
Os efeitos de Unidades de memória em Estado Sólido sobre o

Microsoft SQL Server 2008



Índice

Visão Geral	3
Objetivos	3
Principais resultados	4
O Problema	8
A Tecnologia	8
A Metodologia de Teste	9
Benchmark baseado no TPC-C	13
Etapas de Configuração do Armazenamento	14
Custos de Energia	16
Resultados do Teste	17
Apêndice A — Configuração Detalhada do Servidor	21
Apêndice B — Configuração Detalhada do Armazenamento	23
Sobre o Autor	25

Resumo

Este documento contém os resultados dos testes realizados com o intuito de se conhecer melhor de que forma as unidades de estado sólido da Kingston® afetam positivamente a escalabilidade dos bancos de dados Microsoft® SQL Server® 2008. As conclusões foram aqui apresentadas sob o ponto de vista do usuário final, utilizando a metodologia de benchmarks do padrão TPC-C de um grande banco de dados OLTP. Os resultados mostram um aumento potencial significativo na carga do usuário e escalonamento e apresentam, ao mesmo tempo, melhores tempos de resposta para todas as cargas testadas. Além disso, numa configuração de 15.000 usuários, constatou-se uma economia de energia de 45% com o uso de unidades de estado sólido (SSD) em relação às tradicionais unidades de disco rígido. As unidades de estado sólido da Kingston oferecem uma melhor experiência de usuário a um custo - se não similar - inferior, com uma complexidade significativamente reduzida quando comparadas a subsistemas normais de disco.

Visão Geral

Existe uma ideia generalizada que, para conseguir maior desempenho com as tradicionais opções de armazenamento existentes, é necessário aumentar o número de discos rígidos (independentemente de sua capacidade) para se obterem níveis de desempenho aceitáveis para determinada carga de trabalho. Frequentemente são utilizadas centenas de discos, mesmo em um banco de dados razoavelmente pequeno em relação ao tamanho das unidades combinadas, devido às necessidades gerais de I/O da infraestrutura.

Além disso, há ainda uma grande complexidade inerente à grande quantidade de discos utilizados em uma carga de trabalho, com um número cada vez maior de peças móveis, maior necessidades de energia, ruído, maior potencial de quebra de máquinas etc. Com a tecnologia de unidades de estado sólido (SSD), um servidor pode fornecer o mesmo grau de performance, ou superior, com uma infraestrutura substancialmente menor. A tecnologia SSD também pode acrescentar a toda a equação uma significativa economia de energia, graças aos diferenciais de mecânica física, calor e infraestrutura de suporte necessários.

Este documento enfoca um cenário específico que compara as unidades de estado sólido com a utilização de discos SCSI (SAS) acoplados em série, no cenário de uma plataforma de banco de dados do Microsoft SQL Server 2008. O banco de dados objeto desta análise é de um porte razoavelmente grande - um conjunto de dados OLTP (On-Line Transaction Processing) de 100GB, criado e testado com um benchmark que se assemelha ao teste do TPC-C (Transaction Processing Council), aplicando os testes Benchmark Factory[®] para banco de dados da Quest. O uso dos padrões Benchmark Factory em vez dos testes comuns de subsistemas de discos I/O oferece melhores indicadores sobre a maneira como o usuário final será afetado por eventuais alterações em sua arquitetura de unidades de memória.

Apesar de haver um número cada vez maior de artigos técnicos on-line sobre a tecnologia da unidade de estado sólido, este artigo é único, pois ele é escrito a partir da perspectiva do usuário, com uma configuração que segue muitas das melhores práticas do SQL adotadas para a separação de uso dos discos. Por exemplo, o servidor testado foi configurado com sistemas operacionais, banco de dados e registros de transações independentes. Além disso, os níveis de RAID foram selecionados com base naquilo que é hoje adotado na prática para os bancos de dados, e não com o simples intuito de apresentar melhores resultados para este teste de benchmark.

Objetivos

Os principais critérios escolhidos para fundamentar esta análise foram:

1. Apresentar os efeitos da tecnologia SSD na escala e tempo de resposta para o usuário final.
2. Efetuar o teste junto ao usuário final, com um teste de benchmark que pode ser repetido, com base em cargas de trabalho do padrão TPC-C.
3. Utilizar um aplicativo de banco de dados comum, amplamente instalado no atual mercado.

4. Manter, ao máximo, as configurações padrão, da forma como são geralmente configuradas pelos usuários finais da plataforma de banco de dados do mercado. Isso impede a ocorrência de frequentes armadilhas observadas em diferentes comparações de testes de benchmark, que são os "ajustes de percurso".
5. Limitar os testes a uma carga de trabalho OLTP o mais próximo possível do mundo real. A mesma cópia básica do banco de dados básico foi utilizada em todos os testes. Embora o local da base de dados tenha sido o mesmo nas três interações do teste, a única variável foi o subsistema do disco. Isso assegura que o teste de benchmark possa ser duplicado por qualquer pessoa que tiver acesso ao mesmo hardware e software.
6. Mostrar a significativa economia de energia que pode resultar da tecnologia SSD em comparação com a tradicional tecnologia SAS, com o mesmo desempenho.

Principais Resultados

Foram analisados três cenários principais para o subsistema de armazenamento do banco de dados:

1. 48 unidades SAS configuradas como RAID10 externo com um Controlador RAID PERC 6/E.
2. 8 Unidades de Estado Sólido configuradas como RAID10 externo com um Controlador RAID PERC 6/E.
3. 4 Unidades de Estado Sólido configuradas como RAID10 interno, com um Controlador RAID PERC 6/I.

O desempenho e a escalabilidade geral da infraestrutura do banco de dados foi avaliado com base em duas variáveis principais:

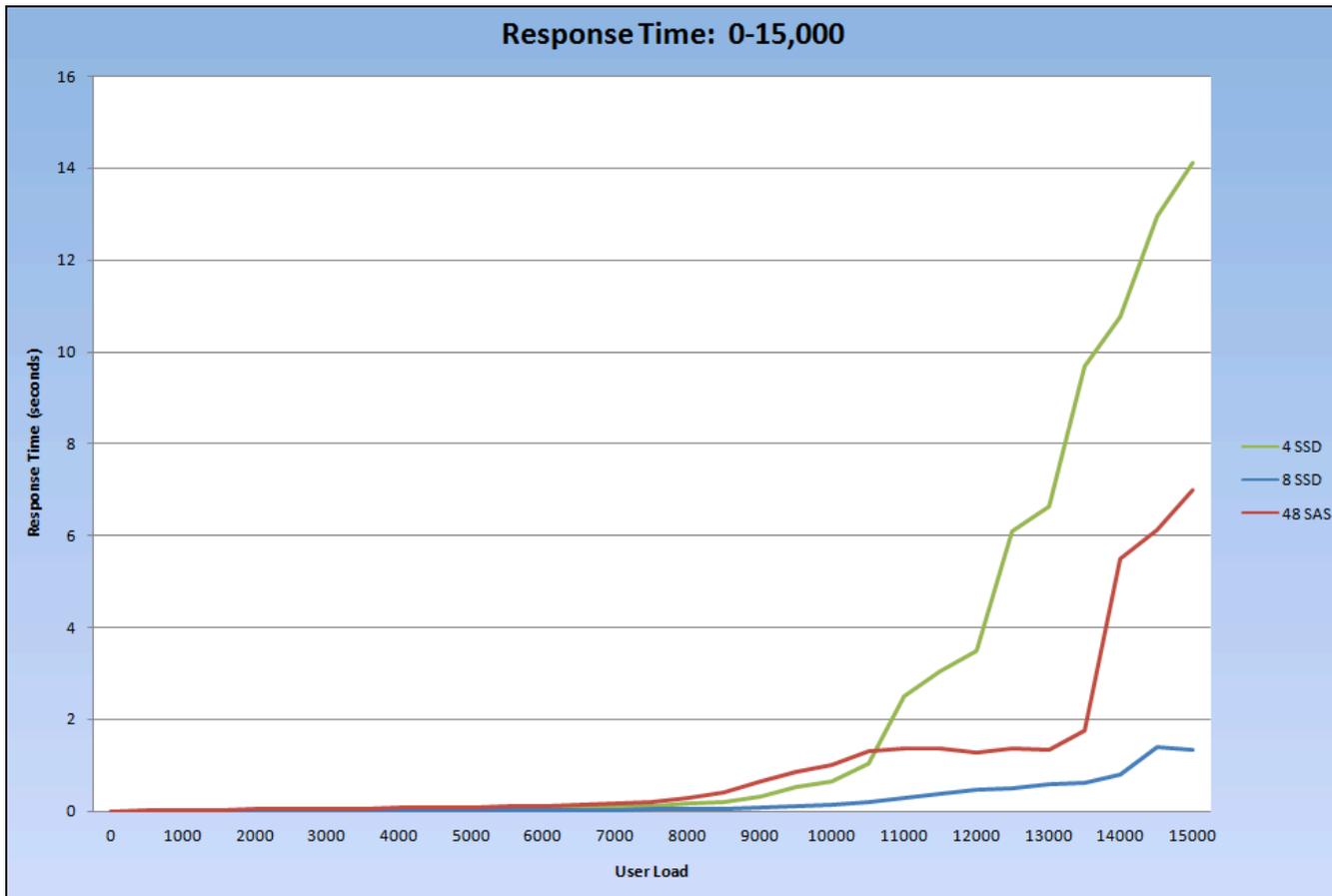
1. Tempo de Resposta
2. Contagem Total de Usuários (antes de atingirem tempos de resposta inaceitáveis)

Essas duas métricas de desempenho foram usadas para avaliar a escalabilidade das três opções de armazenamento.

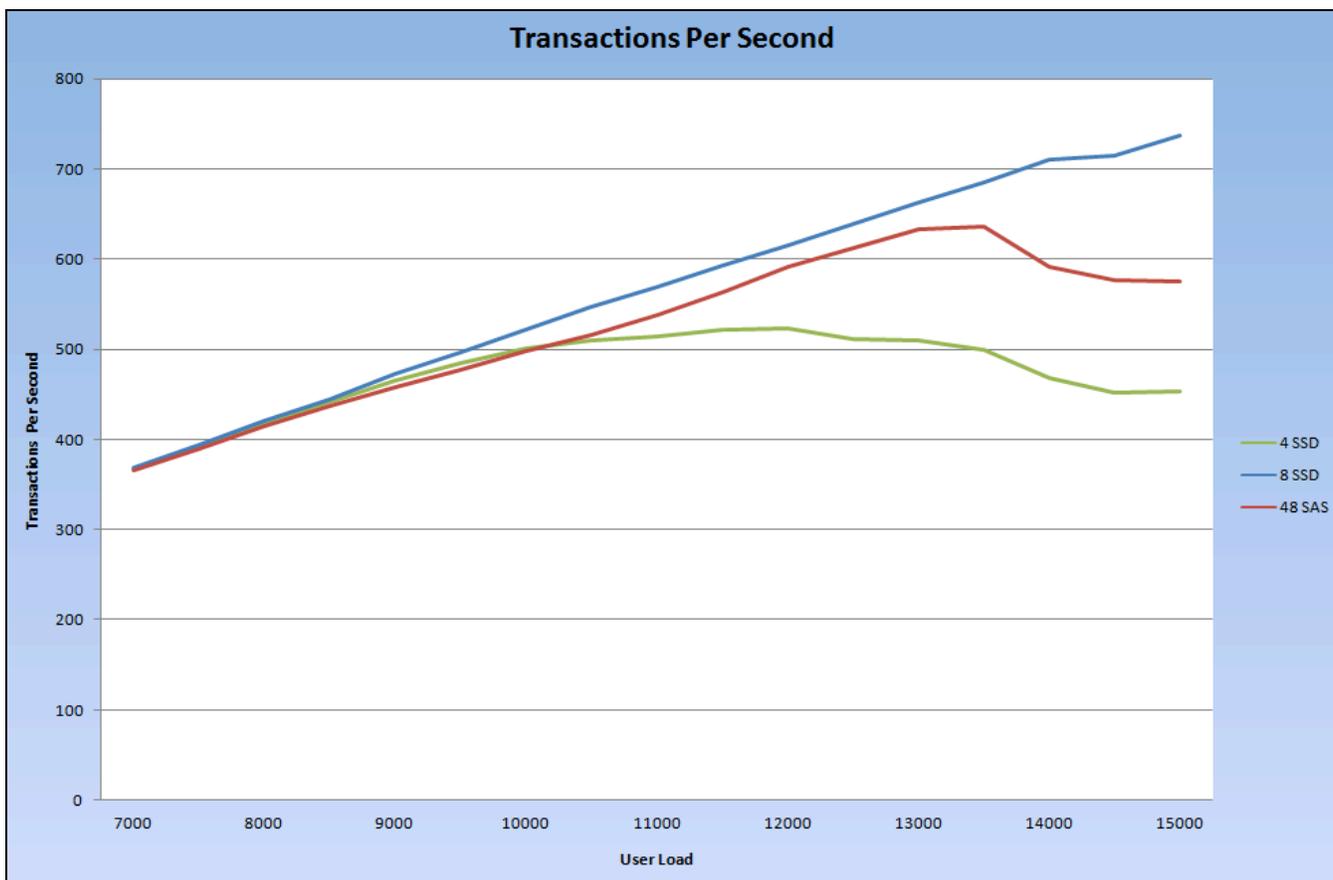
A métrica mais importante para o usuário final é o tempo de resposta. Quando o tempo de resposta médio de um banco de dados OLTP é de mais de um segundo, o desempenho é considerado inaceitável pelo usuário final. Por isso, a contagem total de usuários totais se baseia em um corte balizado no tempo de resposta superior a um segundo.

Para um número aproximado de 10.000 usuários, a infraestrutura SAS apresenta um tempo de resposta médio de um segundo. A infraestrutura SSD continua a escalar em 40% adicionais, totalizando 14.000 usuários antes de atingir o limite de um segundo de tempo de resposta.

Comparativos que mostram o "tempo de resposta" para cada plataforma de armazenamento:

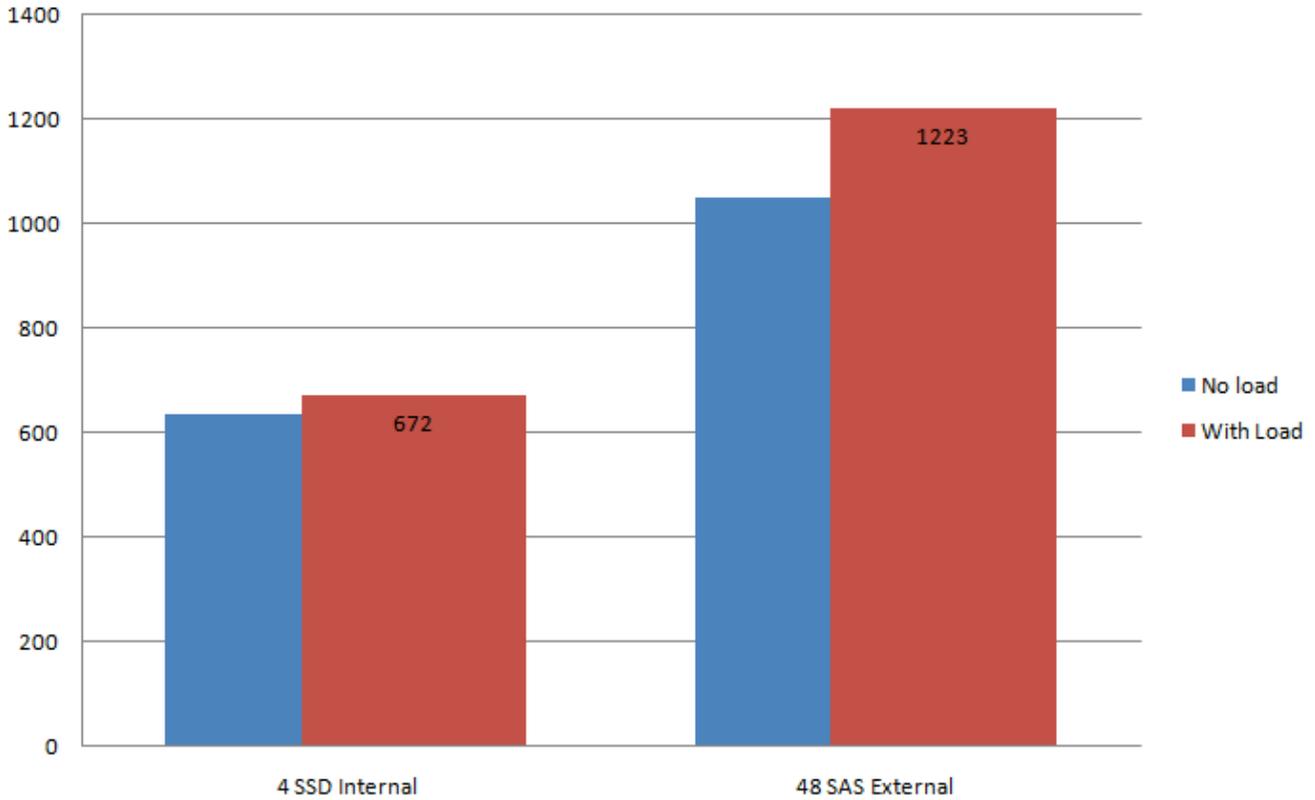


Diretamente relacionadas com o tempo de resposta, as transações totais por segundo apresentadas pela infraestrutura são menores com a infraestrutura SAS que na infraestrutura SSD.



A economia de energia decorrente do uso de (4) SSDs que apresentam o mesmo desempenho que (48) SASs estão detalhados abaixo. O teste comprovou uma redução do gasto de energia da ordem de 45%, utilizando-se o mesmo banco de dados com um número pré-estabelecido de usuários. Considerando a média nacional de gastos de energia nos Estados Unidos de US\$0,10 por quilowatt, isso representa uma economia de energia de US\$486 por ano para esta infraestrutura de servidor.

Power Usage in Watts



O Problema

O desempenho do servidor de banco de dados está, via de regra, diretamente relacionado com o subsistema do disco. Pelo fato de as memórias e CPUs terem atualmente um custo baixo, a grande maioria dos servidores de banco de dados do mercado dispõem de uma profusão de CPUs e memórias, mas carecem de capacidade de I/O para escalonar adequadamente. O Microsoft SQL Server 2008 é considerado um aplicativo de “elevado I/O”, e como tal, os aplicativos que utilizam o I/O com grande intensidade normalmente geram alto tráfego de armazenamento. Quando um banco de dados precisa escalonar para atender às necessidades de milhares, ou até de milhões de usuários, são necessários dezenas, se não centenas de discos rígidos. Essa contagem de disco frequentemente não é necessária para o cálculo da capacidade, mas para o desempenho. Cada disco pode suportar, individualmente, uma certa capacidade de I/O.

À medida que o armazenamento fica saturado, os tempos de resposta do armazenamento sobem para níveis inaceitáveis.

A Tecnologia

A nova linha de drivers de estado sólido SSDNow série E da Kingston utiliza unidades de estado sólido X25-E Extreme SATA da Intel. Essas unidades, que são NAND Flash baseadas em SLC, são atualmente consideradas as unidades SSD de melhor desempenho no mercado.

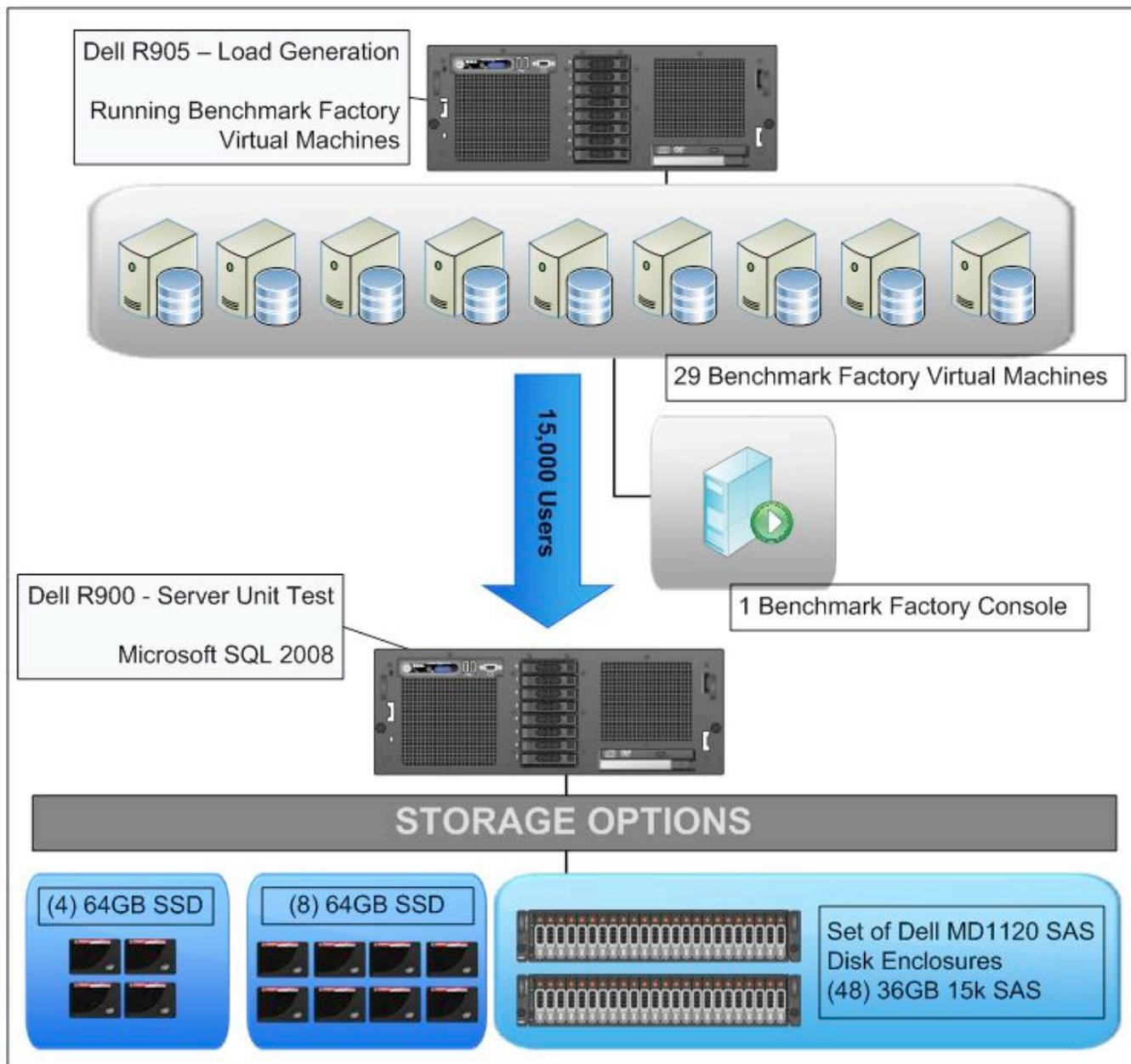
Há várias versões de SSDs disponíveis hoje, incluindo SLC (single-level cell) e MLC (multi-level cell). Pelo fato de a tecnologia SSD ser tão nova, frequentemente as duas são confundidas. Na infraestrutura de servidores, são necessárias unidades Flash baseadas em SLC para maior tolerância e para suportar maiores cargas de gravação I/O.

Uma discussão sobre as diferenças entre as tecnologias do SLC e do MLC estão além do escopo deste documento.

A Metodologia de Teste

Foram realizados vários procedimentos, ao mesmo tempo metódicos e focados, para garantir uma excelente comparação dos resultados. O foco dos testes foi um caso de utilização único: um grande banco de dados OLTP de 105GB. Só foram selecionadas as mais novas tecnologias para testar a infraestrutura do servidor e da unidade de memória.

Visão geral da infraestrutura de testes:



Consulte o Apêndice A para detalhes sobre a infraestrutura do servidor.

Como servidor do banco de dados a ser submetido aos testes (SUT), foi escolhido um Dell R900 Six-Core Quad-Socket com 64GB de RAM e um total de 24 cores de 2.4GHz). Apesar da quantidade de CPUs e de memória ter sido considerada mais elevada que o necessário, esta escolha foi intencional, para garantir que o foco deste estudo fosse mantido sobre a infraestrutura de drivers.

Para o sistema operacional a executar o Microsoft SQL Server 2008 Enterprise como plataforma de banco de dados, foi escolhido o Microsoft Windows Server 2008 Enterprise versão de 64 bits. O Microsoft SQL Server foi limitado a utilizar 2GB da memória do servidor, para garantir que o armazenamento fosse o ponto central do teste de benchmark. Para assegurar que o CPU, a memória, os discos do sistema e a rede não representassem uma sobrecarga durante os testes, foi selecionado o Spotlight da Quest sobre o SQL Server Enterprise.

A instalação do sistema operacional e do banco de dados foi feita em dois discos tradicionais Seagate SAS 73GB 10k SAS como RAID1. Essas unidades foram montadas diretamente no backplane do servidor de testes R900, em um controlador de armazenamento Dell PERC 6/i. Foram realizados vários testes com o sistema operacional instalado em outros conjuntos de dados antes de se tomar a decisão de se utilizar um simples conjunto de dados interno RAID1. Não foi detectado qualquer impacto na escala ou no tempo de resposta do banco de dados que estava sendo testado.

Um ponto comum de ajuste para SQL, TEMPDB, também foi localizado no conjunto simples de dados RAID1 do sistema operacional. Em alguns cenários, principalmente nos problemas da arquitetura do banco de dados, o TEMPDB, pode representar uma sobrecarga. Em uma instalação padrão, o TEMPDB pode terminar em uma unidade que não está tecnologicamente preparada para altas performances (normalmente o drive C:\ local). Para assegurar que o TEMPDB não representasse nenhum problema neste estudo, os testes foram executados com o TEMPDB transferido para um armazenamento extremamente distante. Não houve nenhum impacto em mover o TEMPDB para um ponto de armazenamento distante, o que permitiu que banco de dados fosse mantido em seu estado padrão.

Pela associação do I/O sequencial com os logs transacionais, um tradicional conjunto de discos RAID SAS não representou qualquer sobrecarga para os testes. Os resultados com outras cargas de trabalho podem variar. Pela maneira como os logs de transações fazem a gravação sequencial no disco, os benefícios do SSD não são tão profundos quanto o são usados com cargas de trabalho da banco de dados. Com isso em mente, foi utilizado um conjunto RAID1 de discos SAS de 73GB para todas as cargas de transações dos testes. Mesmo com uma quantidade muito grande de usuários nos testes, o acréscimo de mais de 40% da utilização do disco foi raramente percebida.

O tradicional armazenamento SAS de banco de dados incluiu dois gabinetes Dell MD1120 completamente carregado com discos SAS RPM Seagate de 15k com 36GB. Estes gabinetes foram então conectados a um controlador externo padrão PERC 6/E SAS RAID com 256MB de cache e configurado como RAID10 em todos os 48 discos (dois conjuntos de 24 discos).

As opções de armazenamento do banco de dados SSD incluíram dois cenários de configuração. A principal opção foi a utilização de oito SSDs de 32GB da Kingston, conectados externamente ao Dell R900 configurado como RAID10, contra o mesmo controlador de armazenamento Dell PERC 6/E como os MD1120s utilizado. A opção secundária foi a de utilizar quatro SSDs de 64GB montados diretamente no blackplane interno do R900 do Dell PERC 6/i.

Foram escolhidos dois cenários com drivers de estado sólido para garantir a imparcialidade do teste

(utilizando o mesmo controlador para um teste principal), mas também para mostrar o poder que os SSDs têm em substituir os tradicionais discos rígidos na complexidade de um dispositivo de armazenamento, com equipamentos extras e controladores.

A seguir, uma visão geral detalhada das diferentes opções de armazenamento do teste.

	48 Seagate 15k SAS	8 Kingston "E" 32GB	4 Kingston "E" 64GB
Tipo de Drive/Conexão	SAS/SAS, 15k	SSD/SATA	SSD/SATA
Controlador de Armazenamento	Dell PERC 6/E	Dell PERC 6/E	Dell PERC 6/i
Drives totais no teste	48	8	4
Nível do RAID	RAID 10	RAID 10	RAID 10
Capacidade Básica por drive	36GB	32GB	64GB
Volume Formatado	116GB	116GB	116GB

Consulte o Apêndice B para detalhes sobre a infraestrutura de armazenamento.

Os principais pontos de controle dos testes incluíram:

- Os benchmarks foram realizados em uma instalação climatizada controlada e com energia regulada localizada no centro da cidade de St. Louis, Missouri.
- O servidor que gerou a carga de trabalho permaneceu inalterado entre um teste e outro.
- Não foram feitas nenhuma modificação no SQL, banco de dados etc., entre um teste e outro.
- Foram empregados exatamente os mesmos scripts de testes entre os diferentes testes executados.
- Todas as estruturas de dados RAID foram configuradas com um tamanho de trilha de 64k, com um conjunto read-ahead ajustado para "adaptive" e cache de gravação configurado para "write back."

Os testes foram executados com um ferramenta Benchmark Factory para Bancos de Dados fornecida pela Quest Software. Embora se trate de uma ferramenta comercial, foi escolhida a Benchmark Factory (BMF) pelo foco da ferramenta em testes de usuários finais, em vez de um rudimentar teste de I/O do armazenamento dos subsistemas. É difícil encontrar uma comparação extraída do mundo real que mostre como um servidor irá reagir junto ao usuário final sem que se execute um estudo a partir desta perspectiva do usuário.

A ferramenta Benchmark Factory permite uma excelente escalonamento da carga para gerar a infraestrutura dos testes. Podem ser criadas várias instâncias de "carga" com as máquinas executando um simples agente do BMF. Foi utilizado um console central da Benchmark Factory para prescrever os testes e coletar os resultados dos agentes.

O foco do teste foi o tempo de resposta, utilizando-se um script de testes do tipo TPC-C padronizado no pacote de instalação da ferramenta BMF. A utilização deste teste padrão irá permitir que outras pessoas dupliquem o teste deste documento (caso desejem) para comprovar sua validade. Sabe-se que ajustes em diferentes fatores no servidor em teste pode aumentar ou reduzir os resultados que os testes da BMF podem oferecer. O foco deste artigo é a comparação entre as unidades utilizando exatamente a mesma configuração e a mesma infraestrutura.

Visão Geral da ferramenta Benchmark Factory:

- Foi utilizado o Benchmark Factory 5.8 da Quest para gerar a carga de trabalho do padrão TPC-C com até 15.000 usuários simultâneos. A infraestrutura do teste foi pré-testada e considerada apta para gerar mais de 30.000 usuários.
- As cargas de trabalho do Benchmark Factory foram executadas a partir de um servidor dedicado Dell R905 Quad-Core Quad-Socket, executando o VMware vSphere 4.0 com 29 máquinas virtuais como agentes. Uma máquina virtual Windows XP foi posteriormente instalada como o “console” do Benchmark Factory, para controlar aquelas máquinas agentes e reportar resultados.

Testes de Benchmark baseados no padrão TPC-C

A utilização do Processamento OLTP (Online Transaction Processing) se caracteriza por um alto nível de I/O, geralmente com grande volume de pequenas transações. Um benchmark padrão amplamente conhecido pela indústria é o benchmark “TLC-C”, do Transaction Processing Council. Este documento se baseia no teste realizado dentro do padrão “TPC-C”, integrado ao software Benchmark Factory da Quest. Apesar dos benchmarks neste documento não terem sido certificados pelo Conselho, eles refletem fielmente as normas do teste do “TCP-C”. O TPC-C do Benchmark Factory permite um alto nível de escalonamento, de testes e relatórios do ponto de vista do usuário final.

O benchmarking do padrão TPC-C emulam o uso de bancos de dados OLTP. O teste reproduz o status dos pedidos de compras, a entrada de novos pedidos etc., e pode ser comparado a entradas reais de bancos de dados de múltiplas camadas. Apesar de não ser perfeito, ele constitui um dos testes mais realistas existente, altamente propício à repetição. O teste simula vários depósitos que são monitorados, controlados, estocados, e têm suas entregas agendadas. As transações individuais podem gerar uma ou mais transações adicionais, como ocorre na vida real. Não são executadas quaisquer operações em lotes (batch), mas é possível estabelecer um ponto de controle no teste de benchmark para simular facilmente uma operação de batch.

As métricas de TPS geradas pelo Benchmark Factory incluem apenas transações de novos pedidos. O valor real do TPS seria maior se fossem contadas todas as transações. Entretanto, por este teste tratar de comparações e não de valores máximos de números, esse fato é certamente aceitável. Os dados gerados pelo benchmark representam a demanda do mundo real dos usuários finais. Pelo fato de o TPC-C já estar disponível há algum tempo, e ser considerado um teste de benchmark de alta credibilidade, os administradores de TI podem reproduzir este teste para validar os resultados apontados neste documento.

As Transações por Segundo (TPS) ou Transações por Minuto (TPM) são métricas amplamente aceitáveis. Contudo, esses valores não foram alvo do foco de nossos testes. A relação dos valores e a conexão destas métricas à latência dos usuários foram as principais áreas de concentração. O tempo de resposta afeta diretamente a percepção dos usuários sobre o desempenho de um banco de dados.

Os testes TPC-C de banco de dados OLTP gera um alto tráfego de armazenamento — a a latência do armazenamento é um fator crítico de sucesso nesse teste.

Etapas de Configuração do Armazenamento

Além de seguir os padrões exigidos das etapas de instalação, foram tomadas algumas medidas para garantir que a comparação fosse a mais isenta possível. O tamanho de stripe do conjunto, o alinhamento da partição, a idade do disco SSD e o tamanho da partição foram configurados exatamente da mesma forma em cada subsistema dos discos escolhidos para o banco de dados.

Tamanho do Stripe do Storage Controller

Todas os conjuntos (arrays) de armazenamento foram configurados com o tamanho padrão do stripe de 64k. O "read-ahead" do disco foi ajustado para "adaptive" e o cache da gravação foi configurado para "write-back" em todas as instâncias dos conjuntos, ao logo de todo este trabalho.

Dimensionamento das Partições

Todas as partições foram dimensionadas exatamente da mesma maneira, em 116GB. Este tamanho se baseou no tamanho máximo de um conjunto SSD RAID10 de quatro discos, já que ele foi o menor volume real.

Criação e Alinhamento da Partição

Para cargas elevadas de I/O, tanto nos conjuntos SAS como SSD, o alinhamento das partições é extremamente importante. Todos os arrays foram alinhados exatamente da mesma forma, com a ferramenta diskpart da Microsoft, como definido abaixo:

1. Num prompt de comando, execute diskpart.exe
2. Digite list disk
3. Digite select disk # (onde # é a quantidade de discos com os quais se está trabalhando)
4. Digite create partition primary align=32 (no caso deste documento, também utilizamos size=119805 como uma variável)
5. Digite active para ativar a partição.
6. Saia do prompt de comando
7. Use o Disk Manager do Windows para formatar e atribuir uma letra de drive à partição que agora está alinhada adequadamente. Para os testes deste artigo, toda a formatação foi feita com partições NTFS e tamanhos de alocação padrão.

Pré-Condicionamento do SSD

Para garantir a imparcialidade desse teste, as unidades SSD requerem uma atenção especial quando usadas em testes de benchmark. Um SSD "novo" é mais rápido na gravação do que ele será após um uso intenso. Embora este seja um item conhecido, fácil de ser explicado, sem este pré-condicionamento os resultados deste conjunto SSD poderia levar a interpretações equivocadas.

Num processo de produção, este pré-condicionamento NÃO é necessário. Ele só é necessário para fornecer resultados fiéis em um estudo da importância deste que gerou o presente artigo técnico. Os processos que foram seguidos incluíram as etapas abaixo:

1. Cada SSD foi individualmente pré-condicionado, embora conectado simultaneamente (não em RAID arrays)
2. Inicialize o disco e formate-o no NTFS do Windows.
3. Com a ferramenta IOMeter, execute um teste sequencial de leitura de 128k em todo o espaço de endereços do drive. Isso irá forçar todos os blocos do disco a ter algum tipo de conteúdo, não permitindo que o SSD não contenha um espaço excepcionalmente grande reservado para as gravações.
4. Delete o arquivo .tst do drive
5. Repita o teste acima por mais três vezes para eliminar qualquer espaço de reserva, excluindo o .tst e executando o teste novamente.
6. Novamente utilizando o IOMeter, foi executado um teste contínuo de leitura/gravação de 128k por 4 horas em todo o espaço de endereço do drive para condicionar a unidade.

Os Custos de Energia

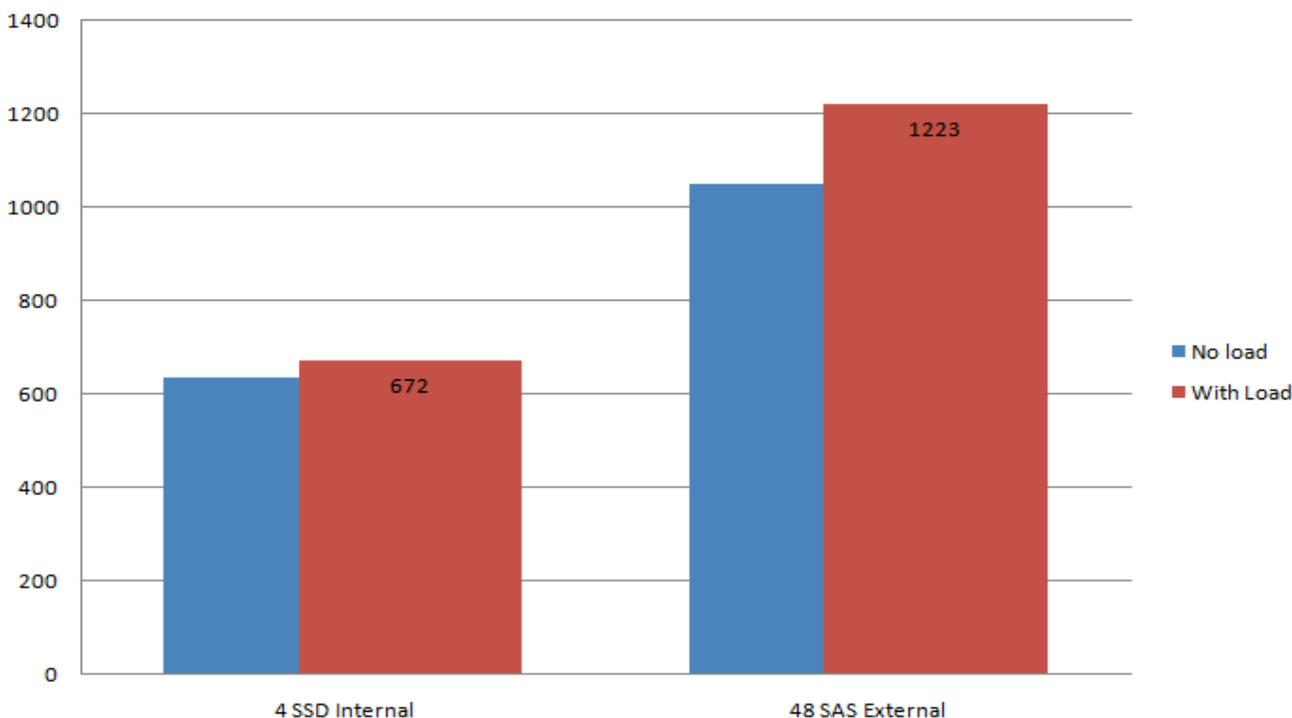
Quando as unidades de estado sólido substituem o tradicional disco rígido, mesmo em uma tradicional infraestrutura de servidores, a economia de energia obtida pode ser bem impressionante. Para conferir maior clareza, foi estudado o cenário mais provável de utilização. O estudo compara 48 unidades a 4 SSDs internos - que mais se assemelham em termos de desempenho - com base nos resultados do estudo de performance.

O servidor submetido ao teste, o Dell R900, o sistema operacional padrão e os discos de log de transações foram incluídos em ambos os cenários enfocando o gasto de energia. Os SSDs foram adicionados àquela base para um dos dois estudos; também no segundo estudo foram incluídos dois gabinetes MD1120 e 48 discos SAS 2.5".

O custo médio de eletricidade, em janeiro de 2009 nos Estados Unidos girou em torno de US\$0,10 por quilowatt-hora, conforme informa um relatório do governo dos EUA emitido pela agência de informações sobre energia EIA (Energy Information Administration).

Foi comprovada uma economia de 45% com a simples mudança para unidades de estado sólido para a mesma carga de trabalho. Um único servidor poderia economizar 551 watts de energia por hora, ou 13 quilowatts por dia ou 4.826 quilowatts por ano. Ao custo de US\$0,10 por quilowatt, uma economia seria de US\$482 por ano, pela simples substituição de unidades disco rígido tradicionais para as unidades de estado sólido.

Power Usage in Watts



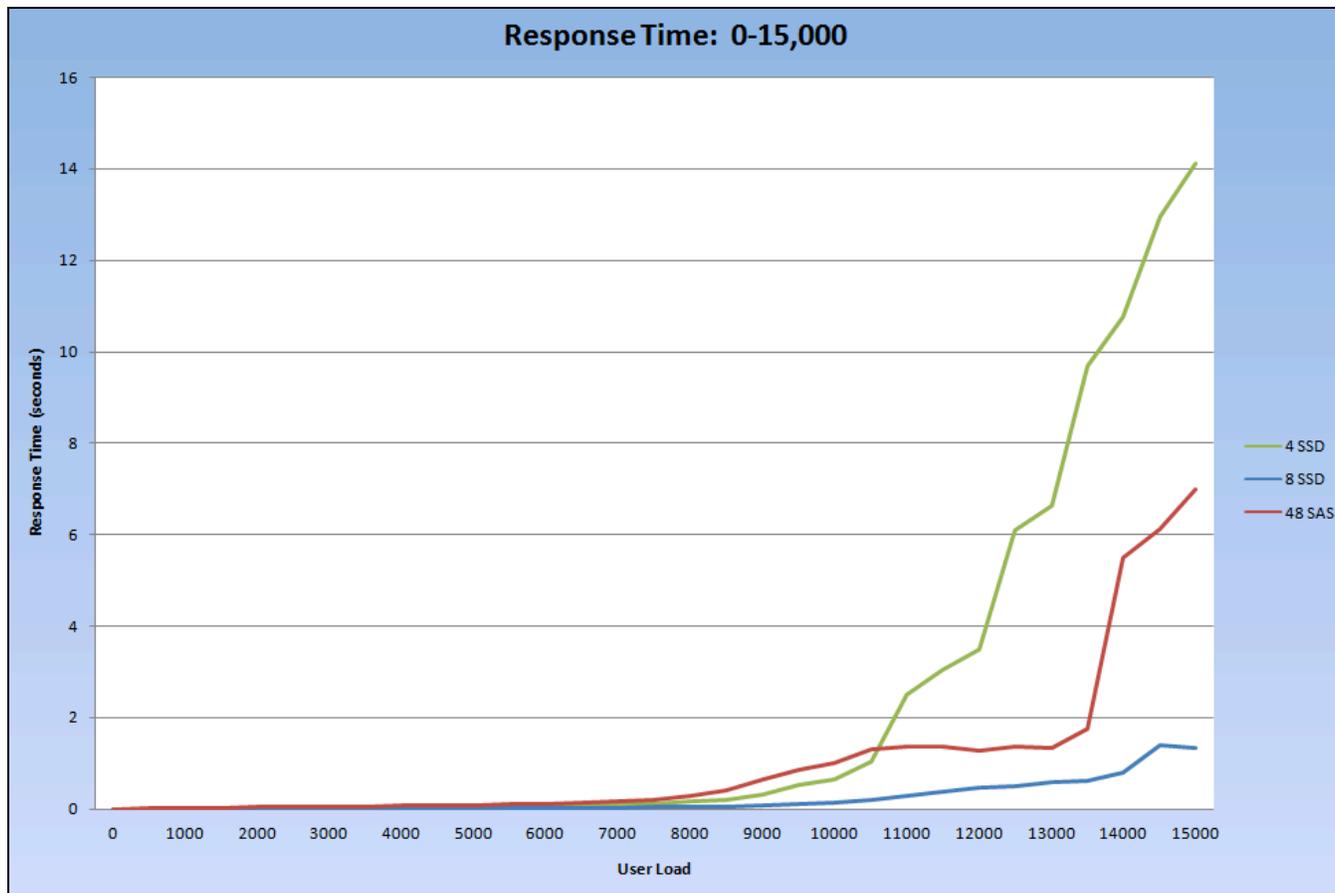
Os Resultados do Teste

Um dos principais componentes da capacidade do Microsoft SQL Server 2008 de oferecer um bom nível de desempenho para um aplicativo OLTP de carga de trabalho é a latência do armazenamento. Assumindo que a plataforma do host não tenha sobrecargas em sua CPU ou na memória, o desempenho do usuário passa a se basear na velocidade com que o armazenamento e a rede respondem aos pedidos. A correta configuração do conjunto de armazenamento reduz a espera no tempo de resposta; no entanto, a tecnologia de unidades de estado sólido reduz significativamente o tempo médio de resposta por transação no ambiente.

A latência em aplicativos OLTP provoca problemas de desempenho e altos tempos de resposta para o usuário final. Longos períodos de latência podem fazer com que o servidor do banco de dados desencadeie algumas atividades de salvar recursos, ou, ainda pior, pode fechar o aplicativo.

Os dados da rede, da CPU e da memória não estão presentes nos resultados dos testes, pois todos os três não representam fatores para os resultados. O hardware, “em excesso” foi propositalmente escolhido para garantir que a sobrecarga fosse concentrada no subsistema do armazenamento.

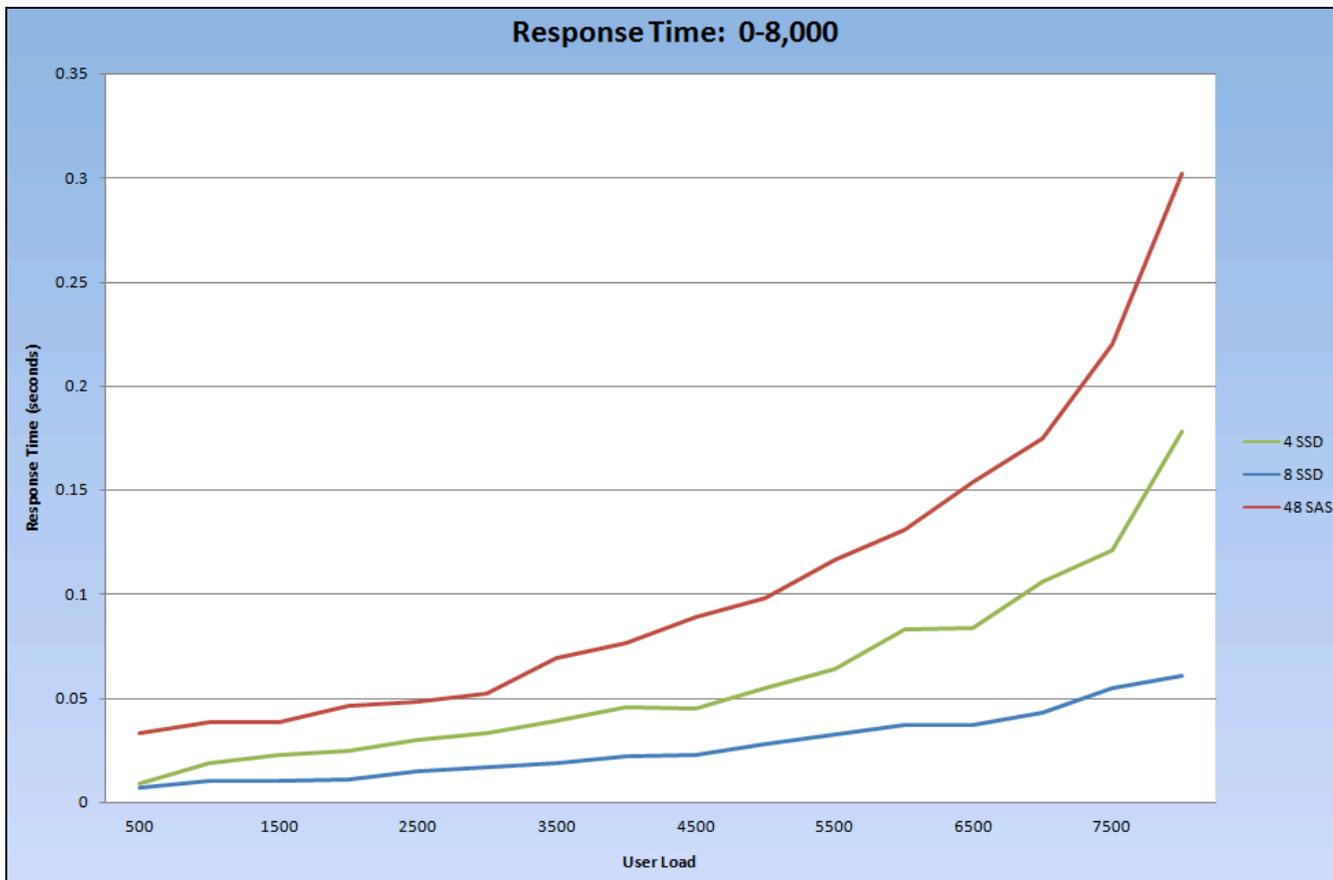
Esta tabela relaciona o tempo de resposta de 0 a 15.000 usuários em 500 ciclos de teste. Quando o teste atingiu 10.000 usuários, o tempo de resposta dos 48 discos do conjunto SAS passa a atingir níveis inaceitáveis.



É interessante observar que a configuração interna básica de 4 unidades SSD equivale ao conjunto de 48 unidades SAS até que a infraestrutura de 4 unidades encontra uma sobrecarga no nível de 10.500 usuários. Com pouco menos do que 10.500 usuários, essa configuração resultaria em um sistema mais rápido, com uma infraestrutura consideravelmente menor.

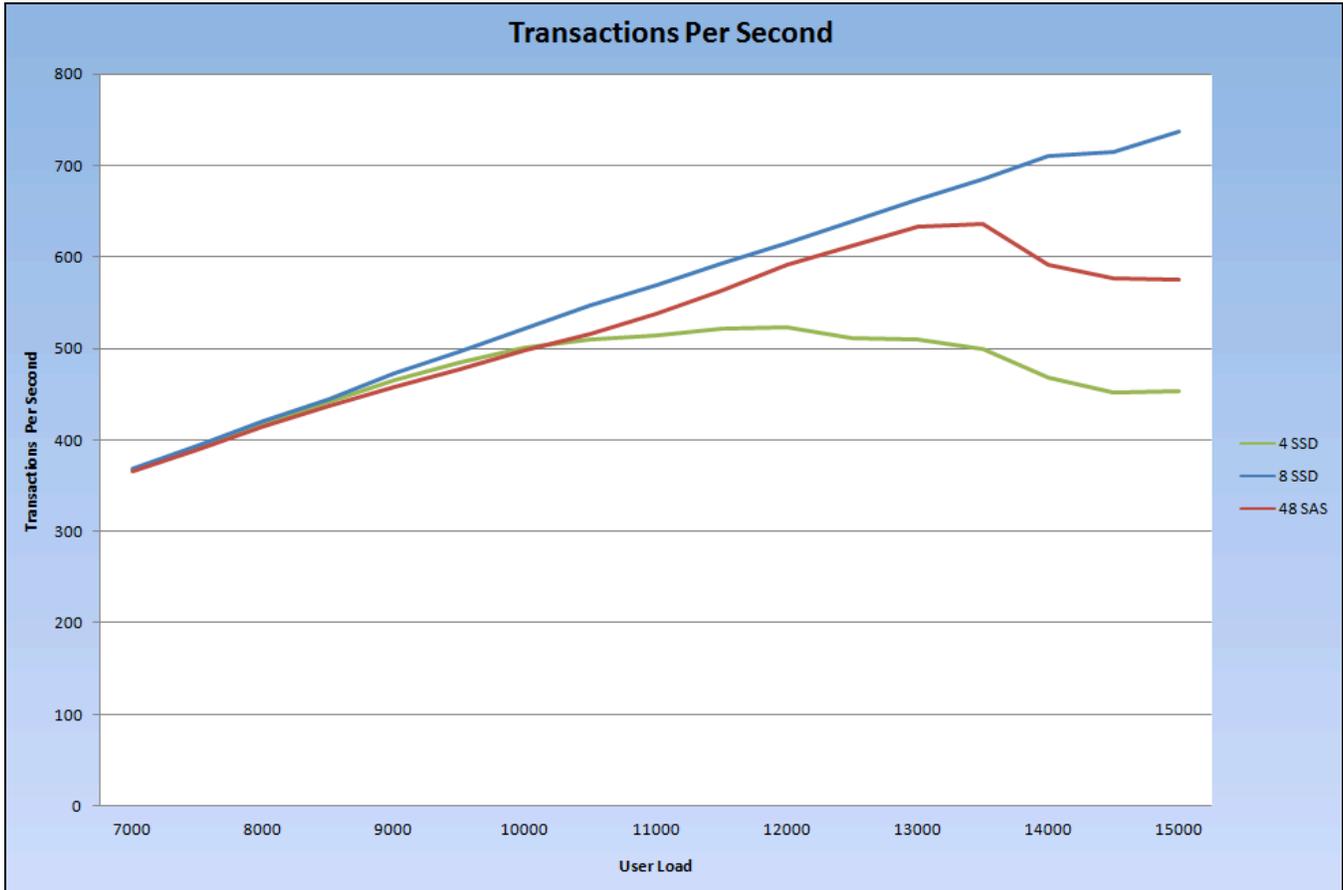
A infraestrutura de 8 unidades SSD continua a escalonar com tempos de resposta aceitáveis com até os 15.000 usuários, enquanto uma configuração SAS similar à da configuração de 4 unidades SSD aumenta bruscamente para valores de tempo de resposta.

Para uma visão mais detalhada, esta tabela mostra o tempo de resposta com uma carga de usuários que vai de 0 a 8.000. Mesmo quando a performance para o usuário final seja considerada “aceitável”, as opções com unidades de estado sólido apresentam um desempenho médio de 50% mais rápido em todas as cargas de usuários.



O tempo de resposta não é limitado à métrica de latência do usuário. A melhora do tempo de resposta provoca uma redução nas transações realizadas por segundo dentro da infraestrutura do banco de dados.

Conforme observado nos gráficos de tempo de resposta, a configuração de 4 unidades começa a declinar com aproximadamente 10.500 usuários. A configuração SAS, no nível de 13.500 usuário também começa a declinar. Contudo, a configuração de 8 SSDs continua a aumentar quase linear até atingir 15.000 usuários.



Apêndice A – Configuração Detalhada do Servidor

Tabela 1: Servidor submetido ao Teste (servidor efetivo do banco de dados).

Dell R900 Six-Core Quad-Socket	
Detalhes da Plataforma	
Fabricante/Modelo	Dell R900
Versão do Bios	1.1.10
Configurações do Bios	Padrão
CPUs	
Fabricante/Modelo	Intel Xeon E7450
Velocidade por Core	2,4GHz
Número de Sockets	4
Número de Cores por Scket	6
Número Total de Cores	24
Threads de Hardware por Core	1
Módulos de Memória	
Tipo	PC2-5300 FB-DDR2
Tamanho do Módulo	2048MB
Número de Módulos	32
Memória Total do Sistema	64GB
Subsistema da Rede	
Fabricante/Modelo	Broadcom Corporation BCM5708C
Tipo	PCI Express
Unidade	4.6.110.0
Quantidade NIC Total	4
Sistema Operacional	
Nome	Windows Server 2008 Enterprise 64 bits
ID do Build	6001
Service Pack	1
Sistema de Arquivos (todos os conjuntos de unidades)	NTFS
Kernel	Multiprocessador ACPI x64
Idioma	Inglês

Tabela 2: Servidor de Geração de Carga

Dell R905 Quad-Core Quad-Socket	
Detalhes da Plataforma	
Fabricante/Modelo	Dell R905
Versão do Bios	3.0.2
Configurações do Bios	Padrão
CPUs	
Fabricante/Modelo	AMD Opteron 8356
Velocidade por Core	2,3GHz
Número de Sockets	4
Número de Core por Socket	4
Número Total de Cores	16
Threads de Hardware por Core	1
Módulos de Memória	
Tipo	PC2-5300 FB-DDR2
Tamanho do Módulo	2048MB
Número de Módulos	32
Memória Total do Sistema	64GB
Subsistema da Rede	
Fabricante/Modelo	Broadcom Corporation BCM5708
Tipo	PCI Express
Unidade	4.6.110.0
Quantidade NIC Total	4
Subsistema do Disco	
Tipo de Unidade	SAS, 10k
Modelo da Unidade	Fujitsu MBB2073RC
Montagem do Armazenamento	Backplane, R905 interno
Controlador de Armazenamento	PERC 6/i
Firmware do Controlador	6.1.1-0047
Unidade do Controlador	Native ESX
Firmware da Unidade	D406
Unidades Totais	8
Nível do RAID	RAID5
Sistema Operacional	
Nome	VMware vSphere 4.0
Sistema de Arquivos (todos os conjuntos de unidades)	VMFS/ext3

Apêndice B – Configuração Detalhadas do Armazenamento

Tabela 1: Principais opções de armazenamento do teste.

	48 Seagate 15k SAS	8 Kingston SNE125-S2/32GB	4 Kingston SNE125-S2/64GB
Tipo de Drive/Conexão	SAS/SAS, 15k	SSD/SATA	SSD/SATA
Controlador de Armazenamento	Dell PERC 6/E	Dell PERC 6/E	Dell PERC 6/i
Firmware do Controlador	6.1.1-0047	6.1.1-0047	6.1.1-0047
Unidade do Controlador	2.23.00.64	2.23.00.64	2.23.00.64
Firmware	D506	045C8621	045C8790
Total de Unidades no Teste	48	8	4
Nível do RAID	RAID10	RAID 10	RAID 10
Capacidade Básica por Unidade	36GB	32GB	64GB
Volume Formatado	116GB	116GB	116GB

Tabela 2: Armazenamento do Sistema Operacional (o mesmo para todos os testes)

	2 Seagate 10k SAS
Tipo de Unidade	SAS, 10k
Tipo de Conexão da Unidade	SAS
Modelo da Unidade	Seagate ST973402SS
Montagem do Armazenamento	Backplane, R900 interno
Controlador de Armazenamento	Dell PERC 6/i
Firmware do Controlador	6.1.1-0047
Unidade do Controlador	2.23.00.64
Firmware da Unidade	S22C
Total de Unidades no Teste	2
Nível do RAID	RAID 1
Capacidade Básica por Unidade	73GB
Volume Formatado	67GB

Tabela 3: Armazenamento do Log de Transações

	2 Seagate 10k SAS
Tipo de Unidade	SAS, 10k
Modelo da Unidade	Fujitsu MBB2073RC
Montagem do Armazenamento	Backplane, R900 interno
Controlador do Armazenamento	Dell PERC 6/i
Firmware do Controlador	6.1.1-0047
Unidade do Controlador	2.23.00.64
Firmware da Unidade	D406
Total de Unidades no Teste	2
Nível do RAID	RAID 1
Capacidade Básica por Unidade	73GB
Volume Formatado	67GB

Sobre o Autor

Jason Campagna é renomado líder nas áreas de virtualização corporativa, computação baseada no servidor e otimização de tecnologias através da inovação.

Ele é o principal arquiteto técnico (Chief Technical Architect) da Centric Consulting, uma empresa de consultoria cujo foco está na área de computação baseada no servidor e na virtualização. A Centric Consulting está sediada em St. Louis, no estado norte-americano de Missouri. O singular enfoque da companhia é "só tecnologia/sem vendas", oferecendo arquitetura, projetos de engenharia para grande parte das empresas classificadas no Fortune 1000, para o Governo, e também para os setores de finanças, saúde e educação.

Ligue grátis para: (888) 747-4700 ou entre em contato pelo email whitepapers@servercentric.com para mais informações.

