



ORACLE

PTR/INT Oracle Solaris 11
Internals

Titulo: PTR/INT Oracle Solaris 11 Internals

Clave: D90819GC10

Duración: 05 días

Contents

- 1 Introduction** Overview 1-2 Course
Goals 1-3 Course Agenda: Day 1 1-4 Course Agenda: Day 2 1-6 Course Agenda: Day 3 1-8 Course Agenda: Day 4 1-9 Course Agenda: Day 5 1-10 Introductions 1-13

Your Learning Center 1-14

- 2 Oracle Solaris 11 Operating System: Introduction** Objectives 2-2
Lesson Agenda 2-3 Operating System Basics 2-4
Lesson Agenda 2-6 Defining Processes 2-7
Lesson Agenda 2-8
SPARC 32-Bit Address Space 2-9
SPARC 64-Bit Address Space 2-10
x86 32-Bit Address Space: 64-Bit OS 2-11
x86 64-Bit Address Space: 64-Bit OS 2-12
Lesson Agenda 2-13
Process Structures 2-14
Lesson Agenda 2-16
Kernel Mode Entry 2-17
Lesson Agenda 2-19
System Calls 2-20
Interrupts 2-22
Lesson Agenda 2-23
clock() Routine 2-24
LWP Accounting Scalability 2-26
LWP Accounting 2-27
Callout Queue 2-28
callout Structure 2-29

Callout Scalability 2-30
Lesson Agenda 2-32

Kernel Memory Allocator	2-33	
Lesson Agenda	2-35	
SunOS Evolution	2-36	
Solaris 9 +10: Features	2-37	
Solaris 11 OS: New Features	2-38	
Lesson Agenda	2-40	
Tools	2-41	
adb and kadb Tools	2-42	
mdb Tools	2-44	
kmdb Tools	2-49	
DTrace	2-53	
Lesson Agenda	2-55	
Probe Descriptions and Clauses	2-56	
Practice 2 Overview: Introducing DTrace		2-58
Summary	2-59	

3 Multithread Architecture

Objectives	3-2	
Lesson Agenda	3-4	Common
Terminology	3-5	
Lesson Agenda	3-7	
Multiprocessor Architectures	3-8	
Threads in the Solaris 9 and Solaris 10 OS	3-9	
Kernel Threads	3-10	
Lesson Agenda	3-12	
Process Structures	3-13	
Fields from the proc Structure	3-14	
Fields from the user Structure	3-15	
Fields from the Kernel Lightweight Process Structure (klwp_t)		3-16
Fields from the Kernel Thread Structure (kthread_t)		3-17
Fields from Kernel Thread Structure	3-18	
cpu Structure	3-19	
Lesson Agenda	3-21	
Interrupts	3-22	
Interrupt Threads	3-23	
Interrupt Threads Priorities	3-24	
Lesson Agenda	3-25	
Locks	3-26	
Mutex Locks	3-28	

Adaptive Mutex	3-29
Spin Mutex	3-30
Acquiring a Mutex Lock	3-31
Turnstiles	3-32
Semaphores	3-33
Multiple-Reader, Single-Writer Locks	3-35
Condition Variables Structure	3-36
Sleep Queue Properties	3-37
Practice 3 Overview: Multithread Architecture	3-38
Summary	3-39

4 Hardware Memory Management

Objectives	4-2
Lesson Agenda	4-3
Main Memory	4-4
Virtual Memory	4-5
Process Address Space	4-6
Memory Terminology	4-7
Lesson Agenda	4-9
System Memory Model	4-10
Lesson Agenda	4-11
Virtual-to-Physical Address Translation	4-12
Lesson Agenda	4-13
x86 – 32-Bit MMU	4-14
Page Table Entry	4-15
Page Table Entry (PTE)	4-16
x86 with Physical Addressing Extension (PAE)	4-17
AMD 64-Bit MMU	4-18
Translation Lookaside Buffer (TLB)	4-19
Large Page Sizes	4-20
Spitfire Memory Management Unit (SFMMU)	4-21
Translation Storage Buffer (TSB) Properties	4-22
Translation Storage Buffer	4-23
Table Translation Entry	4-24
ctx Structure	4-25
hme_blks Block	4-27
Lesson Agenda	4-28
Cache	4-29
Cache Compared to Memory	4-30
Cache Hit Rate	4-31
Defining Types of Caches	4-32
Virtual Address Cache	4-33
Physical Address Cache	4-34
Cache Aliasing on a Virtual Cache	4-35
Direct-Mapped Cache	4-36
Set-Associative Cache	4-37
Set-Associative Cache Properties	4-38
Harvard and Unified Caches	4-39

Write-Through and Write-Back Cache	4-40
Cache Snooping	4-41
I/O Cache	4-42
Lesson Agenda	4-43
Hardware Address Translation Layer	4-44
Practice 4 Overview: Hardware Memory Management	4-45
Summary	4-46

5 Software Memory Management

Objectives	5-2
Lesson Agenda	5-3
SunOS VM1: Features	5-4
Process Address Space	5-5
VM1 Virtual Memory System Layers	5-6
VM2 Virtual Memory System Layers	5-7
What Has Not Changed in Phase 1	5-8
mmap(2) System Call	5-9
System Calls and Services	5-10
madvise(3C) Routine	5-11
NUMA Locality	5-13
Latency Topology: Example	5-15
Ladder Topology: Example	5-16
Address Space Layer	5-17
as Structure	5-20
Physical Pages	5-21
page Structure	5-22
sf_hment Structure	5-23
memseg Structure	5-24
Lesson Agenda	5-26
Virtual Memory Segment Drivers	5-27
seg Structure	5-28
seg_ops Structure	5-29
File I/O Shared-Mapping Segment Driver	5-30
Device Segment Driver (segdev)	5-31
Kernel Memory Segment Driver (segkmem)	5-32
vnode Segment Driver (segvn)	5-33
segvn_data Structure	5-34
vpage Structure	5-36
Anonymous Memory	5-37
AVL Trees	5-38
Mapping Structures	5-39
Process Memory Data Structures	5-40
segkp Driver	5-41
segkp Driver Strategy	5-42
Kernel Physical Mapping Segment Driver (segkpm)	5-43
Practice 5 Overview: Software Memory Management	5-44
Summary	5-45

Objectives	6-2
Lesson Agenda	6-3
Motivation for VM2	6-4
What Is in VM2 Phase 1?	6-8
Lesson Agenda	6-9
What Has Not Changed in Phase 1?	6-10
VM1 Virtual Memory System Layers	6-11
VM2 Virtual Memory System Layers	6-12
VM2 Phase 2 and Phase 3	6-13
VM2: The Big Picture	6-14
VM2 Update 1	6-15
Proposed Process Mappings	6-16
Lesson Agenda	6-17
Criteria for Memory Selection: mnodes	6-18
mnode	6-19
mnode MDB: Example	6-20
Lesson Agenda	6-21
Criteria for Memory Selection: Tiles	6-22
Tiles	6-23
Tile Sizes	6-24
System Tiles: Example	6-25
Tile Data Structures	6-26
Example in mdb	6-27
Criteria for Memory Selection: tilelets and tilechunks	6-28
Tilelets	6-29
System Tilelets: Example	6-31
Example in mdb	6-32
tilechunk_t	6-33
Tilechunks: Example	6-34
System Tilechunks: Example	6-35
tilechunk_t mdb: Example	6-36
Physical Address of Tilechunk Maps	6-37
::pachunk – mdb Example	6-38
tileset_t	6-39
::tileset mdb Example	6-40
Criteria for Memory Selection: Kernel Cage	6-41
VM2	6-42
Kernel Cage in VM2	6-43
Lesson Agenda	6-45
Criteria for Memory Selection: Typed Page Credits	6-46
Typed Page Credits	6-47
Capture	6-48
Bounds Predictor	6-50
Criteria for Memory Selection: memgrp	6-52
memgrp	6-53

memgrp – MDB Example	6-54	
System Structure	6-55	
System Structure: mdb Example	6-56	
Criteria for Memory Selection	6-57	
Page Size Codes	6-58	
::size: mdb Example	6-59	
Page Allocation Credits	6-60	
crd Structure	6-61	
crd: mdb Examples	6-63	
Wallet	6-64	
Fed	6-65	
Fed: mdb Example	6-66	
Breadline	6-67	
Breadline: mdb Examples	6-68	
Soupline	6-69	
FLR and TCM	6-70	
SAC	6-71	
SAC: mdb Example	6-72	
FLB	6-73	
FLB: mdb Examples	6-74	
End-to-End Credit Auditing: mdb Example		6-75
Global Page Size Statistics: mdb Examples		6-76
Reverse Map Entry	6-77	
RM: Implementation	6-78	
RM: Important Flags	6-80	
RM: mdb Examples	6-81	
RMG	6-83	
RMG: mdb Examples	6-84	
Sparse Data Structures	6-85	
Sparse Data Structures: mdb Examples		6-86
Sparse Data Structures	6-88	
Criteria for Memory Selection: Review		6-89
Lesson Agenda	6-90	
VM2 Allocation	6-91	
Allocation Credits	6-92	
Wallet Types	6-93	
Freelist Buckets	6-94	
Allocation: Glue Interfaces		6-95
Allocation: Locality	6-96	
Allocate Credits: Allocation Parameters		6-97
Allocation: Allocate crd_ts	6-98	
Allocation: Breadlines	6-99	
Allocation: Allocate crd_ts	6-100	
Allocation: Allocate rm_ts	6-102	

Allocate rm_ts: flr_iterate_slot()	6-103
Allocation: Allocate rm_ts	6-104
Allocate RMs: flr_iterate_tile()	6-105
Allocation: Allocate page_ts	6-106
Lesson Agenda	6-107
Predictor	6-108
Predictor Data Types	6-109
Predictor Threads	6-111
Predictor Components	6-112
Predictor Sampler	6-113
Predictor Analyzer	6-114
Predictor States	6-115
Predictor Action Engine	6-116
Predictor mdb dcmds	6-118
VM2 Structures	6-123
Lesson Agenda	6-124
Procedure for a VM2 Quick Check	6-125
Practice 6 Overview: VM2	6-127
Summary	6-128

7 Paging and Swapping

Objectives 7-2 Lesson	
Agenda 7-3 Paging:	
Overview 7-4 Paging In 7-5	
Page Replacement	7-6
Parts of the Page Daemon	7-7
Clock Algorithm	7-8
Defaults for SunOS 5.5.1 Through SunOS 5.8	7-9
Paging Parameters: Updates 7-11 schedpaging() Routine	7-12
pageout()Routine	7-13
pageout_scanner() Routine	7-14
checkpage(pp, whichhand) Routine	7-15
Lesson Agenda	7-16
Swapper	7-17
Swapper (sched.c)Operation	7-18
not_swappable() Macro	7-19
Desperate Swapper Memory Conditions	7-20
sched() Routine, Part 1	7-21
sched() Routine, Part 2	7-22
CL_SWAPOUT() Macro and Routines	7-23
CL_SWAPOUT() Routines	7-24
CL_SWAPIN() Macro and Routines	7-27
rt_swapin()and fx_swapin() Routines	7-28
Lesson Agenda	7-29
Virtual Address Lookup	7-30

8 The swapfs File System

Objectives 8-2 Lesson Agenda

8-3

Problems with Anonymous Memory in SunOS 4.x 8-4 Anonymous
Memory in SunOS 5.x 8-5

Lesson Agenda 8-6

Swap Management Structures 8-7

anoninfo Structure 8-8 swapinfo Structures

8-9 anon_map Structure 8-10 anon_hdr

Structure 8-12

anon Structure in Sun OS 5.x 8-13

Lesson Agenda 8-14

Swap Space Management 8-15

Swap Area: Example 8-17

Lesson Agenda 8-18

Advantages of swapfs File Systems 8-19

Practice 8 Overview: The swapfs File System 8-20 Summary 8
21

9 Scheduling Objectives 9-2

Lesson Agenda 9-3 Scheduling

Features 9-4 Real-Time

Scheduling 9-5

SDC Scheduling Class 9-14

Lesson Agenda 9-18

Dispatch Priorities 9-19

dispq_t Structure 9-20

State Diagram 9-21

Class Array 9-22

Lesson Agenda 9-23

Process Structures 9-24

tsproc_t Structure (View 1) 9-25

tsproc_t Structure (View 2) 9-26

rtproc_t Structure 9-27

classfunc Structure 9-28

Lesson Agenda 9-29

Kernel Mode Priority Assignment 9-30

Timesharing Dispatch Parameter Table (View 1) 9-31

Timesharing Dispatch Parameter Table (View 2) 9-32

Timesharing/Interactive Dispatch Parameter Table 9-33

Real-Time Dispatch Parameter Table	9-35
prionctl(1) Command	9-36
dispadmin(1) Command	9-37
Calculating a Thread's Priority	9-39
Lesson Agenda	9-40
Kernel Scheduling-Related Variables	9-41
Kernel Scheduling-Related Functions	9-42
ts_tick() Routine	9-43
ts_tick() Routine (View 2)	9-44
rt_tick() Routine	9-45
preempt() Routine	9-46
CL_PREEMPT()	9-47
setbackdq(kthread_t *tp)	9-48
ts_preempt(tspp) Routine	9-49
rt_preempt() Routine	9-50
setfrontdq() and setbackdq() Routines	9-51
disp() and swtch()	9-52
disp() Routine	9-53
dispgetwork()	9-54
CMT	9-55
Lesson Agenda	9-56
Priority Inversion	9-57
Callout Queue Processing	9-58
Bounded Priority Inversion	9-59
Unbounded Priority Inversion	9-60
Priority Inheritance	9-61
Blocking Chains	9-62
turnstile_t Structure	9-63
Practice 9 Overview: Scheduling	9-64
Summary	9-65

10 Process Lifetime Objectives 10-2

Lesson Agenda	10-3
Process Creation	
Routines	10-4
Process Structures	10-5
Process Creation System Calls	10-6
Fork	
Return values	10-7
Fork Extensions	10-8
posix_spawn(3C)Routine	10-9
cfork()	
	10-10

cfork() Routine	10-11	
getproc() and pid_allocate() Routines	10-12	
pid_allocate() Routine	10-13	
getproc() Routine (View 1)	10-14	
getproc() Routine (View 2)	10-15	
Process Structures After getproc()	10-16	
cfork() and as_dup() Routines	10-17	
Return to the cfork() Routine	10-18	
Process Structures After as_dup()	10-19	
forklwp() and lwp_create() Routines	10-20	
thread_create() Routine	10-22	
Process Structures After thread_create() and forklwp()		10-24
cfork() and CL_FORKRET() Routines	10-25	
Return to cfork()	10-26	
Lesson Agenda	10-27	
exec Routines	10-28	
exec_common() Routine	10-29	
execsw Array	10-31	
gexec() Routine	10-32	
Executable and Linking Format (ELF)	10-33	
Elf32_Phdr Program Header Table	10-34	
Elf64_Phdr Program Header Table	10-36	
Process Segments	10-37	
elfexec() Routine	10-38	
elfexec() Routine (View 1)	10-39	
Auxiliary Vector	10-40	
elfexec() Routine (View 2)	10-41	
mapelfexec() Routine	10-42	
execmap() Routine	10-44	
The Initial Process Stack	10-46	
proc_exit(why, what) Routine	10-47	
waitid(idtype, id, ip, options) Routine	10-50	
waitid(idtype, id, ip, options) Routine	10-51	
waitid() Routine (View 1)	10-52	
waitid() Routine (View 2)	10-53	
Practice 10 Overview: Process Lifetime	10-54	
Summary	10-55	

11	Signals	Objectives	11-2	Lesson	
	Agenda	11-3	Solaris 10 OS Signals		
		11-4	Kernel Signal Bitmasks	11-8	
	Kernel Signal Bitmask: k_sigset_t				11-9
	Interrupt and Trap Signals			11-10	
	Lesson Agenda		11-11		
	Signal Delivery		11-12		
	Signal Actions		11-13		
	Signal Actions: Ignoring		11-15		
	Signal Actions: Holding		11-16		
	Kernel-Signal-Related Variables		11-17		
	Signal Mask Routines		11-19		
	Assigning a Signal Disposition			11-21	
	sigaction(sig, *actp, *oactp) System Call				11-23
	sigaction(sig, *actp, *oactp): System Call				11-25
	sigtoproc(p, t, sig) Routine		11-26		
	issig(why) Routine		11-27		
	issig_forreal() Routine		11-28		
	fsg() Routine		11-29		
	psig() Routine		11-30		
	sendsig() Routine		11-31		
	Calling the Handler		11-33		
	Practice 11 Overview: Signals			11-34	
	Summary		11-35		
A	Appendix A	Contents	A-2		
	sysdc_update()		A-3		
	sysdc_update_pri()		A-4		