

Zacks Kernel-News

Strenge Richtlinien für Treiber in 2.4

Neue Features sollen nur dann in den alten Kernel 2.4 kommen, wenn sie in Version 2.6 ausreichend erprobt sind, und zwar auch dann, wenn sie ursprünglich für 2.4 entwickelt wurden. Das zeigt das Beispiel von Sam Hopkins, der Ende April versuchte, seinen ATA-über-Ethernet-Treiber (AoE) in den Kernel 2.4 zu bringen.

2.4-Maintainer Marcelo Tosatti wollte den Treiber ohne „formellen Antrag“ über die Kernel-Mailingliste nicht einmal ansehen. Jeff Garzik und andere bekannte Entwickler fanden Hopkins' Treiber zwar gut, waren allerdings auch

der Ansicht, dass er ihn vorher auf 2.6 portieren muss. Hopkins begann die Entwicklung unter 2.4, weil ihm Kernel 2.6 damals nicht stabil genug erschien. Nun arbeitet er an einem Port für 2.6.

Marcelo Tosatti hält den Umweg über die Portierung für die beste Möglichkeit, den Treiber in den 2.4-Baum zu bringen. Gleichzeitig warnt er vor den 2.4-Kompatibilitätsfunktionen: Unter 2.6 erfülle manches API nur den Zweck, die korrekte Kompilierung und Ausführung von 2.4-Code zu ermöglichen, obwohl 2.6 das bessere Interface bietet. ■

Kernel-Listserver auf Blacklist

Der Server, der die Linux-Kernel-Mailingliste hostet, ist kurzzeitig auf der schwarzen Liste von Spamcop gelandet. Für die Kernel-Mailingliste ist Spam ein sehr ernsthaftes Problem, da die Server Mails von nicht-registrierten Nutzern annehmen. Das ist der einfachste und eindeutigste Weg, auf dem Anwender Bugs an Kernelentwickler melden können.

Diese Offenheit zwingt die Listen-Admins jedoch zur ständigen Wachsamkeit. Sie sind laufend damit beschäftigt, die Suchmuster zur Spamerkennung anzupassen, um zu vermeiden, dass die

Leser der Liste durch Spam überflutet werden. Dabei leisten sie gute Arbeit. Trotz allem kommt ab und zu Spam durch, was jetzt dazu führte, dass Spamcop den Listen-Server als Spamquelle ausgemacht hat.

Natürlich hagelte es Beschwerden und innerhalb eines Tages war der Rechner »vger.kernel.org« wieder aus der Blacklist verschwunden. Gerade wegen solcher falschen Positivmeldung sind Blacklists wie Spamcop heftig umstritten. Eine andere Lösung, etwa eine Anpassung des SMTP-Protokolls, ist jedoch nicht in Sicht. ■

Diskussion um Binärmodule

Seit einiger Zeit nimmt die Kernel-Mailingliste Diskussionen um binäre Module sehr ungnädig auf. Ein Vorschlag, wie sich solche Module debuggen ließen, ohne den Quellcode zu kennen, veranlasste Bartłomiej Zolnierkiewicz zu einer grundsätzlichen Stellungnahme gegen Binärmodule.

Aus vier Gründen seien diese abzulehnen: Erstens verlangsamen sie indirekt die Kernelentwicklung, da die Prüfung von fehlerhaften Bug-Meldungen umständlich sei. Zweitens widersprechen sie dem Geist von Linux.

Drittens schaden sie dem Ansehen von Linux, da der Anwender dem gesamten Kernel die Schuld gibt, auch wenn nur ein System zusammenbricht. Viertens entspräche ein Kernel, der binäre Module enthält, einem modifizierten Kernel mit geheim gehaltenen Änderungen. Benutzer könnten also keine kostenlose Hilfe aus öffentlichen Foren erwarten.

Als Konsequenz schlägt Bartłomiej vor, die Diskussion binärer Module auf der Kernel-Mailingliste zu beenden und eine separate Liste dafür zu gründen. ■

Problem mit Signalbehandlung

Ulrich Drepper, der Maintainer der Libc, wies die Kernelentwickler auf ein Problem in neueren 2.6-Versionen hin: Nicht-privilegierte Benutzer haben die Möglichkeit, alle Slots der Message Queue zu belegen, sodass diese für andere User blockiert ist. Ein systemweiter Grenzwert halte einzelne Nutzer nicht davon ab, den gesamten Pool für sich zu nutzen.

Marcelo Tosatti fand heraus, dass weitere Bereiche von diesem Problem betroffen sind. Offenbar ist ein nicht-privilegierte Benutzer in der Lage, die Liste der ausstehen-

den Signale sowie eine Reihe anderer Daten komplett für sich in Anspruch zu nehmen. Andrew Morton wies darauf hin, dass einige Infrastrukturelemente in »kernel/user.c« helfen könnten, das Problem in den Griff zu bekommen. Marcelo Tosatti implementierte daraufhin einen neuen Grenzwert für »RLIMIT_SIGPENDING«, der die pro Benutzer ausstehenden Signale einschränkt.

Grenzwerte pro Benutzer werden wahrscheinlich sehr bald in 2.6 eingeführt – weil Tosatti daran mitarbeitet, vermutlich auch in 2.4. ■

Ausführen oder nicht ausführen?

Zurzeit verursacht der Non-Executable-Stack bei einigen Entwicklern Kopfschmerzen. Das Konzept wurde eingeführt, um die Gefahr der berüchtigten Buffer-Overflows einzudämmen. Die Lösung unter Linux ist bisher, die Ausführung des Programm-Stacks zu unterbinden, sodass der Angriff einfach fehlschlägt, ungeachtet des Codes, den der Angreifer eingeschleust hat.

Diese strikte Regelung – ausführbar oder nicht – ist einigen Entwicklern jedoch zu wenig praxisorientiert. Sie fordern spezifische Ausnahmen für diese Regel. Im Augenblick ist es möglich, einen ausführbaren Stack in einer Programmdatei vor dem Aufruf zu aktivieren. Es sollte aber ebenfalls möglich sein, diese Berechtigung zur Laufzeit eines Programms zu ändern. ■

Kernel 2.6.6

Der Anfang Mai von Linus Torvalds freigegebene Kernel 2.6.6 gilt als Maintenance Release. Trotzdem fanden gegenüber der Release 2.6.5 einige Neuerungen den Weg in die offizielle Quellen. Zwei Projekte von Jens Axboe fallen dabei besonders auf: der Laptop Mode und ein I/O-Scheduler mit dem ganz schön selbstbewussten Namen Completely Fair Queing Disk I/O (CFQ). CFQ verrin-

gert die maximal möglichen Latenzzeiten beim Lesen und Schreiben. Der Laptop-Mode fasst Lese- und Schreibzugriffe auf die Festplatte nach Möglichkeit zusammen, um die Festplatte möglichst selten aus dem Standby-Schlaf zu wecken. Die Änderungen gegenüber 2.6.6-rc3 betreffen unter anderem die Unterstützung für NTFS, XFS, FAT und CIFS, die verbessert wurde. ■

Changelogs richtig interpretieren

Kernel-Changelogs sind mitunter schwer zu lesen, vor allem dann, wenn ein Eintrag darin auf zurückgezogene Patches verweist. In diesem Fall kann der gesamte Changelog-Eintrag beispielsweise so aussehen: »Cset exclude: davej@suse.de|ChangeSet|20020403195622«.

Wer Logs liest, weil er wissen möchte, was in einer bestimmten Release in den Kernel gelangt ist, erfährt auf diese Weise, dass ein früherer Changelog-Eintrag zu diesem Zeitpunkt ungültig wurde. Es ist aber unmöglich zu erkennen, um welchen Eintrag es sich handeln könnte. Zum

Glück gibt es aber doch einen Trick, um den Changelog-Eintrag für das ausgeschlossene Patch zu identifizieren. Im Beispiel »Cset exclude« etwa hängt man den Teil »davej@suse.de|ChangeSet|20020403195622« an die URL »http://linux.bkbits.net:8080/linux-2.5/cset@«; dabei ändert man

die Angabe »2.5« der betroffenen Kernelreihe entsprechend. In unserem Beispiel lautet die URL, die das ausgeschlossene Patch wieder zu Tage fördert: »http://linux.bkbits.net:8080/linux-2.5/cset@davej@suse.de|ChangeSet|20020403195622«. ■

Missverständliche Optionen

Vor kurzem hat Björn Helgaas die Kernel-Konfigurationsoptionen im PCI Media Control Interface (MCI) konsolidiert. Er entfernte diese dabei aus architekturenspezifischen Stellen, zum Beispiel dem i386-Zweig. Insgesamt begrüßten viele Entwickler diese Arbeiten. Wegen der Unterschiede zwischen den Architekturen besteht jedoch die Möglichkeit, dass die Benutzer unterschiedlicher Hardware Optionen falsch auslegen: Zum Beispiel wird der Begriff Vec-

tor benutzt, um die MSI-Option im Kernel-Konfigurationssystem zu identifizieren, während unter IA64 der gleiche Begriff externe Interrupts bezeichnet. Manche Benutzer könnten also annehmen, dass sie diese Option auf IA64-Maschinen immer aktivieren müssen, was nicht der Fall ist. Von diesen Namensproblemen abgesehen glaubt Helgaas, dass es architekturenspezifische Informationen im MSI-Code gibt, die ins Verzeichnis »arch« gehören. Tom

Long Nguyen und andere Mitarbeiter von Intel arbeiten zurzeit daran, um einen Teil dieses Problems zu beheben. Danach wird es für den Anwender einfacher sein, MSI auf der x86_64-Architektur einzusetzen. Diese zum Teil sehr missverständlichen Optionen entstehen – wie der Kernel selber ja auch – im Laufe einer Entwicklung, bei der immer wieder neue Teile zum Sourcecode hinzukommen und andere wieder verworfen werden müssen. ■

Vanilla-Kernel

Stephen Hemminger ärgerte sich auf der Kernel-Mailingliste über Distributoren, die das System derart auf ihren eigens gepatchten Kernel abstimmen, dass ein selbst gebauter Kernel aus den Originalquellen nicht mehr funktioniert. Als kompatible Distributionen nannten ihm andere Developer Fedora, Slackware, verschiedene Debian-Versionen sowie Gentoo, die alle entweder direkt Originalquellen mitliefern oder mit diesen lauffähig sind. (uwo/agr) ■