# Zacks Kernel-News

Zack Brown

## Einbruch in Kernelquellen-Server

Unbekannten Angreifern ist es gelungen, in den Bitmover-Server einzubrechen und einen Root-Exploit in den Linux-Kernel-Sourcetree einzuschleusen, der im CVS gespiegelt wird.

Der betroffene Code hätte zwar nie den direkten Weg in eine offizielle Kernel-Release gefunden, weil er sich im CVS-Mirror und nicht im Bitkeeper-Tree befunden hat. Wenn Bitmover-Mitarbeiter aber den Angriff im November nicht entdeckt hätten, wäre es dennoch denkbar gewesen, dass der Code in Linux eingeflossen wäre – wenn Entwickler den Code aus dem CVS in eigene Patches integriert hätten.

Der manipulierte Codeabschnitt war winzig – nur ein paar Zeilen – und hätte von Linus bei der Durchsicht der Patches leicht übersehen werden können. Nach Rücksprache mit Linus und anderen, hat Bitmover die Server heruntergefahren, um die Lücke aufspüren und schließen zu können.

Der Angriff kam zu einem kritischen Zeitpunkt, nur ein paar Augenblicke vor der Release der nächsten stabilen Version. Hätte der manipulierte Code den Weg in nur eine Version der Stable-Reihe gefunden, hätte er sich theoretisch auf Millionen von Computern weltweit verbreiten können. ■

## Fehlertolerantes verteiltes Dateisystem

Das Self-Stabilizing Replication File System Project (SrFS) der Ben-Gurion-Universität im israelischen Negev ist der Versuch ein verteiltes Dateisystem auf die Beine zu stellen, das gegen lokale Fehler tolerant ist. Das Projekt unterscheidet sich von Coda-Systemen, die diese Form der Selbststabilisierung nicht beherrschen.

Außerdem basiert SrFS im Gegensatz zu Coda nicht auf einem zentralen Server, der die Daten an alle Knoten liefert, sondern beruht auf einer Gruppe von mehreren Peer-Systemen, die sich die Daten gegenseitig zuschieben, bis alle Knoten einen identischen Stand haben. Unterschiedliche Konvergenzalgorithmen lösen lokale Diskrepanzen und Fehler wieder auf. Man hofft damit ein schnelles, robustes und zudem fehlertolerantes System erstellen zu können. Allerdings befand sich das Projekt zum Zeitpunkt der Veröffentlichung noch in der Proof-of-Concept-Phase.

Es gibt Überlegungen, vor allem von Pavel Machek, SrFS und Coda künftig zu einem einzigen Kerneltreiber zusammenzufassen und die Unterschiede zwischen beiden ausschließlich im Userspace-Code darzustellen. Das könnte sinnvoller sein als zwei Treiber, die ähnliche Aufgaben übernehmen. ■

## Umstrittener »devfs«-Ersatz

Das USDE-System (User-space System Device Enumeration) ist ein Ersatz für »devfs« und ähnelt dem seit einiger Zeit verfügbaren »udev«. Wie dieses auch listet das von Mark Bellon entwickelte USDE Device-Dateien dynamisch auf und kann dazu benutzt werden, mit verschiedenen, am laufenden System angeschlossenen Devices zu interagieren.

USDE kann außerdem dazu dienen, einen flexiblen Richtliniensatz für jedes dieser Devices zu formulieren. Leider ähnelt es »udev« so sehr, dass von Seiten dieses Projekts heftige Kritik kam. Insbesondere hat sich Greg KH sehr kritisch geäußert, obwohl er das USDE-Projekt nicht ganz und gar verurteilt.

Das Dilemma besteht darin, dass es in vielen Fällen zwar sinnvoll ist, wenn sich mehrere Projekte um die Lösung von ähnlichen Problemen bemühen. Aber gleichzeitig sollte man gerade in einer Community von Freiwilligen versuchen, doppelte Arbeit zu vermeiden. ■

## Erhöhung der Gruppen-Anzahl

Rusty Russell, Tim Hockin und andere haben daran gearbeitet, die Anzahl der zulässigen Gruppen in einem System stark zu erhöhen. Zwar ist das Feature bei Kernelentwicklern nicht sonderlich begehrt, jedoch gibt es vor allem bei Cluster-Bauern Bedarf dafür.

Das größte Problem bei der Erhöhung des festen Limits für die Gruppenanzahl besteht darin, die verschiedenen Datenstrukturen zu bändigen. Eine Erhöhung der maximalen Anzahl von Gruppen auf etwas über 200 wäre vielleicht etwas einfacher, die vorliegenden Patches erhöhen das Limit jedoch auf einige tausend.

Linus Torvalds hat einige frühen Versionen schlicht abgelehnt, aber er scheint jetzt eher geneigt zu sein, diese Erweiterung in Betracht zu ziehen. Da einige Unternehmen offensichtlich Interesse haben, ist davon auszugehen, dass Rusty und die anderen das Patch bis zu einer akzeptablen Lösung weiterentwickeln werden. ■

## Stability Freeze

Im Oktober hat Linus Torvalds ein Stability Freeze verordnet, während dessen nur kritische Bugs behoben werden, also solche, die beispielsweise Sicherheitsprobleme, Datenkorruption, das Einfrieren des Systems oder andere ernst zu nehmende Probleme hervorrufen. Sogar scheinbar ganz triviale Patches, die lediglich den Code umformatieren ohne Änderungen vorzunehmen, lehnt Linus jetzt ab. Einfache Feature-Erweiterungen werden ohne vorherige Prüfung zurückgewiesen.

Außerdem sieht es so aus, als wollte Linus den Maintainer-Posten des 2.6er Tree sofort nach der Release von 2.6.0 auf Andrew Morton übertragen, wenn nicht sogar vorher. Andrew wird seit einigen Monaten als neuer 2.6-Maintainer gehandelt, aufgrund seiner hervorragenden Arbeit am Virtual-Memory-Subsystem und in anderen Bereichen. Er pflegt außerdem eigene 2.6-Patches, die er Linus regelmäßig zur Integration übergibt. Theoretisch wird die neue Unstable-Reihe von der Stable-Reihe abgespalten, sobald die Übergabe an den neuen Maintainer vollzogen ist. Aber in der Vergangenheit folgte dieser Schritt immer erst einer längeren Einsatzphase der Stable-Reihe, während der die letzten Bugs beseitigt wurden. Sollte das diesmal wieder so sein und Andrew trotzdem sofort den 2.6er Kernel übernehmen, würde das dazu führen, dass Linus einige Zeit gar keinen eigenen Linux-Tree pflegt. So eine Auszeit erscheint den meisten Beobachtern aber eher unwahrscheinlich.

Man kann deshalb davon ausgehen, dass er den 2.7er Tree sofort nach der Release von 2.6.0 abspalten wird. Der 2.6er Tree kann den Stabilisierungsprozess dann fortsetzen, während 2.7 sofort mit der Aufnahme von neuen Features und dem Umorganisieren von Problembereichen beginnt. ■

## ISCSI – SCSI über TCP

Roman Zippel hat die erste Implementierung des ISCSI-Protokolls für Linux abgeschlossen, das den Austausch von SCSI-Daten über TCP ermöglicht. Obwohl diese Implementierung noch nicht alle Aspekte der Spezifikation vollständig unterstützt, ist der ISCSI-Treiber einsatzfähig, sodass die Entwicklung bei entsprechendem Interesse der User fortgesetzt werden kann. Das Interface ist allerdings etwas veraltet; es basiert auf Dateien in »/proc« statt das SysFS-Interface zu nutzen. Das ist teilweise darauf zurückzuführen, dass der Treiber zunächst für den 2.4-Kernel geschrieben wurde, und teilweise darauf, dass Roman bemüht war, einfach schnell nur die grundlegenden Funktionen bereitzustellen. Vor allem im Storage-Bereich dürfte ein großes Interesse daran bestehen, SCSI über TCP bald produktiv einzusetzen. ■

## Kernel-Binaries als Debian-Paket

Wichert Akkerman, der ehemalige Debian Project Leader, hat ein Build Target für die Kernelquellen formuliert: Die Erstellung eines ».deb«-Pakets mit kompiliertem Kernel-Binary. Es wurde mit dem neuen »kbuild«-System erstellt und weist noch einige Ecken und Kanten auf. Außerdem wurde es nach dem Stability Freeze eingereicht. Wichert hat aber trotzdem noch einige Versionen seines Patches zur Begutachtung eingeschickt. Weil es schon ein Build-Target für die Erstellung einer ».rpm«-Datei gibt, ist ein ».deb»-Target wohl nicht aussichtslos. ■

## Neuer Zufallszahlengenerator

Mit »/dev/frandom« von Elli Billauer gibt es eine neue Quelle für Zufallszahlen. »frandom« ist schneller als das herkömmliche »/dev/urandom«, es ist aber nicht für den Gebrauch bei Kryptographie- oder Sicherheitslösungen konzipiert. Allerdings ist der Weg in den Kernel für »frandom« etwas beschwerlich, weil eine ähnliche Funktionalität, wenn auch nicht ganz so effektiv, ebenso komplett im Userspace zu realisieren ist. ■

## Neue Virtualisierungs-Software Xen

Ian Pratt und andere Mitglieder der University of Cambridge Computer Laboratory Systems Research Group haben die erste stabile Version ihres Xen Virtual Machine Monitor veröffentlicht.

Dieses Tool ermöglicht die Ausführung von mehreren Betriebssystemen auf einer einzelnen Maschine – laut Aussagen der Entwickler bereits jetzt mit vernünftiger Geschwindigkeit. Um das System zu nutzen, muss jedoch der Betriebssystemkernel portiert werden; anschließend sollte jede beliebige Usersoftware funktionieren. Schon jetzt sollte jede neuere Linux-Distribution unter Xen laufen, Kernel 2.4.22 ist portiert.

Zu den Einschränkungen gehört die Tatsache, dass Xen derzeit nicht rekursiv ist; es ist also noch nicht möglich, Xen als eines der Betriebssysteme zu laden, die unter Xen laufen.

Außerdem ist Xen momentan stark an die x86-Hardwareplattform gebunden, Portierungen auf andere Architekturen sind aber prinzipiell möglich. Theoretisch ist die Performance von Xen höher als die von VMware oder User Mode Linux. Das System ist auch bereits so stabil, dass es die Entwickler selbst als Standardausstattung für Web- und Datenbankserver einsetzen. Das Projekt wurde von Intel und Microsoft gesponsert. *(uwo)* ■