

Stručný obsah

Část I – Úvod

1. O Linuxu obecně	31
2. Historie Linuxu	37
3. Možnosti nasazení	41

Část II –Vnější rozhraní jádra

1. Úvod	47
2. Start systému	53
3. Správa modulů	59
4. Systémová volání	65
5. Signály	75
6. Souborové operace	83
7. Souborové systémy	111
8. Mapování souboru do paměti	129
9. Roury a zprávy	137
10. Sockety	149
11. Procesy a vlákna	157
12. Spouštění programů	175
13. Synchronizační objekty	183
14. Souborový systém procfs	191
15. Souborový systém sysfs	197
16. Některé další speciální souborové systémy	205
17. Netlink	209
18. udev, HAL	211
19. inotify	219
20. kevent	223
21. FUSE	225
22. Další informace, budoucnost	229

Část III –Vývoj ovladačů

1. Úvod	249
2. Licenční aspekty	253
3. Základy vývoje pro jádro	259
4. Ladění a testování	273
5. Základní operace	295
6. Paměť, její adresace a správa	309
7. Komunikace s hardwarem	329
8. Synchronizace přístupu	335
9. Obsluha přerušení	345

10. Měření času, časovače	353
11. Odložené zpracování	365
12. DMA	371
13. Objektový model jádra	379
14. Oznamování událostí do uživatelského prostoru	387
15. Načítání firmwaru	389
16. Bloková zařízení	393
17. Síťová zařízení	405
18. Zařízení na sběrnici PCI	413
19. Zařízení na sběrnici USB	423
20. Tvorba uživatelského rozhraní	435
21. Ovladače souborových systémů	443
22. Souborové systémy pomocí FUSE	455
23. Zajímavosti, plány do budoucna	465

Část IV – Pohled dovnitř jádra

1. Úvod	471
2. Zdrojové kódy jádra	475
3. Architektura jádra	485
4. Správa paměti	493
5. Správa procesů a vláken	515
6. Paměť procesu	535
7. Systémová volání a signály	549
8. Přerušení a výjimky	557
9. DMA	565
10. Časovače	569
11. Soubory a souborové systémy	579
12. Síťová komunikace	591
13. Plánování I/O	599
14. Disková cache	605
15. Odkládání na disk	613
16. Řízení spotřeby	625
17. Další informace	635
18. Plány do budoucna	643

Část V – Přílohy

A. Přehled signálů	653
B. Přehled systémových volání	657
C. Literatura	667

Obsah

Předmluva

27

ČÁST I

Úvod

KAPITOLA 1

O Linuxu obecně **31**

Co je to Linux 31

Linux a unixové systémy 31

Linux a POSIX 32

Linux a GNU 33

Verze jader 34

 Vývojový cyklus 35

KAPITOLA 2

Historie Linuxu **37**

Vznik Linuxu 37

Řady a verze 37

Důležité změny ve verzích řady 2.6 38

Budoucnost 39

KAPITOLA 3

Možnosti nasazení **41**

Servery 41

Desktopy 41

Výpočetní clustery, superpočítače 42

Embedded Linux 42

Upravená jádra 43

ČÁST II

Vnější rozhraní jádra

KAPITOLA 1

Úvod 47

Jádro a programy	48
Utility pro jádro	49
Získání a kompilace jádra	50

KAPITOLA 2

Start systému 53

Hardwarová kontrola systému	53
LinuxBIOS	53
Načtení a start zavaděče	54
Načtení a rozbalení jádra	55
Inicializace jádra	55
Spuštění programu init	56
Spouštění procesů	57

KAPITOLA 3

Správa modulů 59

Informace o modulech	59
Zavádění a odebírání modulů	60
Program modprobe a aliasy modulů	60
Parametry modulů	61
Závislosti mezi moduly	62
Příznaky tainted	63

KAPITOLA 4

Systémová volání 65

Princip systémových volání	65
Jak to funguje	66
Do režimu jádra a zpět	67
Předávání parametrů	68
Reakce na výsledek	68
Přerušená volání	69
Záhadná knihovna linux-gate.so	71
Oprávnění procesu	72

KAPITOLA 5

Signály 75

Princip	75
Druhy signálů	75
Posílání a doručování signálů	76
Reakce na signály	77
Blokace signálů	81
Problémy a jejich řešení	82

KAPITOLA 6

Souborové operace 83

Otevírání a zavírání souborů	84
Otevírací operace	84
Zavření otevřeného souboru	86
Čtení a zápis	87
Základní operace	87
Pozice v souboru, synchronizace	88
Volání sendfile a splice	89
Čekání na události	91
select/pselect	91
poll/ppoll	93
epoll	94
Řídicí operace	96
Funkce fcntl()	96
Funkce ioctl()	98
Zamykání souborů	99
flock()	100
lockf()	100
fcntl()	101
Asynchronní souborové operace	102
POSIX-kompatibilní AIO v uživatelském prostoru	102
Implementace AIO v jádře	105
Další operace na souborech	107
Operace souborového systému	107
Rozšířené otevírání	109

KAPITOLA 7

Souborové systémy 111

Virtuální souborový systém	111
Řetězec pro přístup k souborům	112

Konkrétní souborové systémy	113
ext2	113
ext3	114
reiserfs	114
xfs	115
procfs	115
sysfs	115
iso9660	115
fat	116
ntfs	116
nfs	117
tmpfs	117
FUSE	118
Připojování souborových systémů	118
Volání mount()	118
Volání umount() a umount2()	119
Informace o připojených systémech	120
Informace o souborech	121
Význam atributů	121
Funkce pro práci s atributy	123
Rozšířené atributy	124
Přístupové seznamy	125

KAPITOLA 8

Mapování souboru do paměti **129**

Princip mapování	129
Mapování pomocí mmap()	129
Namapování	130
Odmapování a další souvislosti	132
Zamykání paměti	132
Sdílená paměť System V IPC	133
Klíče	133
Otevření a namapování paměti	133
Odmapování a zrušení segmentu	134
Další operace	135

KAPITOLA 9

Roury a zprávy **137**

Anonymní roury	137
Jednoduchá roura do podprocesu	137
Roura s oběma konci	138
Komunikace s podprocesem	139
Komunikace mezi vlákny	139
Řízení procesu událostmi	139
Využití roury při obsluze signálu	140
Pojmenované roury	142

Komunikace pomocí zpráv	144
POSIX Message Queues	144
Zprávy System V IPC	145

KAPITOLA 10

Sockety **149**

Spojová komunikace	149
Na straně klienta	149
Na straně serveru	151
Unixová doména	153
Datagramová komunikace	154
Datagramový filtr	155
Porovnání rour a socketů	156

KAPITOLA 11

Procesy a vlákna **157**

Proces, vlákno, úloha	157
Plánování úloh	157
Plánovače	158
Priority běhu	159
Preemptivita jádra	159
Proces a jeho vlastnosti	160
Životní cyklus procesu	161
Implementace správy procesu	162
Plánování procesů	165
Pracujeme s vlákny	166
Životní cyklus vlákna	167
Základy práce s vlákny	167
Ukončování a úklid	169
Další operace s vláknem	170
Atributy vlákna	171
Volání clone()	172

KAPITOLA 12

Spouštění programů **175**

Spouštíme program	175
Funkce volané při ukončení programu	176
Jak jádro spouští program	177
Spouštění binárních programů typu ELF	178
Handler binfmt_misc	179
Spouštěcí domény	180

KAPITOLA 13

Synchronizační objekty 183

O futexech	183
Vzájemné vyloučení – mutex	183
Použití mutexu	184
Parametry mutexu	185
Zámek pro čtení a zápis – rwlock	186
Spinlock	187
Bariéra	187
Podmínková proměnná	188
Semafor	189
Sdílení synchronizačních objektů	190

KAPITOLA 14

Souborový systém procfs 191

Organizace procfs, základní informace	191
Informace o jednotlivých procesech	191
Ovládání jádra přes procfs	193
Další informace poskytované přes procfs	195

KAPITOLA 15

Souborový systém sysfs 197

Jádro a moduly	197
kernel	197
module	198
Třídy zařízení	198
block	199
class	199
Sběrnice a zařízení	200
bus	200
devices	201
Ostatní podstromy	203
firmware	203
fs203	
power	203

KAPITOLA 16

Některé další speciální souborové systémy 205

tmpfs	205
debugfs	206

configfs	206
usbfs	207
KAPITOLA 17	
Netlink	209
KAPITOLA 18	
udev, HAL	211
udev	211
Statické soubory zařízení	211
devfs a jeho neduhy	211
Princip systému udev	212
Konfigurace udev	214
HAL	215
Úkoly HAL	215
Technologie D-BUS	216
Model zařízení	216
Konfigurace	217
KAPITOLA 19	
inotify	219
Jak to funguje	219
Použití inotify	220
Omezení inotify	221
KAPITOLA 20	
kevent	223
Princip fungování	223
KAPITOLA 21	
FUSE	225
Vlastnosti FUSE	225
Použití souborového systému FUSE	226
Tvorba souborového systému pro FUSE	227
Některé existující souborové systémy	228
KAPITOLA 22	
Další informace, budoucnost	229
Volání ptrace	229
Rozhraní pro události (eventfd)	231

Signály	231
Časovače	232
Obecné události	232
Řízení spotřeby	233
Řízení kmitočtu procesorů	234
Zastavování disků	235
Uspávání a probouzení systému	236
Uspávání a probouzení obecně	237
Uspávání do paměti	237
Uspávání na disk	237
Jak uspávat	238
Odkládací prostor a zpožděná alokace paměti	239
Odložitelná paměť	239
Odkládací oddíly	240
Odkládací soubory	241
Zpožděná alokace paměti	241
Klávesa SysRq	243
TIPC	244
Adresace	245
Rozhraní	245

ČÁST III

Vývoj ovladačů

KAPITOLA 1

Úvod 249

Proč vyvíjet součást jádra 249

Co patří do jádra 250

Omezení platící v jádře 251

KAPITOLA 2

Licenční aspekty 253

Co říká GPL 253

Odvozené a souborné dílo 254

Uživatelský prostor a GPL 256

NDA a GPL 257

Exportované symboly 257

KAPITOLA 3

Základy vývoje pro jádro 259

Základní pravidla	259
V čem psát?	261
Standard psaní kódu	262
Jak kompilovat	264
Zavádění zkompilevaného modulu	265
Datové typy	265
Celočíselné typy	266
Struktury	266
Uniony a výčtové typy	267
Pole a ukazatele	267
Speciální typy	268
Seznamy	268
Další doporučení	270
Optimalizační makra	270
Mechanismy uložení dat	271
Algoritmy v jádře	271

KAPITOLA 4

Ladění a testování 273

Statické kontroly	273
Chyby a varování	273
Kontrola návratových hodnot	274
Adresy a ukazatele	274
Nástroj sparse	275
Výpisy jádra	276
Buffer výpisů jádra a práce s ním	276
Funkce printk()	277
Lepší práce se zprávami	278
Více informací	279
Sledování systémových volání	279
Chybová hlášení jádra	280
Úrovně hlášení	280
Využití v ovladačích	281
Vnitřní ladicí režimy	282
Kprobes	283
lockdep	284
Sledování jádra za běhu	285
Souborový systém debugfs	286
Práce s debugfs	286
Další možnosti debugfs	288

Simulace chyb	288
Ovládání	289
Přidání vlastních bodů selhání	289
Ladicí programy	290
kdb	290
kgdb	291
linice	292
Jádro v uživatelském prostoru	292
Instalace a zprovoznění	293
Použití	293

KAPITOLA 5

Základní operace **295**

Základní elementy modulu	295
Inicializace a úklid	296
Soubor Makefile	298
Parametry modulu	299
Jednoduché znakové zařízení	300
Major a minor čísla	301
Souborové operace	301
Registrace zařízení	304
Volání ioctl()	306
Kódování příkazů	306
Kontrola oprávnění	308

KAPITOLA 6

Paměť, její adresace a správa **309**

Adresní prostory	309
Fyzická adresa procesoru	309
Fyzická adresa sběrnice	310
Logická adresa jádra	310
Virtuální adresa jádra	311
Virtuální adresa procesu	311
Mapování logických adres	311
Paměťové stránky	312
Převody mezi adresami	313
Alokace paměti	314
Alokace logické paměti	314
Alokace stránek logické paměti	317
Využití lookaside cache	317
Zásobníky paměti	319
Alokace virtuální paměti	320

Paměťové a řetězcové operace	320
Kopírování dat v paměti	321
Další operace	321
Přenos mezi jádrem a procesem	321
Kontrola přístupu	322
Přenos jedné hodnoty	322
Přenos úseku paměti	323
Další operace v uživatelském prostoru	323
Oblasti virtuální paměti	324
Mapování paměti pro sdílení	325
Jednorázové namapování	325
Mapování při výpadku stránky	326
Mapování na základě adresy a délky	326

KAPITOLA 7

Komunikace s hardwarem **329**

Porty I/O	329
Operace na portech – starý způsob	330
Operace na portech – moderní způsob	331
Paměť zařízení	332
Bariéry	333

KAPITOLA 8

Synchronizace přístupu **335**

Spinlock	335
Spinlock s rozlišením čtení/zápis (rwlock)	337
Vzájemné vyloučení – mutex	338
Semafor	339
Semafor s rozlišením čtení/zápis	340
Dokončení (completion)	340
Atomické proměnné	341
Read-Copy-Update (RCU)	342
Sekvenční zámek (seqlock)	343
Sekvenční čísla a mutex	343
Bitové operace	344

KAPITOLA 9

Obsluha přerušení **345**

Kanály přerušení	345
Registrace obslužné rutiny (handleru)	346
Implementace handleru	348

Dvě poloviny obsluhy přerušení	349
Dočasný zákaz přerušení	349
Lokální zákaz všech přerušení	350
Detekce kanálu přerušení	350

KAPITOLA 10

Měření času, časovače	353
------------------------------	------------

Jádro a čas	353
Vyjádření času, systémový tik	354
Datové typy, převody, aritmetika	354
Zjišťování aktuálního času	357
Čekání	358
Vynucené naplánování úlohy	358
Uspání úlohy	359
Aktivní čekání	359
Čekání na událost (splnění podmínky)	360
Časovače	362

KAPITOLA 11

Odložené zpracování	365
----------------------------	------------

Tasklet	365
Pracovní fronta	367
Společná pracovní fronta	367
Samostatná pracovní fronta	368
Spouštění v kontextu procesu	369

KAPITOLA 12

DMA	371
------------	------------

Fungování DMA	371
Alokace prostředků	372
Alokace paměti	373
Abstraktní vrstva pro DMA	374
Zásobník paměti DMA	374
Proudové mapování	375
Mapování scatter/gather	376
Ovládání řadiče DMA	377

KAPITOLA 13

Objektový model jádra	379
------------------------------	------------

Proč objekty?	379
Architektura objektového modelu	380

Seznam – klist	380
Objekt – kobject	381
Třída – kobj_type	382
Seznam objektů – kset	383
Subsystém – subsystem	384
Atributy	384
Stínové adresáře	385
Druhy objektů	385
Sběrnice (bus)	385
Ovladač (device driver)	385
Zařízení (device)	386

KAPITOLA 14

Oznamování událostí do uživatelského prostoru **387**

Funkce v objektech	388
Explicitní posílání událostí	388
Příjem událostí	388

KAPITOLA 15

Načítání firmwaru **389**

Koncepce načítání firmwaru	389
Načítání v ovladači ze souboru	389
Načítání z uživatelského procesu	390
Načítání prostředky jádra	390
Příklad načítání	391

KAPITOLA 16

Bloková zařízení **393**

Vlastnosti blokových zařízení	393
Princip přístupu v jádře	393
Příprava ovladače	395
Zpracování požadavků	397
Ukončování požadavku	399
Operace s frontou	400
Vnitřní struktura požadavků	400
Zpracování požadavků bez fronty	401
Plánovače I/O	402
No-op scheduler	402
Deadline scheduler	402
Anticipatory scheduler	402
CFQ scheduler	403

Další operace	403
Přímý přístup	403
Výměnná média	403
Geometrie disku	404

KAPITOLA 17

Síťová zařízení **405**

Síťové zařízení v systému	405
Chování ovladače síťového zařízení	405
Jednoduchý ovladač	407
Ověření funkce	409
Práce v dotazovém režimu	410

KAPITOLA 18

Zařízení na sběrnici PCI **413**

O sběrnici	413
Identifikace zařízení	414
Informace o podporovaných zařízeních	414
Přístup do zařízení přes PCI	415
Přerušování a DMA	417
Přístup do konfiguračních registrů	417
Základ ovladače pro zařízení PCI	418
Funkce pro použití v ovladači	419
Přerušování signalizovaná zprávou	420
Zapnutí a vypnutí MSI	420
Zapnutí a vypnutí MSI-X	421

KAPITOLA 19

Zařízení na sběrnici USB **423**

USB v linuxovém jádře	423
Identifikace zařízení	425
Přenosy přes USB	425
Koncové body, rozhraní, konfigurace	426
Bloky požadavků (URB)	427
Vytvoření URB a zpracování dat	427
Přípravné činnosti	429
Identifikace zařízení	429
Příprava pro virtuální zařízení	430
Inicializace a odstranění virtuálního zařízení	431

Registrace ovladače	432
Přenosy bez URB	433

KAPITOLA 20

Tvorba uživatelského rozhraní **435**

Využití souborového systému sysfs	435
Rozhraní pro snadný přístup	436
Využití funkce ioctl()	437
Řízení speciálními daty	439
Speciální systémová volání	439
Funkce syscall()	440
Makra pro systémová volání	440

KAPITOLA 21

Ovladače souborových systémů **443**

Virtuální souborový systém (VFS)	443
Implementace souborového systému	444
Jednoduchý systém s pevnými soubory	445
Souborové operace	445
Datové struktury souborů	447
Základní operace souborového systému	448
Inicializace ovladače	449
Vyzkoušení ovladače	449
Další operace v souborovém systému	450
Vytvoření i-node	451
Vytvoření adresářové položky	451
Vytvoření souboru	452
Příprava superbloku a kořenového adresáře	453
Ověření funkce	454

KAPITOLA 22

Souborové systémy pomocí FUSE **455**

Základy implementace souborového systému	456
Některé důležité operace	457
Jednoduchý systém s pevnými soubory	457
Další operace	460
Práce na nižší úrovni FUSE	462

KAPITOLA 23

Zajímavosti, plány do budoucna **465**

Odstranění taskletů	465
----------------------------	------------

Funkce fault()	466
Předalokace bloků (fallocate())	467

ČÁST IV

Pohled dovnitř jádra

KAPITOLA 1

Úvod **471**

Proč znát fungování	471
----------------------------	------------

Jak se měnila implementace	472
-----------------------------------	------------

KAPITOLA 2

Zdrojové kódy jádra **475**

Vanilla a distribuční jádra	475
------------------------------------	------------

Záplaty (patche)	475
-------------------------	------------

Vytváření záplat	476
------------------	-----

Posílání záplat vývojářům	476
---------------------------	-----

Aplikace záplat na zdrojové kódy	477
----------------------------------	-----

Získání souborů	479
------------------------	------------

Práce se systémem git	479
------------------------------	------------

Použití gitu se zdrojovými kódy jádra	480
---------------------------------------	-----

Struktura adresářů	483
---------------------------	------------

KAPITOLA 3

Architektura jádra **485**

Rozdělení součástí a subsystémů	485
--	------------

Modularita jádra	487
-------------------------	------------

Implementace modulu	488
---------------------	-----

Zavádění a uvolňování modulů	490
------------------------------	-----

Zavádění na žádost, automatické zavádění	491
--	-----

KAPITOLA 4

Správa paměti **493**

Stránkování paměti	494
---------------------------	------------

Manipulační funkce	494
--------------------	-----

Zóny paměti	495
--------------------	------------

Zóny na architektuře x86	496
--------------------------	-----

Zóny na architektuře x86_64	496
-----------------------------	-----

Trvalé mapování stránek v horní paměti	496
--	-----

Dočasné mapování stránek v horní paměti	498
---	-----

Alokace stránek	499
Algoritmus buddy	499
Rezervované stránky	500
Cache pro stránky	500
Proces alokace	500
Sloučené stránky	501
Uvolňování stránek	502
Alokátory SLAB, SLUB a SLOB	503
Alokátor SLAB	503
Alokátor SLUB	504
Alokátor SLOB	505
Alokace nesouvislých úseků	505
Alokace paměti	505
Uvolnění alokované paměti	506
Virtuální mapování	507
NUMA	507
NUMA v linuxovém jádře	508
Řešení nedostatku paměti	509
Reverzní mapování	510
Základní mechanismus uvolňování	510
OOM killer	512

KAPITOLA 5

Správa procesů a vláken **515**

Úlohy, procesy, vlákna	515
Vlastnosti a stav úlohy	515
Stav úlohy	516
Příznaky úlohy, další datové struktury	517
Vztahy mezi úlohami	518
Vytváření a rušení úloh	520
Kopírování úlohy	520
Ukončení běhu úlohy	521
Úlohy a plánovač(e)	522
Základní principy	522
Fungování plánovače	523
Priority a časová kvanta	524
Úloha idle	525
Přepínání kontextu	525
Implementace přepnutí kontextu	526
Požadavky na přepnutí kontextu	527
Přesuny úloh, vyvažování zátěže	528
Přesun úlohy mezi frontami	528
Plánovací skupiny a domény	529
Vyvažování zátěže	530
Vlákna/procesy jádra	532
Implementace vláken jádra	532

KAPITOLA 6

Paměť procesu 535**Deskriptor paměti 535****Oblasti virtuální paměti 536**

Seznam VMA a červeno-černý strom 537

Operace s VMA 538

Práce s adresním prostorem procesu 538

Vytváření adresního prostoru nového procesu 539

Vytváření adresního prostoru pro spouštěný program 539

Systémová volání 540

Manipulační volání 540

Mapovací volání 541

Zamykání a další operace 542

Řešení výpadků stránek 543

Výpadky řešené funkcemi VMA 546

Výpadek anonymní stránky 546

Výpadek při zápisu do stránky COW 546

KAPITOLA 7

Systémová volání a signály 549**Průchod volání jádrem 549**

Základní činnost v rámci volání 549

Práce prováděné při návratu z volání 551

Virtuální volání 552

Signály 552

Signály v jádře 552

Poslání signálu z procesu 553

Poslání signálu z jádra 554

Reakce na signál 554

Synchronní reakce 556

KAPITOLA 8

Přerušování a výjimky 557**Přerušování a jejich obsluha 557**

Obsluha přerušování 558

Výjimky a jejich obsluha 560**Softwarová přerušování a tasklety 561**

Vygenerování a obsluha přerušování 561

Démon ksoftirqd 562

Tasklety 562

Pracovní fronty 563

Samostatné fronty 563

Společná fronta 564

Získávání entropie 564

KAPITOLA 9

DMA 565**Správa paměti pro DMA 565**

Koherentní paměť 565

Nekoherentní paměť 566

Proudové mapování 566

DMA pro sběrnici ISA 566

Ovládání řadiče DMA 567

Rezervace kanálu DMA 567

KAPITOLA 10

Časovače 569**Obecně o časovačích 569**

Druhy časů 569

Časovače podle účelu 570

Časovače pro jádro 570

Báze časovačů 570

Funkce po manipulaci s časovači 571

Obsluha vypršelých časovačů 572

Časovače pro uživatelský prostor 572

Rozhraní pro časovače 572

Správa časovačů 573

Makro CLOCK_DISPATCH 573

Nastavování časovačů 574

Události při vypršení 574

Časovače s vysokým rozlišením 574

Implementace hrtimers 575

Rozhraní pro použití 575

Obsluha vypršelých časovačů 576

Variabilní délka tiků procesorů 576**Odložitelné časovače 577**

KAPITOLA 11

Soubory a souborové systémy 579**VFS a jeho základní funkce 579****Souborové systémy a jejich připojování 580**

Připojování souborového systému 580

Odpojování souborového systému 581

Vyhodnocování souborových cest 582**Typy souborů 583****Operace se soubory 584**

Otevření a zavření souboru 584

Čtení, zápis, posun pozice	585
Další operace	586
Zamykání souborů	587
FUSE	588
KAPITOLA 12	
Síťová komunikace	591
Socket a jeho správa	591
Socket a jeho operace	591
Vytvoření a zrušení socketu	592
Další socketové operace	593
Cesta paketů jádrem	594
Implementace rodiny protokolů	594
Implementace protokolu	595
Jak dál?	595
Netlink	595
Generické rozhraní	596
Odesílání dat	597
KAPITOLA 13	
Plánování I/O	599
Funkce blokové vrstvy	599
Příprava blokových požadavků	599
Vkládání do fronty	600
Připojování a odpojování fronty	600
Plánovače a jejich druhy	601
No-op scheduler	601
Deadline scheduler	601
CFQ scheduler	602
Anticipatory scheduler	602
KAPITOLA 14	
Disková cache	605
Princip a chování	605
Implementace	606
Operace čtení ze souboru	607
Hledání v cache	608
Vektory stránek a značky	608
Operace zápisu	609
Démon pdflush	610
Vlákna démonu	611
Činnost démonu	611
Úlohy pro démon	611

KAPITOLA 15

Odkládání na disk 613**Odkládací zařízení a soubory 613**

Datové struktury pro odkládání 614

Vytváření oblastí, připojování a odpojování 615

Mechanismus odkládání 617

Hledání volného slotu 617

Odkládání stránek 618

Načítání stránek 619

Cache pro odkládání 621

Pomocný démon kswapd 622

KAPITOLA 16

Řízení spotřeby 625**Uspávání a probouzení systému 625**

Podpora v ovladačích 625

Zmrazování a rozmrazování procesů 626

Celkový rámec uspávání 628

Uspávání na disk 629

Změna frekvence procesorů 630

Architektura ovladače 630

Nízkoúrovňový ovladač 631

Řídící objekt (governor) 632

ACPI 632

Podpora pro uspávání 633

Zjišťování a nastavování hodnot 633

Generování událostí 634

KAPITOLA 17

Další informace 635**Šifrování, hashe, komprese 635**

Kryptografické rozhraní 635

Správce algoritmů 636

Algoritmy 636

Transformační objekty 637

Vnější rozhraní 637

Šablony 638

Video For Linux 2 639

Vnější a vnitřní rozhraní 639

Fungování ovladače 640

Specifické operace 641

Přenos dat 641

KAPITOLA 18

Plány do budoucna 643

Plánovač CFS	643
Souborový systém LogFS	644
Ovladače v uživatelském prostoru	645
Miniovladač v jádře	646
Ovladač v uživatelském prostoru	646
Ostatní věci a obecné trendy	647
Slovo závěrem	648

ČÁST V

Přílohy

PŘÍLOHA A

Přehled signálů 653

PŘÍLOHA B

Přehled systémových volání 657

Obecná volání	657
Správa procesů	658
Plánování procesů a vláken	659
Oprávnění procesu, využití prostředků	659
Signály a jejich obsluha	660
Základní souborové operace	661
Informace o souborech, přístupová práva	662
Události na souborových deskriptorech	663
Další souborové operace, IPC	664
Souborové systémy	664
Správa paměti a mapování	665
Zastaralá volání (jen pro informaci – nepoužívat)	666

PŘÍLOHA C

Literatura 667**Rejstřík 669**

Předmluva

Pryč jsou doby, kdy Linux byl používán výhradně studenty univerzit s technickým zaměřením a počítačovými nadšenci. Většina uživatelů dnes plně využívá přítulnosti, kterou moderní linuxové distribuce nabízejí – pohodlnou instalaci a konfiguraci, dostupnost pokročilých nástrojů pro kancelářskou práci, přehrávání multimédií, hraní her.

Důsledkem tohoto vývoje ovšem je, že existuje poměrně značné množství dokumentace a knih, které se věnují výhradně uživatelské stránce používání Linuxu, ale programátoři, kteří chtějí začít s vývojem programů, který by vyžadoval hlubší nízkoúrovňové znalosti systému (ať už systémových knihoven, nebo přímo jádra operačního systému) nemají vždy k dispozici dostatek kvalitní literatury, a to často ani cizojazyčné.

Kniha, kterou držíte v ruce, je v tomto ohledu zcela unikátní publikací – jedná se o první knihu v českém jazyce, která se velmi detailně zabývá problematikou nízkoúrovňové architektury operačního systému Linux.

Tematicky je rozdělena do dvou částí. První se zabývá systémem těsně nad jádrem, tedy systémovými voláními, ladicími technikami a obecně rozhraním mezi uživatelskými procesy a linuxovým jádrem. Čtenář se zde kromě jiného také dozví podrobnosti o signálech, API k souborovému systému (včetně linuxových specialit, které se na jiných UNIXových systémech nevyskytují).

Druhá část se velmi podrobně zabývá architekturou linuxového jádra jako takového a poskytuje čtenáři široké spektrum informací. To poskytuje velmi pevné základy každému, kdo se chce stát vývojářem linuxového jádra, nebo chce mít alespoň detailní přehled o tom, jak linuxové jádro funguje pod pokličkou. Příjemný bonus kniha přináší i v tom, že část zabývající se psaním ovladačů zařízení nepředpokládá u čtenáře hluboké znalosti hardwarových architektur a detailní principy jejich fungování, ale vždy nejprve dojde k vysvětlení příslušných pojmů a zasazení do příslušného kontextu.

Knihu lze doporučit jak všem zvědavým uživatelům, kteří chtějí porozumět tomu, co se děje v zákulisí jejich operačního systému, tak i programátorům, kteří plánují napsat svůj první ovladač zařízení pro Linux či svůj první souborový systém. Díky jejímu širokému záběru (a přitom detailnímu zpracování jednotlivých částí) v ní však i velmi zkušení systémoví programátoři zcela jistě naleznou mnoho nových informací a podnětů.

Mgr. Jiří Kosina

Linux kernel team

SUSE Labs, Novell Inc.