

Monday, April 7. 2008

Backups irrelevant?

It's nice that he references to ZFS in his article, but I think his opinion stated in "#PostgreSQL to Scale to 1 Billion Users, Dr Evil would be proud" is dangerous: Backup is irrelevant for those of you who care about this discussion. LVM/ZFS snapshots are the rule of the land. Well, when the first disaster hits your datacenter (and such disasters begin with a failed disk too much), you will recognize that you are toast without tapes or at least with a disk-based backup on a different storage as a last line of defence. Never rely on a single pool of disk blocks on a single machine for disaster recovery, even when it gives you features like snapshot to freeze the disk state at one or multiple points in time.

Posted by Joerg Moellenkamp in The IT Business at 10:49

rotfl, good luck

Anonymous on Apr 7 2008, 13:06

Backups nur auf Disk zu machen verbietet sich doch schon angesichts der aktuellen "Green-IT" Diskussion. Zudem muss ich deutlich mehr Geld investieren um eine vergleichbare Kapazität mit Disks hinzubekommen. Gerade im Enterprise-Umfeld wird es noch sehr, sehr lange Tapes geben.

Anonymous on Apr 7 2008, 14:33

Erzählst Du das bitte mal den Kunden, die durch bestimmte Hersteller getriebene B2D fuer eine furchtbar tolle Idee halten?

Anonymous on Apr 7 2008, 15:44

Vielleicht reichte es ja auch aus, den B2D-Herstellern zu erzählen, dass man Festplatten bei Nichtbenutzung auch abschalten kann?

Anonymous on Apr 7 2008, 22:03

Hmm, ist trotzdem aus meiner Sicht ein ziemlicher Hack: Elektronik läuft trotzdem weiter, hoher Einschaltstrom beim Anlaufen der Platten und meines Wissens sind Festplatten fuer Server auch nur fuer eine bestimmte Anzahl Einschaltzyklen spezifiziert.

Anonymous on Apr 8 2008, 06:12

Relying solely on disk backup is stupid. This begins with disk systems being buggy, too. Having good old tape backups with tar files handy was a godsend when recently our backup 3510 got magically inconsistent after a harddisk failed. Should never happen, but so it goes.

Anonymous on Apr 8 2008, 09:39

Jetzt habe ich auch mal den referenzierten Artikel gelesen:

Das Zitat hier ist schon sehr sehr verkürzend.

Es geht nicht darum, dass Backup an sich irrelevant sei. Der Autor behauptet vielmehr, dass Online-Backup-Features in Datenbank-Serversoftware irrelevant seien, weil man sie durch Snapshots auf dem Storage Layer ersetzen könne.

Ich persönlich habe trotzdem gerne Online-Backup in der Datenbank, aber auf jeden Fall ist die Aussage eine komplett andere.

Anonymous on Apr 8 2008, 11:05

Ja, Du hast da recht. Ich hatte das zunächst anders gelesen. Ich nehme die Kritik in diesem Punkt an jenem Artikel zurück. Problem ist: es gibt Kunden die wirklich so denken

Anonymous on Apr 8 2008, 21:35

Die Behauptung, daß Online-Backup durch Snapshots im Storage-Layer überflüssig wird, stimmt so in der Form auch nicht.

Wenn z.B. ZFS eines Tages die Transaktions-API exportiert und die relevanten Datenbanksysteme dieses für ihre Schreiboperationen nutzen, dann kann man davon sprechen - vorher nicht.

Im Fall von DB2 beispielsweise wird regelmäßig ein so erstelltes "Backup" auf eine Kompletprüfung der Datenbank laufen, die schon bei wenigen Hundert GB etliche Stunden dauern kann.

Ohne wirklich präzise das physische Transaktionsverhalten der Datenbank und der Snapshot-Funktion des Storage-Layers verstanden zu haben, sollte man nicht von Online-Backup weggehen.

Im Grunde ist der ganze Artikel ziemlich oberflächlich. Warum sollte die Verfügbarkeit von ZFS für MacOS X darüber entscheiden, daß "makes Apple start to look like a viable platform for servers"? Aha, und vor ZFS waren alle Betriebssysteme incl. Solaris nicht "viable platforms", oder wie?

Das sollte man alles nicht so ernst nehmen..
Anonymous on Apr 9 2008, 09:38

Eine der vornehmsten Aufgaben eines Datenbanksystems ist es, den Datenbestand auf Platte unter allen Umständen konsistent zu halten.

U.a. auch bei einem Stromausfall, und ein Snapshot auf dem Storage-Layer hat genau den selben Effekt: in einem Moment hören schlagartig alle Schreiboperationen auf dem Snapshot auf. Der Storage-Layer muss nur sicherstellen, dass das für alle Volumes der Datenbank exakt zum selben Zeitpunkt passiert. (damit haben manche Storage Manager schon ein Problem.)

Richtig ist, dass bei einem solcherart angefertigten Backup genau wie bei einem Systemausfall unter Umständen der Wiederanlauf deutlich länger dauert, weil das Datenbanksystem eine Menge Aktionen aus dem Log wieder einspielen oder ungeschehen muss.

Ein Datenbanksystem aber, die sich zusätzlich zum Log-Replay beim Wiederanlauf nach einem Systemausfall erstmal stundenlang mit Prüfungen beschäftigt, würde ich allerdings nicht so gut finden. Ganz egal, wie ich Backup mache.

Ein Nachteil an der Snapshot-Methode ist allerdings, dass ich die Datafiles des Datenbanksystems sichern muss, die unter Umständen erheblich größer sein können als ein Dumpformat.
Anonymous on Apr 9 2008, 10:59

Korrekt, wobei es langsam drehende Platten gibt, die für eine hohe Anzahl von Anlauf/Stopp-Prozessen gebaut sind (so in Richtung Notebookplatten). Die schalten auch einige leistungshungrige Komponenten der Elektronik ab. Das Problem liegt m.E. eher dann in den MTBFs und den Fehlerquoten der Platten. Da muss man dann gleich Raids und Prüfsummen auffahren, damit das interessant wird. Die Stromspitzen beim Hochfahren bekommt man vielleicht durch verstetzes Hochfahren in den Griff - trotzdem bleibt es ein merkwürdiges Rummgemache.
Wichtig wäre vielleicht die Robustheit der Bänder vs. die Nichtrobustheit der Festplatten, wenn man sie - aus welchen Gründen auch immer - aus der Backupmaschine herausnimmt (hier kommt der Maschinenbauer durch). Bänder überstehen einen Fall aus 2m Höhe. Festplatten würden mir da Sorgen machen - insbesondere wenn die Lager schon ein paar Jahre auf dem Buckel haben.
Anonymous on Apr 9 2008, 12:53

"Der Storage-Layer muss nur sicherstellen, dass das für alle Volumes der Datenbank exakt zum selben Zeitpunkt passiert"

Das "nur" ist aber schon eine sportliche Herausforderung, zudem sind in der Praxis leider auch die renommiertesten Datenbanksysteme in dieser Hinsicht nicht so fehlerfrei wie man sich das wünschen würde. "Restore from backup" habe ich schon mehr als einmal herstellerseitig in Bezug auf logische Fehler in Volumes gehört nach Hardwarefehlern, Stromausfall (typisch: Murks bei Installationsarbeiten im Serverraum) oder auch nur "erst das Stagesystem, dann den Server runtergefahren".

Meine besten Wünsche demjenigen, der versucht, alleine mit Snapshots auf Storageebene auszukommen. Was da alles perfekt funktionieren muß, damit das garantiert geht...
Anonymous on Apr 9 2008, 14:35