

# **SAN versus NAS für Oracle (11g) Datenbanken**

**Ing. Christian Pfundtner  
DB Masters GmbH  
Gerasdorf, bei Wien, Österreich**

## **Schlüsselworte**

SAN, NAS, Oracle Datenbank, Infrastruktur, Performance, iSCSI, NFS, dNFS, Jumbo Frames

## **Einleitung**

Welche ist die beste Storageanbindung für eine Oracle Datenbank? Lange Zeit war die Antwort klar: „SAN!“. Immer häufiger stellt sich in der Zwischenzeit die Frage, ob diese Antwort immer noch unbestritten ihre Gültigkeit hat, oder ob nicht auch „NAS“ die richtige Antwort sein kann. In diesem Vortrag vergleichen wir SAN mit NAS auf verschiedenen Ebenen – von den Spezifikationen über die Performance bis hin zum Kostenvergleich.

Im ersten Teil unseres Vortrags werden die Protokolle SAN, iSCSI und NFS in Bezug auf Overhead und Durchsatz analysiert und miteinander verglichen. Der zweite Teil taucht tiefer in den Performance Aspekt ein und bringt einige Real Live Performance Ergebnisse, die zeigen, was „State of the Art“ Performance ist. Der Abschluss wird durch einen Kostenvergleich SAN versus NAS gebildet und enthält auch einen kleinen Ausblick auf die nächsten Entwicklungen in diesem Bereich.

## **SAN versus NAS**

### **Teil 1: Vergleich der Protokolle**

#### ***Technologie***

SAN ist ein BLOCK orientiertes Protokoll, das sich vor allem durch einen kleinen Protokolloverhead auszeichnet – allerdings ist auch die maximale Datenmenge pro Paket mit 2112 Bytes nicht gerade üppig im Vergleich zu einem Oracle Datenbank Block, der aktuell bei den meisten Datenbanken 8k groß ist. Mittels SAN werden von der Storage bereitgestellte LUNs an einen Server angebunden und dort als DEVICE bereitgestellt. Was man mit diesem Device macht, ist dem SAN und der Storage egal, es werden einfach nur READ und WRITE I/Os abgehandelt. Durch das Anlegen eines Filesystems oder einer ASM Diskgroup wird diese LUN formatiert und damit in der Storage der ganze Bereich reserviert (sofern das nicht schon beim Erzeugen der LUN erfolgt ist).

Bei NAS wird als Transportprotokoll TCP/IP verwendet, welches im Vergleich zu SAN einen größeren Overhead aufweist – allerdings kann man diesen Umstand durch den Einsatz von Jumbo Frames egalisieren, so dass der Overhead für den Transport von Oracle Datenbank Blöcken vergleichbar wird. Im NAS Bereich werden je nach Anforderung zwei unterschiedliche auf TCP/IP aufsetzende Protokolle genutzt: NFS und iSCSI.

NFS ist ein FILE orientiertes Protokoll. Die Storage stellt ein Filesystem bereit, das am Server gemounted wird. Der Zugriff erfolgt für die Applikation wie auf jedes andere Filesystem auch, nur werden im Hintergrund die Filezugriffe an die Storage weitergeleitet. Ein großer Vorteil von NFS ist die Möglichkeit, ein Filesystem von einem Server auf einen anderen Server umzuhängen –

beispielsweise von einem Linux Server auf einen Solaris oder AIX Server. Fragestellungen wie „kennt der andere Server auch das Filesystem bzw. den Logical Volume Manager?“ stellen sich bei NFS erst gar nicht!

Bei iSCSI handelt es sich mehr oder weniger um SAN via TCP/IP. Es werden die von der Storage bereitgestellten LUNs mittels TCP/IP an den Server angebunden, wo diese wieder als DEVICE sichtbar werden. iSCSI nutzt man dann, wenn man unbedingt LUNs benötigt (z.B.: man will/muss ASM nutzen) aber eine NAS Storage nutzen möchte.

## ***Bandbreiten***

Im SAN Bereich gilt aktuell 8Gbit als State of the Art, erste Adapter mit dieser Bandbreite sind seit 2008 verfügbar, durchgesetzt haben sich diese aber erst in den letzten 2-3 Jahren. In der Praxis findet man hin und wieder noch neu gekaufte Lösungen mit 4Gbit (was allerdings eher eine Ausnahme darstellt).

Obwohl für den NAS Bereich 10Gbit Technologien schon seit 2002 verfügbar sind – aber praktisch nur bei den Internet Backbone Providern Einzug gefunden haben – setzt sich die 10Gbit Anbindung der Server erst seit ca. 2 Jahren langsam durch. Bei NAS Storages ist 10Gbit schon etwas länger verfügbar, allerdings war der „Aufpreis“ bis vor kurzem noch merklich.

### **Conclusio Anbindungsgeschwindigkeit**

Lange Zeit lag SAN (2Gbit, 4Gbit) mit der Performance vor NAS (1Gbit). Aktuell sind die beiden ebenbürtig, auch wenn NAS mit 10Gbit doch schon etwas schneller ist. Sieht man sich aber die nächste Zukunft an, so wird NAS mit 40Gbit deutlich vor SAN mit 16Gbit liegen!

## ***Administration***

Ein weiteres Kriterium, beim Vergleich zwischen SAN und NAS sollte nicht unbeachtet bleiben: „Wie sieht das Ganze im administrativen Bereich aus?“

Unter der Voraussetzung, dass man eine NAS Infrastruktur gleich professionell betreibt wie eine SAN Infrastruktur (sprich die NAS Infrastruktur ist getrennt vom Produktions-/Benutzer LAN), gibt es einige Punkte, die die beiden grundlegend unterscheiden!

### **SAN aus Administrationssicht**

Bei SAN werden die LUNs in der Storage definiert, mittels Zoning dem/den Servern zugewiesen, wo diese als LUNs/Devices sichtbar werden. Oft kommen noch Produkte wie EMC PowerPath, usw. zum Einsatz, die ebenfalls installiert, konfiguriert und gepflegt werden müssen. Das Device muss man im OS entsprechend konfigurieren (Logical Volume Manager, Filesystem oder ASM,...) bis dann für die Datenbank und/oder die Applikationen Platz zur Verfügung steht.

Will man eine LUN erweitern, ist dies bei manchen Lösungen im Betrieb möglich, wobei hier wieder alle beteiligte (LVM, Filesystem,...) administriert werden wollen. Ein verkleinern einer LUN ist meines Wissens im laufenden Betrieb überhaupt nicht möglich.

Spannend wird das ganze allerdings dann, wenn man eine größere Reinvestition plant – beispielsweise die Anschaffung einer neuen Storage. Der Storagehersteller gibt typischerweise eine Kompatibilitätsmatrix für die SAN Komponenten vor (SAN Fabric, HBAs,...). Spätestens jetzt wird es

bei vielen Firmen aufwändig, weil man nicht nur alle Komponenten überprüfen, upgrades planen und vieles mehr muss - Nein, es ist noch viel schlimmer! Teilweise sind von den nötigen Upgrades auch Server und Infrastrukturateile betroffen, die nicht direkt etwas mit der neuen Storage zu tun haben, nur weil diese zufälligerweise auf einer SAN Fabric hängen, an die die neue Storage angeschlossen werden soll und wo auch Ihre „noch genutzte“ Storage hängt.

Die Schätzungen für den Aufwand eines „Rollouts“ einer neuen SAN Storage bei einem unserer Kunden mit knapp 200 Servern lagen bei mehreren Mann-Monaten Aufwand, wobei sich auch einige Down Times nicht vermeiden ließen.

### **NAS / NFS aus Administrationsicht**

Bei NAS gibt man ein NFS Share auf der Storage frei – bis hier her ist der Aufwand mit SAN vergleichbar. Die Einbindung am Server ist nur noch ein NFS Mount – deutlich geringer als bei SAN! NAS / NFS Bereiche im laufenden Bereich vergrößern und auch verkleinern können die meisten Lösungen out of the Box. Doch wie sieht es bei der Einbindung einer neuen Storage aus? Nun, einfach anstecken und loslegen – abgesehen von der initialen Konfiguration der Storage selbst, ist kein weiterer Aufwand nötig!

### **NAS / iSCSI aus Administratorsicht**

Stellt eine Mischung aus SAN und NAS dar. iSCSI verhält sich mehr oder weniger wie SAN, bis auf eine entscheidende Ausnahme: Kommt eine neue Storage dazu, ist der Aufwand wie bei NAS / NFS praktisch gleich null.

### **Conclusio: SAN versus NAS aus Administratorsicht**

Der Gewinner ist eindeutig: NAS

## ***Ausblick***

### **Wie sieht die „Zukunft“ bei SAN und NAS aus?**

Nun die Zukunft ist schon da! Im NAS Bereich wurde der 40Gbit sowie der 100Gbit Standard am 17. Juni 2010 verabschiedet, SAN ist mit den Mitte 2011 16Gbit etwas später dran gewesen. Auch bei der Verfügbarkeit von Infrastruktur Komponenten sieht es ähnlich aus. 40/100Gbit gibt es seit einigen Jahren schon zu kaufen (wieder eher nur für Backbone Provider), 16Gbit SAN Lösungen sind noch extreme Exoten.

Bei beiden gilt: „Im Moment einfach noch zu teuer“.

## **Teil 2: Performancevergleich**

Dieser Abschnitt geht auf Real Live Performance Ergebnisse im Vergleich zwischen SAN und NAS aber auch auf das Thema Oracle direct NFS (dNFS) bei verschiedenen Systemen ein und analysiert die Ergebnisse.

### **Fangen wir bei Oracle 10g an, da diese Version immer noch sehr verbreitet ist.**

Wenn man SAN und NAS unter Verwendung von JumboFrames gegeneinander antreten lässt, ist das Ergebnis noch recht eindeutig. SAN liegt im Durchsatz merklich vorne und auch die CPU Auslastung ist im SAN Bereich merklich niedriger als im NAS Bereich. Das hat mehrere Gründe, der wichtigste Grund ist jedoch: SAN Hostbus Adapter sind intelligenter und entlasten die CPUs – hier helfen auch spezielle Netzwerkkarten mit „TCP/IP Offloading“ nur wenig.

Wird – warum auch immer – auf JumboFrames verzichtet, ist das Ergebnis noch viel, viel deutlicher zu Gunsten von SAN. Dies ist auch der Grund, warum bei vielen noch im Kopf ist: „Performante Datenbanken gehen nur mit SAN“.

### **Spannender wird der Vergleich, wenn man Oracle 11g mit dNFS einsetzt.**

In einem direkten Performance Vergleich mit dNFS mit JumboFrames dreht sich das Blatt auf einmal deutlich. Ja, die CPUs müssen im NAS Bereich immer noch mehr arbeiten als im SAN Bereich, jedoch haben aktuelle CPUs einfach so viel mehr Power als die CPUs auf den SAN HBAs, dass dieser Overhead kaum mehr auffällt. Selbst unter optimalen Bedingungen schafft man es im SAN Bereich selten über 750MB/sec (bei 8Gbit SAN) zu erreichen – mit 10Gbit NAS sind jedoch 1050 MB/sec möglich!

Beide Werte wurden auf dem gleichen Server gegen die gleiche Storage einmal über SAN und einmal über NAS mittels Oracle Calibrate I/O Messungen ermittelt. Im SAN Bereich ASM Diskgroup mit 8 LUNs) und im NAS Bereich 2 NFS Shares!

Gleichzeitig ist die CPU Auslastung im NFS Bereich sogar merklich niedriger als im SAN Bereich. Das liegt daran, dass bei dNFS die Oracle Prozesse selbst die – extrem optimierten - NFS Clients sind und den Datenbank Block direkt aus/ins Buffer Cache stellen können, ohne dass dieser noch umkopiert werden muss. Im SAN Bereich muss ein Block in der Regel zumindest einmal kopiert werden, weil der Zugriff auf die SGA für die OS Level Treiber nicht möglich ist → dieses kopieren erklärt den höhere CPU Aufwand.

### **Wie viel bringt der Einsatz von dNFS wirklich?**

Ein Performance Test mittels RMAN Backup einer knapp 2 TB großen SAP Datenbank mit 8 Channels zeigt es beeindruckend:

- Ohne dNFS: 93 Minuten Backupzeit (ca. 300MB/Sec)
- Mit dNFS: 48 Minuten Backupzeit (ca. 600MB/Sec)

Weitere Informationen zu diesem und anderen dNFS Benchmarks finden Sie auf unserer Homepage [www.dbmasters.at](http://www.dbmasters.at) im Know-How Bereich unter „Oracle Direct NFS“

### **Wie viel bringt der Einsatz von JumboFrames wirklich?**

Hier liegt kein genauer Benchmark vor, aber sehr wohl praktische Erfahrungen. Ohne den Einsatz von JumboFrames war die Performance bei I/O Peaks um den Faktor 2-3 niedriger als mit dem Einsatz von JumboFrames. Dies kann man leicht damit begründen, dass ohne JumboFrames ein 8k Oracle Datenbank Block auf 6 Netzwerk Pakete zerlegt werden muss → jedes Netzwerk Paket muss bestätigt werden und erzeugt dabei mindestens einen Interrupt. Solange das System nicht zu viel I/O macht, ist der Unterschied noch nicht sehr groß – vielleicht 10 bis 20 Prozent. Sobald die I/O Last steigt, kann man von einem Faktor 2-3 ausgehen!

## **Teil 3: Preisvergleich SAN versus NAS**

Wie sieht der Preisvergleich aus, wenn man „auf der grünen Wiese“ beginnen kann?

Natürlich ist immer wieder das Thema, dass bestehende Infrastruktur sinnvoll weiter genutzt werden soll – doch manchmal rechnet sich auch der Umstieg von der einen auf die andere Technologie.

Aktuell kostet eine Infrastruktur Lösung mit 10Gbit NAS um 10 bis 20% weniger als eine 8Gbit SAN Implementierung – und die Preisdifferenz steigt. 40Gbit NAS ist aktuell deutlich günstiger als 16Gbit SAN, nur gibt es im Moment viel zu wenig Anbieter – vor allem im Storage Bereich, daher hinkt dieser Vergleich noch etwas.

## **Kontaktadresse:**

Ing. Christian Pfundtner, OCA, OCP, OCE, OCM, ACE  
DB Masters GmbH  
Stammersdorfer Strasse 463  
A-2201 Gerasdorf bei Wien  
Österreich

Telefon: +43 699 150 37 884  
Fax: n/a  
E-Mail [cp@dbmasters.at](mailto:cp@dbmasters.at)  
Internet: [www.dbmasters.at](http://www.dbmasters.at)