

Novosti u novoj generaciji Linux kernela 2.6

Matija Nalis,
Vodatel
<mnalis@vodatel.hr>



Sadržaj

1. Priča do sada...
2. Platforme
3. Datotečni sustavi
4. Mreža
5. Interaktivnost i skalabilnost
6. Ulazno/izlazni podsustav

1. Prijenosnici i senzori
2. Podignuti limiti
3. Sigurnost
4. Razno
5. Testovi i održavanje
6. Problemi
7. Ostali korisni URLovi

Priča do sada...

Linux 1.0 je krenuo kao OS za jednoprocesorske i386 kompatibilne mašine, da bi verzija 1.2 donijela podršku za prve dodatne arhitekture (Alpha, Sparc, Mips). Linux kernel 2.0 uvodi nove arhitekture, a što je najbitnije i podršku za multiprocesorske sustave (SMP). 2.2 donosi poboljšanja u skalabilnosti i značajne promjene u VFSu, a 2.4 ubacuje Linux u svijet desktopa (podrška za ISA Plug-and-Play, USB, PCMCIA).

Platforme

- Razne nove arhitekture
 - Voyager, ES7000, PC9800,...
 - Naniže (Embedded)
 - Integracija uCLinux-a
 - Hitachi's H8/300, NEC v850, sh, Motorola embedded m68k (u PalmPilota do Palm III npr.),
 - Naviše
 - x86-64 (AMD Hammer), ppc64
 - Big Iron (s390x,..)
- User-mode Linux

Datotečni sustavi

- Podrška za više datotečnih sustava
 - ext2, ext3, reiserfs, jfs, xfs, minix, romfs, iso9660, udf, msdos, vfat, ntfs (ro), adfs, amiga ffs, apple macintosh hfs, BeOS befs (ro), bfs, efs (ro), cramfs, free vxfs, os/2 hpfs, qnx4fs, sysvfs, ufs, ...
- NFSv4, AFS, InterMezzo – mrežni datotečni sustavi
- Poboljšanja postojećih datotečnih sustava (manje općeniti lockovi, ubrzanja)
- Ext3 - Orlov allocator, dir_index
- POSIX ACLovi

Mreža

- kernel IPsec
- Bridging firewall
- RFC 2960 - SCTP (Stream Control Transmission Protocol)
- RFC3173 (IP Payload Compression)
RFC3041 (IPv6 Privacy Extensions)
RFC2473 (IPv6 in IPv6 tunnels)

Interaktivnost i skalabilnost

- Bolja interaktivnost, manja kašnjenja
- O(1) scheduler
 - brzina neovisna o broju procesa
 - bolja SMP skalabilnost
 - bolje korištenje HyperThreadinga (razlikovanje fizičkih i virtualnih procesora)
- CONFIG_PREEMPT
 - prekidanje procesa usred kernela
 - dijelovi kernela su još uvijek pod lockovima
- Još manje BKLova, granularniji lockovi
- Rik van Riel – rmap (Reverse Mapping) - bolje ponašanje VMa pod opterećenjem
- Processor affinity

Ulazno/izlazni podsustav

- BIO (Block I/O) – iznova napisan
- LVM2 – DeviceMapper
- (U)DMA za CD pisanje / čitanje audio traka
- IO elevators – elevator=anticipatory, deadline
- Kernel Asynchronous I/O (AIO) podrška
- ATA TCQ (Tagged Command Queueing)
- human interface layer (touchscreens, braille readers, ...)
- ALSA - Advanced Linux Sound Architecture
- USB 2.0
- bluetooth

Prijenosnici i senzori

- power management ACPI (spuštanje brzine CPUa)
- Software suspend-to-disk
- Lmsensors – uključen u kernel
- IPMI (Intelligent Platform Management Interface)
 - <http://openipmi.sourceforge.net>
 - <http://www.intel.com/design/servers/ipmi/spec.htm>

Podignuti limiti

- Datotečni sustavi i blok uređaji - dignuti limiti maksimalne veličine datoteka i cijelog datotečnog sustava (do 16TB IA32, 8EB 64bit)
- UID i GID – sa 65 tisuća na 4 milijarde
- Broj procesa (PID) sa 32 tisuće na 1 milijardu
- Broj uređaja (/dev) s 255 uređaja s 255 poduređaja svaki, na 4 tisuće uređaja s milion poduređaja svaki

Sigurnost

- root više ne mora biti svemoćan
- LSM (Linux Security Module)
- NSA Security-Enhanced Linux (SELinux) - <http://www.nsa.gov/selinux>

Razno

- Poboljšani sustav za komapajliranje kernela – brži, efikasniji, čitljiviji
- brži syscallovi (SYSENER) – zahtjeva ažurirani glibc
- IPVS (IP Virtual Server) - <http://www.linuxvirtualserver.org/>
- Windows Logical Disk Manager ("Dynamic Disks").
- kernel podrška za Native Posix Threading Library
 - <http://people.redhat.com/drepper/nptl-design.pdf>
 - Fast userspace mutexes (Futexes).
 - <http://people.redhat.com/drepper/futex.pdf>
 - <http://ds9a.nl/futex-manpages>
- Više podrške za PCI hotplug uređaje
- Debug opcije, uključen kksymoops

Testovi i održavanje

- Bitkeeper
- Linux Test Project
 - <http://ltp.sourceforge.net/>
- OSDL (Open Source Development Labs)
 - <http://www.osdl.org/>
 - Nightly kernel regression testing
 - Bugzilla - <http://bugme.osdl.org/>

Problemi

- Potpuno novi sustav modula – `module_init_tools`
- Zastarjeli pozivi (`bdflush`,...)
- Paziti prilikom preuzimanja starog (2.4) `.config` (`CONFIG_VT` i sl.)
- RPM, `export LD_ASSUME_KERNEL=2.2.5`
- Direktno bootanje s diskete više nije podržano
- Zbog $O(1)$ schedulera nice od -10 (kao kod Debianovog X servera ponekad) ima puno veći značaj nego do sada, moguće preskakanje mp3-eva ili sl. – popraviti prioritete
- imena etherneteta ponekad imaju drugi redosljed:
<http://xenotime.net/linux/doc/network-interface-names.txt>

Ostali korisni URLovi

- <http://www.kerneltraffic.org/>
- LKML - <http://www.tux.org/lkml>
- ACLovi - <http://acl.bestbits.at>
- Robert Love - <http://tech9.net/rml/schedutils>
- ALSA - <http://www.alsa-project.org/>
- uClinux - <http://www.uclinux.org/>
- NFSv4 <http://nfs.sourceforge.net/>
- Async IO - <http://lse.sf.net/io/aio.html>
- Bridging firewall - <http://bridge.sourceforge.net>
- “Read the Source, Luke”