

Contents

	<i>Page</i>
Preface	xxi
UNICOS System Administration Publications	xxi
Related Publications	xxii
Ordering Cray Research Publications	xxv
Conventions	xxv
Reader Comments	xxvii
Introduction to System Administration [1]	1
Accounting [2]	3
Cray Research System Accounting (CSA)	3
Concepts and Terminology	4
Files and Directories Overview	5
Structures of the acct and tmp Directories	5
Shell Scripts and C Binaries	7
Unprocessed Data Files	7
Data Files Being Processed	8
Processed Data Files	9
Reports	10
Daily Operation Overview	10
Setting up CSA	11
The csarun Command	16
Daily Invocation	16
Error and Status Messages	17
States	17

	<i>Page</i>
Restarting <code>csarun</code>	19
Verifying and Correcting Data Files	20
Fixing <code>wtmp</code> Errors	20
Verifying Data Files	21
Editing Data Files	21
Files and Directories	23
<code>/usr/adm/acct</code> Directory	23
<code>/etc</code> Directory	27
<code>/etc/config</code> Directory	28
CSA Data Processing	29
Data Recycling	33
How Sessions Are Terminated	33
Why Recycled Sessions Should Be Scrutinized	34
How to Remove Recycled Data	34
Adverse Effects of Removing Recycled Data	36
NQS Requests and Recycled Data	38
Tailoring CSA	39
System Billing Units (SBUs)	40
Daemon Accounting	51
Setting up User Exits	51
Charging for NQS Jobs	52
Tailoring CSA Shell Scripts and Commands	53
Using <code>at</code> to Execute <code>csarun</code>	53
Allowing Nonsuper Users to Execute CSA	54
Using an Alternate Configuration File	55
Disk Usage Reporting (<code>diskusg</code>)	55
Per-process Accounting Data	56
Base Accounting Record	56

	<i>Page</i>
End-of-job Accounting Record	59
Multitasking Accounting Record	60
SDS Accounting Record	61
MPP Accounting Record	61
Performance Accounting Record	61
Multitasking Incentives	62
Memory Integrals	62
Reducing Charges	64
Socket Accounting	65
Device Accounting	65
Categories of Devices	65
Structures and Device Names	66
Configuration Changes	67
System Header Files	67
Using Device Accounting (Devacct(8))	68
Switching / and /usr File Systems	70
Logging Information	70
Boot Log	70
cron Log	71
Dump Log	71
New User Log	72
su Log	72
OLDSu Log	73
System Logs	73
Error Log	74
Multilevel Security (MLS) Log	75
System Activity Log	76

	<i>Page</i>
Message Log	76
Accounting Logs	76
NQS Log	78
Standard UNIX Accounting	79
Files and Directories	80
Daily Operation	81
Setting up the Accounting System	82
Setting up a User Exit	83
Converting Standard UNIX Accounting to CSA Accounting	83
the runacct Command	84
Failure Recovery for runacct	87
Restarting runacct	88
Fixing Corrupted Files	88
Fixing wtmp Errors	89
Fixing tacct Errors	89
Updating Holidays	90
Reports	90
Daily Report	91
Daily Usage Report	91
Daily Command and Monthly Total Command Summaries	93
Last Login Report	94
Accounting Files	94
Front-end Formatting	97
Why Use Front-end Formatting	97
Preparing to Use a Formatter	98
CSA Front-end Formatting	99
Generic Front-end Formatting	99
Data Consolidation	100

	<i>Page</i>
Required Data Variables	100
Default and Optional Data Variables	104
Data File Format	120
csagfef Source Scripts	121
Automated Incident Reporting (AIR) [3]	137
AIR Components Overview	137
AIR Configuration File	137
AIR Coordinator Daemon	138
AIR Monitoring Functions	138
AIR Report Generator	138
Initiation and Administration	139
AIR Configuration	140
Basic Syntactic Rules	140
Configuration Keywords	141
File Delineation Keywords	141
Basic Operational Keywords	142
Monitored Products Keywords	144
Monitoring Function Specification	146
Installation Tool Configuration	147
Return Tags	149
Sample Configuration File	150
Configuration File Tuning and Validation	159
Monitoring Functions	160
TCP/IP	161
NQS	162
Online Tapes	163
Disk-integ	164

	<i>Page</i>
URM	164
Adding Products and Functions	165
Creating Functions	166
Integrating the Functions	168
Configuring the Functions	169
Function Configuration through the Installation Tool	171
Validating Configuration	173
Production	173
Final Verification	174
Using the Report Generators	174
Record Types	174
Output	176
Using the <code>airprconf</code> Command	176
Using the <code>airdet</code> Command	176
Using the <code>airtsum</code> Command	178
Using the <code>airsum</code> Command	180
Log File Analysis	183
Summary	188
Fair-share Scheduler [4]	191
Design Objectives	191
Fair-share Feature Summary	192
Components of Fair-share	194
User Database (UDB)	194
Support Functions	196
User and Administrator Displays	197
Administrator Controls	198
Fair-share Hierarchy	198

	<i>Page</i>
Share Normalizing	199
Process Scheduling	201
Fair-share Limits Node (Lnode)	201
Fair-share and NQS	202
Fair-share and URM	202
Using the Fair-share Scheduler	202
Setting up a Fair-share Hierarchy	203
Creating Resource Groups	205
Allocating Shares to Users	207
Setting up System UDB Entries	209
Idle Account	209
UnKnown or Unknown Account	210
Other System Resource Accounts	210
Setting up Share by Account Mode	211
Activating the Fair-share Scheduler	212
Setting Scheduling Options and Flags	213
Modifying Fair-share Scheduler Settings	215
Enabling Resource Group Administrators	216
Disabling the Fair-share Scheduler	217
Costs, Usage, and Background Users	218
Monitoring the Fair-share Scheduler	219
Using the <code>shriview(1)</code> Command	219
Using the <code>shrmon(8)</code> Command	221
Using the <code>shrtree(8)</code> Command	222
Customizing the Fair-share Scheduler (User Exit)	224
Fair-share User Exit Example	225
Tuning the Fair-share Scheduler	227
Fair-share Parameters	227

	<i>Page</i>
shradm(8) Options Affecting Cost	227
-U Option (<i>MAXUSAGE</i>)	228
-D Option (<i>priority Decay Rates</i>)	229
-X Option (<i>maxushare</i>)	231
-Y Option (<i>mingshare</i>)	231
-Z Option (<i>sharemin</i>)	232
-R Option (<i>delta</i>)	234
-K Option (<i>usage Decay Rate</i>)	234
Example Parameter Settings	234
Fair-share and the Memory Scheduler	236
Priority-based Scheduling and I/O Resources	236
Using CPU Quotas	237
Additional Reference Material	238
File System Quotas [5]	239
Components of the File System Quota Feature	239
Enabling the Quota Feature	240
Changing the <i>NQUOTA</i> Value	240
Quotas and Data Migration	241
Configuring Quotas	241
Format 1: Relative File Name	242
Format 2: Absolute File Name	242
Format 3: Quota Control Groups	242
Determining Defaults and Special Users	243
Creating a Quota Control Source File	244
The <i>qudu(8)</i> Method of Source Creation	245
Changing Defaults	245
Setting Specific Quotas	246

	<i>Page</i>
Manual Source File Creation	248
Generating the Quota Control File	248
Activating Quota Enforcement	249
Setting Current Usage Information	251
Usage Accumulation Rules	251
Inode Usage	251
File Usage	251
Administering the Quota Enforcement Feature	252
System Startup	252
Back-up File Example	258
Quota Enforcement Control	258
Adding Users	258
Deleting Users	259
Creating or Extending Files	259
Viewing Quota Control	259
Additions to Login Profile	261
User-level Behavior	261
Nonstandard Configuration Options	262
Nonresident Quota Control Files	262
File System Groups under One Quota Control File	263
Aggregate Quotas	263
Using the Oversubscription Option (Soft Quota)	264
Behavior of the Oversubscription Mechanism	264
Supported Algorithms	265
Exponential Algorithm	265
Linear Algorithm	266
Algorithm Comparison	267
Changing ID Class Control	268

	<i>Page</i>
Quota Enforcement Across a Network	268
File System Space Monitoring [6]	269
Operation of the Space Monitor	269
Interprocess Communication	270
The Monitor Daemon, <code>fsdaemon</code>	270
File System Monitoring	271
Critical and Warning Command Processing	272
Request Processing	273
Status Display	275
Example 1:	275
Using the <code>fsdaemon</code> and <code>fsmon</code> Commands	277
The Log File	278
Informative and Error Messages	279
Installation and Operation Information	280
Installation	281
Operation	281
Permissions	282
Testing	282
Related Files	282
System Activity Monitoring [7]	283
Standard UNIX System Activity Package (<code>sar</code>)	283
System Activity Counters	284
CPU Time Counters	285
<code>lread</code> and <code>lwrite</code> Counters	285
<code>bread</code> and <code>bwrite</code> Counters	285
<code>phread</code> and <code>phwrite</code> Counters	285
<code>swpin</code> and <code>swpout</code> Counters	285

	<i>Page</i>
xswapin and xswapout Counters	286
switch and syscall Counters	286
runque, runocc, swpocc, and swpque Counters	286
iget, namei, and dirblk Counters	286
readch and writtech Counters	287
rcvint and xmtint Counters	287
rawch, canch, and outh Counters	287
clists Counter	287
I/O Activities	287
Logical Device Cache	288
Inode, File, Text, and Process Tables	288
sysrda, syswra, and syslstio Counters	288
pkin, pkout, and pkbad Counters	288
shuffle, textlock, datlock, and punlock Counters	288
Exchange Counts	288
System Activity Commands	289
sar(8) Command	289
sag(8) Command	290
timex(1) Command	290
mppview(8) Command	290
xmppview(8) Command	291
Daily Report Generation	291
Facilities	291
Suggested Operational Setup	292
Source Files and Scripts	293
Derivations	294
Cray Research System Activity Monitoring (sam) Package	295
samdaemon Process	297
csam(8) Utility	298

	<i>Page</i>
csam Commands	298
Record/replay Function	299
xsam Utility	303
X11 Window Settings	303
xsam Windows	304
Console Window and Available Commands	305
xsam Commands	309
Cray Research System Activity Reporting (tsar(8)) Package	318
sdc Command	319
tsar Command	320
Data Collection	320
CPU Data	321
System Calls	322
Process Management	322
Memory Management	323
System Table Management	323
System I/O, General	324
System I/O, Caching	325
General System Data	326
Disk Data	328
Tape Data	329
TCP/IP Data	330
Terminal Data	330
NFS Data	330
Network Interface Data	331
IPC Data	332
Restricted Data Collection	332
Data File Format	333
Header Records	333
Definitions Record	333

	<i>Page</i>
Meta-data Record	333
Data Records	334
tsar(8) Modes	335
Compilation-only Mode	335
Online Mode	335
Playback Mode	335
tsar(8) Source Scripts	335
BEGIN Section	336
END Section	336
RESTART Section	336
RECONFIG Section	336
function Section	336
Body	337
Example Source Scripts	337
tsar(8) Language Description	337
Statements	337
Operators	339
Built-in Functions	340
Built-in Variables	343
Operational Setup	343
Difference between sdc(8) and sdcx	343
Boot Time Data	344
Shutdown Data	345
crontab(1) Entries	345
Examples	346
sdc(8) Data Collection	346
tsar(8) Data Collection	346
tsar(8) Report Formatting	347
Limitations	349

	<i>Page</i>
Disk Usage Monitoring	351
Unified Resource Manager (URM) [8]	353
URM Features	354
Summary of URM Commands	355
Installing URM	355
Verifying Automatic Startup of URM Daemon	356
Configuring Initial URM Values	357
Individual Session Initiator Configuration Changes	358
Verifying Security Parameters	359
Enabling Service Providers	359
Configuring URM	359
Authorized Administrators	360
Authorized Hosts	362
Machine Load Evaluation Rates	362
Machine Target Values	363
Individual Session Initiator Targets	365
Individual Session Initiator Defaults	365
URM Control Settings	366
SDS Residency Time	367
Local URM Configuration File	368
Rank Boost to Previously Running Batch Jobs	368
Weighting Factors for the Batch Selector	368
Auto-configuration Settings	369
Resetting to Default URM Configuration	371
Importing URM Configuration	371
Activating URM Configuration Changes	371
Using URM with NQS	372
URM Administrator Tasks	372

	<i>Page</i>
Using URM Log Files	373
Monitoring URM Log Files	374
Turning Off URM Logging	374
Moving the URM Log Files	374
Viewing URM Results	375
Viewing Machine Load	375
Viewing All Users	375
Viewing Jobs of a Given User	376
Changing a Job's Minimum Rank	376
Changing URM Configuration Based on Time of Day (cron)	376
Customizing URM (User Exits)	377
URM User Exits	378
URM Job-ranking Functions	378
URM Data Structures	380
URM User Exit Example 1	381
URM User Exit Example 2	383
URM User Exit Example 3	385
URM User Exit Example 4	386
Troubleshooting URM	388
URM Daemon Failures	388
URM and NQS	389
URM and the Fair-share Scheduler	390
Share Priority Weight	390
Share Entitlement Weight	391
Usage Weight	391
URM Architecture	392
Selection Server (urmd)	392
Query and Command Server (rmgr)	393

	<i>Page</i>
SDS Management (sdsmgr)	393
URM Resources	395
URM Checkpointing	396
Configuring URM Checkpointing	396
Managing Restart Images	398
Checkpointing at Shutdown	400
Tuning URM	400
Tuning URM Control Settings	401
Machine Target Values	401
Monitoring Cycles	405
Group Scheduling Control	408
Load Smoothing Factors	410
Tuning URM Job Selection Criteria	414
Batch Jobs	415
Interactive Jobs	420
Appendix A Automatic Incident Reporting Tests	423
Index	425
Figures	
Figure 1. /usr/adm/acct and tmp directory structures	6
Figure 2. CSA program data flow	30
Figure 3. Directory structure of the adm login	80
Figure 4. Share normalizing	200
Figure 5. Example of system resource division for fair-share	204
Figure 6. Hardware configuration options for sam	296
Figure 7. The xsam console window	306
Figure 8. The xsam device display window	316

	<i>Page</i>
Figure 9. Example of different smoothing factors	412
Figure 10. Default smoothing factor for memory load	413
 Tables	
Table 1. Possible effects of removing recycled data	38
Table 2. Base accounting record fields by function	56
Table 3. End-of-job accounting record fields by function	59
Table 4. Required data variables	101
Table 5. pacct base record variables — per-process values	105
Table 6. pacct base record variables - total values	105
Table 7. pacct base record variables - prime time values	107
Table 8. pacct base record variables - nonprime time values	107
Table 9. pacct secondary data storage (SDS) record variables - total values	108
Table 10. pacct SDS record variables - prime time values	109
Table 11. pacct SDS record variables - nonprime time values	109
Table 12. pacct end-of-job record variables	110
Table 13. pacct device I/O record variables - total values	110
Table 14. pacct device I/O record variables - prime time values	111
Table 15. pacct device I/O record variables - non-prime time values	112
Table 16. pacct massively parallel processing (MPP) record variables - total values	113
Table 17. pacct MPP record variables - prime time values	114
Table 18. pacct MPP record variables - nonprime time values	114
Table 19. pacct multitasking record variables - total values	114
Table 20. pacct multitasking record variables - prime time values	115
Table 21. pacct multitasking record variables - nonprime time values	115
Table 22. pacct performance record variables - total values	115
Table 23. pacct performance record variables - prime time values	116
Table 24. pacct performance record variables - nonprime time values	116

	<i>Page</i>
Table 25. Tape accounting variables	117
Table 26. NQS accounting variables	117
Table 27. Connect time accounting variables	119
Table 28. System billing units (SBU) variables	119
Table 29. Built-in variables	131
Table 30. Share division among resource groups	206
Table 31. Share division within resource groups	208
Table 32. Field usage of the exponential algorithm	266
Table 33. Field usage of the linear algorithm	267
Table 34. CPU data	321
Table 35. System call data	322
Table 36. Process management data	322
Table 37. Memory management data	323
Table 38. System table management data	324
Table 39. General system I/O	324
Table 40. System I/O - caching	325
Table 41. General system data	327
Table 42. IOS-E physical disk data	328
Table 43. Tape drive data	329
Table 44. TCP/IP performance data	330
Table 45. Terminal activity data	330
Table 46. NFS data	331
Table 47. Network interface data	331
Table 48. IPC activity data	332
Table 49. Records written	334
Table 50. tsar command built-in variables	343