



Uma abordagem de rede unificada ao storage iSCSI com controladores Broadcom

Por Dhiraj Sehgal, Abhijit Aswath e Srinivas Thodati

Em ambientes baseados em Internet SCSI (iSCSI) e 10 Gigabit Ethernet, a implantação de servidores Dell™ PowerEdge™ com controladores de rede convergente Broadcom® NetXtreme II® pode ajudar os departamentos de TI a eliminar o gargalo de redes, reduzir o consumo de energia e maximizar a eficiência do data center.

O uso do protocolo Internet SCSI (iSCSI) aumentou significativamente em data centers empresariais principalmente pelo fato de ele oferecer diversas vantagens em relação às tecnologias tradicionais de storage Fibre Channel. Como o iSCSI fornece um método simples para transportar comandos, dados e mensagens de status do SCSI nas redes TCP/IP padrão, ele permite que as organizações aproveitem as infraestruturas e as bases de conhecimento existentes e, ao mesmo tempo, usem componentes conhecidos e econômicos. Ele também é baseado em padrões, o que facilita a adoção do setor e ajuda a garantir a interoperabilidade. E seu desempenho pode ajustar a escala com um caminho de tecnologia comum: a Ethernet.

Mas, com o crescimento e com a evolução da utilização de rede, o uso difundido do iSCSI também criou desafios. Por exemplo, muitas organizações estão executando um número crescente de aplicativos de conteúdo avançado e com grande volume de dados ou ambientes virtualizados em seus servidores, o que pode aumentar significativamente o tráfego na rede baseada em iSCSI em um data center. Para ajudar a manter esse crescimento, muitos departamentos de TI começaram a mudar de controladores de rede Ethernet Gigabit (GbE) para controladores Ethernet de 10 Gigabit (10 GbE). Entretanto, sob cargas de

trabalho pesadas, os controladores tradicionais podem consumir uma quantidade significativa de potência de processamento do servidor, reduzindo a quantidade disponível para aplicativos essenciais. Além disso, a abordagem tradicional normalmente exige o uso de controladores de rede diferentes para lidar com tipos diferentes de tarefas (como rede de alta velocidade, armazenamento e organização por clusters), o que leva a um número crescente de dispositivos, aumentando o custo e a complexidade, além de limitar a capacidade dos departamentos de TI de aproveitar os controladores 10GbE.

A família Broadcom NetXtreme II de controladores de interface de rede convergentes 10GbE (C-NICs) foi projetada para lidar com esses problemas. Disponíveis nos servidores Dell PowerEdge como LAN nas Placas-mãe (LOMs), placas de mezanino e placas de rede (NICs) de suporte, esses controladores oferecem a velocidade e a eficiência necessárias para oferecer suporte a cargas de trabalho pesadas de tráfego na rede, permitindo que os departamentos de TI simplifiquem seus sistemas e forneçam recursos para organização por clusters, storage e rede nas infraestruturas TCP/IP e Ethernet existentes, ajudando, ao mesmo tempo, a reduzir o consumo de energia e permitindo o uso altamente eficiente de recursos de processamento.



“Como controladores convergentes, as C-NICs Broadcom NetXtreme II podem lidar com vários tipos de tarefas de rede simultaneamente.”

Facilitando a carga de processamento de rede

Como controladores convergentes, as C-NICs Broadcom NetXtreme II podem lidar com vários tipos de tarefas de rede simultaneamente. Um servidor tradicional normalmente é equipado com quatro controladores GbE, juntamente com outros adaptadores de storage, em alguns casos. A execução simultânea de rede, storage e organização por clusters em taxas altas normalmente exige vários adaptadores e um grande número de ciclos de CPU. A abordagem da C-NIC Broadcom, por outro lado, permite que os administradores usem uma única porta 10GbE para vários tipos de tráfego. Como alternativa, se os administradores optarem por executar apenas um tipo de tráfego, eles poderão fazê-lo sem implantar hardware e software personalizados, o que ajuda a simplificar a implantação e o gerenciamento de TI. Além disso, as C-NICs NetXtreme disponíveis nos servidores Dell PowerEdge podem oferecer suporte a Microsoft® Windows®, Linux®, VMware® e outras plataformas, ajudando a maximizar a flexibilidade da implantação.

Para ajudar a reduzir a carga do processamento

do cabeçalho do pacote no servidor host, as C-NICs NetXtreme II oferecem a funcionalidade do adaptador de barramento do host (HBA) iSCSI com a tecnologia iSCSI offload engine (iSOE) (consulte a Figura 1). Descarregando o processamento de cabeçalho do iSCSI dos processadores host para HBAs, esses controladores podem ajudar a otimizar a utilização do processador do servidor e, ao mesmo tempo, aumentar o desempenho e a taxa de transferência do storage voltado para arquivos, storage voltado para blocos, backups, transações de bancos de dados e aplicativos distribuídos totalmente integrados, como cargas de trabalho de computação de alto desempenho. A tecnologia iSOE foi desenvolvida para liberar núcleos de processadores host e recursos de memória, além de aumentar a E/S por segundo (IOPS), inclusive, no ambiente de teste Broadcom detalhado na próxima seção deste artigo, a ativação de até 400.000 IOPS em taxas de linha de 10GbE em uma única porta Ethernet, reduzindo a utilização do processador significativamente.

O iSOE permite que as C-NICs NetXtreme II lidem, de forma eficaz, com o processamento de TCP/IP e iSCSI. Descarregando as pilhas de TCP/IP e iSCSI, o controlador não precisa competir com o software da camada superior, como e-mail ou aplicativos da Web para ciclos de CPU: o desempenho iSCSI não é afetado pela carga de trabalho do aplicativo. A convergência do storage de blocos e do processamento de rede em uma infraestrutura TCP padrão ajuda a eliminar a necessidade de um adaptador de armazenamento diferente e de um cabeamento adicional, enquanto fornece desempenho e confiabilidade comparáveis ao Fibre Channel por um custo significativamente reduzido.

Modelo de controlador Broadcom	Portas	Número de peça do Broadcom	Número de peça da Dell	Tipo de cartão	Barramento de E/S	Interface física	Recursos de iSCSI
BCM5709	Dois GbE	BCM95709A0907G	430-3254	PCIe padrão	PCIe 1.0	1000Base-T	Boot iSCSI
BCM5709	Dois GbE	BCM95709A0916G	430-3263	Mezanino	PCIe 1.0	1000Base-T	Boot iSCSI e HBA
BCM5709	Dois GbE	BCM95709SA0908G	430-3310	Mezanino blade	PCIe 1.0	SerDes	Boot iSCSI e HBA
BCM5709	Dois GbE	BCM95709A0907G	430-3261	PCIe padrão	PCIe 1.0	1000Base-T	Boot iSCSI e HBA
BCM5709	Quatro GbE	BCM95709A0906G	430-0800	PCIe padrão	PCIe 1.0	1000Base-T	Boot iSCSI e HBA
BCM57710	Um 10GbE	BCM957710A1022G	430-2834	PCIe padrão	PCIe 1.0	10GBase-T	—
BCM57710	Dois 10GbE	BCM957710A1021G	430-2836	Mezanino	PCIe 1.0	10GBase-T	Boot iSCSI e HBA
BCM57711	Dois 10GbE	BCM957711A1123G	430-0674	Mezanino blade	PCIe 2.0	10GBase-KX4	Boot iSCSI e HBA
BCM57711	Dois 10GbE	BCM957711A1113G	430-0710	PCIe padrão	PCIe 2.0	SFP+	Boot iSCSI e HBA

Figura 1. Broadcom NetXtreme II C-NICs habilitados para iSOE disponíveis nos servidores Dell PowerEdge

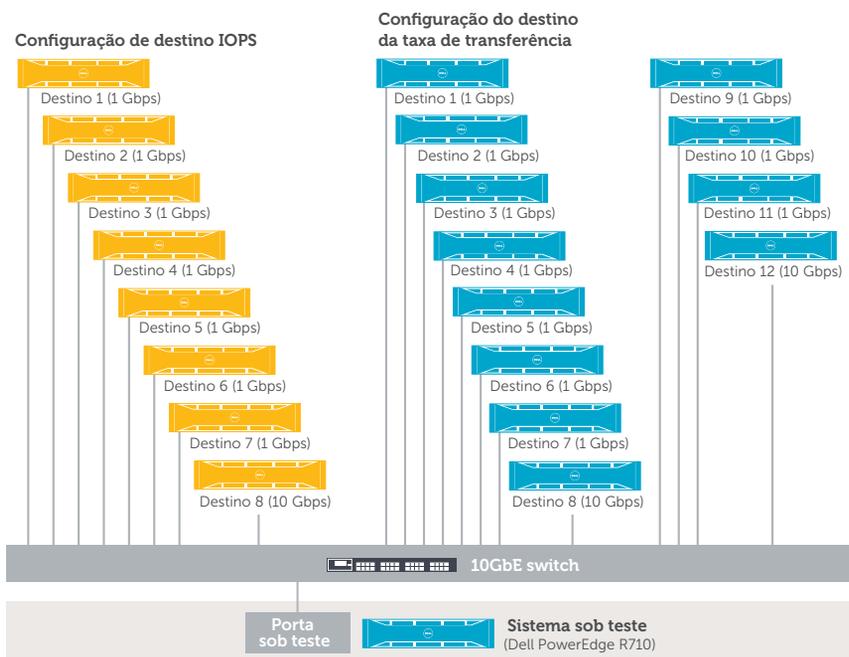


Figura 2. Configuração de alto nível usada no ambiente de teste

A facilidade do gerenciamento e a redução de custos são objetivos contínuos no data center e, para isso, a família NetXtreme II ajuda a simplificar a administração de controladores no ambiente por meio do aplicativo de gerenciamento Broadcom Advanced Control Suite (BACS) 3, que fornece uma plataforma única para o gerenciamento de rede e de E/S HBA iSCSI. O BACS 3 inclui relatórios de status rápidos de adaptadores e controladores LAN em um sistema, testes de rede para confirmar a conectividade a uma estação remota e estatísticas detalhadas de desempenho dos adaptadores e controladores.

Avaliando a abordagem convergente

Para ajudar os departamentos de TI a compreenderem os aprimoramentos e os benefícios da abordagem convergente, em outubro de 2009 os laboratórios de desempenho da Broadcom executaram uma série de testes para analisar o consumo de energia, a utilização do processador, a taxa de transferência e a eficácia do processador (IOPS por ciclo de CPU) de um C-NIC Broadcom NetXtreme II e um NIC de terceiros com um iniciador de software iSCSI.

A figura 2 mostra o ambiente de teste, que foi projetado para avaliar o valor máximo do desempenho que poderia ser esperado pelas organizações neste tipo de implantação. O sistema testado foi um servidor Dell PowerEdge R710 com dois processadores Intel® Xeon® X5570 de quatro núcleos a 2,93 GHz, com 12 GB

de RAM, e o SO Microsoft Windows Server® 2008 Enterprise Edition. A configuração incluía um C-NIC 10GbE NetXtreme II BCM57711 de porta dupla com a funcionalidade HBA iSCSI ativada, juntamente com um NIC de terceiros com uso de um iniciador de software iSCSI. Os testes usaram 8 destinos para a medição de IOPS e 12 destinos para a medição de taxa de transferência; os sistemas de destino eram servidores PowerEdge R710 com dois processadores Intel Xeon X5570 de quatro núcleos a 2,93 GHz, 2 GB de RAM, o SO Red Hat® Enterprise Linux 5.2 e o software iSCSI Enterprise Target (IET) 0.4.16.

Os testes utilizaram a ferramenta Iometer de medição e caracterização do subsistema de E/S, que foi projetada para avaliar o desempenho dos aplicativos de armazenamento e de rede, executar testes de estresse em dispositivos de armazenamento e de rede e prever o desempenho dos aplicativos de armazenamento e de rede. O ambiente de teste usou a versão 2006.07.27 dessa ferramenta, configurada para 128 E/Ss excepcionais. O consumo de energia foi medido como a média de três amostras de um segundo tiradas na entrada de energia.

Como mostrado nesses testes, o uso dos C-NICs NetXtreme II como HBAs iSCSI pode fornecer várias vantagens nos ambientes de data center, inclusive ajuda para reduzir o consumo de energia e aumentar a eficiência do processamento. No ambiente de teste, o C-NIC NetXtreme II no modo HBA iSCSI proporcionou economias no nível



Otimização do armazenamento

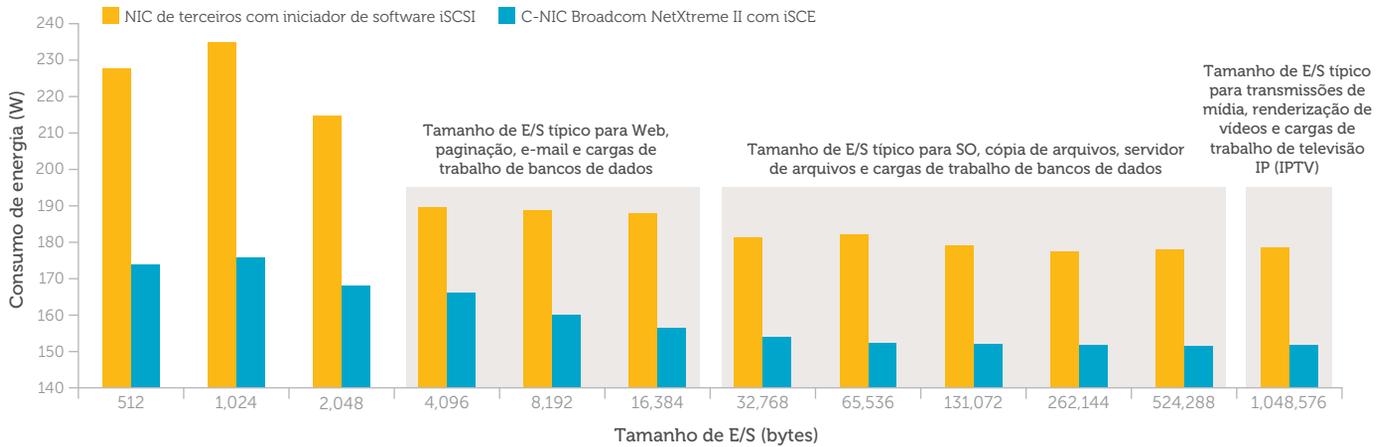


Figura 3. Consumo de energia em tamanhos de E/S diferentes para o NIC de terceiros e o C-NIC Broadcom NetXtreme II

do sistema de aproximadamente 60 W por porta em comparação com o NIC de terceiros (consulte a Figura 3) — uma diferença que pode levar a economias de custo significativas. Por exemplo, considere um data center com 2.500 servidores, cada um com quatro portas 10GbE. A apropriação das mesmas características de carga no mesmo período do ambiente de teste, usando esse C-NIC NetXtreme II no lugar do NIC

de terceiros, poderia reduzir o consumo de energia em 600 kW apenas no nível da porta; ao adicionar um exemplo do valor da Eficiência do uso de energia (PUE) de 1,8 e assumir as economias correspondentes no consumo de energia na infraestrutura de suporte, a redução total atingiria 1.080 kW. Por um custo de energia de US\$ 0,088/kWh, depois disso, as economias de custo ao longo de um ano somariam US\$

832,550. Os departamentos de TI também poderiam aproveitar essas economias enormes de energia para ajudá-los a ajustar o tamanho correto da infraestrutura de suporte do data center, ajudando a reduzir o custo total de propriedade e aprimorar a eficiência da PUE.

Além de reduzir o consumo de energia, o C-NICNetXtreme II reduziu a utilização do processador em aproximadamente 30% em

grandes tamanhos de E/S (2 a 64 KB) e forneceu três vezes o IOPS por ciclo de CPU do NIC de terceiros, possibilitando o processamento altamente eficiente das cargas de trabalho de storage (consulte a Figura 4). Essa operação eficiente, por sua vez, pode proporcionar desempenho aprimorado de IOPS para os usuários finais de cargas de trabalho de storage e, ao mesmo tempo, liberar capacidade de processamento

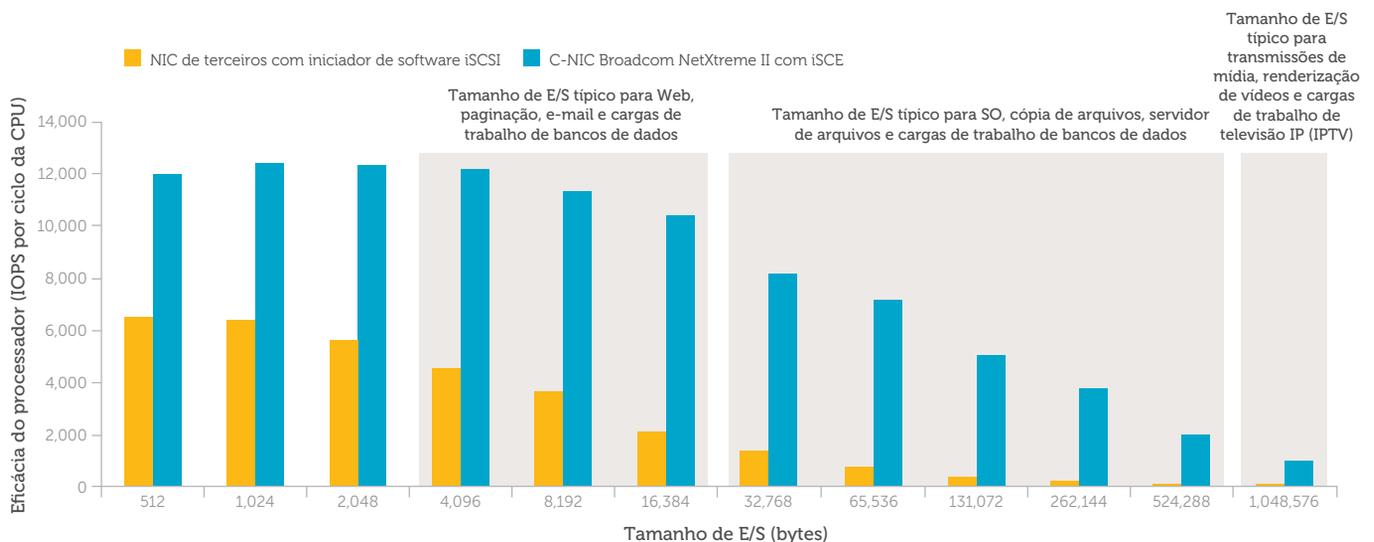


Figura 4. Eficácia do processador em tamanhos de E/S diferentes para o NIC de terceiros e o C-NIC Broadcom NetXtreme II

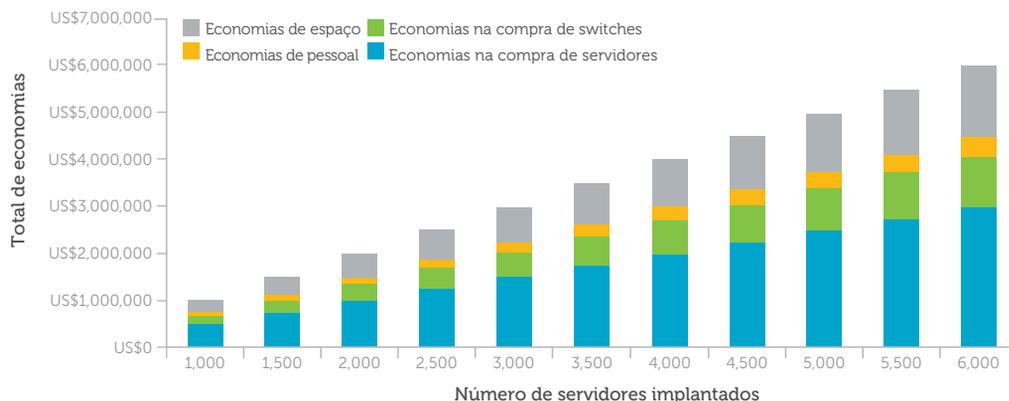


Figura 5. Exemplo de economias únicas possíveis com o planejamento adequado e a utilização eficiente do processador com os C-NICs Broadcom NetXtreme II

para oferecer suporte a aplicativos adicionais ou uso aprimorado de aplicativos de alto desempenho: tudo isso sem aumentar o custo total das operações.

Essa operação eficiente também pode ajudar os departamentos de TI durante a instalação, permitindo que eles planejem a capacidade de carga aprimorada. Em um data center com 5.000 servidores, por exemplo, as economias anteriores de planejamentos apropriados e utilização aprimorada poderiam atingir US\$ 5 milhões (consulte a Figura 5).¹

Reduzindo os custos por meio da rede convergente

A convergência de várias funções em uma fabric usando os servidores Dell PowerEdge com C-NICs Broadcom NetXtreme II pode ajudar a simplificar o funcionamento da rede, aumentar a eficiência do processador e reduzir o custo total de propriedade. Além disso, permitindo que os departamentos de TI aproveitem a infraestrutura Ethernet existente e evitem a necessidade de placas HBA autônomas, esses controladores convergentes ajudam a diminuir o custo da aquisição, implantação e gerenciamento do iSCSI. Dessa forma, não é necessário manter uma infraestrutura de armazenamento separada, além da obtenção de treinamento especializado.

A família Broadcom NetXtreme II também pode permitir que as organizações aproveitem a

rede 10GbE iSCSI e, ao mesmo tempo, desfrutem a flexibilidade da convergência de dados e do tráfego de storage em uma única rede ou o uso de uma rede dedicada para cada. Essa abordagem promete agilizar a adoção da rede iSCSI e ajudar os departamentos de TI a buscarem várias iniciativas essenciais, desde a computação ecologicamente correta até a virtualização e consolidação, além da constante necessidade de manter os custos baixos e maximizar a eficiência do data center. **PS**



Mahesh Natarajan é gerente sênior da linha de produtos dos controladores Ethernet na Broadcom.

Abhijit Aswath é gerente sênior da linha de produtos do software de controlador Ethernet na Broadcom.

Srinivas Thodati é gerente sênior de marketing de produtos dos servidores PowerEdge M-Series na Dell.

Saiba mais



Controladores Ethernet da Broadcom:
broadcom.com/products/Ethernet-Controladores



Servidores Dell PowerEdge:
dell.com/poweredge

¹ Com base em uma redução de 30% na contagem de servidores de uma média de 30% de redução na utilização de processadores. Os cálculos estimam um custo médio do servidor de US\$ 5.000, uma vida útil de servidor de 3 anos, 1 switch para cada 2,5 servidores, um custo médio de switch de US\$ 1.500, 12 servidores por rack, 10 pés quadrados de espaço no rack a um custo comercial de US\$ 1.000 por pé quadrado, 1 administrador para cada 250 servidores e um salário médio de administrador de US\$ 60.000.