

Kernel-Log

Linux 5.1: Performance-Verbesserungen und neue Speichertechnik

Mit dem neuen Linux-Kernel wird es attraktiver, Datenträger mittels asynchronem I/O (AIO) anzusprechen, wie es in der Windows-Welt schon lange üblich ist. Außerdem geht es einem in der Anfangszeit von Linux zentralen Dateiformat jetzt an den Kragen. Die Entwickler wollen zudem die Unterstützung für Grafikerne von ARM verbessern.

Von Thorsten Leemhuis

Das jetzt erhältliche Linux 5.1 geht eine altbekannte Schwachstelle des Kernels an, denn die neue Version bringt endlich eine zeitgemäße Infrastruktur, um Daten asynchron zu lesen oder zu schreiben. Bei diesem „Asynchronous I/O“ (AIO) können Programme eine Übertragung initiieren und dann andere Dinge erledigen, während der Kernel die I/O-Operation ausführt. Unter Windows arbeiten Programme häufig auf diese Weise, weil sie Performance und Reaktionsfreude verbessern kann. Die meiste Linux-Software macht sich AIO nicht zunutze; das liegt zum Teil an historischen Gründen, vor allem aber an der bisherigen AIO-Infrastruktur des Kernels, die allerlei Tücken birgt und daher seit Langem in der Kritik steht.

Das neue „io_uring“ soll Linux hier endlich nach vorne bringen. Der Austausch von I/O-Operationen und -Daten mit dem Userspace erfolgt dabei über separate Ringbuffer. Damit Anwendungsentwickler die neue Infrastruktur leicht verwenden können, stellt der zuständige Entwickler eine Userspace-Bibliothek bereit. Mit dem Gesamtpaket positioniert sich der Kernel letztlich gegen Lösungen wie das Storage Performance Development Kit (SPDK), die den Kernel bei I/O-Operationen umgeht – ähnlich, wie das bekanntere und in Kernel-Kreisen ebenfalls kritisch beäugte Data Plane Develop-

ment Kit (DPDK), das den Netzwerkstack des Kernels ausbootet.

Fest im Speicher

Linux 5.1 kann die Speicherkapazität nichtflüchtiger Speichermodule (NVDIMMs) jetzt so einbinden, dass Programme oder der Kernel diesen Bereich wie normalen Arbeitsspeicher ansprechen können. Durch dieses von Intel-Entwicklern vorangetriebene Feature müssen Anwendungen nicht mehr groß angepasst werden, um Persistent Memory (PMEM) zu verwenden, bei dem Daten auch beim Kappen der Stromversorgung nicht verloren gehen; zu solch raren und meist teuren Speichermodule zählt etwa Intels im April vorgestelltes Optane DC Persistent Memory.

Anwendungen können mit Fanotify jetzt einfacher und zuverlässiger überwachen, ob sich etwas an Dateisystemen ändert, denn die Entwickler sind einige altbekannte Schwachstellen des bisherigen Interface angegangen. Bei Btrfs kann man beim Komprimieren mit Zstd (Zstandard) jetzt den Kompressionslevel vorgeben, um entweder Performance- oder Packdichte zu steigern.

Entrümpeln

Linus Torvalds hat eigenhändig Vorbereitungen getroffen, um die Unterstützung für das Dateiformat a.out (Assembler Output) zu entfernen. Dieses wurde in der Anfangszeit von Linux für ausführbare Dateien und Objektdateien genutzt. Vor mehr als 25 Jahren lernte Linux dann ELF (Executable and Linking Format), das schnell zum dominierenden Format aufstieg. Daher war auch niemand aufgefallen, dass der Kernel-Code für Speicherabzüge (Core Dumps) von a.out-Executables nicht mehr korrekt arbeitet. Torvalds hat ihn daher kurzerhand bei 5.1 entfernt und den a.out-Support auf x86-Systemen lahmgelegt. Alles sieht danach aus, dass der bald rausfliegt. Sicher ist das aber noch nicht, denn das wäre eine „Regression“. Solche versucht Torvalds tunlichst

zu vermeiden – selbst abgekanzelte Funktionen bleiben daher im Kernel, solange nur eine Person glaubhaft (und zeitnah) belegt, das Feature mit aktuellen Linux-Versionen produktiv (und nicht nur im Museum oder so) einzusetzen.

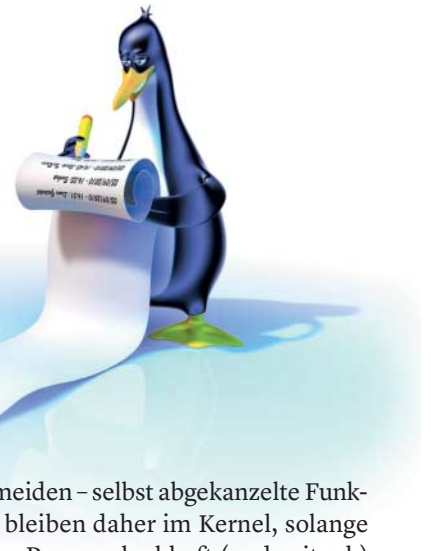
Einen Meilenstein gab es bei den Umbauten zur Beseitigung des Jahr-2038-Problems, das 32-Bit-Prozessoren betrifft: Die Kernel-Entwickler haben einen Schwung neuer Funktionsaufrufe (Syscalls) eingebaut, die alten ähneln, aber 64-Bit-Zeitwerte verstehen. Entwickler von Systembibliotheken und Anwendungen können ihre Software darauf anpassen, damit sie 2038 nicht plötzlich denkt, es sei 1970 oder gar 1901; das ist auch langsam überfällig, schließlich werden 32-Bit-ARM-Kerne aktuell noch in Autos oder anderen Embedded Systems verwendet, die wahrscheinlich auch in 19 Jahren noch im Einsatz sind.

Beim Kernel Live Patching (KLP), mit dem sich kleinere Kernel-Änderungen (etwa Sicherheitskorrekturen) im Betrieb vornehmen lassen, kann ein Sammel-Patch jetzt alle vorher angewendeten Live-Patches ersetzen. Das ist einem „Atomic Replace“ genannten Ansatz zu verdanken, der die Entwicklung kumulativer Live-Patches erheblich erleichtert.

Memory Compaction, das im Hintergrund der Fragmentierung des Arbeitsspeichers entgegenwirkt, verspricht jetzt deutlich größere Erfolge zu erzielen; dabei soll es den Prozessor sogar weniger belasten als zuvor und kann so die Performance verbessern.

Flexible Sicherheit

Große Fortschritte gab es bei den seit über fünfzehn Jahren vorangetriebenen Bemühungen, Linux Security Modules (LSMs) wie SELinux, AppArmor, Smack durch „Stacking“ parallel zu nutzen. Nachdem das mit einzelnen kleineren LSMs schon klappte, sind jetzt die wesentlichen Vorarbeiten in Linux 5.1 eingeflossen, durch die das bald mit allen LSMs gelingen soll.



Dabei entstand der neue Boot-Parameter `lsmod`, über den man das oder die aktiven LSMs jetzt beim Systemstart festlegen kann. Ein Stacking funktioniert bei 5.1 aber nur mit einigen der kleineren, wenig bekannten LSMs; langfristig gelingt es dann aber vielleicht, auf einem mit SELinux arbeitenden System etwa die Schutztechniken des Paketformats Snap nutzen zu können, das dazu AppArmor erfordert.

Netzwerk

Nachdem einzelne Treiber bereits „Air-time Fairness“ boten, wird diese Technik nun vom WLAN-Stack des Kernels unterstützt. WLAN-Treiber, die sich das zunutze machen, können das Funkspektrum besser ausnutzen und damit den Summendurchsatz in WLAN-Funkzellen steigern: Langsame Gegenstellen, also alte oder welche mit aktuell schlechter Verbindung, bekommen durch die Technik seltener eine Sendegelegenheit, sodass flottere Clients häufig zum Zug kommen.

Auch Version 5.1 des Linux-Kernel steigert wieder die Menge der unterstützten Hardware, denn die Entwickler haben Dutzende neue Treiber eingebaut und Hunderte verbessert. Dadurch läuft Linux jetzt beispielsweise von Haus auf dem Raspberry Pi 3 Model A+. Außerdem unterstützt es Intel Wi-Fi-6-Chips (802.11ac) der Serie Wireless-AX 22260 sowie Mediateks Wi-Fi-4-Chips (802.11n) MT7603E und MT7628/7688.

Neu dabei ist auch ein Treiber für den KI-Inferencing-Prozessor Habana Goya, der auf Beschleunigerkarten für den PCIe-Slot sitzt und mit Nvidias Tesla-Grafikkernen für Server konkurriert. Im Vorfeld wurde die Aufnahme des Treibers kontro-



Durch viele neue und verbesserte Treiber unterstützt der neue Kernel jetzt etwa den Raspberry Pi 3 Model A+ von Haus aus.

vers diskutiert, weil sich einige Entwickler von Kernel-Grafiktreiber als zuständig betrachteten, aber umgangen wurden. Ferner verhalte auch ihre mit schlechten Erfahrungen bei Grafiktreibern begründete Forderung, den Goya-Treiber außen vor zulassen, solange Open-Source-Programme und -Bibliotheken fehlen, um ihn selbst und seine Interaktion mit dem Userspace zu testen.

Schöner Booten

Einige Systeme mit moderner Intel-Grafik starten mit 5.1 schneller und eleganter, denn Intels Entwickler haben den „Fastboot“-Support in ihrem Grafiktreiber aktiviert. Einzelne Distributionen realisieren damit einen Startprozess auf modernen PCs, der Bildaussetzer, Flackern oder harte Übergänge vermeidet, wodurch er letztlich so geschmeidig verläuft wie bei modernen Macs.

PCs Radeon-GPUs der Vega-Generation laufen mit 5.1 sparsamer; das ist dem Support für eine „ZeroCore Power Mode“ oder „BACO“ (Bus Active, Chip Off) genannte Funktion in AMDs Grafiktreiber zu verdanken, durch die der Kernel den Grafikern weitgehend schlafen legen kann, wenn es nichts zu tun gibt. Der quelloffenen Nvidia-Treiber Nouveau beherrscht jetzt Heterogenous Memory Management (HMM) und kann dadurch etwa Speicherinhalte zwischen Arbeits- und Grafikspeicher spiegeln, was Programmieren den Rückgriff auf GPUs zu Rechenzwecken erleichtert.

ARM hat den Grafiktreiber Komedas beigesteuert, der den Mali-D71 und dessen Nachfolger unterstützt. Das sind „Display Processors“, die vornehmlich für die Bildausgabe gedacht sind und keine 3D-Einheit enthalten. Letztere sind leistungsfähigeren Mali-Modellen vorbehalten, die geläufiger sind und sich auf so manchem Einplatinencomputer finden. Solche GPU unterstützt Linux bislang nicht von Haus aus. Das sollen die Kernel-Treiber Lima (für Utgard-GPUs) und Panfrost (Midgard und Bifrost) ändern, die zur Aufnahme bei Linux 5.2 vorgesehen sind; darauf bauen gleichnamige OpenGL-Treiber zur 3D-Beschleunigung auf, die Mesa 19.1 mitbringen wird. Bei der Entwicklung dieser Treiber hat ARM untätig danebengestanden. Sie entstanden mithilfe von Hardware- und Treiber-Analysen (Reverse Engineering), daher unterstützen sie nur ausgewählte GPUs und Funktionen. Wenn alles wie gewohnt läuft, dürfte die um diese Treiber erweiterte Kernel-Version Mitte Juli erscheinen. (thl@ct.de) **ct**

Anzeige