

altec ComputerSysteme Entwicklungsdokumentation

Image-Dateien

Datum: Oktober 12, 2010

Autor: René Bellmer



altec ComputerSysteme GmbH

<http://www.altec-cs.de>

Inhaltsverzeichnis

1. Einleitung.....	3
2. Zusammenhang zwischen Datenträger und Image-Datei.....	3
3. Informationen zum ShortImage.....	4
4. MD5-Prüfsummen.....	6
5. Image-Informations-Dateien (nur PowerSoftware).....	6

1. Einleitung

Nachfolgend werden Informationen zur Behandlung von Image-Dateien bei alten CopyStations erörtert. Grundsätzlich arbeiten alle CopyStation mit Roh-Image-Daten, welche im Kapitel 2 erklärt werden. Die Funktionsweise eines verkürzten Images (ShortImage) wird im Kapitel 3 durchleuchtet. Die Sicherstellung von Images kann von allen CopyStations mit Hilfe von MD5 Prüfsummen bewerkstelligt werden, wie es in Kapitel 4 beschrieben wird. Neuere CopyStation, die mit Hilfe der PowerSoftware arbeiten, bieten zudem mehr Sicherheit über die Verwendung zusätzlicher Image-Informations-Dateien. Die Vorteile dieser Zusatzdateien werden im Kapitel 5 erörtert.

2. Zusammenhang zwischen Datenträger und Image-Datei

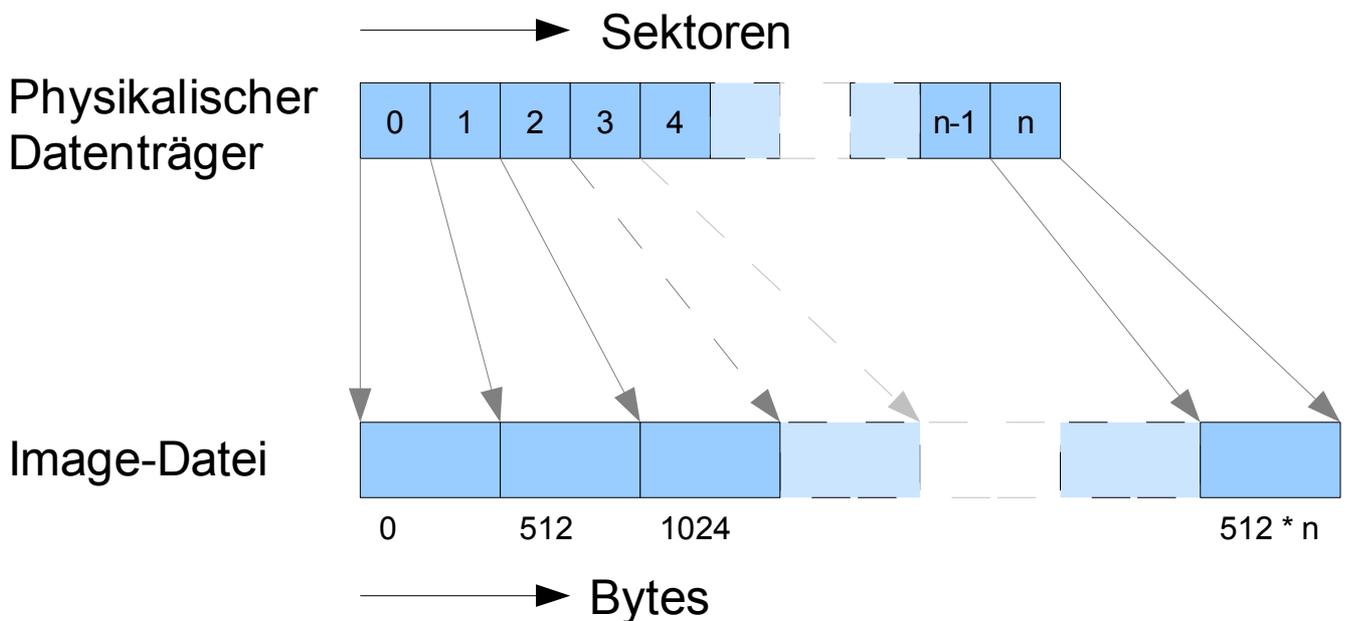


Bild 1: Speicherung von Daten auf einem physikalischen Datenträger und in einer Image-Datei

Auf einem physikalischen Datenträger werden die einzelnen Bits bzw. Bytes in sogenannten Sektoren adressiert. Diese Sektoren haben meist eine Größe von 512 Bytes, können aber in seltenen Fällen auch größer sein. Ein Sektor stellt die kleinste adressierbare Einheit auf einem Datenträger da. Die Zuordnung bzw. Adressierung der Sektoren geschieht linear beginnend bei 0 und wird wie der Zugriff auch durch das entsprechende Interface spezifiziert (z.B. SCSI, ATA, SD/MMC). Wie die Daten auf dem

physikalischen Medium (z.B. Flash-Speicher oder Magnetscheiben) hinterlegt sind, ist über das Interface nicht transparent bzw. für die Nutzung auch nicht von Interesse. Die Zuordnung zwischen physikalischer Anordnung und Adressierung der Daten über das Interface wird in der Regel über einen Controller abgewickelt. Die Anzahl der Sektoren und deren Größe spiegelt des Weiteren die Kapazität eines Datenträgers wieder.

Image-Dateien, wie sie bei altec CopyStations Verwendung finden, speichern das Abbild eines Datenträgers. Bei der Speicherung werden die einzelnen Bytes des Datenträgers in derselben Reihenfolge in eine Datei geschrieben wie sie auf dem Datenträger adressiert werden. Ein Image ist somit die exakte digitale Kopie des physikalischen Datenträgers (Aus Sicht des Interfaces). Bild 1 zeigt den Zusammenhang zwischen der Adressierung auf dem Datenträger und dem zugehörigen Image.

3. Informationen zum ShortImage

Zum Verständnis der Arbeitsweise eines ShortImages ist zunächst die Datenstrukturierung eines Dateisystems näher zu durchleuchten. Ein sehr weit verbreitetes (vor allem bei Flash-Speicher-Medien) und einfaches Dateisystem ist FAT. Das Dateisystem ist notwendig um Daten hierarchisch in Dateien zu organisieren und über die Sektoren des physikalischen Datenträgers zu verteilen.

Bild 2 zeigt eine stark vereinfachte Organisation der Daten bei FAT auf einem physikalischen Medium (Reservierte Bereiche, Bootloader etc. sind zum vereinfachten Verständnis weggelassen). Zu sehen sind die wichtigsten Strukturen. Der MBR (Master Boot Record) und PBR (Partition Boot Record) legen die Grundstruktur fest. Im MBR, der optional für Wechseldatenträger ist, werden Partitionsdaten wie die Art des Dateisystem und die Größe festgelegt. Im für das Dateisystem spezifischen PBR werden die weiteren Daten deklariert. Bei FAT wird hier vor allem die Größe der FAT-Tabelle und des Clusterbereichs deklariert. Im Cluster-Bereich befinden sich die eigentlichen Nutzdaten von Dateien und Verzeichnissen. Die kleinste Adressierungseinheit ist ein Cluster, der minimal aus einem Sektor besteht. Die FAT-Tabelle speichert die Verknüpfungen zwischen Clustern und deren Benutzung.

Beim ShortImage wird nun die FAT-Tabelle analysiert und der letzte benutzte Cluster gesucht. Ausgehend vom Sektor 0 wird das Image nun bis einschließlich dieses Clusters in eine Datei geschrieben. Bildlich wird der nicht von Daten genutzte Teil eines vollen Images beim ShortImage von hinten abgeschnitten bzw. gekürzt. Bild 2 zeigt wie von einem physikalischem Datenträger mit FAT-Dateisystem ein ShortImage erstellt wird.

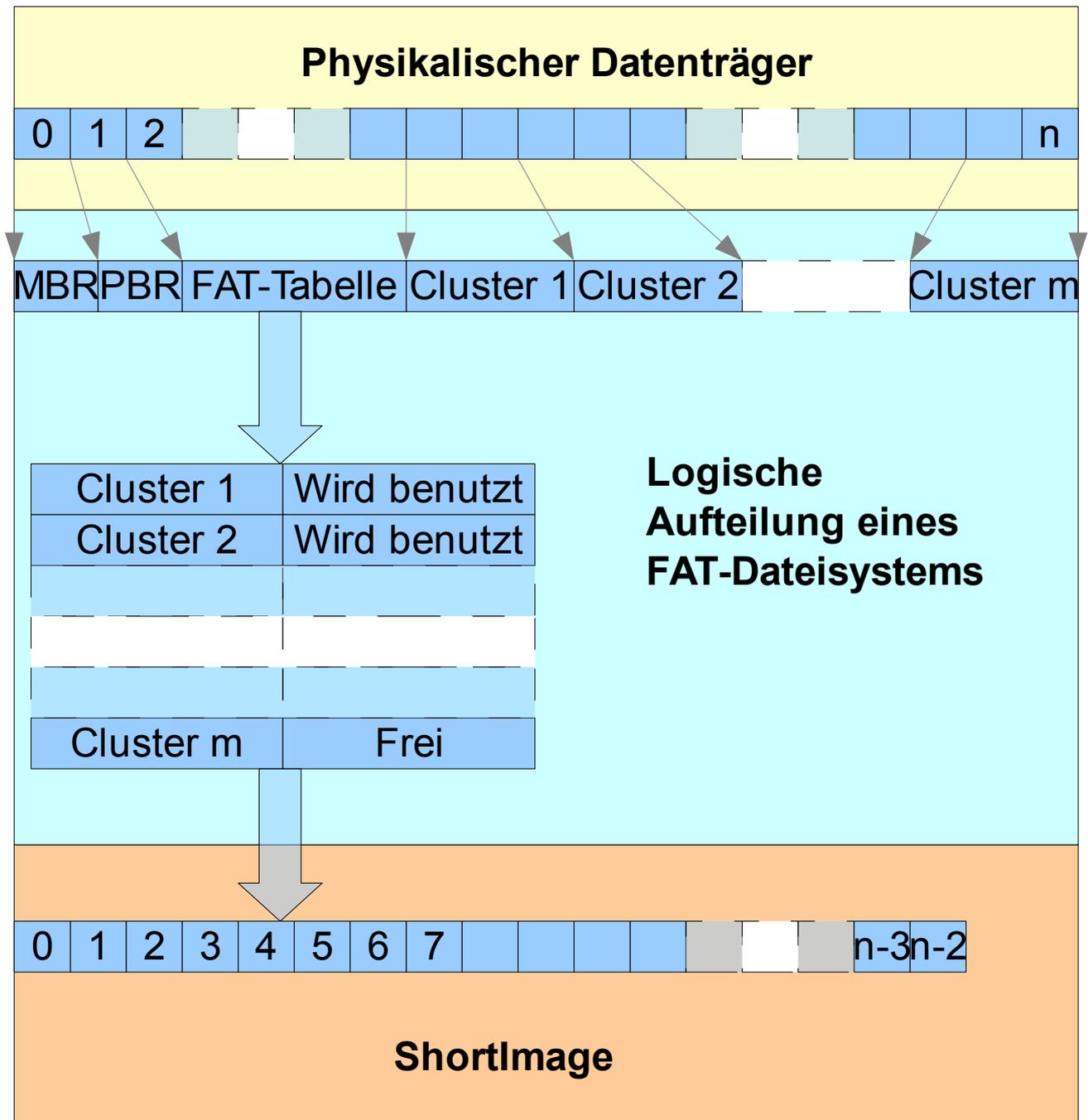


Bild 2: Zusammenhang zwischen physikalischen Laufwerk, FAT-Dateisystem und ShortImage

4. MD5-Prüfsummen

Die Software für die HighSpeed CopyStation und die PowerSoftware unterstützen den Image-Abgleich mit MD5-Prüfsummen. Die Implementierung und das Handling für den Benutzer unterscheiden sich in den beiden Varianten.

Die MD5-Prüfsumme ist ein standardisiertes Verfahren um aus beliebig langen digitalen Daten einen 128-Bit-Wert zu generieren. Dieser Wert ist eine Signatur für das Image, welche in einer Image-Informations-Datei (PowerSoftware) oder in der Job-Dateien (HSCS und PowerSoftware) hinterlegt wird. Ändert sich das Image (durch z.B. einen Virus, Datenträgerdefekte) so wird bei der Neuberechnung der Prüfsumme ein anderer Wert generiert. Die Neuberechnung der Prüfsumme findet beim Laden eines Jobs (HSCS) oder beim Starten eines Jobs (PowerSoftware) statt. Die CopyStation kann somit einen Abgleich durchführen und entsprechende Fehlermeldungen ausgeben.

Die MD5-Prüfsumme bildet eine große Datenmenge (Image-Datei mit mehreren GByte) auf einen kleinen Wert ab (16 Bytes). Solche Verfahren werden als Hash-Funktion bezeichnet. Sie können in der Theorie nicht garantieren, dass zwei unterschiedlichen Images auch unterschiedlich MD5-Prüfsummen generieren. In der Praxis gilt die Verwendung von MD5-Prüfsummen zur Erkennung von Datenübertragungsfehler von Dateien jedoch als sehr sicher.

5. Image-Informations-Dateien (nur PowerSoftware)

Die PowerSoftware, die für die IQ CopyStation, die USB CopyStation, die FlashCard CopyStation III und das Lite CopyModule verfügbar ist, speichert neben den reinen Image-Dateien eine zusätzliche Datei. Diese Image-Informations-Dateien hinterlegt mehrere Informationen über das Quellmedium, welche in einer Roh-Image-Datei verloren gehen. Ins Besondere speichert die Image-Informations-Datei, die Ursprungsgröße eines Mediums bei Verwendung eines ShortImage, die ursprüngliche Sektorgröße und eine MD5-Prüfsumme. Weitere Daten die hinterlegt werden dienen der Fehlersuche wie z.B. die Version der PowerSoftware, die Medienseriennummer, Erstellungsdatum und CopyStation-Type.