

Data and Storage Management

Gedächtnisprotokoll vom 21. Februar 2019

Note: 1,0

1 Virtualisierung

Frage: Was ist die Motivation von Virtualisierung?

Antwort: Performancesteigerung, einfacheres Management, effizientere Speichernutzung, Ausfallsicherheit, Konsolidierung von Speicher.

Frage: Auf welchen Ebenen kann das passieren?

Antwort: Von unten nach oben: Disk, Block, File, FS, Tape.

2 File system Virtualisierung

Frage: Jetzt haben Sie schon die File System Virtualisierung angesprochen. Was versteht man darunter? Warum macht man das?

Antwort: Darunter versteht man das Zugreifen auf ein entferntes Dateisystem, das ich bei mir lokal eingebundet bekomme. Es scheint für mich, als wäre es lokal verfügbar. Das macht man, weil man beispielsweise kein RAID mit sich herumtragen möchte, aber die Zuverlässigkeit davon mobil abrufen können möchte.

Frage: Wie sieht es mit den anfangs genannten Motivationen aus? Werden die erfüllt? Wenn ja, wie? Wenn nein, warum?

Antwort: Performance hängt halt davon ab, wie das entfernte Dateisystem physisch realisiert ist. Wenn da auch nur eine Platte dran hängt, ist keine Steigerung zu erwarten. Wenn da aber ein SAN angeschlossen ist, sind mittlere Zugriffszeiten nur noch von der Netzgeschwindigkeit abhängig. Management ist remote immer gut, der Speicher kann geteilt und dadurch womöglich besser genutzt werden, Ausfallsicherheit ist auch Netzabhängig, und Konsolidierung ist wieder abhängig von dem, was auf der anderen Seite hängt. *Frage:* Wieso ist das Netzwerk eventuell ein Bottleneck?

Antwort: Weil wir auf der Setie zwischen mir und dem Gateway File I/O über Ethernet, TCP/IP haben. Das sind Protokolle, die nicht für den Transport von Blöcken optimiert sind. Variable Headerlänge, best effort, maximum segment size, congestion control window, ...

3 SAN

Frage: Was kann den auf der anderen Seite des Gateways zum Einsatz kommen?

Antwort: Im SAN wird üblicherweise FC gesprochen.

Frage: Wiese gerade FC? Was ist das?

Antwort: Das ist ein Protokoll für Lichtwellenleiter, das für Block I/O optimiert ist. Es hat eine deutlich zu bevorzugende Performance hinsichtlich des Anwendungsfalls Dateiübertragung.

Frage: Jetzt haben wir 2 verschiedene Infrastrukturen, gegeben dass Ethernet und FC nicht kompatibel sind. Kann man da Abhilfe schaffen?

Antwort: Mit Protokollen wie iFCP, iSCSI oder FCIP kann man FC Pakete durch IP Netze routen.

Aufgabe: Erklären Sie die einzelnen!

Antwort: iFCP wird genutzt, um ein SAN aufzubauen, aber die Infrastruktur ist TCP/IP basiert. Bei FCIP werden FC Pakete unverändert in ein Ethernet TCP/IP Paket eingepackt und können somit durch IP Netze getunnelt werden. iSCSI ist ein Protokoll, das nicht zwangsweise FC überträgt. Hier werden jegliche SCSI Pakete über TCP/IP geroutet.

Frage: Was hat man davon?

Antwort: Man kann bestehende Ethernet-Infrastruktur, die bereits angeschafft wurde und beim Aufrüsten günstiger als FC-Infrastruktur ist, nutzen, um sein SAN aufzubauen.

4 Blockvirtualisierung

Frage: Reden wir über Blockvirtualisierung. Was ist das?

Antwort: Eine Virtualisierungsschicht, die dazu genutzt werden kann, mehrere Festplatten zu einer logischen zusammenzufassen, oder umgekehrt. Man kann eine Platte auch zu mehreren logischen aufteilen.

Frage: Wie kann man hier Ausfallsicherheit rausholen?

Antwort: Beispielsweise mit einer RAID Konfiguration.

Frage: Wie kann man hier Performance rausholen?

Antwort: Auch mit einer RAID Konfiguration.

Frage: Wie groß ist denn so ein Block herkömmlicher Weise?

Antwort: Um die 500KB bis zu advanced format, 4KB.

Frage: Kann Ihnen eine RAID Konfiguration dabei helfen, die Performance zu steigern, wenn ich nun etwas schreiben will, das 4KB groß ist? Das landet doch dann auf einer einzelnen Platte, nicht?

Antwort: Öhm joa joo..... hä?

Nachhilfe: Stellen Sie sich vor, sie haben 4 Platten. Ich schreibe nun 4KB. Was kann ich machen, um die Performance zu steigern?

Antwort: Ahh, man kann den 4KB Block im RAID Controller zerlegen und jede Platte 1KB schreiben lassen. Diese Blöcke müssen dann beim Lesen des 4KB Blocks wieder aggregiert werden.

5 RAID

Frage: Welche RAID Konfigurationen kennen Sie?

Antwort: RAID 0, RAID 1, RAID 10, RAID 01, RAID 2, RAID 3, RAID 4, RAID 5, RAID 6, RAID53, RAID 55...

Frage: Okay, okay! Welchen dieser würden Sie bevorzugen, wenn es Ihnen nur um Performance geht?

Antwort: RAID 0.

Frage: Und für Ausfallsicherheit?

Antwort: RAID 1.

Frage: Das verschenkt ja aber auch eine Menge Speicher und ist deswegen Teuer. Was kann man noch machen?

Antwort: RAID 5.

Frage: Wie funktioniert das?

Antwort: Man schreibt Stripes auf die Platten und hängt dahinter eine Parität an. Die Paritäten sind über alle Platten verteilt. Fällt eine Platte aus, kann man die verlorenen Daten anhand der Paritäten zurückrechnen.

Aufgabe: Vergleichen Sie mal die Performance davon mit RAID 1.

Antwort: Lesen ist bei beiden gleich gut, gegeben die Platten haben gerade keine weiteren Aufträge. Beim Schreiben kommt bei RAID 5 die Write Penalty hinzu. Diese entsteht dadurch, dass beim Überschreiben von alten Daten, eben diese von der Parität abgezogen werden müssen. Dann werden die neuen Daten geschrieben und die Parität neu berechnet und geschrieben.

Frage: Kommt das immer vor?

Antwort: Bei Fullstripe-Writes interessieren uns die alten Daten und die alte Parität nicht. Die werden ja eh alle überschrieben.

6 Disk-Virtualisierung

Frage: Reden wir über Disk-Virtualisierung. Was ist das? Warum macht man das?

Antwort: Disk-Virtualisierung macht man, um logische Blockadressen auf physische Sektoren, Cylinder und Heads abzubilden. Der Controller, der dies verarbeitet kann Wear-Leveling implementieren, oder bei HDDs tote Sektoren vermeiden und somit Datenrettung betreiben. Nach außen hin präsentiert er eine einheitliche Schnittstelle.

Frage: Was haben Sie dort für Möglichkeiten, die Performance zu steigern?

Antwort: Short-Stroking, Caching und Command Queuing.

Frage: Wann lohnt sich was besonders?

Antwort: Short-Stroking nur, wenn man wenige Daten hat, die auf die äußeren Sektoren der Platter passen. Der Rest lohnt sich immer, nicht...?

Frage: Da gibt es Rahmenbedingungen. Beispielsweise beim Command Queuing. Die Daten, die zuerst rein kommen, müssen ja warten, bis sie geschrieben werden. Das ist doch eine Verzögerung, oder nicht?

Antwort: Die Platte kann ja sofort ein Signal geben, dass die Daten geschrieben wurden und behält sie in der Queue, bis sich eine sinnvolle destage-Reihenfolge ergibt.

Frage: Da müssen Sie aber aufpassen, dass ihnen nicht der Strom ausfällt. Wann lohnt sich Command-Queuing besonders?

Antwort: Wenn die Daten eben nicht warten müssen, also viel Traffic besteht und sie schnell rausgeschrieben werden können.