

Schnellerer Oracle Datenzugriff durch RamSan-220 Solid State Disks

Woody Hutsell

Texas Memory Systems

Quelle: *“Oracle 9i Database Performance Guide and Reference”, Version 1 (9.01), Seite 2:*

„Bei jeder Datenbank ist das I/O Subsystem ein Engpass für Leistung, Datenvollständigkeit und die Verfügbarkeit des Systems schlechthin. Eine Schwäche in irgendeinem dieser Bereiche kann das Datenbanksystem instabil, nicht skalierbar oder unverlässlich machen.“

Die Lösung: Die RamSan-220 Solid State Disk

Der schnelle Zugang zu den Firmendaten ist für kleine und große Unternehmen gleich bedeutend. Mechanische Festplatten begrenzen das I/O Subsystem. Solid state disks (SSDs) verbessern das I/O Subsystem und stellen eine ausgezeichnete Alternative dar.

Seit mehr als zwanzig Jahren hat Texas Memory Systems (TMS) modernste SSD Systeme für spezialisierte Anwendungen der US Verteidigungsindustrie entwickelt. Das System RamSan-220 ist ein Entwicklungsschritt aus diesem Anwendungsgebiet. Es wurde entworfen, um die zunehmenden I/O Nachfragen der leistungsstarken Speichersysteme zu befriedigen. Das RamSan-220 System bietet in diesem Marktsegment den großen Unterschied in Wert und Leistung. Mit bis zu vier schnellen Fiber Channel Schnittstellen und 32 Gbyte SDRAM Speicher kann ein einzelnes RamSan System Speicherengpässe für ein gesamtes Rechenzentrum beseitigen. Speicherkapazität und Schnittstellenanzahl sind wichtige Faktoren. Der Schlüssel, um Speicherengpässe zu beseitigen, ist jedoch die I/O Leistung. Mit vier Fiber Channel Verbindungen liefert RamSan-220 200.000 IOPS. Die Zugriffszeit eines RamSan-220 Systems ist mehr als 250mal schneller als die der bestehen RAID Systeme.

Vom technischen Standpunkt aus ist das RamSan-220 System ein logisches Laufwerk, welches an Servern oder Network Storage Systemen angeschlossen wird. Je nach Konfiguration hat das System 1 bis 64 LUN's. Jedes LUN erkennt das Betriebssystem als logisches Laufwerk. Es ordnet dem Laufwerk einen Buchstaben zu und kann mit jedem möglichem Dateisystem formatiert werden. Das RamSan-220 System wird mittels Fiber Channel an Servern oder Network Storage Systemen angeschlossen. Der Hauptunterschied zwischen einem RamSan-220 und anderen Laufwerken ist, daß die Zugriffszeit für eine RamSan-220 SSD nur 20 Mikrosekunden beträgt, die schnellsten RAID System-Zugriffszeiten (Latenz und Suchvorgangzeiten) jedoch mehr als 5 Millisekunden benötigen. Kombiniert man niedrige Latenz und hohe Bandbreiten System-Platinen, befähigt man das System, I/O Engpässe für Orakel Datenbanken zu beseitigen.

Bei dem Versuch, die Leistung von Speichermedien zu erhöhen, wenden System Administratoren meist folgende drei Strategien an:

- Sie erweitern das System durch neue Festplatten. Das Hinzufügen von Festplatten zu JBOD (just a bunch of disks) oder RAID Systemen ist zwar ein Weg, die Speicherleistung zu verbessern. Indem man die Zahl der Festplatten erhöht, kann sich das I/O Sub System einer Datenbank über mehrere Festplatten

verteilen. Wie jedoch auch die nachstehend angeführten Strategien führt dies nicht wirklich zu einer Beseitigung der Engpässe.

- Sie verschieben die am häufigsten verwendeten Daten auf eine separate Festplatte. Diese Strategie liefert das beste I/O Verhältnis, welches bei Einzelfestplatten möglich ist. Offensichtlich ist jedoch, dass die I/O Fähigkeit einer einzelnen Festplatte sehr begrenzt ist. Bestenfalls liefert eine einzelne Festplatte 300 I/Os pro Sekunde. Das RamSan-220 System ist jedoch in der Lage 200.000 I/Os pro Sekunde zu liefern!
- Sie installieren ein RAID System. Oft wechseln System Administratoren ihre JBOD Systeme gegen bessere RAID Systeme aus. Die dadurch meist viel bessere Leistung wird erzielt, indem ein Cache Controller vor dem eigentlichen Laufwerk eingesetzt wird. Dadurch können zusätzliche Daten über mehrere Festplatten verteilt werden. Diese Strategie liefert eine wesentlich höhere Leistung als herkömmliche Systeme, speziell dann, wenn große Cache Komponenten eingesetzt werden.
- Das beste, heute am Markt befindliche RAID System liefert nur 5.000 I/Os pro Sekunde. Trotzdem kosten solche Systeme mindestens eine Million Dollar. Das RamSan-220 System ist nicht nur um Vieles günstiger, es liefert außerdem 200.000 I/Os pro Sekunde.

Viele Administratoren haben sich bisher aus Kostengründen nicht für SSD als Lösung zur Leistungsverbesserung entschieden. Der drastische Fall der Speicherpreise sowie die Zunahme der Prozessorgeschwindigkeiten hat jedoch dazu geführt, dass SSD Systeme heute eine wirklich kosteneffektive Lösung für Unternehmen jeder Art und Größe geworden sind. Der SSD Markt ist einer der am schnellsten wachsenden Marktsegmente der Speicherindustrie.

Überlegungen zum Wechsel zu SSD Systemen

Um das Problem der I/O Wartezeit zu lösen, muß der Administrator erst feststellen, in welchem Segment seiner Orakeldatenbank das höchste I/O Niveau auftritt und dadurch I/O Wartezeiten verursacht. Die folgenden Datenbank Segmente sollten dabei besonders beachtet werden:

Die gesamte Datenbank. Es gibt einige Datenbanken, die alle ihre Daten auf einem RamSan-220 System laufen haben. Diese Datenbanken sind durch folgende Eigenschaften ausgezeichnet:

- Hohe, gleichzeitige Zugriffsrate auf spezielle Datenbanken. Jene Datenbanken, die von vielen Mitarbeitern zur gleichen Zeit benutzt werden, sollten alle Daten auf der SSD speichern. Diese Strategie minimiert den Engpass zur Anwendung, während gleichzeitig die Ausnutzung der Server und des Netzwerks optimiert wird. Die I/O Wartezeit wird kleiner, weil Server und Netzwerk Bandbreiten besser ausgenutzt werden.
- Hoher Zugang, jedoch verteilt auf viele Datenbank Tabellen. Es gibt Datenbanken, wo es nicht möglich ist, jene Dateien zu identifizieren, welche sehr

oft verwendet werden. Meistens sind solche Datenbanken sehr grosse Index Server.

- Kleine und mittlere Datenbanken. Die Implementierungskosten für ein RAID System sind oft sehr hoch. Es ist in diesen Fällen meist ökonomischer, ein SSD System zu installieren. Ein RamSan-220 System mit 32GB zur Datenbankspeicherung kostet nicht mehr als ein Enterprise RAID System.
- Der Schlüssel zum Erfolg eines Unternehmens ist auch die Performance seiner Datenbank. Sehr viele Datenbanken helfen den Firmen mehr Geld zu verdienen - oder auch weniger zu verlieren. Wenn Anfragen schneller bearbeitet werden können, erhöht dies die Zufriedenheit der Kunden. Das RamSan-220 System kann genau dies und hilft dem Unternehmen dadurch mehr Profit zu erzielen.

Redo Logs. Redo Logs sind einer der wichtigsten Faktoren in der Schreibleistung bei Orakel Datenbanken. In dem Moment, wo ein Schreibzugriff auf die Datenbank eintritt, legt die Orakel Datenbank einen Redo Eintrag an. Jeder Redo Eintrag wird zwei mal geschrieben. Orakel empfiehlt den Gebrauch von mindestens zwei Redo Logs, damit im Falle eines Ausfalls noch ein Backup Redo Log vorhanden ist. Der Schreibvorgang ist erst in dem Moment abgeschlossen, wenn der Schreibvorgang auch im Redo Log abgeschlossen ist.

Während des Datenbankbetriebes sind Redo Logs die Quelle konstanter I/O Vorgänge. Es ist wichtig, daß Redo Logs auf der möglichst schnellsten Festplatte gespeichert werden. Es ist verständlich, dass ein Verschieben der Redo Logs auf eine SSD, wie es das RamSan-220 System darstellt, die gesamte Leistung der Datenbank verbessert.

Index. Ein Index ist eine Datenstruktur, die den Zugang zu Datenbankaufzeichnungen beschleunigt. Ein Index wird normalerweise für jede Tabelle in einer Datenbank angelegt. Diese Indizes werden immer dann aktualisiert, wenn Aufzeichnungen hinzugefügt werden und wenn die gekennzeichneten Daten einer Aufzeichnung geändert werden. Wird von der Datenbank gelesen, entsteht zuerst eine Anfrage an den Index, damit Orakel schneller an die korrekte Aufzeichnung gelangen kann. Ausserdem können viele Benutzer jeden möglichen Index gleichzeitig lesen. Die Anfragen an das Laufwerk werden durch häufige, kleine Transaktionen charakterisiert. Unter diesen Umständen sind herkömmliche Laufwerke nicht imstande, Anfragen schnell genug zu beantworten. So entstehen I/O Wartezeiten.

Durch die Speicherung von Indizes auf einer RamSan-220 SSD kann die Leistung der gesamten Anwendung erhöht werden. Eine hohe Anzahl von gleichzeitigen Benutzern verursacht schnelle Datenbankzugriffe bei Online-Transaktionsverarbeitungen (OLTP) Systemen. Weil Indizes von den vorhandenen Daten neu erstellt werden können, sind sie stets Oracle Komponenten, die auf SSD's verschoben werden sollten.

Temporäre Segmente. Temporäre Segmente werden verwendet, um temporäre Daten während bestimmter Oracle Operationen zu unterstützen. Die Tabellen unterstützen komplizierte Anfragen, Anmeldungen und Indexkreationen. Weil temporäre Segmente viele verschiedene Operationen unterstützen, werden sie schnell fragmentiert. Interne Tests von Texas Memory Systems haben ergeben, dass die Orakeldatenbankleistung schnell vermindert wird, wenn Daten fragmentiert sind.

Komplizierte Datenbank Operationen werden schneller durchgeführt, wenn das temporäre Segment auf eine SSD verschoben werden kann. Oftmalige Anfragen an temporäre Segmente bringen herkömmliche Festplatten schnell an ihr Limit.

Rollback Data. In jenen Datenbanken mit einer hohen Anzahl gleichzeitiger Benutzer, sind oft die Rollbacksegmente Ursache von Datenstaus. Rollbackdaten entstehen durch Oracle Transaktionen in denen Einträge geändert werden. Das heißt, wenn ein Löschbefehl gegeben wird, werden alle ursprünglichen Daten im Rollbacksegment gespeichert, bis der normale Betrieb fortgeführt wird. Wenn die Operation rückgängig gemacht wird, werden diese Daten vom Rollbacksegment zurück zu jenem table verschoben, von dem sie vorher entfernt wurden.

Weil die Rollbacksegmente bei jedem Updatebetrieb genutzt werden und ausserdem die Zahl der Rollbacksegmente begrenzt ist, wird empfohlen die Rollbacksegmente auf einer SSD zu speichern. Dies führt zu schnelleren Schreibzugriffen, während die Update Operation läuft und stellt ausserdem die Rollbacksegmente schneller für den folgenden Updatebetrieb zur Verfügung.

Häufig verwendete Tabellen. Es wird geschätzt, daß nur 5 % -10 % jener Daten, die in den OLTP Systemen gespeichert sind, häufig verwendet werden. Diese Tabellen sind aber gewöhnlich der größte Prozentsatz aller Datenbanktätigkeit und folglich der I/O Aktivität. Wenn viele Benutzeranfragen eine Tabelle betreffen, sind diese wahrscheinliches für unterschiedliche Aufzeichnungen und Attribute in dieser Tabelle bestimmt. Infolgedessen ist die Tätigkeit auf dieser Tabelle vom Zufall bestimmt und nicht geordnet. Herkömmliche Laufwerke sind notorisch schlecht, um zufällige Anfragen schnell zu beantworten. Tatsächlich fällt die Höchstleistung eines Laufwerks um bis zu 95 %, wenn sie zufällige Anfragen beantworten muss. Wenn eine Tabelle sehr oft Anfragen erhält, entstehen Anfrage Warteschlangen, wobei neue Anfragen buchstäblich auf die Festplatte warten, bis die vorhergehenden Anfragen erledigt wurden. Diese Warteschlangen sind ein sicheres Zeichen für eine I/O Wartezeit in Ihrem System.

Es ist sinnvoll, häufig verwendete Tabellen auf ein RamSan-220 System zu verschieben. Durch zufällige Anfragen wird die Leistung eines RamSan-220 nicht verändert. Außerdem haben SSD's grundsätzlich schnelleren Zugriff als herkömmliche Laufwerke. Folglich wird die Leistung einer Anwendung bis zu 10x verbessert, wenn häufig verwendete Tabellen auf das RamSan-220 System verschoben werden.

Sie möchten mehr wissen?

Ein ausführlicheres White Paper mit einer kompletten Beschreibung, wie man die Prozessorleistung und die I/O Wartezeiten feststellt und die am häufigsten benutzten Dateien und Tabellen Ihres Systems erkennt, erhalten Sie unter: <http://www.superSSD.com/whitepapers.htm>.

Um mehr über SSD's zu erfahren oder um Hilfe beim Erkennen der I/O Engpässe in Ihren Systemen zu erhalten, kontaktieren Sie bitte Woody Hutsell, Director of Marketing for Texas Memory Systems, woody@superSSD.com.

Weitere Informationen über SSD's und SSD Applikationen finden Sie auf unserer Webseite: www.superSSD.com.

Texas Memory Systems präsentiert sein RamSan-220 System auf der CeBIT Show 2003 in Hannover, 12 – 19 März. Besuchen Sie uns in Halle1, USA SAN Pavillon, Stand 6d12, Booth A5