Kernel-Log

Linux 4.18: Raspi-3B-Support und neue Firewall-Technik

Der neue Linux-Kernel unterstützt USB 3.2. Verbessert wurde auch die Live-Migration für VMs, die Netzwerk-Hardware des Hosts nutzen. Dank Rauswurf des Dateisystems Lustre schrumpft der Kernel abermals. Neu dabei ist auch ein Treiber für den Nachfolger der Grafikeinheit des Raspberry Pi.

Von Thorsten Leemhuis

Ein Highlight des Mitte August erschienenen Linux-Kernel 4.18 sind erste Teile des Bpfilter, einer neuen Paketfiltertechnik für Firewalls. Aber keine Angst, Sie brauchen sich nicht um Ihr Iptables- oder Nftables/nft-Know-how zu sorgen: Bpfilter ersetzt lediglich den Unterbau im Kernel, auf den das altbekannte Iptables und sein designierter Nachfolger aufbauen. Das soll beide beschleunigen.

Noch liegt dieses Ziel aber in weiter Ferne, denn die Entwickler haben bei 4.18 nur erste Teile des Fundaments für Bpfilter gelegt. Das enthält eine weitere größere Neuerung, die mittel- und langfristig womöglich größere Bedeutung bekommt: eine Infrastruktur, um im Rahmen der Linux-Quellen entwickelte Werkzeuge in Kernel-Module zu verpacken. Das ist für Helferprogramme gedacht, die der Kernel wie ein normaler Userspace-Prozess ausführt; sie laufen daher mit geringeren Privilegien, sodass Angreifer darüber nicht gleich das ganze System übernehmen können. Das verhilft dem eher monolithischen Linux-Kernel letztlich zu einer Infrastruktur, die prinzipiell eine Modularisierung in der Art von Microkerneln ermöglicht. Details dazu finden sich in c't 15/2018, S. 34.

Support für Raspi 3B

Die neue Linux-Version ist die erste, die von Haus aus den Raspberry Pi 3B und den 3B+ halbwegs ordentlich unterstützt. Zu verdanken ist das unter anderem dem Support für den GPIO-Expander der zwei Kleinstcomputer, durch den sich jetzt die Eingabe- und Ausgabekontakte steuern lassen, die WLAN, Bluetooth, HDMI-Erkennung und Aktivitäts-LED verwenden. Diese und weitere Fortschritte beim Raspi-3B-Support sind für Distributionen wie Debian, Fedora & Co. wichtig, die Raspis mit ihren regulären ARM-/ARM64-Kerneln unterstützen wollen. Raspbian & Co. juckt das weniger, denn sie haben solche Treiber schon länger in ihre Kernel eingebaut.

Linux 4.18 bringt zudem allerlei Grundlagen zum Support des Qualcomm-Prozessors Snapdragon 845, der in einigen mit Windows ausgelieferten ARM-Notebooks steckt. Das hat Linus Torvalds interessiert aufhorchen lassen: Der Linux-Erfinder hofft schon länger auf Notebooks mit ARM-Prozessoren, die eine halbwegs mit x86-Geräten vergleichbare Leistung liefern und zugleich von Linux ordentlich unterstützt werden. Noch fehlen dazu aber mehrere Treiber, von denen einige aber schon bei 4.19 folgen sollen; darunter auch einer für die Adreno-6x0-Grafik des SoC (System-on-Chip).

Eine Reihe kleiner Umbauten verbessert den Schutz vor der Prozessor-Sicherheitslücke Spectre v1. Die Entwickler haben zudem den Spectre-v4-Schutz für AMD-Prozessoren optimiert. Außerdem schützt Linux nun endlich von Spectre v2 betroffene 32-Bit-ARM-CPUs vor dem Ausnutzen der zu Jahresanfang bekannt gewordenen Lücke. Nach wie vor schwelende Schwierigkeiten rund um den Betreuer des ARM32-Codes sind der Grund, warum der offizielle Kernel erst jetzt Gegenmaßnahmen erhalten hat, obwohl solche schon länger kursieren. Dem 32-Bit-x86-Code von Linux fehlen nach wie vor Gegenmaßnahmen für die Prozessorlücke Meltdown; nach einer längeren Entwicklungsphase sollen diese aber wahrscheinlich in Linux 4.19 einfließen.

In einem Container definierte und dort mit Root-Rechten ausgestattete Anwender dürften dort jetzt eigenmächtig Dateisysteme per FUSE (Filesystem in Userspace) mounten. Aus Sicherheitsgründen ist es aber weiter nicht möglich, vom Kernel-Code unterstützte Dateisysteme wie Btrfs, Ext4, FAT oder XFS einzuhängen: Das soll Ausbrüche mithilfe von Image-Dateien verhindern, bei denen Angreifer die Dateisystemstrukturen manipuliert haben, um vom Container aus den Host-Kernel aus dem Tritt zu bringen. Apropos Btrfs: Eine Optimierung an Btrfs Send/Receive hat bei gezielten Tests zu einem dramatischen Performance-Zuwachs geführt. Außerdem lassen sich leere Snapshots jetzt mit einem simplen `rmdir` selbst ohne Root-Rechte entfernen.

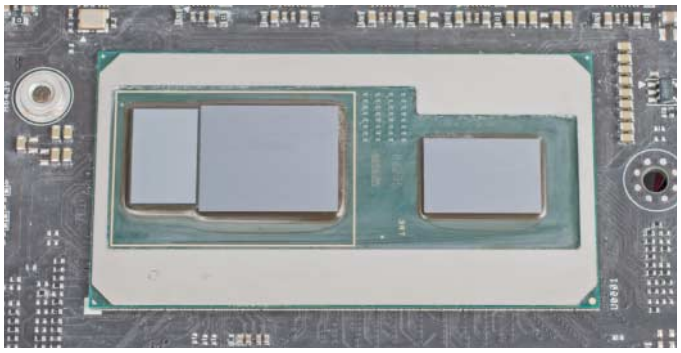
AMD-Treiber für Intel-CPU

Der Amdgpu-Treiber unterstützt nun auch den Grafikprozessor AMD Radeon RX Vega M, der auf Intel-AMD-Kombiprozessoren der Core-i-8000er-Familie sitzt. Ein passender 3D-Treiber für diese als „Kabylake-G" bekannte GPU steckt in aktuellen Mesa-Versionen.

Der für aktuelle AMD-GPUs zuständige Grafiktreiber spricht nun auch eine neue, noch nicht offiziell angekündigte Generation von Vega-Chips namens „Vega20" an. Ferner lässt sich AMDs GPU-Computing-Lösung ROCm jetzt mit GPUs der Vega-Generation nutzen, um damit allgemeine Berechnungen durchzuführen (GPGPU/General-purpose computation on Graphics Processing Units).

Erstmals dabei ist der Grafiktreiber V3D, der die als VideoCore V (VC5) und VideoCore VI (VC6) bekannten Grafikeinheiten von Broadcom anspricht. Sie sind die Nachfolger des VideoCore IV (VC4), der in den SoCs aller bisherigen Raspberry-Pi-Modelle steckt. Der neue Treiber

stammt vom Hauptentwickler des Open-Source-Grafiktreiberstacks für Raspis. Das hat zu dem schon länger kursierenden Gerücht geführt, künftige Raspis würden eine modernere VideoCore-Einheit bekommen.



Linux 4.18 unterstützt die Grafikeinheit der Intel-AMD-Kombiprozessoren.

Schnellere Datentransfers

Linux unterstützt jetzt die mit USB 3.2 definierte Dual-Lane-Übertragung, die die maximale Datentransferrate von USB-C-Verbindungen auf 20 GBit/s verdoppelt.

Der Hardware-Monitoring-Treiber K10temp liefert die CPU-Temperatur auch bei AMD-CPUs der Generationen Stoney Ridge and Bristol Ridge. Die Treiber für Eingabegeräte unterstützen erstmals den Valve Steam Controller. Durch diese und hunderte ähnliche Verbesserungen spricht Linux 4.18 letztlich über 250 Geräte oder Geräteklassen mehr an als sein Vorgänger; 42 davon sind PCI/PCIe-Geräte.

Datenbank-Beschleuniger

Das neue Writecache Target im Device Mapper (DM) ist für Systeme gedacht, bei denen möglichst geringe Latenzen gefragt sind, wenn Datenbanken oder andere Performance-kritische Anwendungen Daten speichern. Für dieses Ziel bindet das neue DM-Target persistente Speichermodule als Schreibcache ein, um die Daten später in weniger zeitkritischen Situationen auf andere Datenträger zu überführen; alternativ funktioniert das auch mit SSDs.

Nach vielen Jahren Entwicklungsarbeit bringt Linux 4.18 jetzt einen Funktionsaufruf für „Restartable Sequences“ mit. Der Syscall ist für mit mehreren Threads arbeitende Anwendungen gedacht, die mit möglichst wenig Overhead eine geteilte Datenstruktur verändern wollen. Mit den Restartable Sequences gelingt das ohne klassisches Locking, indem der Userspace-Code bei der Ausführung darauf hofft, dass ihm kein anderer Thread in die Quere kommt. Falls das doch passiert, informiert der Kernel das Programm, damit das passend reagieren kann.

Laufende VMs umziehen

Der neue Treiber Net-Failover ermöglicht eine vom Hypervisor kontrollierte Live-Migration bei Virtual Machines (VMs), die Teilfunktionen des Netzwerkchips im Wirt verwenden. Letzteres gelingt per SR-IOV (Single-root input/output virtualization), das in der VM eine Virtual Function (VF) der NIC im Host bereitstellt. Solch eine Assoziation mit echter Hardware bietet Performance-Vorteile, verkompliziert die Netzwerkkonfiguration des Gast-Betriebssystems aber signifikant, wenn die VM in der Lage sein soll, im Betrieb auf einen anderen Host umzuziehen; der neue Treiber erleichtert das Prozedere.

Weniger Staus in Netzwerken und dadurch bessere WLAN-Performance verspricht die neue Selective Acknowledgment (SACK) Compression im TCP-Stack. Durch sie verzögert der Kernel das Senden von SACK-Paketen, um diese dann komprimiert zu verschicken, falls sie denn überhaupt noch nötig sind. Durch eine Erweiterung der noch jungen Kernel-eigenen TLS-Unterstützung (Kernel TLS/KTLS) kann der Kernel die Datenverschlüsselung mit TLS (Transport Layer Security) jetzt auch an Netzwerkschnittstellen delegieren, wenn der Chip und sein Treiber das beherrschen.

Das neue TCP Zerocopy Receive verschafft Anwendungen einen direkten Zugriff auf empfangene TCP-Datenpakete. Das erfordert einen eigenen Codepfad in der Anwendung, kann beim Paket-Handling aber den Kopiervorgang zwischen dem Speicherbereich des Kernels und dem der Anwendung vermeiden. Das erspart Prozessor und Arbeitsspeicher ein wenig Arbeit, was die Performance verbessert und Ressourcen für andere Aufgaben freimacht. Das Ganze klappt aber nur unter sehr bestimmten Umgebungsbedingungen (u. a. einer MTU von 4K), daher zielt das Ganze vornehmlich auf Systeme, bei denen das letzte Quäntchen Performance gefragt ist.

Viel Know-how und Programmierarbeit erfordert auch das neue „AF_XDP“. Über dieses Feature können sich Anwendungen in die noch junge Netzwerk-Schnellstraße XDP (Express Data Path) einklinken, um die weitere Handhabung des Pakets zu beeinflussen. Letztlich soll das High Performance Packet Processing ermöglichen, das für eine flexibel programmierbare Netzwerk-Infrastruktur wichtig ist (Software Defined Networking/SDN). Der Ansatz ist unter anderem für Lösungen wie das Data Plane Development Kit (DPDK) gedacht und offeriert einen Eingriffspunkt, wie ihn die Programmierschnittstelle AF_PACKET beim normalen Netzwerkstack bietet.

Erneut geschrumpft

Wie jüngst bei Linux 4.17 schrumpft auch mit 4.18 der Umfang der Kernel-Quellen: Die neue Version ist knapp hunderttausend Zeilen schlanker. Das ist erst das vierte Mal in der Geschichte der modernen Linux-Entwicklung, dass der Quellcode mit einer neuen Version schrumpft. Wie schon bei Version 4.17 ist das vor allem Aufräumarbeiten zu verdanken. Diesmal gab es die im Staging-Zweig – einem vor allem für Treiber gedachten Sonderbereich des Kernels, der Entwicklern helfen soll, bislang extern entwickelten Code mit bekannten Qualitätsmängeln im Rahmen der normalen Linux-Entwicklung auf Vordermann zu bringen. Im Fall des bei 4.18 entfernten und für die Schrumpfkur hauptverantwortlichen Lustre ging diese Hoffnung nicht auf, weil es beim Zusammenspiel mit den Entwicklern des Cluster-Dateisystems gehakt hat. Wenn sich das ändert, könnte das vor allem bei Rechnerverbunden zum High Performance Computing (HPC) eingesetzte Dateisystem irgendwann wieder zum Kernel stoßen.

Zur Integration in Linux 4.19 stehen aber erstmal andere Dinge an – darunter beispielsweise Support für den nächsten WLAN-Standard IEEE 802.11ax und Treiber für die USB-WLAN-Sticks AMD FRITZ! AC 430 und 860. Diese Kernel-Version erscheint aller Wahrscheinlichkeit nach Mitte Oktober; sofern Torvalds seine Andeutungen wahr macht, folgt zum Jahreswechsel dann Linux 5.0.

(thl@ct.de) c't