID-0 setup,
414-415
 - 8-way NUMA RAID
setup, 412-414
 - AS scheduler, 415-417
 - CFQ I/O scheduler,
tuning, 417-420
 - single-CPU single-disk
setup, 410-411
- I/O subsystem, 228-229
- kernel performance
 - benchmarks, 179-184
- networks, 455, 468
 - benchmarks, 456-457
 - IRQ affinity, 462-464
 - kernel enhancements,
457-468



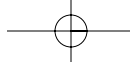


- NAPI (network device driver API), 464-466
- NetBench, 468-470
- Netperf3 benchmark, 471-472
- process affinity, 462
- SendFile API, 458-460
- SPECWeb99 benchmark, 475-476
- TOE (TCP Offload Engine), 467-468
- TSO (TCP Segmentation Offloading), 460-462
- VolcanoMark benchmark, 473-474
- Zerocopy, 458-459
- print servers, 358
- web servers, tuning, 337-340
- workloads, benchmarking, 132-143, 147-168
- performance analysis, 58-59
 - CPU utilization, 59-69
 - I/O utilization, 73-78
 - memory utilization, 69-73
 - network utilization, 78-83
 - system tracing, 87-88
 - OProfile command, 93-101
 - PI (Performance Inspector), 102-112, 125-130
 - requirements, 88
 - strace command, 90-93
 - top utility, 88-90
- performance counters, PI (Performance Inspector), 103
- Performance Inspector (PI). *See* PI (Performance Inspector)
- performance monitoring, 131
- performance tuning. *See* tuning
- Physical Address Extension (PAE), 22
- physical memory banks, 22
- PI (Performance Inspector), 102-112, 125-130
 - Above Idle tool, 102-105
 - dynamic kernel patches, 103
 - instruction tracing, 103, 110-111
 - JLM (Java Lock Monitor) tool, 102, 125-128
 - JPROF tool, 102, 112
 - performance counters, 103
 - picpnr tool, 129
 - post tool, 129
 - ptt tool, 130
 - pttstats tool, 102, 105, 107, 130
 - swtrace tool, 128-129
 - System Trace Data Collection tool, 102
 - tprof tool, 107, 110
- picpnr tool (Performance Inspector), 129
- pipes, 289
 - IPC (Interprocess Communication), 26, 299
 - named pipes, 26
- planning
 - file systems, 234
 - pre-installation planning, Linux, 3-8
- point of sale (POS) terminals, 164
- policies, resource management policies, schedulers, 187-193
- pools of requests, I/O schedulers, 223
- positioning time, processes, 224
- post tool (Performance Inspector), 129
- PostMark application benchmarks, 164
- pre-installation planning, Linux, 3-8
- print servers, 343
 - performance, 358
- printk parameter (kernel), 511
- PRIO_BONUS_RATIO parameter (scheduler), tuning, 194
- proc file system, 21
 - kernel parameters, tuning, 508-509
 - memory utilization, 70-71
 - virtual memory parameters, tuning, 209
- process affinity column (NetBench), 470
- process management, 24-25
- processes, 24-25
 - affinity, 462
 - hierarchies, viewing, 88, 90
- IPC (interprocess communication), 289
 - dynamic modification, 298-299
 - FIFOs, 289
 - identifiers, 292
 - message queues, 289, 295-296
 - pipes, 26, 289, 299
 - semaphores, 27-28, 289-294
 - shared memory segments, 289, 297-298
 - signals, 25
 - SysV (System V), 26, 289-292
- kernel parameters, 510
- managing, databases, 368-369
- positioning time, 224
- read anticipation process, AS (anticipatory I/O scheduler), 227-228





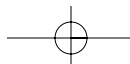
- scheduling, 24
 - timeslices, 25
 - virtual memory
 - address spaces, 196-197
 - address translation
 - process, 204-205
 - processor affinity, SMP (Symmetrical Multiprocessing) system, 29
 - processors
 - Linux, support, 19
 - memory caches, 39-40
 - monitoring, Above Idle tool (PI), 104-105
 - multiprocessing
 - clusters, 38, 43-45
 - NUMA (Non-Uniform Memory Architecture), 38, 48-54
 - processor mixing, 38
 - SMP (Symmetrical Multiprocessing), 36-39
 - registers, 39
 - schedulers
 - 2.6 Linux scheduler, 189, 191
 - load balancing, 191
 - NUMA (non-uniform memory access), 188
 - single processor systems, 187
 - SMP (Symmetrical Multiprocessing) system, 188
 - SMT (symmetric multi-threading), 189
 - tuning, 192-193
 - service processors, 35
 - SMP (Symmetrical Multiprocessing) system, 28-29, 188
 - utilization, monitoring, 59-69
 - word sizes, 39
 - procs interface, kernel parameters
 - CONFIG, 517
 - debugging, 511
 - file system, 513-514
 - ICMP, 515-516
 - IP fragmentation, 519
 - NEIGHBOR, 516
 - net core, 515
 - networks, 515-519
 - processes, 510
 - profiling, 511
 - routing, 517
 - shared memory, 510
 - signals, 510
 - system, 511
 - TCP, 518-519
 - tuning, 508-509
 - virtual memory, 511-513
 - prof command, 302
 - profiling
 - applications, 302-303
 - kernel parameters, 511
 - programming languages, choosing, 309
 - programs, designing, 305-309
 - clients, 323-324
 - code, 309
 - file I/O, 322-323
 - servers, 309-310
 - sockets, 311-314
 - synchronization, 318-321
 - threads, 315-317
 - timing routines, 310-311
 - protocol layering, clients, 351-352
 - protocol stack (network), 277-278
 - protocols
 - network file system protocols, 345-350
 - network protocol stack, 277-278
 - proxy_arp parameter (kernel), 517
 - proxy_delay parameter (kernel), 516
 - proxy_qlen parameter (kernel), 516
 - ps command, memory utilization, 71-72
 - PTT (per-thread time), 102
 - monitoring, pttstats tool (PI), 105, 107
 - ptt tool (Performance Inspector), 130
 - pttstats tool (Performance Inspector), 130
 - Pulavarti, Badari, 507
 - Python programming language, 309
- Q-R**
- queue length, databases, 363
 - RAID (Redundant Array of Inexpensive Disks), 7, 30-31
 - configuring, 7-8
 - levels, 31
 - multiple hard drives, 4
 - raw I/O
 - databases, 371
 - file systems, 251-252
 - read anticipation process, AS (anticipatory I/O scheduler), 227-228
 - read buffers, TCP, kernel parameters, 281-283
 - read-only clients, 353-354
 - read_batch_expire parameter (AS I/O scheduler), 225
 - read_expire parameter (AS I/O scheduler), 224
 - read_expire parameter (I/O scheduler), 221
 - Reaim benchmarks, 147
 - real-root-dev parameter (kernel), 511

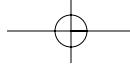


- Redundant Array of
 Inexpensive Disks
 (RAID). *See* RAID
 (Redundant Array of
 Inexpensive Disks)
 reentrancy, 310
 reference counted objects, 197
 registers, processors, 39
 ReiserFS, 5, 239, 259-260
 creating, 261-262
 external logs, 262
 kernel configuration
 support, 260
 layout, analyzing, 432-434
 mounting, 262-263
 reiserfsck utility, 264
 tuning, 260, 264, 440-441
 utilities, 264
 reiserfsck utility
 (ReiserFS), 264
 relationships, Ext2 file
 system, 254
 reliability, IBM zSeries
 mainframes, 47
 replacement policy, virtual
 address spaces, 202-203
 replacing pages, 207
 requests (web), 332-333
 application servers, 380
 availability, 378
 byte range requests, 333
 I/O request locks,
 databases, 373
 resize2fs utility (Ext2 file
 system), 259
 resize_reiserfs utility
 (ReiserFS), 264
 resource management
 policies, schedulers, 187
 2.6 Linux scheduler,
 189-191
 load balancing, 191

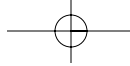
 NUMA (non-uniform
 memory access), 188
 single processor
 systems, 187
 SMP (symmetric
 multiprocessing), 188
 SMT (symmetric
 multithreading), 189
 tuning, 192-193
 response times, FFSB
 (flexible file system
 benchmark), 410
 responses (web), 331-333
 chunked encoding, 333
 databases, tuning, 363
 responsibilities, kernel, 20
 retrans_time parameter
 (kernel), 516
 Rexford, Jennifer, 333, 341
 rmap (reversed mapping), 205
 rmem_default parameter
 (kernel), 515
 rmem_max parameter
 (kernel), 280, 515
 rmen_default parameter
 (kernel), 280
 root partition, placement, 4-5
 rotating logs, 11
 rotational speed, disks,
 242-243
 routing kernel parameters, 517
 rp_filter parameter
 (kernel), 517
 rtsig-max parameter
 (kernel), 510
 rtsig-nr parameter
 (kernel), 510
 run rules, benchmarks, 176
 run.itrace command, 110-111
 runlevels, Slackware, 14
 runtime information
 collection, tracing, 173

S
 SACK (Selective
 Acknowledgements),
 TCP, enabling, 285
 Samba servers, tuning,
 356-357
 SANs (storage area
 networks), 344
 NAS (network attached
 storage), compared, 345
 SAP Standard Application
 Benchmark, 168
 sar command
 CPU utilization, 67-69
 I/O utilization,
 monitoting, 76, 78
 scalability, application
 servers, 378
 SMP scaling, 390-391
 schedulers, 187
 2.6 Linux scheduler,
 189, 191
 CPU scheduler, 212-214
 I/O schedulers,
 217-219, 407
 16-way RAID-0 setup,
 414-415
 8-way NUMA RAID
 setup, 412-414
 AS (anticipatory
 I/O scheduler),
 222-228
 AS scheduler, 415-417
 benchmarks, 408-409
 CFQ (Completely Fair
 Queuing) I/O
 scheduler, 225
 CFQ I/O scheduler,
 417-420
 deadline I/O schedulers,
 219-220
 noop I/O scheduler, 226

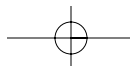


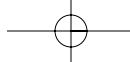


- performance, 409
- performance issues, 226
- pools of requests, 223
- single-CPU single-disk
 - setup, 410-411
 - tuning, 221-222
- load balancing, 191
- NUMA (non-uniform memory access), 188
- single processor
 - systems, 187
- SMP (symmetric multiprocessing), 188
- SMT (symmetric multithreading), 189
- thread schedulers, 316
- tuning, 192-193
- scheduling processes, 24
- SCSI technology, 41
- SDET benchmarks, 147-150
- secure_redirects parameter (kernel), 517
- seek_cost parameter (I/O scheduler), 221
- Selective Acknowledgements (SACK), TCP, enabling, 285
- sem parameter (kernel), 510
- semaphores, 289, 318-321
 - IPC (Interprocess Communication), 27-28
 - parameters, 292-294
- SendFile API, network
 - performance, 458-460
- SendFile/Zero-copy column (NetBench), 469
- send_redirects parameter (kernel), 517
- sequential read
 - benchmarks, 409
- sequential read performance, AS scheduler, 415-416
- sequential write
 - benchmarks, 409
- servers, 35
 - application servers, 375-376
 - availability, 378
 - clustering, 391-394
 - commercial workload tuning, 479-480
 - concurrency, 379
 - connection pools, 502-503
 - data integrity, 378
 - databases, support, 378
 - distribution, 379
 - enterprise
 - application, 383
 - failover, 378
 - fault tolerance, 378
 - file systems, 390
 - hardware, 387-388
 - heterogeneousness, 379
 - hyperthreading, 504
 - interactivity, 379
 - J2EE, 384-385
 - J2EE (Java 2 Enterprise Edition), 376-377
 - Java, 376
 - JVM (Java Virtual Machine), 386-389
 - memory usage, 389
 - multiple-applications-in-one, 380
 - multiple-services-in-one, 380
 - native code, 386
 - networking, 389
 - pass-by-reference, 502
 - performance
 - requirements, 377-390
 - requests, 380
 - scalability, 378
 - security, 378
 - SMP scaling, 390-391
 - software performance stack, 382-383
 - string manipulation, 390
 - synchronization, 389
 - thread pools, 501-502
 - threads, 389
 - topologies, 394-398
 - transactions, 380
 - tuning, 398-402
 - vendor server
 - portability, 378
 - vendors, 380
 - web servers, 390
 - clusters, 38, 43
 - HA (high availability) clusters, 43-45
 - HPCs (high performance clusters), 43
 - database servers, 361
 - designing, 309-310
 - DNS servers, 35
 - enterprise servers, 42
 - file servers, 343
 - benchmarks, 408
 - NAS (network attached storage), 344
 - network file system protocols, 345-350
 - tuning, 354-358
 - I/O subsystems, 40-41
 - mail servers, benchmarks, 409
 - memory, 39-40
 - multiprocessing
 - clusters, 38, 43-45
 - NUMA (Non-Uniform Memory Architecture), 38, 48-54
 - processor mixing, 38
 - SMP (Symmetrical Multiprocessing), 36-39
 - topologies, 37-38

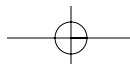


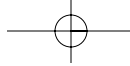
- network storage
 - servers, 344
- NUMA servers, 38, 48-54
 - hardware, 51
 - hierarchy, 51
 - high bandwidth
 - interconnect, 49
 - interconnects, 50, 52
 - latency ratios, 52
 - MPIO (Multipath I/O), 53
 - nodes, 50-51
 - timers, 54
- print servers, 343
 - performance, 358
- processor mixing, 38
- Shark servers, 48
- SMP servers, 36-39, 369
 - mutual exclusion, 36
 - topologies, 37-38
- systems
 - blades, 48
 - IBM zSeries
 - mainframes, 45-48
 - NUMA (Non-Uniform Memory Architecture), 48-54
- web servers, 331
 - architecture models, 336-337
 - benchmarks, 408
 - concurrency, 336
 - dynamic content
 - generation, 331
 - Ethereal packet capture
 - tool, 334-335
 - HTTP (HyperText Transfer Protocol), 331
 - network behavior, 333
 - requests, 332-333
 - responses, 331-333
 - transactions, 335-336
 - tuning, 337-340
- service processors, 35
- services, kernel, 20
- servlets (Java), 376
- SFS benchmarks, 160
- SGI, XFS (Next-Generation File System), 239, 268-271
- sh command, 339
- shared memory
 - kernel parameters, 510
 - semaphores, 27
- Shared Memory
 - Multiprocessing (SMP). *See* SMP (Shared Memory Multiprocessing)
- shared memory segments, 289
- IPC (interprocess communication)
 - parameters, 297-298
- shared system resources,
 - mutual exclusion, 36
- shared_media parameter (kernel), 517
- Shark servers, 48
- Sharoff, Narasimha, 507
- shmall parameter (kernel), 510
- shmmax parameter (kernel), 510
- shmni parameter (kernel), 510
- signals
 - IPC (Interprocess Communication), 25
 - kernel parameters, 510
- simulators, 174
- slab allocators, 208
- Slackware, 14-15
- sleep averages, tuning,
 - scheduler, 192-193
- SMB/CIFS network file system, 346
- smblog=1 column (NetBench), 469
- SMP (Symmetrical Multiprocessing) system, 28, 36-39
 - application servers,
 - scaling, 390-391
 - cache coherency, 29
 - concurrency, 28
 - data serialization, 28
 - locks, 28-29
 - mutual exclusion, 36
 - processor affinity, 29
 - schedulers, 188
 - servers, 369
 - topologies, 37-38
- SMT (symmetric multithreading),
 - schedulers, 189
- sockets, 311, 313-314
- software
 - performance evaluation
 - methodologies, 173
 - analysis methodologies, 175-179
 - benchmarking
 - methodology case study, 175
 - kernel performance
 - benchmarks, 179-184
 - numerical analysis, 174
 - simulators, 174
 - tracing, 173
 - workload characterization, 174
 - tracing, swtrace tool (Performance Inspector), 128-129
- software performance stacks,
 - application servers, 382-383
- Solaris, 201
- somaxconn parameter (kernel), 281



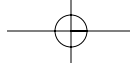


- SPEC (Standard Performance Evaluation Corporation), 185
- SPEC SDET benchmarks, 147-150
- SPEC SFS benchmarks, 160
- SPECjAppServer benchmarks, 168
- SPECjvm benchmarks, 163
- SPECweb benchmarks, 166-168
- SPECweb SSL benchmarks, 166-168
- SPECWeb99 benchmark, 456
 - content case studies, 475-476
- SPECweb99 benchmarks, results, 184
- SPECWeb99 workload generators, 340
- spinlocks column (NetBench), 470
- starting block numbers, 235
- STARVATION_LIMIT parameter (scheduler), tuning, 194
- starving non-interactive tasks, scheduler, 194
- status flags (file), setting, 247
- storage area networks (SANs), 344
 - NAS (network attached storage), compared, 345
- strace command, system tracing, 90-93
- strings, application servers, 390
- struct page pointers, BIO (block I/O) layer, 219
- subsystems (I/O), 217
 - integration, lack of, 217
 - performance issues, 228-229
 - I/O schedulers, 217-226
- Sun Microsystems, Inc., Java, 376
- Sun ONE application servers, 377
- super blocks, file systems, 236
- super-max parameter (kernel), 514
- super-nr parameter (kernel), 514
- sustained transfer rates, data movement, 243
- swap partition, placement, 4-5
- swappiness parameter (virtual memory), 211
- swapping memory, 23
 - virtual address spaces, 201-203
- switching, context switching, 24-25
- swtrace tool (Performance Inspector), 128-129
- Symmetrical Multiprocessing (SMP). *See* SMP (Symmetrical Multiprocessing)
- Syn Queue length, TCP, controlling, 286
- synchronization
 - application servers, 389
 - file synchronization, 246-248
 - semaphores, 318-321
- synchronized I/O, 246
- sysctl command, kernel parameters, modifying, 299
- sysctl interface, kernel parameters
 - CONFIG, 517
 - debugging, 511
 - file system, 513-514
 - ICMP, 515-516
 - IP fragmentation, 519
 - NEIGHBOR, 516
 - net core, 515
 - networks, 515-519
 - processes, 510
 - profiling, 511
 - routing, 517
 - shared memory, 510
 - signals, 510
 - system, 511
 - TCP, 518-519
 - tuning, 507-508
 - virtual memory, 511-513
- sysfs interface, kernel parameters, tuning, 509
- sysrq parameter (kernel), 511
- system calls
 - kernel, 20
 - tracing, strace command, 90-93
- system initialization
 - BSD, 12-16
 - init process, 12-16
 - initialization table, 13-14
 - System V, 12-16
- system initialization table (/etc/inittab), 12
- system parameters, 511
- system performance analysis, 58-59
 - CPU utilization, 59-62
 - gtop command, 65-67
 - sar command, 67-69
 - top command, 65-67
 - vmstat, 62-65
 - I/O utilization, 73-78
 - memory utilization, 69-73
 - network utilization, 78-83
- System Trace Data Collection tool (Performance Inspector), 102

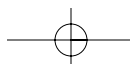


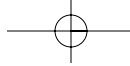


- system tracing, 87-88
 OProfile command, 93-96, 98-99, 101
 PI (Performance Inspector), 102-112, 125-130
 Above Idle tool, 102-105
 dynamic kernel patches, 103
 instruction tracing, 103, 110-111
 JLM (Java Lock Monitor) tool, 102, 125-128
 JPROF tool, 102, 112
 performance counters, 103
 pipcnt tool, 129
 post tool, 129
 ptt tool, 130
 pttstats tool, 102, 105-107, 130
 swtrace tool, 128-129
 System Trace Data Collection tool, 102
 tprof tool, 107-110
 requirements, 88
 strace command, 90-93
 top utility, 88, 90
 System V, 289
 BSD, compared, 12-16
 initialization table, Red Hat, 15-16
 IPC (Interprocess Communication), 26, 290-292
 system initialization, 12-16
- T**
 tables
 decayed frequency tables, 224
 page frame maps, 200
 page tables, 203-205
- targets, benchmarks, setting, 177
 tasks, 24
 interactive tasks, 192-194
 non-interactive tasks, 194
 sleep averages, tuning, 193
 threads, 24
 timeslices, tuning, 193
 TSRUN (TASK_RUNNABLE) state, 213
 virtual memory, 195
 TCP (transport control protocol)
 connection management, kernel parameters, 285-286
 D-SACK, enabling, 285
 FACK (Forward Acknowledgement), enabling, 285
 keep-alive management, kernel parameters, 286-287
 kernel parameters, 518-519
 memory management, kernel parameters, 283-284
 read/write buffers, kernel parameters, 281, 283
 SACK (Selective Acknowledgements), enabling, 285
 Syn Queue length, controlling, 286
 window scaling, enabling, 284
 TCP-W benchmarks, 166-168
 TCP/IP (transport control protocol/Internet protocol), 277-278
 kernel parameters, 281-287
 stacks, 455
 statistics, displaying, 82
- tcp_app_win parameter (kernel), 518
 tcp_dsack parameter (kernel), 285, 519
 tcp_fack parameter (kernel), 285, 519
 tcp_fin_timeout parameter (kernel), 518
 tcp_keepalive_intvl parameter (kernel), 287, 519
 tcp_keepalive_probes parameter (kernel), 287, 519
 tcp_keepalive_time parameter (kernel), 286, 519
 tcp_max_orphans parameter (kernel), 518
 tcp_max_syn_backlog parameter (kernel), 286, 518
 tcp_max_tw_buckets parameter (kernel), 518
 tcp_mem parameter (kernel), 283, 518
 tcp_reordering parameter (kernel), 518
 tcp_retries1 parameter (kernel), 519
 tcp_retries2 parameter (kernel), 519
 tcp_rfc1337 parameter (kernel), 518
 tcp_rmem parameter (kernel), 281, 518
 tcp_sack parameter (kernel), 285, 519
 tcp_stdurg parameter (kernel), 518
 tcp_synack_retries parameter (kernel), 286
 tcp_syn_retries parameter (kernel), 519

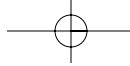


- tcp_timestamps parameter (kernel), 518
- tcp_tw_reuse parameter (kernel), 518
- tcp_window_scaling parameter (kernel), 284, 518
- tcp_wmem parameter (kernel), 282, 518
- thread dumps, JVM (Java Virtual Machine), 492-495
- thread listeners, 315
- thread pools, application servers, 501-502
- thread schedulers, 212-214, 316
- threads, 24
 - application servers, 389
 - default threads, virtual memory, 195
 - designing, 315-317
 - hyperthreading, 504
 - kernel threads, 24
 - PTT (per-thread time), monitoring, 105-107
 - SMT (symmetric multithreading) schedulers, 189
 - thread dumps, 493
 - TSRUN (TASK_RUNNABLE) state, 213
- thread_max parameter (kernel), 510
- three-tier topology, application servers, 395
- tightly coupled multi-processing systems, 28
- timers, NUMA servers, 54
- timeslices
 - processes, 25
 - tuning, scheduler, 193
- timing routines, 310-311
- tiobench, 153-155
- TOE (TCP Offload Engine), 467-468
- Tomcat application servers, 377
- top command
 - CPU utilization, 65-67
 - system tracing, 88-90
- topologies
 - application servers, 394-398
 - multiprocessor systems, 37-38
- Torvalds, Linus, 19
- tprof command, 107, 110
 - JPROF reports, generating, 112
- tracing runtime information collection, 173
- tracing systems, 87-88
 - OProfile command, 93-101
 - PI (Performance Inspector), 102-112, 125-130
 - Above Idle tool, 102-105
 - dynamic kernel patches, 103
 - instruction tracing, 103, 110-111
 - JLM (Java Lock Monitor) tool, 102, 125-128
 - JPROF tool, 102, 112
 - performance counters, 103
 - picntr tool, 129
 - post tool, 129
 - ptt tool, 130
 - pttstats tool, 102, 105, 107, 130
 - swtrace tool, 128-129
 - System Trace Data Collection tool, 102
 - tprof tool, 107, 110
- requirements, 88
- strace command, 90-93
- top utility, 88, 90
- track-to-track seek speed, disks, 242-243
- Trade3 benchmark, commercial workload, 481-492, 496-499
- connection pools, 502-503
- database I/O activity, 495-496
- hyperthreading, 504
- JVM (Java Virtual Machine), 499-501
- pass-by-reference, 502
- thread dumps, 492-495
- thread pools, 501-502
- web servers, 491
- transactions
 - application servers, 380
 - journalled file systems, 238-240
 - web servers, 335-336
- transfer rates, disks, 243
- transport control protocol (TCP). *See* TCP (transport control protocol)
- transport control protocol/Internet protocol (TCP/IP). *See* TCP/IP (transport control protocol/Internet protocol)
- TSO (TCP Segmentation Offloading), network performance, 460-462
- TSRUN (TASK_RUNNABLE) state, 213
- tunable parameters, kernel, 509
 - CONFIG, 517
 - debugging, 511
 - file system, 513-514
 - ICMP, 515-516
 - IP fragmentation, 519
 - NEIGHBOR, 516
 - net core, 515
 - network, 515-519
 - processes, 510
 - profiling, 511





- routing, 517
 - shared memory, 510
 - signals, 510
 - system, 511
 - TCP, 518-519
 - virtual memory, 511-513
 - tune command, file systems, tuning, 234
 - tune2fs utility (Ext2 file system), 259
 - tuning
 - application servers, 398-402
 - commercial workload, 489-504
 - AS (anticipatory I/O scheduler), 224-225
 - benchmarks, 177-179
 - CFQ I/O scheduler, 417-420
 - clients, 350-354
 - code, 301-304
 - algorithms, 305-328
 - commercial workload, 479-480, 489-504
 - compilers as tuning tools, 303-304
 - databases, 362-365
 - file servers, 354-358
 - file systems, 436
 - Ext3 file system, 256, 436-440
 - JFS (Journaled File System), 266-268, 442-443
 - ReiserFS, 260, 264, 440-441
 - XFS, 270-272, 443-445
 - Gigabit Ethernet, Netperf3, 471-472
 - I/O schedulers, 221-222
 - kernel parameters, 507-509
 - networks
 - kernel parameter tuning mechanisms, 278-287
 - network protocol stack, 277-278
 - profiling, 302-303
 - schedulers, 192-193
 - virtual memory, 209-212
 - web servers, 337-340
 - Apache web servers, 338-339
 - event-driven web servers, 339
 - Flash web servers, 339
 - log analysis tools, 340
 - Tux web servers, 339
 - workload generators, 340
 - Tux web servers, tuning, 339
- U**
- ucast_solicit parameter (kernel), 516
 - UNIX, Linux, compared, 19
 - unres_qlen parameter (kernel), 516
 - user address spaces, 21, 197
 - User Ramp Up tests, 489
 - usr partition, placement, 4-5
 - utilities
 - Ext2 file system, 258-259
 - JFS (Journaled File System), 268
 - ReiserFS, 264
 - XFS (Next-Generation File System), 272-273
 - uuidgen utility (Ext2/Ext3 file system), 259
- V**
- var partition, placement, 4-5
 - var/log/messages logs, 11
 - var/log/XFree86.0.logs, 11
 - variables (kernel), 277
 - vectored I/O, databases, 371
 - vendor server portability, application servers, 378
 - vendors, application servers, 380
 - version parameter (kernel), 511
 - vertical clustering, application servers, 392
 - VFS (virtual file systems), 30, 32
 - VGs (volume groups), 31-32
 - viewing
 - address spaces, virtual memory, 196
 - logs, 10
 - virtual file systems, 30, 32
 - virtual memory, 195, 205
 - address spaces, 21, 196-197
 - area structures, 198-199
 - generic view, 196
 - kernel address spaces, 199
 - paging, 201-203
 - swapping, 201-203
 - user address spaces, 197-198
 - CPU scheduler, 212-214
 - highmem interface, 201
 - kernel image section, 195, 200
 - kernel module section, 199
 - kernel parameters, 511-513
 - largepages, support for, 206
 - ObjRMAP (object-based reversed mapping), 206
 - page allocation, 207
 - page replacement, 207
 - page tables, 203, 205
 - paging, 23
 - processes, 195
 - address translation process, 204-205



- rmap (reversed mapping), 205
 - slab allocators, 208
 - tuning, 209-212
 - vm.heap-stack-gap parameter (kernel), 511
 - vmstat command
 - CPU utilization, 62-65
 - memory utilization, 72-73
 - vm_cache_scan_ratio parameter (kernel), 513
 - vm_gfp_debug parameter (kernel), 513
 - vm_lru_balance_ratio parameter (kernel), 513
 - vm_mapped_ratio parameter (kernel), 513
 - vm_passes parameter (kernel), 513
 - vm_vfs_scan_ratio parameter (kernel), 513
 - VolanoMark Java Server
 - benchmarks, 161-162, 183, 456, 473-474
 - volume groups (VGs), 31-32
 - volume managers, 30-31
- W**
- web servers, 331
 - application servers, 390
 - need for, 375
 - architecture models, 336-337
 - benchmarks, 166, 408
 - ECPerf, 168
 - SPECjAppServer, 168
 - SPECweb, 166-168
 - SPECweb SSL, 166-168
 - TCP-W, 166-168
 - commercial workload, Trade3 benchmark, 491
 - concurrency, 336
 - dynamic content generation, 331
 - Ethereal packet capture tool, 334-335
 - HTTP (HyperText Transfer Protocol), 331
 - network behavior, 333
 - requests, 332-333
 - responses, 331-333
 - transactions, 335-336
 - tuning, 337-340
 - Apache web servers, 338-339
 - event-driven web servers, 339
 - Flash web servers, 339
 - log analysis tools, 340
 - Tux web servers, 339
 - workload generators, 340
 - WebDAV extensions (HTTP), 347
 - WebSphere Application Server, 377
 - Trade3 benchmark, commercial workload, 481-504
 - window scaling, TCP, enabling, 284
 - wmem_default parameter (kernel), 515
 - wmem_max parameter (kernel), 280, 515
 - word sizes, processors, 39
 - workload characterization, 174
 - workload generators, web servers, 340
 - workloads
 - benchmarking, 132, 134, 173
 - analysis methodologies, 175-179
 - application benchmarks, 161-168
 - benchmarking methodology case study, 175
 - kernel performance benchmarks, 179-184
 - microbenchmarks, 134-160
 - performance evaluation methodologies, 173-174
 - commercial workloads
 - J2EE model, 480-481
 - Trade3 benchmark, 481-504
 - tuning, 479-480
 - databases, 367-368
 - decision support system (DSS) workloads, 362
 - steady state of operation, 131
 - write buffers, TCP, kernel parameters, 281-283
 - write_starved parameter (I/O scheduler), 221
- X-Z**
- XFS (Next-Generation File System), 5, 268-269
 - allocation groups, 269
 - creating, 270-271
 - external logs, 271
 - fullness, 271
 - kernel configuration support, 269
 - layout, analyzing, 435
 - mounting, 271-272
 - partitions, 269
 - tuning, 270-272, 443-445
 - utilities, 272-273
 - Zerocopy, network performance, 458-459
 - Ziff-Davis, NetBench, 456
 - zones, memory nodes, 22
 - Zope application servers, 377
 - zSeries mainframes, 45-48

