

Comparativo de desempenho do Pervasive PSQL v11

Um artigo Pervasive PSQL
Setembro de 2010

Conteúdo

Resumo executivo	3
O impacto das novas arquiteturas de hardware nos aplicativos	3
O projeto do Pervasive PSQL v11	4
Configurações.....	6
Caches de memória	6
Testes comparativos	7
Testes de operações atômicas.....	8
Conclusão	8

Resumo executivo

Antigamente, os avanços na tecnologia de hardware materializavam-se no aumento da velocidade de processamento. Os aplicativos existentes passavam a rodar mais rápido sem que fosse necessário qualquer esforço de adaptação. Hoje, os avanços na tecnologia da computação materializam-se em aumento do paralelismo, não em maiores frequências de clock. A utilização de múltiplos núcleos abre o caminho para o aumento do paralelismo, mas os aplicativos precisam ser escritos com código paralelo para que possam utilizar os vários núcleos simultaneamente. O Pervasive PSQL v11 Server foi projetado especificamente para melhorar a escalabilidade e o desempenho em máquinas multinúcleo e ambientes multiclente.

Este artigo mostra os resultados dos testes comparativos de desempenho do Pervasive PSQL v11 Server. O Pervasive PSQL v11 foi comparado à última versão da família 10.x, o Pervasive PSQL v10 SP3.

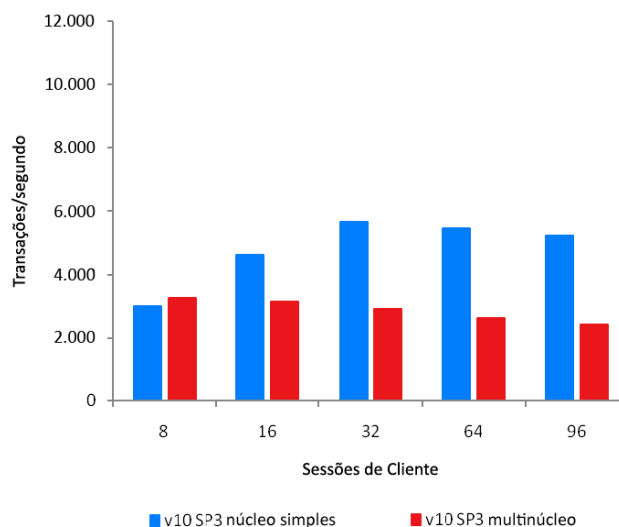
Em máquinas multinúcleo, com vários clientes acessando dados, o Pervasive PSQL v11 teve um desempenho significativamente superior ao Pervasive PSQL v10 SP3, em alguns casos em mais de 300%.

O impacto das novas arquiteturas de hardware nos aplicativos

Atualmente, as máquinas multinúcleo são a norma. Quando você atualiza suas máquinas de núcleo simples, seus aplicativos passam a trabalhar em uma arquitetura de hardware diferente. O desempenho da maioria dos aplicativos é afetada dramaticamente pela arquitetura multinúcleo.

Enquanto examinamos o assunto com mais profundidade, considere o seguinte gráfico de desempenho da transferência do Pervasive PSQL v10 Server de uma máquina de núcleo simples para uma máquina multinúcleo.

Figura 1. Desempenho do Pervasive PSQL v10 Server em máquinas de núcleo simples e múltiplo



Como podemos ver, o desempenho diminui com o aumento do número de sessões de cliente. Esse comportamento não é exclusivo do Pervasive PSQL v10 Server. Todos os aplicativos multiclente complexos que não tenham sido projetados especificamente para utilizar a arquitetura multinúcleo, além da maioria dos aplicativos criados durante a era do núcleo simples, estão sujeitos à redução do desempenho em máquinas multinúcleo.

Por quê? As razões técnicas são complexas e foram detalhadas no artigo [O Dilema Multinúcleo](#), de Dan Woods, chefe de tecnologia da CITO Research. Mas o breve resumo contido neste documento pode esclarecer os motivos por trás da noção pouco intuitiva de que os aplicativos multithreaded ficam mais lentos em sistemas multinúcleo.

Em aplicativos multithreaded que compartilham dados, a sincronização entre os threads consome muitos recursos do sistema em máquinas multinúcleo. O compartilhamento de dados é o flagelo da computação paralela. O acesso de vários threads aos mesmos dados precisa ser sincronizado. Quando sincronizamos o acesso aos dados, a parte do código que faz o acesso não pode ser executada por mais de um thread de cada vez, por isso o código deixa de ser concorrente. Essa parte do código torna-se uma passagem estreita e todos devem fazer fila para passar por ela.

A utilização de cache piora o problema em vez de amenizá-lo. Quando os caches de vários núcleos ou processadores estão apontando para os mesmos dados e um dos núcleos modifica os dados, os caches dos demais núcleos deixam de ser válidos e precisam ser sincronizados com os novos dados. O acúmulo de processamento extra devido a todas essas operações de sincronização é significativo. Em última análise, isso significa que o desempenho em processadores multinúcleo pode ser pior que em processadores de núcleo simples, que não precisam desse tipo de sincronização.

Sempre que possível, cada núcleo deve trabalhar com seus próprios dados. Caso contrário, a necessidade de sincronização leva ao processamento extra que pode reduzir significativamente o desempenho.

O projeto do Pervasive PSQL v11

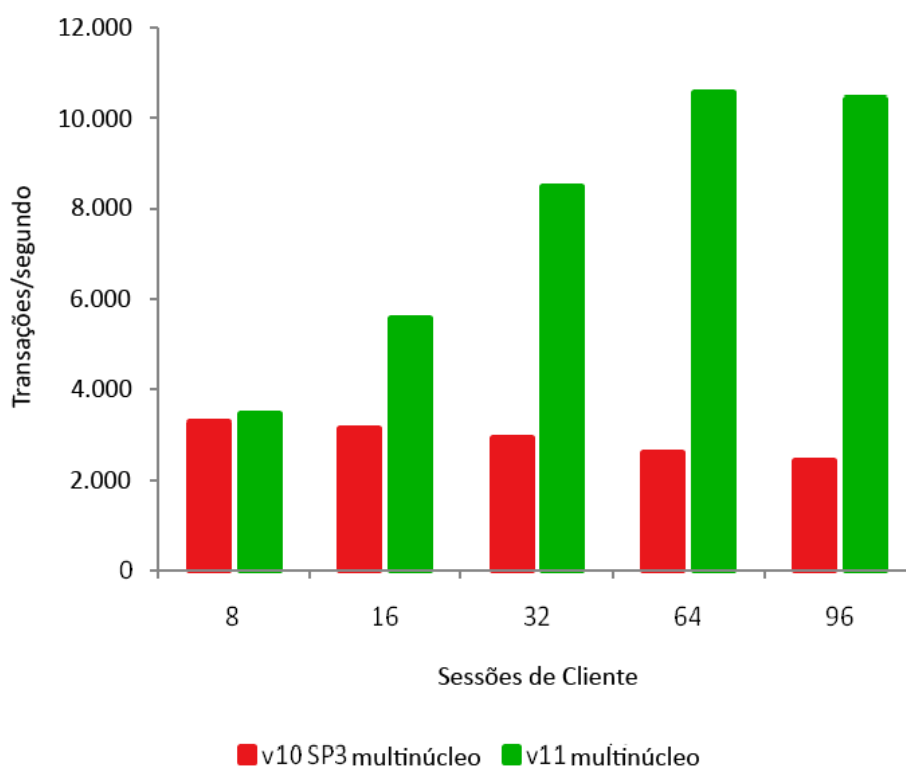
O Pervasive PSQL v11 foi projetado com uma arquitetura que fornece threads (linhas de execução) paralelos que executam atividades semelhantes. A capacidade de realizar processamento paralelo aumenta à medida que mais processadores são acionados. O resultado é que o desempenho do sistema de banco de dados aumenta nos ambientes multinúcleo em que vários clientes acessam um servidor central.

O Pervasive PSQL v11 também apresenta aprimoramentos nos mecanismos de sincronização de baixo nível da interface transacional. Vários usuários podem ler simultaneamente as mesmas páginas de arquivo em cache e suas operações podem ser executadas no servidor em CPUs independentes. Atividades não relacionadas aos usuários, como checkpoints e gerenciamento de logs, também podem utilizar as CPUs extras do servidor.

A escalabilidade do Pervasive PSQL v11 também foi aprimorada por meio de mudanças na arquitetura feitas especificamente para o hardware multinúcleo. Por exemplo, vários usuários acessando arquivos independentes podem ser atendidos por CPUs independentes do servidor. O sistema de banco de dados também pode suportar uma maior carga de usuários sem ficar sobrecarregado, resultando em um desempenho mais estável.

Como essas modificações melhoram o desempenho? Vamos comparar o desempenho em máquinas multinúcleo do Pervasive PSQL v11 e do Pervasive PSQL v10. Os resultados foram obtidos com a interface transacional acessando 16 arquivos completamente carregados no cache de memória.

Figura 2. Desempenho do Pervasive PSQL v11 Server e Pervasive PSQL v10 Server em máquinas multinúcleo



Como podemos ver, as mudanças na arquitetura do Pervasive PSQL v11 proporcionam um aumento significativo do desempenho. Melhor ainda: os seus aplicativos multicliente podem beneficiar-se desse aumento de desempenho sem que seja necessário recompilar nem reprojeter o código.

Configurações

Os testes comparativos de desempenho mostrados neste artigo foram realizados com as configurações descritas na tabela a seguir.

Tabela 1. Configurações utilizadas nos testes comparativos

Processadores na máquina Server	Duas CPUs Intel Xeon E5420 de 2.50GHz com 4 núcleos
Total de núcleos na máquina Server	8*
Memória total na máquina Server	16GB
Sistema operacional da máquina Server	Microsoft 2008 Enterprise Server (Service Pack 2), 64-bit
Software servidor	Pervasive PSQL v11 Server e Pervasive PSQL v10 SP3 Server
Configuração do Pervasive PSQL Server	Instalado com a configuração padrão.
Processadores na máquina Client	Uma CPU Intel Core2 Quad Q9400 de 2.66GHz com 4 núcleos. (duas máquinas físicas acessando o Pervasive PSQL Server com sessões de cliente divididas igualmente entre as máquinas cliente)
Total de núcleos na máquina Client	4 (em cada máquina)*
Memória total na máquina Client	4GB (em cada máquina)
Sistema operacional da máquina Client	Microsoft XP (Service Pack 3) 32-bit
Software cliente	Pervasive PSQL v11 Client para o Pervasive PSQL v11 Server. Pervasive PSQL v10 SP3 Client para o Pervasive PSQL v10 SP3 Server.
Configuração do Pervasive PSQL Client	Instalado com a configuração padrão.
* Como as máquinas encontradas no mercado atualmente têm 4 ou 8 núcleos, a Pervasive Software escolheu essas máquinas e configurações como sendo representativas dos ambientes reais dos nossos clientes. O Pervasive PSQL continuará a ser aprimorado à medida que o número de núcleos aumentar nas máquinas encontradas no mercado.	

Caches de memória

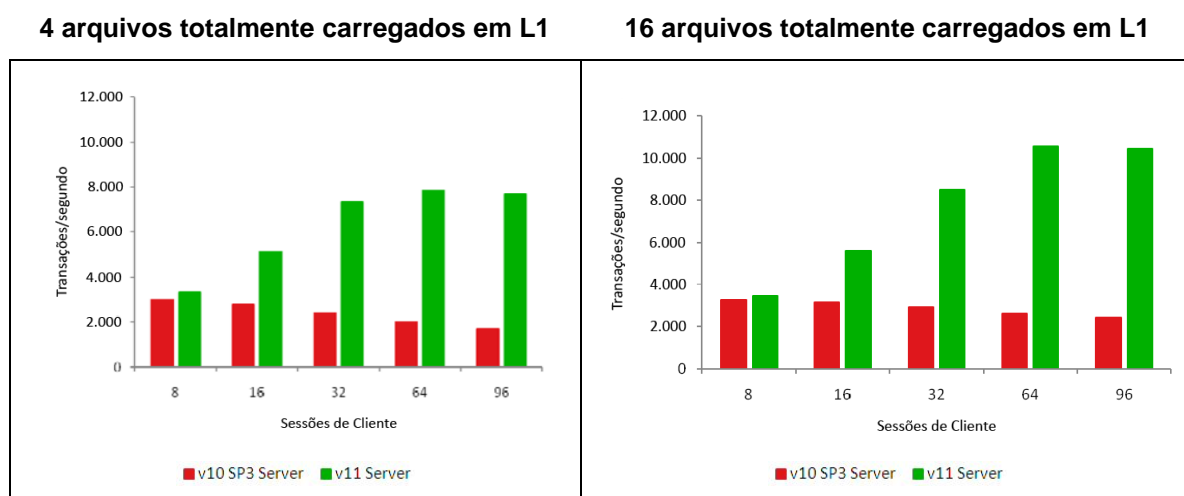
O Pervasive PSQL utiliza dois caches principais, Level1 e Level2, que vamos chamar de L1 e L2. O cache L1 tem um tamanho fixo determinado pela opção **Cache Allocation Size**. Seu tamanho não muda em função das operações do banco de dados. Os testes comparativos foram feitos com os dados totalmente carregados em L1.

Testes comparativos

Para demonstrar o efeito de diferentes cargas devidas às sessões de cliente, a Pervasive Software rodou um sistema de teste com instalações padrão do Pervasive PSQL no servidor e no cliente (veja a seção "Configurações" acima). O sistema de teste simula uma carga de processamento de maneira semelhante ao comparativo padrão TPC-B. Uma única transação executa um READ e um UPDATE sobre um registro de cada um de três arquivos, depois executa um INSERT de registro em um quarto arquivo.

Para conhecer detalhes sobre o comparativo TPC-B do Transaction Processing Performance Council, veja <http://www.tpc.org/tpcb/default.asp>.

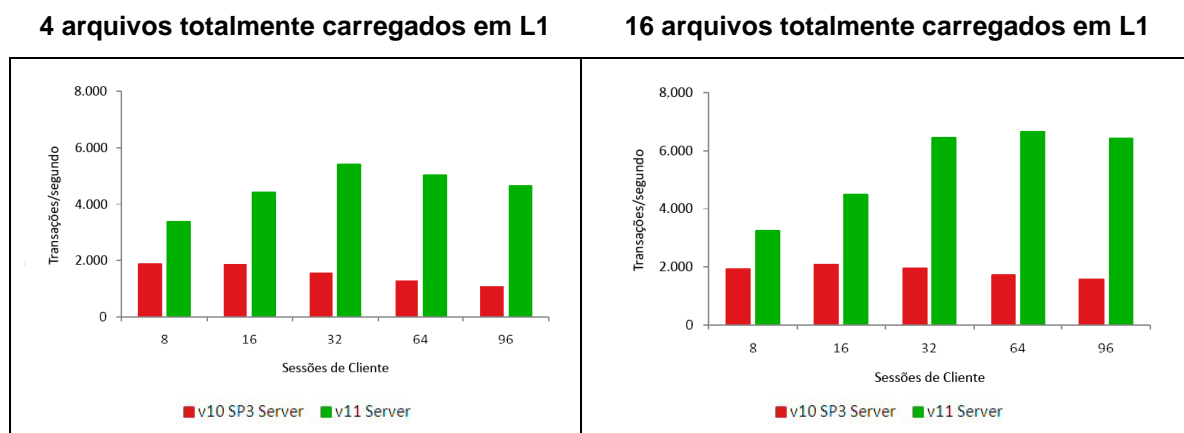
Figura 3. Resultados comparativos da interface transacional (Btrieve)



Aumento de desempenho do Pervasive PSQL v11 por sessões de cliente:

8	10%	8	6%
16	86%	16	77%
32	206%	32	191%
64	297%	64	302%
96	347%	96	328%

Figura 4. Resultados comparativos da interface relacional (ODBC)



Aumento de desempenho do Pervasive PSQL v11 por sessões de cliente:

8	79%	8	69%
16	136%	16	116%
32	240%	32	232%
64	292%	64	284%
96	326%	96	304%

Testes de operações atômicas

Além dos testes comparativos TPC-B, foram feitos testes de operações atômicas READ e UPDATE. Os resultados dos testes de operações atômicas em máquinas multinúcleo foram semelhantes aos dos testes TPC-B. Eles mostraram que o desempenho do Pervasive PSQL v11 é melhor que o do Pervasive PSQL v10 SP3. Para receber os resultados dos testes de operações atômicas, envie uma mensagem para database@pervasive.com ou ligue para 1 800 287 4383 (nos Estados Unidos).

Conclusão

Os números falam por si. O Pervasive PSQL v11 Server apresenta um aumento de desempenho em relação ao Pervasive PSQL v10 Server em máquinas multinúcleo. O aumento típico foi maior que 100%, chegando a ultrapassar 300% em alguns casos.

Quanto maior o número de arquivos de dados utilizados, melhor o desempenho, porque o processamento paralelo aprimorado do Pervasive PSQL v11 aciona vários processadores. Vários usuários podem acessar arquivos independentes e podem ler as mesmas páginas de arquivo em cache ao mesmo tempo porque essas operações são executadas por CPUs independentes no servidor.

Em comparação, o desempenho do Pervasive PSQL v10 Server em máquinas multinúcleo diminui à medida que o número de sessões de cliente aumenta. Todo aplicativo multiusuário complexo criado durante a era dos núcleos simples está sujeito à diminuição do desempenho em máquinas multinúcleo. Naquela época, não era possível prever as muitas considerações de projeto necessárias para promover o melhor desempenho em hardware multinúcleo.

O suporte a processadores multinúcleo é um recurso fundamental do Pervasive PSQL v11 porque o novo hardware multinúcleo altera dramaticamente o desempenho da maioria dos aplicativos. Esse suporte é essencial para a transição de seus aplicativos multiclente para os ambientes multinúcleo.