

Magnetische Festplatten sind dank ihres deterministischen Aufbaus leicht zu benchmarken: Die Performance hängt von der Plattendichte, Reaktionsgeschwindigkeit und Drehzahl ab. Im Fall einer SSD sieht die Lage anders aus. Das liegt – unter anderem – am Aufbau des verwendeten Magnetspeichers. Flash lässt sich nur zellenweise neu beschreiben. Enthält der Zielbereich bereits Nutzdaten, muss der Controller diese vor dem Ablegen der neuen Informationen auslesen und beim Neuschreiben mitberücksichtigen. In der Literatur findet sich dieses Phänomen unter dem Begriff Write Amplification. Im schlimmsten Fall sind Verstärkungsfaktoren von 1 : 128 oder sogar 1 : 256 möglich.

Ob der eher gemächlichen Arbeitsweise der Lese- und Schreibköpfe sind Festplatten beim nichtsequenziellen Lesen von Daten langsam. Sind die Blockgrößen klein, liegen auch moderne HDDs – unter Nichtberücksichtigung des Cache – im Bereich von wenigen Hundert EA-Operationen pro Sekunde.

### Lese in Eile und ...

SSDs dagegen müssen zum Lesen keine mechanischen Prozesse bewerkstelligen. Daraus folgt, dass die Leistung der Platte – primär – von der Arbeitsgeschwindigkeit des Controllers abhängt. Benchmark-Hersteller begegneten dem durch Einführung von Tests, die sehr viele zufällig verteilte Anfragen nach meist 4 KByte kleinen Datenblöcken absetzen. Die Zahl der pro Sekunde abzuarbeitenden Befehle bezeichnet man dann als IOPS, also I/O Operations per Second.

Während die als Zeitangabe dienenden Sekunden eine SI-Einheit darstellen, gibt es für die Operationen keine standardisierte Größe. Veröffentlichte Zahlen beschreiben normalerweise die „rohe“ Zahl

der Schreiboperationen. Deshalb lässt sich die Transferrate durch einfache Multiplikation von Blockgröße und IOPS-Wert ermitteln – ein Zustand, den folgende Einheitengleichung beschreibt:

$$\text{Blocks/s} \times 7 \frac{\text{KByte}}{\text{Block}} = \text{KByte/s}$$

Iometer ist das klassische Werkzeug zum Ermitteln der Festplattengeschwindigkeit. Das quelloffene Produkt entstammt einem internen Projekt aus dem Hause Intel. Als Alternativen bieten sich *IOzone* und *fio* an.

Im Mai 2015 erfreute Samsung die Welt mit einem neuen Phänomen. Die im Allgemeinen sehr schnell arbeitende 840 EVO zeigte im Laufe der Zeit nur schwer erklärbar verlangsamungen. Bei schon länger gespeicherten Daten verlangsamte sich die sequenzielle Leserate auf anämische 30 bis 50 MByte/s.

Als Begründung für dieses zugegebenermaßen seltsame Verhalten führt Samsung die Verschiebung der Spannungsebene an. Bits der schon länger auf dem Medium befindlichen Informationen sind „schwächer“, weshalb das Lesen mehr Zeit in Anspruch nimmt.

Benchmarker haben an dieser Stelle keine Chance. Die Degradationseffekte treten erst nach einigen Wochen auf. Zudem lesen Benchmark-Applikationen meist die von ihnen selbst geschriebenen Datenblöcke. Eine Alterung tritt ob des kurzen Abstands zwischen Lesen und Schreiben nicht auf.

Wer schon das Lesen von auf SSDs befindlichen Daten als komplex empfindet, erlebt beim Schreiben sein blaues Wunder. Aufgrund des in der Einleitung beschriebenen Effekts ist der Schreibprozess – per se – nichtdeterministisch.

Controller-Hersteller verwenden deshalb diverse Optimierungen, die das Schreiben kleinerer Datenmengen stark beschleunigen. Ein Klassiker ist das Vorhalten eines Schreib-Cache, der eingehenden

Kurz erklärt: IOPS als Performance-Angabe

# Tropfen für Tropfen

Tam Hanna



Dank der Verbreitung von SSDs geben Hersteller Performance-Werte immer häufiger als IOPS an.

Dass diese Werte nicht immer viel aussagen, ergibt sich aus der Funktionsweise der Hardware, die sie beschreiben sollen.

de Daten ohne Resortierung aufnimmt. Das eigentliche Remapping erledigt der Controller später. Nutzbare Benchmark-Resultate nehmen deshalb die Form von Kurven an, die die pro Sekunde angelieferten Segmente unter Berücksichtigung der Laufzeit des Testprogramms beschreiben.

Schwerer ist das Berechnen der IOPS bei RAID-Sets. Der Durchsatz hängt hier nicht nur von den einzelnen Medien, sondern auch vom Controller und vom RAID-Level ab. Als Beispiel soll RAID 5 dienen. Beim Lesen benötigt der Controller pro Vorgang nur einen Zugriff. Er weist die für das Datensegment zuständige Festplatte einfach an, die gewünschten Informationen herauszurücken.

### ... schreibe mit Weile

Beim Schreiben tritt ein als Write Penalty bezeichneter Effekt in Erscheinung. Der Controller muss die abzulegenden Daten nicht nur auf das eigentliche Speichermedium schieben, sondern zudem die Paritätsdaten berechnen. Dies vergrößert die IOPS-Zahl

erheblich: Bei fünf Festplatten wären insgesamt 5 IOPS erforderlich, um eine Operation abzubilden.

Ein alter Kalauer im Bereich der Elektronik besagt, dass der Messende oft Mist misst. Im Fall von SSDs ist die Situation ähnlich: Benchmarks sind nur dann vergleichbar, wenn sie im selben Ermittlungsprozess entstehen. Zudem sind hohe IOPS-Raten nicht der einzige Garant für hohe Performance: Besonderheiten des Controllers wie die Zuordnung zwischen physischen und logischen Sensoren verlangsamen die Prozesse mitunter.

Wer SSDs in kritischen Systemen einsetzen möchte, kommt um eine hausinterne Testreihe nicht herum. Diese schützt zwar nicht vor Controller-Bugs, unter Last langsame Halbleiterspeicher lassen sich indes recht zuverlässig „ausmisten“. (sun)

Tam Hanna

ist Gründer der Tamoggemon Holding k.s.

