

Manual do usuário
NetXtreme-E

Broadcom[®] NetXtreme-C e NetXtreme-E
MANUAL DO USUÁRIO

Histórico de revisões

<i>Revisão</i>	<i>Data</i>	<i>Descrição das alterações</i>
NetXtreme-E-UG100	26/02/2018	Versão inicial para 20.6.

© 2018 por Broadcom. Todos os direitos reservados.

Broadcom[®], o logotipo pulse, Connecting everything[®], Avago Technologies e o logotipo A são marcas comerciais da Broadcom e/ou de suas filiais nos Estados Unidos, outros países e/ou na Europa. O termo "Broadcom" se refere a Broadcom Limited e/ou as suas subsidiárias. Para obter mais informações, visite www.broadcom.com.

A Broadcom se reserva o direito de fazer alterações sem aviso prévio em quaisquer produtos ou dados aqui mencionados para aprimorar a confiabilidade, o funcionamento ou o design. As informações fornecidas pela Broadcom são consideradas precisas e confiáveis. No entanto, a Broadcom não assume responsabilidade alguma pela aplicação ou uso dessas informações ou pela aplicação ou uso de qualquer produto ou circuito aqui descrito, nem fornece licença alguma sob seus direitos de patente ou os direitos de terceiros.

Índice

Aprovações reguladoras e de segurança	7
Atendimento a normas	7
Segurança	7
Compatibilidade eletromagnética (EMC).....	7
Conformidade de descarga eletroestática (ESD).....	8
Declaração FCC.....	8
Descrição funcional	9
Link de rede e indicação de atividade	14
BCM957402AXXXX/BCM957412AXXXX	14
BCM957404AXXXX/BCM957414AXXXX	15
BCM957406AXXXX/BCM957416AXXXX	16
BCM957414M4140D.....	17
BCM957412M4120D.....	18
BCM957416M4160	19
Recursos	20
Recursos de software e hardware.....	20
Recursos de virtualização	21
VXLAN.....	22
NVGRE/GRE/IP-in-IP/Geneve	22
Descarregamentos sem estado	22
RSS.....	22
TPA	22
Divisão de cabeçalho/carga	22
Descarregamento da fragmentação de UDP	23
Descarregamento do túnel de transporte sem estado	23
Suporte Multiqueue para OS.....	23
NDIS VMQ	23
VMWare NetQueue.....	23
KVM/Xen Multiqueue	23
Matriz de suporte à configuração SR-IOV.....	24
SR-IOV	24
Particionamento da rede (NPAR).....	24
RDMA over Converge Ethernet – RoCE	24
Combinações suportadas.....	25
NPAR, SR-IOV e RoCE	25
NPAR, SR-IOV e DPDK.....	25
Combinações não suportadas	25

Instalação do hardware	26
Precauções de segurança.....	26
Requisitos do sistema	26
Requisitos de hardware	26
Lista de verificação de pré-instalação.....	26
Instalação do adaptador.....	27
Conexão dos cabos de rede	27
Cabos e módulos suportados	27
Cobre	28
SFP+	28
SFP28	28
Pacotes de software e instalação	29
Sistemas operacionais suportados	29
Instalação de drivers	29
Windows.....	29
<i>Dell DUP</i>	29
<i>Instalação da GUI</i>	29
<i>Instalação silenciosa</i>	29
<i>Instalação de INF</i>	29
Linux.....	30
<i>Instalação de módulo</i>	30
<i>Comandos Ethtool do Linux</i>	30
VMware	32
Atualização de firmware	33
Pacote de atualização da Dell.....	33
Windows.....	33
Linux.....	33
Propriedades avançadas do driver de Windows e mensagens do log de erros	34
Propriedades avançadas do driver.....	34
Mensagens de log de evento	35
Agrupamentos	37
Windows.....	37
Linux.....	37
Configuração de nível de sistema	38
Menu UEFI HII.....	38
Página de configuração principal	38
Propriedades da imagem do firmware	38
Configuração de nível de dispositivo	38
Configuração de NIC.....	38

Configuração iSCSI.....	39
Gerenciamento de configuração abrangente	39
Configuração de nível de hardware	39
Menu de configuração MBA.....	39
Menu principal de inicialização iSCSI	39
Configuração de negociação automática	40
<i>Velocidade do link operacional</i>	43
<i>Velocidade do link do firmware</i>	43
<i>Protocolo de negociação automática</i>	43
<i>Configurações do driver do Windows</i>	43
<i>Configurações do driver do Linux</i>	43
<i>Configurações do driver ESXi</i>	44
Negociação automática da FEC.....	45
Treinamento do link.....	46
Detecção Automática de Mídia	47
Inicialização iSCSI	49
Sistemas operacionais suportados para inicialização iSCSI.....	49
Configuração da inicialização iSCSI	49
Configuração do destino iSCSI.....	49
Configuração dos parâmetros da inicialização iSCSI	50
Configuração do protocolo de inicialização do MBA.....	51
Configuração de inicialização iSCSI	51
<i>Configuração de inicialização iSCSI estática</i>	51
<i>Configuração de inicialização iSCSI dinâmica</i>	52
Ativação da autenticação CHAP	54
Configuração do servidor DHCP para suportar a inicialização iSCSI	54
Configurações de inicialização iSCSI do DHCP para IPv4.....	54
<i>Opção 17 de DHCP, Caminho raiz</i>	54
<i>Opção 43 de DHCP, informações específicas do fornecedor</i>	55
<i>Configuração do servidor DHCP</i>	55
Configuração de inicialização iSCSI do DHCP para IPv6.....	56
<i>Opção 16 de DHCPv6, opção de classe do fornecedor</i>	56
<i>Opção 17 de DHCPv6, informações específicas do fornecedor</i>	56
<i>Configuração do servidor DHCP</i>	56
VXLAN: Exemplos de configuração e de casos de uso	57
SR-IOV: Exemplos de configuração e de casos de uso	58
Caso de uso Linux.....	58
Caso do Windows	59
Caso do VMWare SRIOV.....	60

NPAR – Exemplo de configuração e de caso de uso	62
Recursos e requisitos.....	62
Limitações.....	62
Configuração.....	62
Observações sobre a redução do consumo de memória NIC.....	65
RoCE – Exemplos de configuração e de casos de uso	66
Configuração do Linux.....	66
Requisitos.....	66
Dependências do driver BNXT_RE.....	66
Instalação.....	67
Limitações.....	67
Problemas conhecidos.....	68
Windows.....	68
Modo do kernel.....	68
Verificação do RDMA.....	68
Modo do usuário.....	69
VMware ESX.....	70
Limitações.....	70
Requisitos do driver BNXT RoCE.....	70
Instalação.....	70
Configuração de adaptadores de rede RDMA paravirtualizados.....	70
<i>Configuração de um Virtual Center para o PVRDMA</i>	70
<i>Marcação do vmknic para PVRDMA em hosts ESX</i>	71
<i>Configuração da regra de firewall para o PVRDMA</i>	71
<i>Adição de um dispositivo PVRDMA ao VM</i>	71
<i>Configuração do VM no SO convidado do Linux</i>	72
DCBX – Data Center Bridging	73
Perfil QoS – Perfil da fila QoS padrão.....	73
Modo DCBX = Ativar (IEEE apenas).....	74
Bit de disposição do DCBX.....	74
Perguntas frequentes	78

Aprovações reguladoras e de segurança

As seções a seguir detalham a conformidade ao padrão de Compatibilidade Regulamentar, de Segurança e Eletromagnética (EMC) e de Descarga Eletroestática (ESD) para a NetXtreme-E placa de interface de rede.

Atendimento a normas

Tabela 1: Aprovações reguladoras

Item	Padrão aplicável	Aprovação/certificado
CE/União Europeia	EN 62368-1:2014	CB relatório e certificado
UL/EUA	IEC 62368-1 ed. 2	CB relatório e certificado
CSA/Canadá	CSA 22.2 No. 950	CSA relatório e certificado.
Taiwan	CNS14336 Classe B	–

Segurança

Tabela 2: Aprovações de segurança

País	Tipo/padrão de certificação	Conformidade
Internacional	CB, esquema ICES 003 – dispositivo digital UL 1977 (segurança do conector) UL 796 (segurança de fios PCB) UL 94 (inflamabilidade de peças)	Sim

Compatibilidade eletromagnética (EMC)

Tabela 3: Compatibilidade eletromagnética

Padrão/país	Tipo de certificação	Conformidade
CE/UE	EN 55032:2012/AC:2013 Classe B EN 55024:2010 EN 61000-3-2:2014 EN 61000-3-3:2013	Relatório CE e CE DoC
FCC/EUA	CFR47, Parte 15 Classe B	FCC/IC DoC e padrões FCC e IC de referência de relatório de EMC
IC/Canadá	ICES-003 Classe B	FCC/IC DoC e padrões FCC e IC de referência de relatório
ACA/Austrália, Nova Zelândia	AS/NZS CISPR 22:2009 +A1:2010/ AS/NZS CISPR 32:2015	ACA, certificado RCM, marca
BSMI/Taiwan	CNS13438 Classe B	BSMI, certificado
BSMI/Taiwan	CNS15663	BSMI, certificado
MIC/S. Coreia	KN32 Classe B KN35	Certificado da Coreia MSIP, marca
VCCI/Japão	V-3/2014/04 (em vigor desde 31/03/2015)	Cópia do certificado VCCI on-line

Conformidade de descarga eletroestática (ESD)

Tabela 4: Resumo da conformidade ESD

Padrão	Tipo de certificação	Conformidade
EN55024:2010 (EN 61000-4-2)	Descarga aérea/direta	Sim

Declaração FCC

Este equipamento foi testado e considerado em conformidade com os limites para um dispositivo digital da Classe B, de acordo com a Parte 15 das Regras da FCC. Esses limites foram projetados para proporcionar proteção razoável contra interferência prejudicial em uma instalação residencial. O equipamento gera, utiliza e pode irradiar energia de frequência de rádio. Se não for instalado e utilizado em conformidade com as instruções, o equipamento poderá provocar interferência prejudicial nas comunicações de rádio. Entretanto, não há garantias de que a interferência não ocorrerá em uma instalação específica. Se o equipamento provocar interferência prejudicial na recepção de rádio e televisão (o que pode ser determinado ao se ligar e desligar o equipamento), o usuário será encorajado a tentar corrigir a interferência executando uma ou mais das ações a seguir:

- Reorientar ou recolocar a antena receptora.
- Aumentar a separação entre o equipamento e o receptor.
- Consultar o revendedor ou um técnico experiente de rádio/TV para obter assistência.



Nota: Alterações ou modificações não expressamente aprovadas pelo responsável do fabricante relacionadas à conformidade podem cancelar a autoridade do usuário para operar o equipamento.

Descrição funcional

A Dell suporta 10GBase-T, 10G SFP+ e 25G SFP28 como placas de interface de rede (NICs). Essas NICs são descritas na [Tabela 5](#).

Tabela 5: Descrição funcional

Placa de interface de rede	Descrição
BCM957402A4020DLPC/BCM957402A4020DC/BCM957412A4120D/BCM957412M4120D	
Velocidade	Porta Ethernet dupla de 10 Gbps
PCI-E	Gen 3 x8 ^a
Interface	SFP+ para 10 Gbps
Dispositivo	Controlador Broadcom BCM57402/BCM57412 de 10 Gbps MAC com transceptor SFI de canal duplo integrado de 10 Gbps.
Nome NDIS	Adaptador de porta dupla Broadcom NetXtreme série E 10Gb SFP+ Ethernet PCIe
Nome UEFI	Ethernet Broadcom Dual 10Gb SFP+
BCM57404A4041DLPC/BCM57404A4041DC/BCM957414A4141D/BCM957414M4140D	
Velocidade	Porta Ethernet dupla de 25 Gbps ou de 10 Gbps
PCI-E	Gen 3 x8 ^a
Interface	SFP28 para 25 Gbps e SFP+ para 10 Gbps
Dispositivo	Controlador Broadcom BCM57404/BCM57414 de 25 Gbps MAC com transceptor de canal duplo integrado de 25 Gbps SFI.
Nome NDIS	Adaptador de porta dupla Broadcom NetXtreme série E 25 Gb SFP28 Ethernet PCIe
Nome UEFI	Ethernet dupla Broadcom de 25 Gb SFP 28
BCM957406A4060DLPC/BCM957406A4060DC/BCM957416A4160D/BCM957416M4160	
Velocidade	Porta Ethernet 10GBase-T dupla
PCI-E	Gen 3 x8 ^a
Interface	RJ45 para 10 Gbps e 1 Gbps
Dispositivo	Controlador Broadcom BCM57406/BCM57416 de 10 Gbps MAC com transceptor de canal duplo integrado de 10GBase-T.
Nome NDIS	Adaptador de porta dupla Broadcom NetXtreme série E 10GBASE-T Ethernet PCIe
Nome UEFI	Ethernet Broadcom Dual 10GBASE-T

- a. A NIC suporta PCI-E de velocidade de Gen 3, Gen 2 e Gen 1, entretanto, uma PCI Gen 3 é recomendada para alcançar a produtividade nominal quando 2 portas de links 25G transmitem e recebem tráfego ao mesmo tempo.

Figura 1: Placa de interface de rede BCM957402A4020DC, BCM957412A4120D

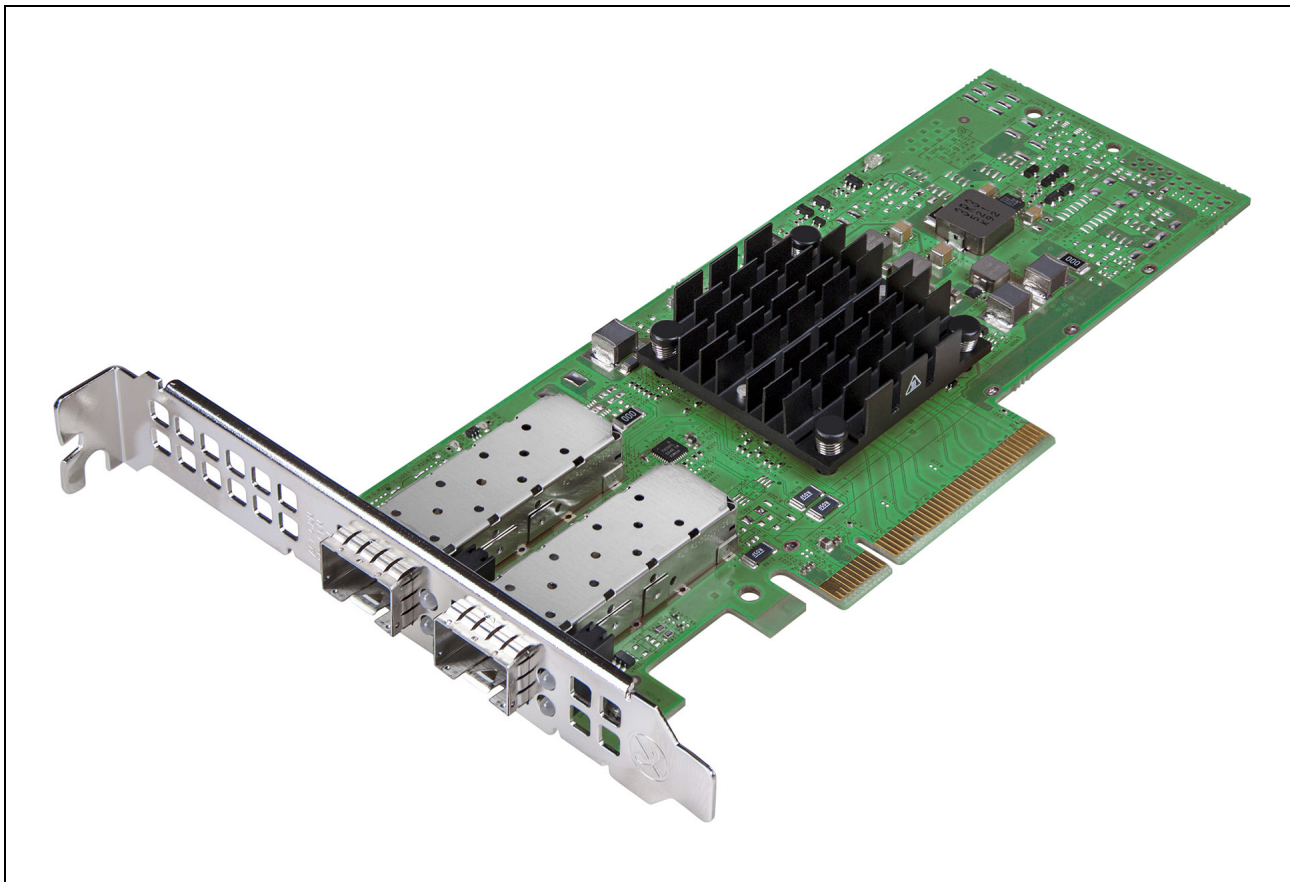


Figura 2: Placa de interface de rede BCM957404A4041DLPC, BCM957414A4141D

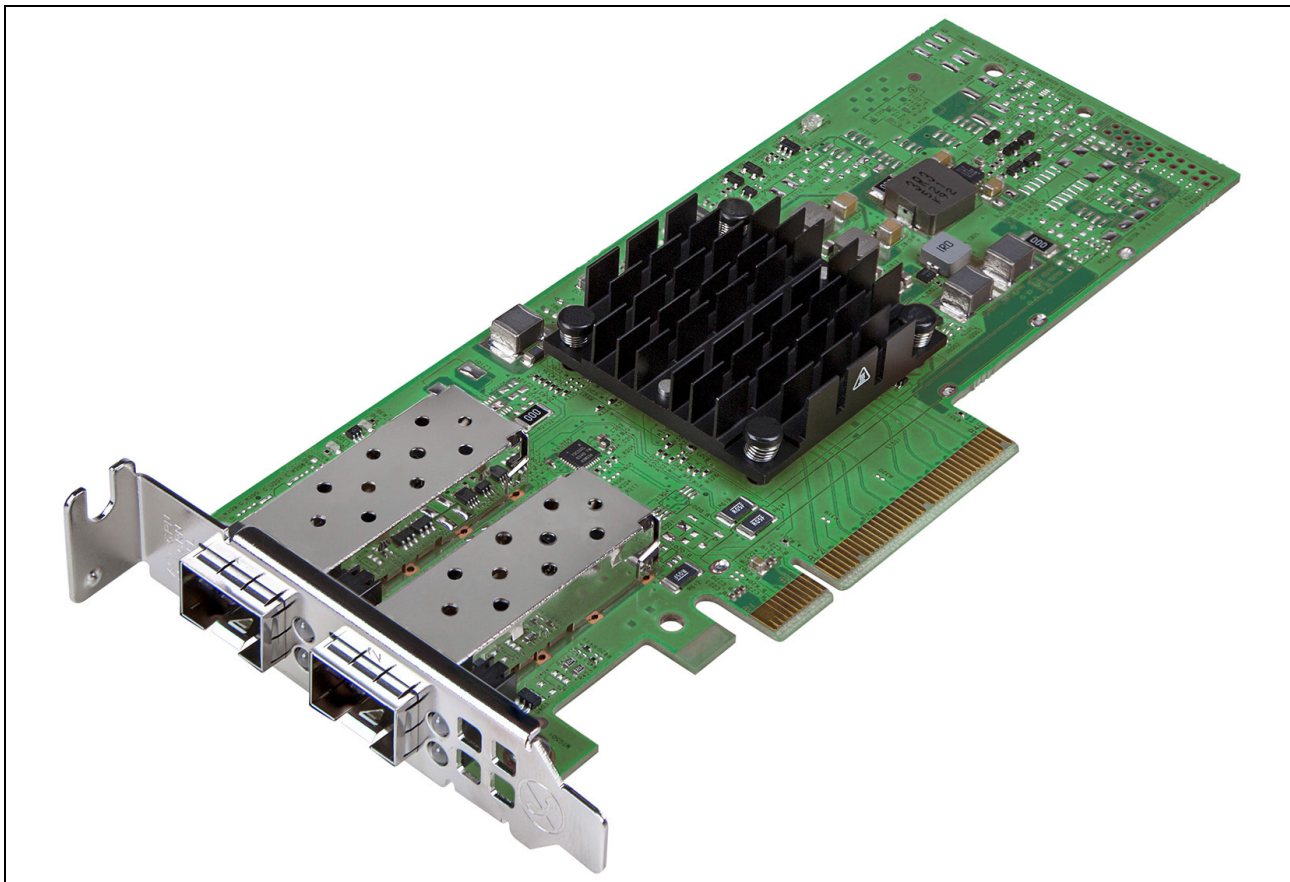


Figura 3: Placa de interface de rede BCM957406A4060DLPC, BCM957416A4160D

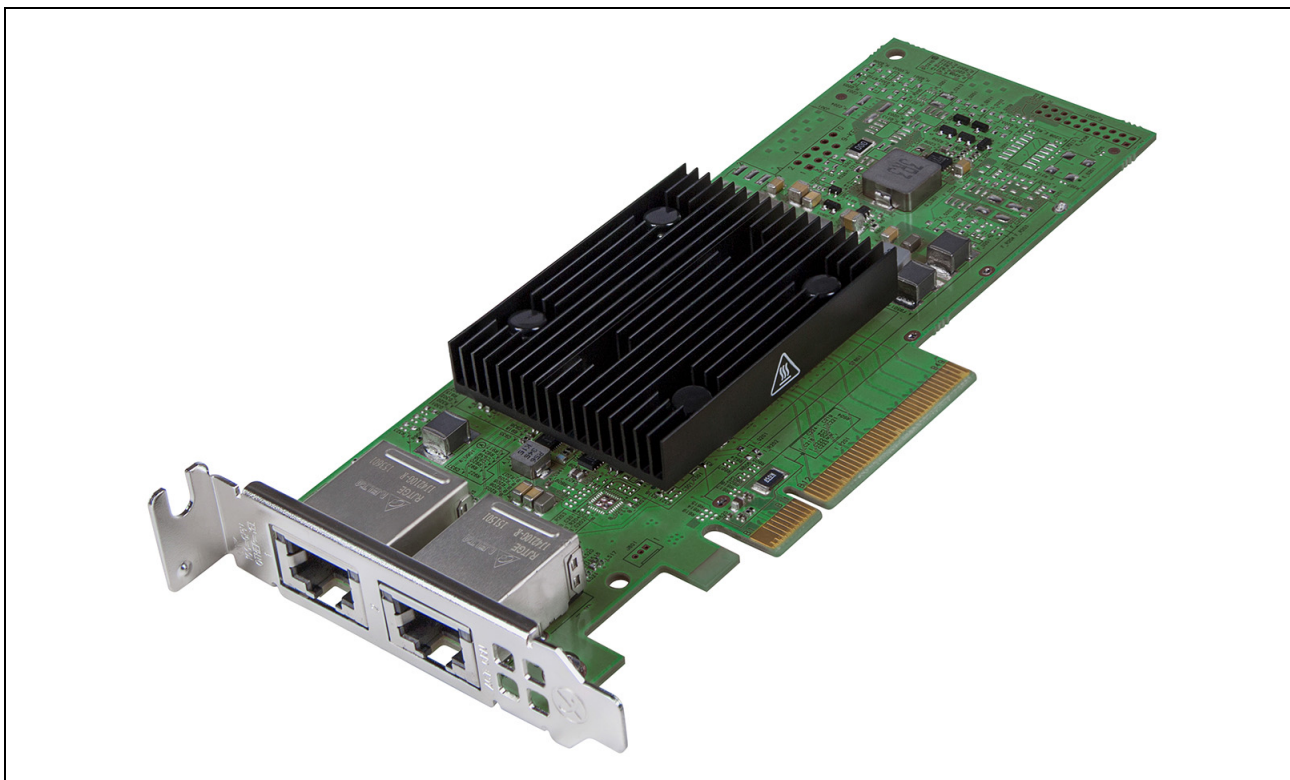


Figura 4: Placa auxiliar de rede (rNDC) BCM957414M4140D

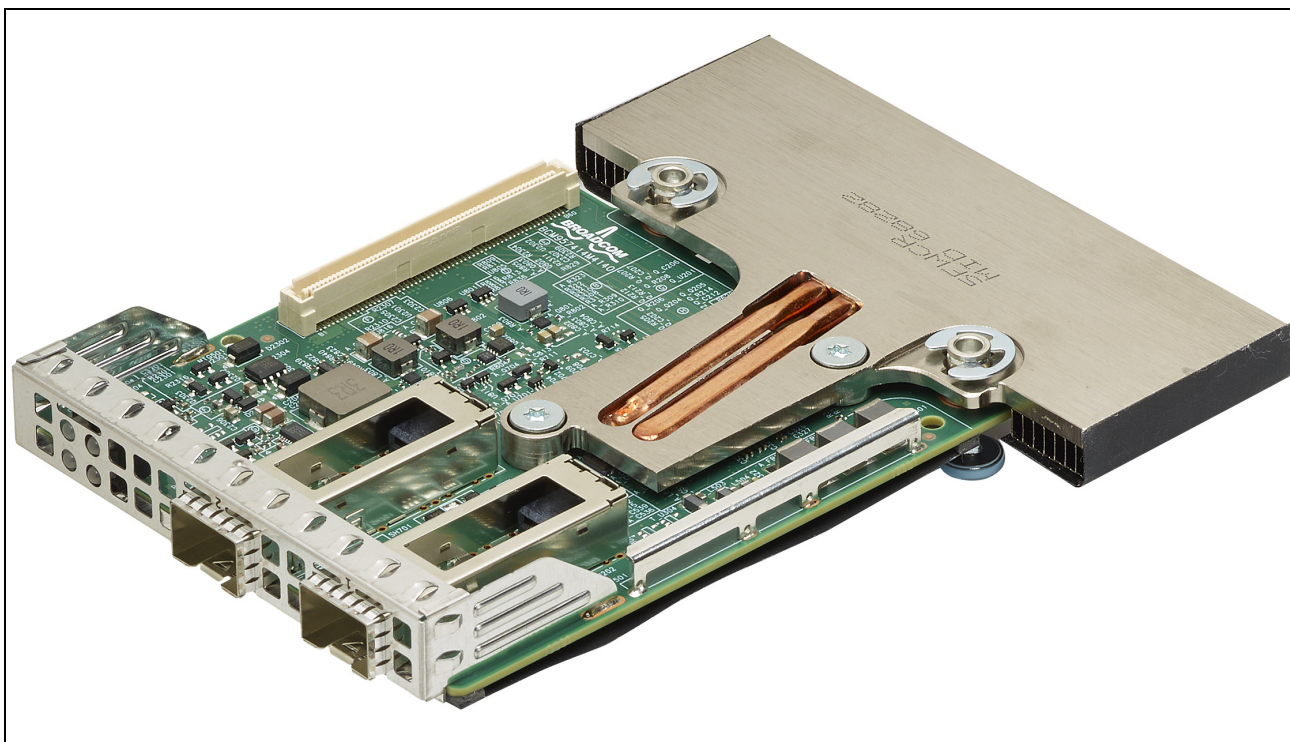


Figura 5: Placa auxiliar de rede (rNDC) BCM957412M4120D

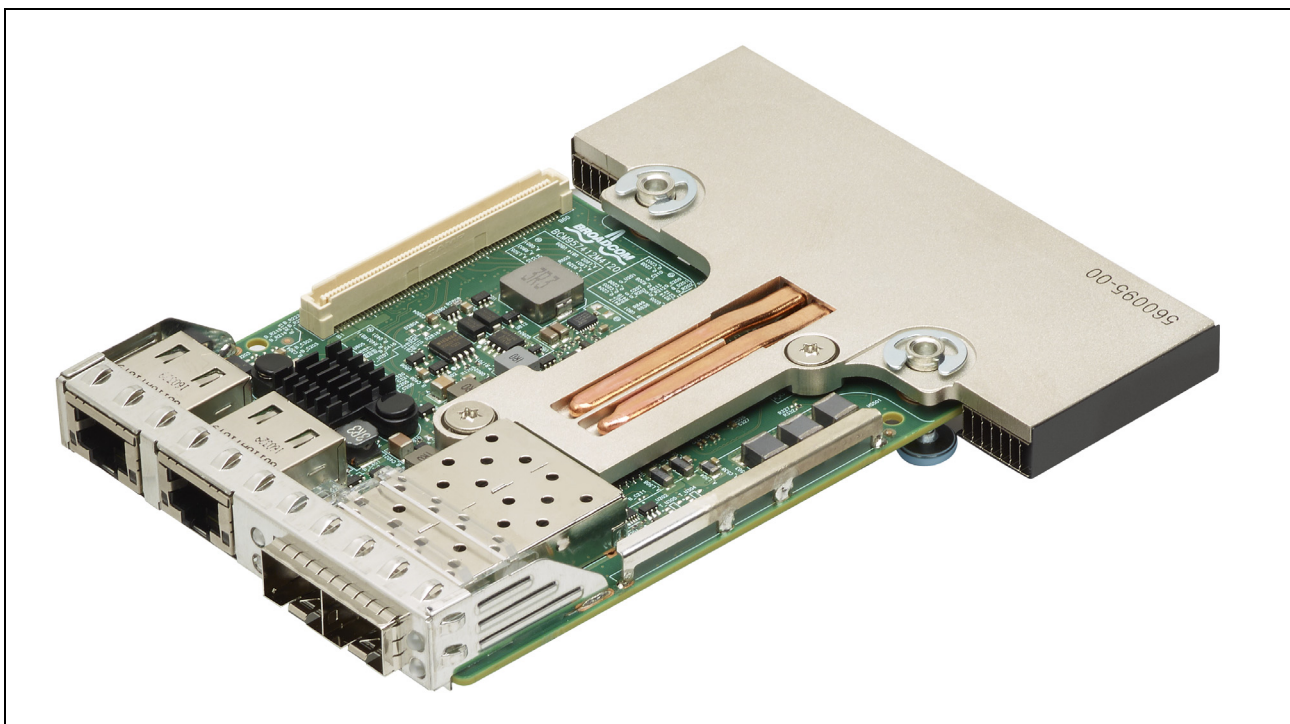
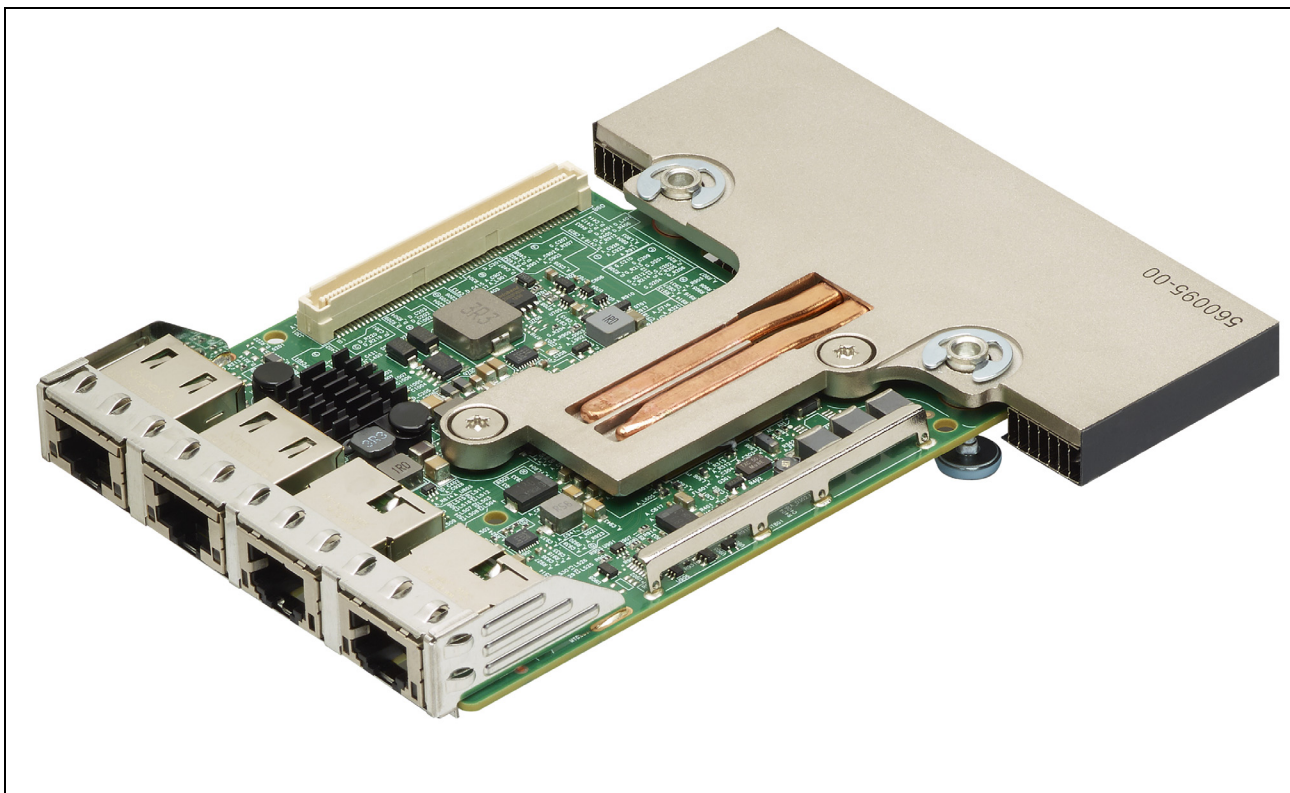


Figura 6: Placa auxiliar de rede (rNDC) BCM957416M4160



Link de rede e indicação de atividade

BCM957402AXXX/BCM957412AXXX

A porta SFP+ tem dois LEDs para indicar atividades de tráfego e velocidade de link. Os LEDs são visíveis pelo corte no suporte, como mostrado na [Figura 7](#). A funcionalidade do LED é descrita na [Tabela 6](#).

Figura 7: BCM957402AXXX/BCM957412AXXX Locais do LED e do link de atividade

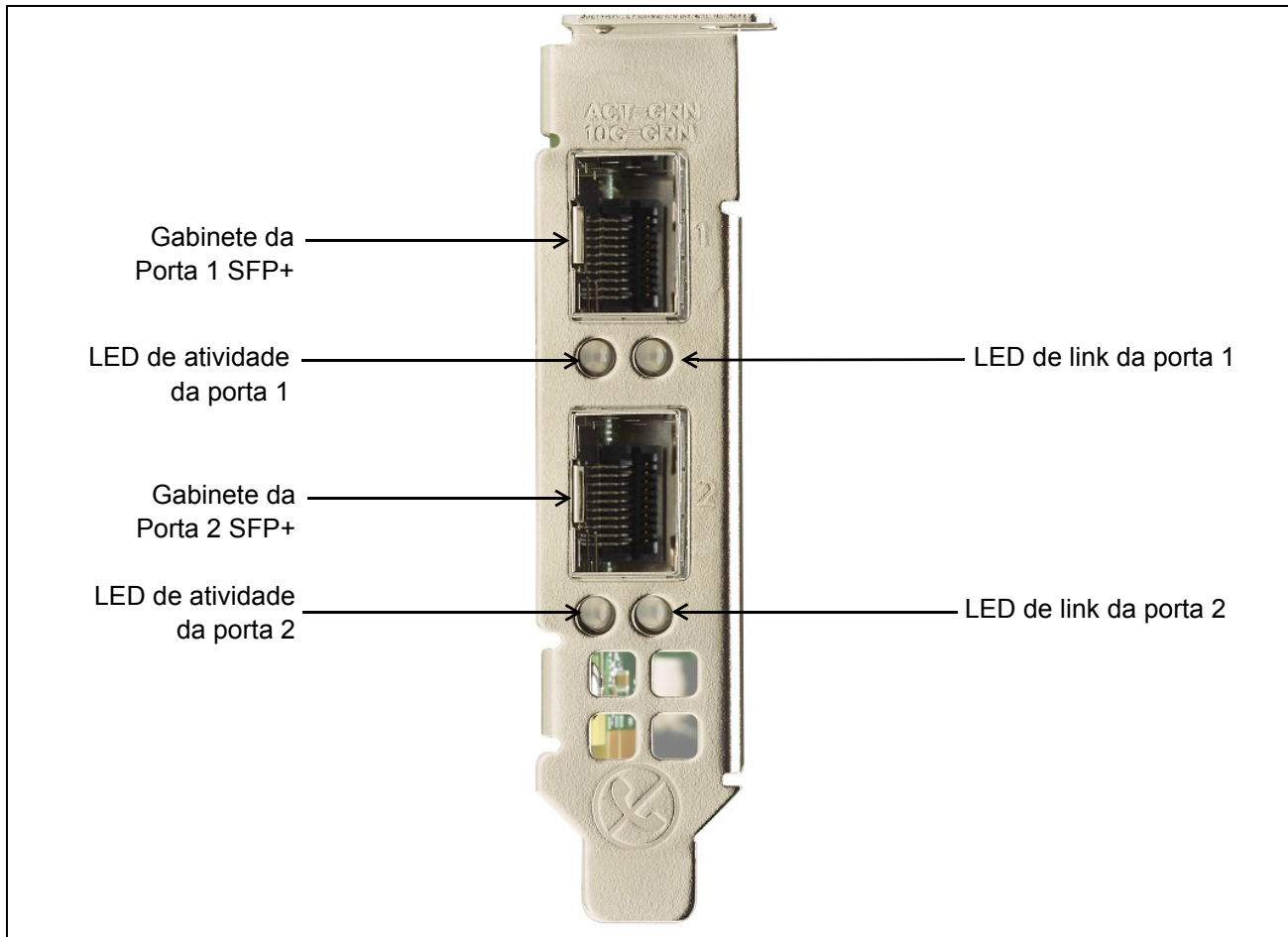


Tabela 6: BCM957402AXXX/BCM957412AXXX Locais do LED e do link de atividade

Tipo de LED	Cor/comportamento	Nota
Atividade	OFF	Sem atividade
	Verde piscando	Atividade de fluxo de tráfego
Link	OFF	Sem link
	Verde	Link 10 Gbps

BCM957404AXXXX/BCM957414AXXXX

A porta SFP28 tem dois LEDs para indicar atividades de tráfego e velocidade de link. Os LEDs são visíveis pelo corte no suporte, como mostrado na [Figura 8](#). A funcionalidade do LED é descrita na [Tabela 7](#).

Figura 8: BCM957404AXXXX/BCM957414AXXXX Locais do LED e do link de atividade

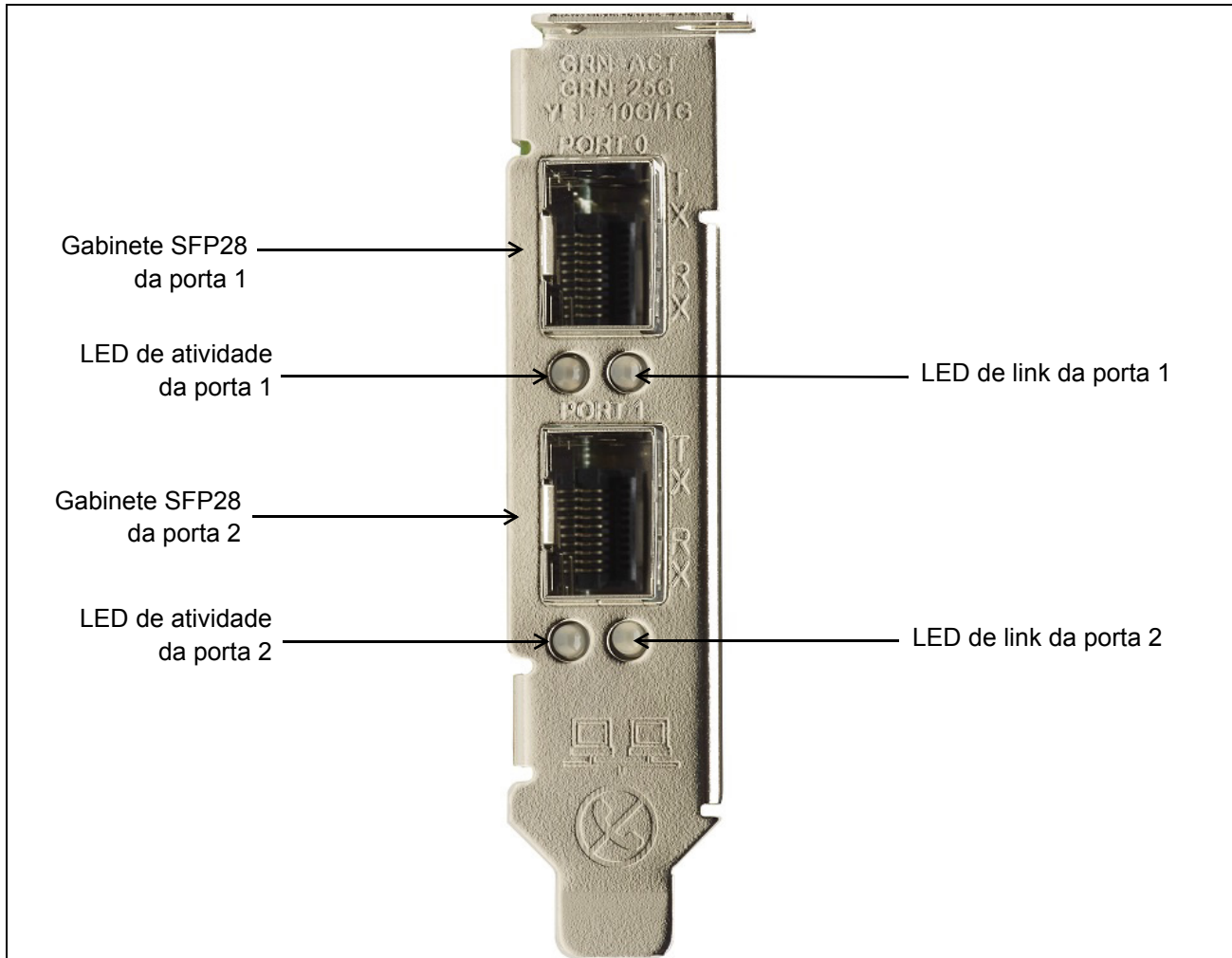


Tabela 7: BCM957404AXXXX/BCM957414AXXXX Locais do LED e do link de atividade

Tipo de LED	Cor/comportamento	Nota
Atividade	OFF	Sem atividade
	Verde piscando	Atividade de fluxo de tráfego
Link	OFF	Sem link
	Verde	Link 25 Gbps
	Amarelo	Link 10 Gbps

BCM957406AXXXX/BCM957416AXXXX

A porta RJ-45 tem dois LEDs para indicar atividades de tráfego e velocidade de link. Os LEDs são visíveis pelo corte no suporte, como mostrado na [Figura 9](#). A funcionalidade do LED é descrita na [Tabela 8](#).

Figura 9: BCM957406AXXXX/BCM957416AXXXX Locais do LED e do link de atividade

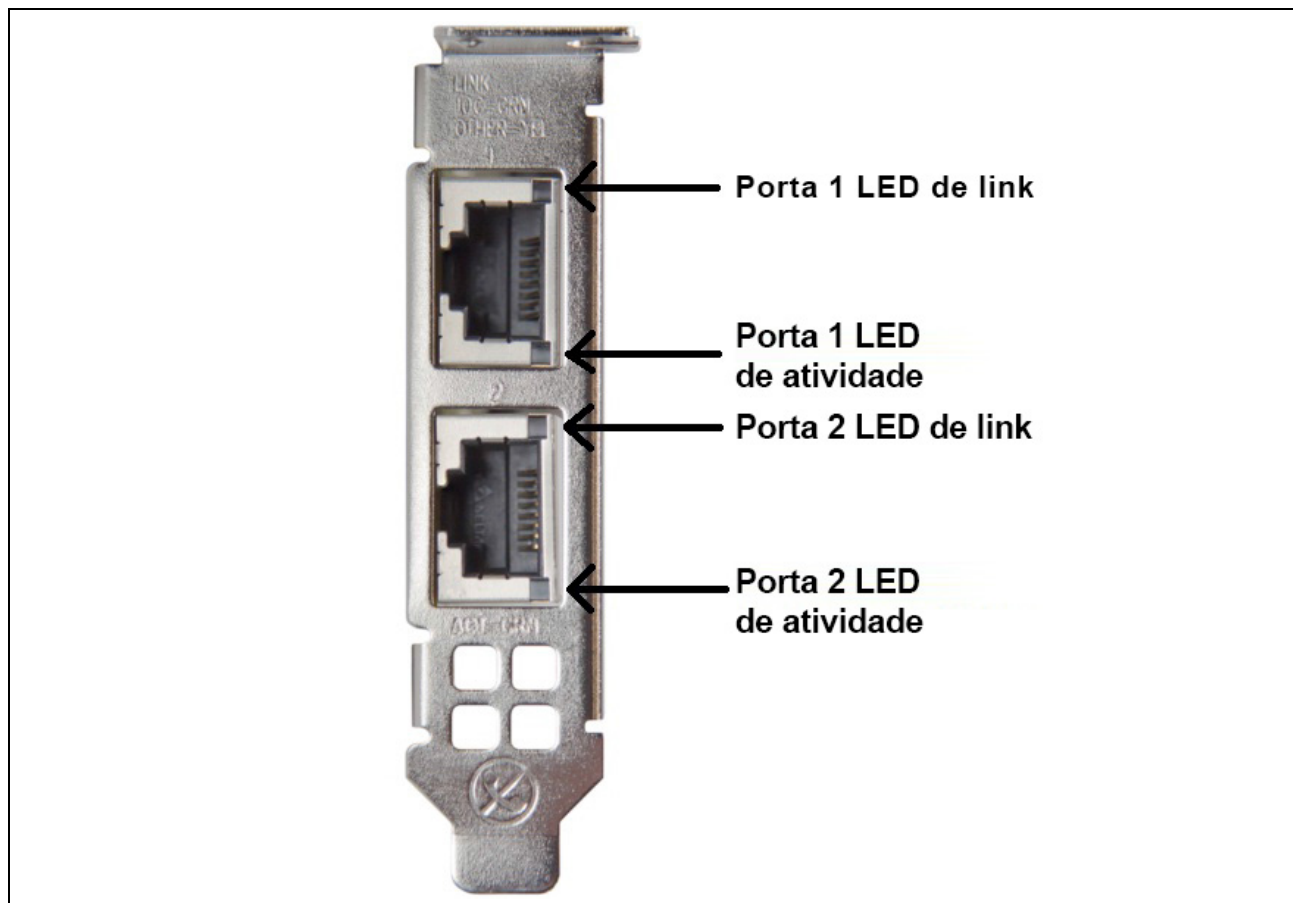


Tabela 8: BCM957406AXXXX/BCM957416AXXXX Locais do LED e do link de atividade

Tipo de LED	Cor/comportamento	Observações
Atividade	OFF	Sem atividade
	Verde piscando	Atividade de fluxo de tráfego
Link	OFF	Sem link
	Verde	Link 10 Gbps
	Âmbar	Link 1 Gbps

BCM957414M4140D

A porta SFP28 tem dois LEDs para indicar atividades de tráfego e velocidade de link. Os LEDs são visíveis pelo corte no suporte, como mostrado na [Figura 10](#).

Figura 10: Placa auxiliar de rede (rNDC) BCM957414M4140D, locais do LED de atividade e link

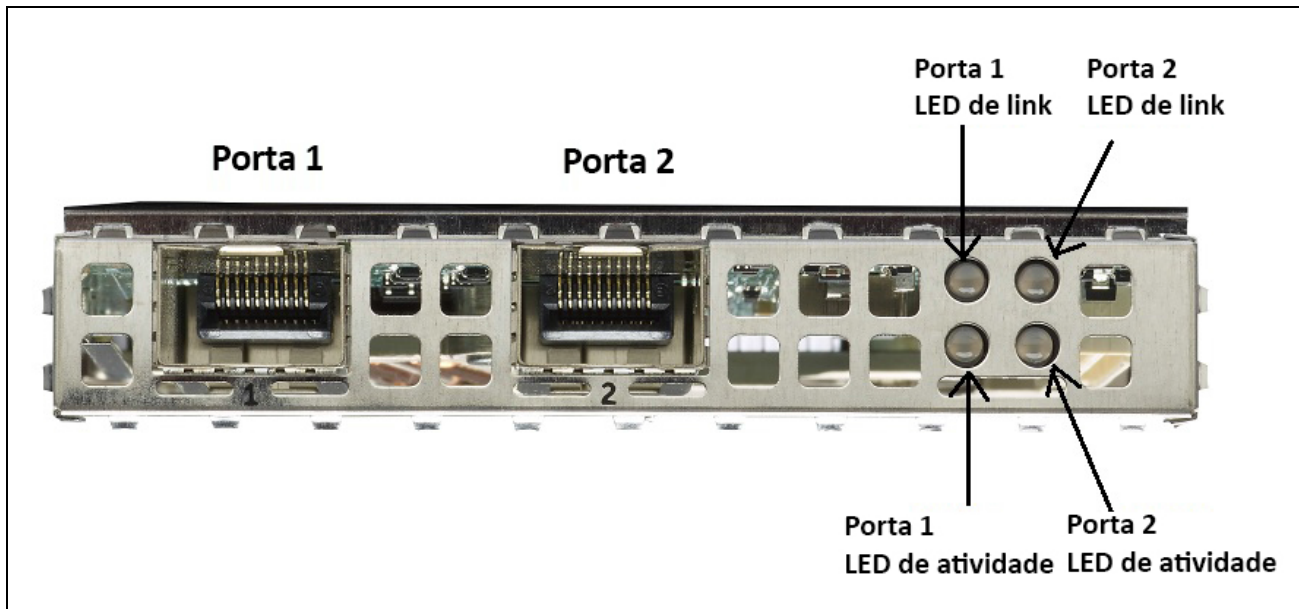


Tabela 9: Placa auxiliar de rede (rNDC) BCM957414M4140D, locais do LED de atividade e link

Tipo de LED	Cor/comportamento	Observações
Atividade	OFF	Sem atividade
	Verde piscando	Atividade de fluxo de tráfego
Link	OFF	Sem link
	Verde	Link 25 Gbps
	Amarelo	Link 10 Gbps

BCM957412M4120D

Este rNDC tem portas SFP+ e RJ-45, cada uma com dois LEDs para indicar as atividades de tráfego e a velocidade do link. Os LEDs são visíveis conforme mostrado na [Figura 11](#).

Figura 11: Placa auxiliar de rede (rNDC) BCM957412M4120D, locais do LED de atividade e link

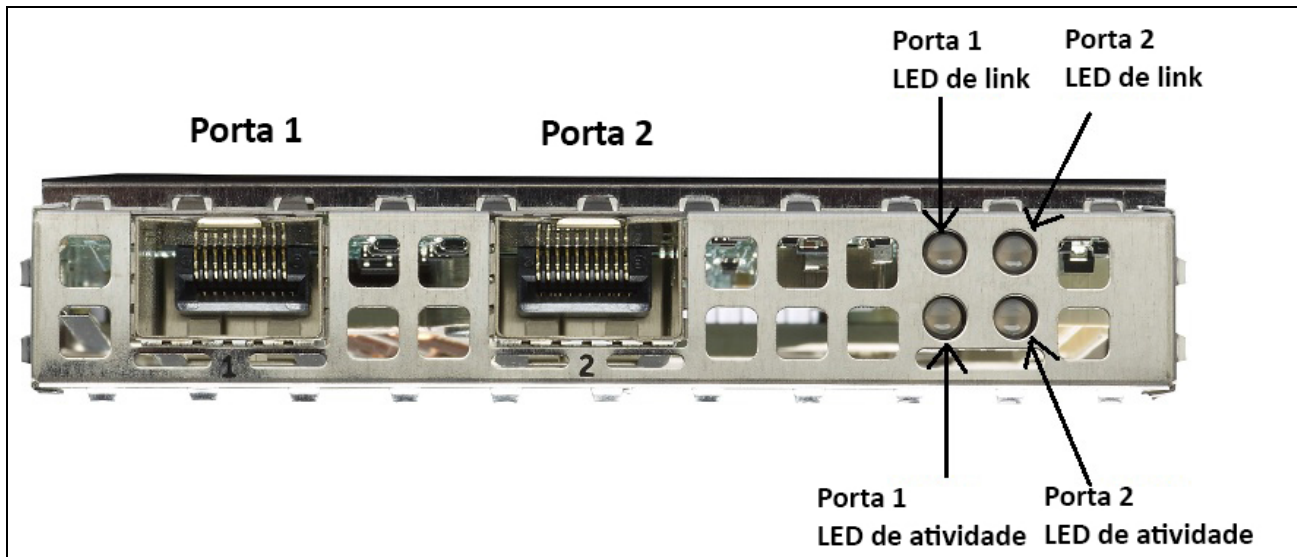


Tabela 10: Placa auxiliar de rede (rNDC) BCM957412M4120D, locais do LED de atividade e link, Porta SFP+ 1 e 2

Tipo de LED	Cor/comportamento	Observações
Atividade	OFF	Sem atividade
	Verde piscando	Atividade de fluxo de tráfego
Link	OFF	Sem link
	Verde	Link 10 Gbps

Tabela 11: Porta 1000BaseT 3 e 4

Tipo de LED	Cor/comportamento	Observações
Atividade	OFF	Sem atividade
	Verde piscando	Atividade de fluxo de tráfego
Link	OFF	Sem link
	Verde	Link 1 Gbps
	Âmbar	Link a 10/100 Mbps

BCM957416M4160

Este rNDC tem portas RJ-45 10GBaseT e 1000BaseT, cada uma com dois LEDs para indicar as atividades de tráfego e a velocidade do link. Os LEDs são visíveis conforme mostrado na [Figura 12](#).

Figura 12: Placa auxiliar de rede (rNDC) BCM957416M4160, locais do LED de atividade e link

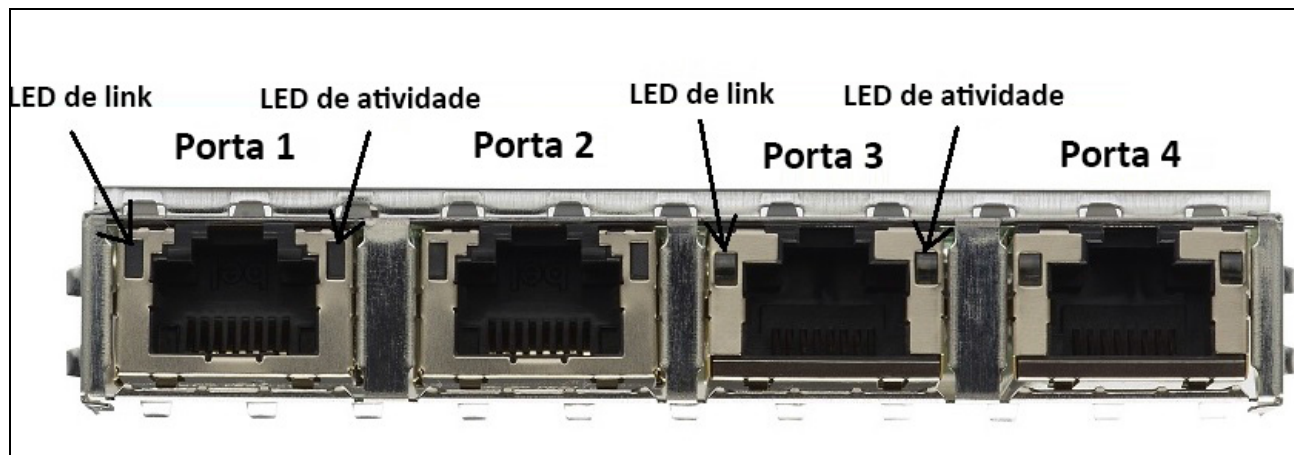


Tabela 12: Placa auxiliar de rede (rNDC) BCM957416M4160, locais do LED de atividade e link, Porta 10GBaseT 1 e 2

Tipo de LED	Cor/comportamento	Observações
Atividade	OFF	Sem atividade
	Verde piscando	Atividade de fluxo de tráfego
Link	OFF	Sem link
	Verde	Link 10 Gbps
	Âmbar	Link 1 Gbps

Recursos

Consulte as seguintes seções para conhecer os recursos do dispositivo.

Recursos de software e hardware

Tabela 13 fornece uma lista de recursos de interface do host.

Tabela 13: Recursos de interface do host

Recurso	Dados
Interface do host	PCIe v3.0 (Gen 3: 8 GT/s; Gen 2: 5 GT/s; Gen 1: 2.5 GT/s).
Número de vias PCIe	Conector PCI-E Edge: x8.
Dados de produtos vitais (VPD)	Suportado.
ID de roteamento alternativo (ARI)	Suportado.
Redefinição do nível de função (FLR)	Suportado.
Relatório de erro avançado	Suportado.
ECNs PCIe	Suporte a Dicas de processamento TLP (TPH), relatórios de tolerância à latência (LTR) e esvaziamento/enchimento de buffer otimizado (OBFF).
Vetor de intervalos MSI-X por fila	1 por fila RSS, 1 por NetQueue, 1 por Fila de máquina virtual (VMQ).
Descarregamento da soma de verificação de IP	Suporte para o lado de transmissão e de recepção.
Descarregamento da soma de verificação de TCP	Suporte para o lado de transmissão e de recepção.
Descarregamento da soma de verificação de UDP	Suporte para o lado de transmissão e de recepção.
Descarregamento de envio de grande porte de TCP NDIS	Suporte a LSOV1 e LSOV2.
Receive Segment Coalescing (RSC) de NDIS	Suporte a ambientes Windows.
Descarregamento de segmentação TCP (TSO)	Suporte para ambientes Linux e VMware.
Descarregamento de recepção de grande porte (LRO)	Suporte para ambientes Linux e VMware.
Descarregamento de recepção genérica (GRO)	Suporte para ambientes Linux e VMware.
Escalonamento do lado de recepção (RSS)	Suporte para ambientes Windows, Linux e VMware. Até 8 filas/porta suportadas para RSS.
Divisão de cabeçalho/carga	Permite que a pilha de TCP/IP de software receba pacotes TCP/IP com dados de cabeçote e de carga divididos em buffers separados. Suporta ambientes Windows, Linux e VMware.
Quadros Jumbo	Suportado.
Inicialização iSCSI	Suportado.
Particionamento de NIC (NPAR)	Suporta até oito funções físicas (PFs) por porta ou até 16 PFs por silício. Essa opção é configurável no NVRAM.
Protocolo RoCE (RDMA over Converge Ethernet)	O BCM5741X suporta o RoCE v1/v2 para Windows, Linux e VMware.

Tabela 13: Recursos de interface do host (Continuação)

Recurso	Dados
Data Center Bridging (DCB)	O BCM5741X suporta DCBX (especificação IEEE e CEE), PFC e AVB.
NCSI (Interface de banda lateral do controlador de rede)	Suportado.
Wake on LAN (WOL)	Suportado no rNDC com as interfaces 10GBase-T, SFP+ e SFP28.
Inicialização PXE	Suportado.
Inicialização UEFI	Suportado.
Controle de fluxo (Pausa)	Suportado.
Negociação automática	Suportado.
802.1q VLAN	Suportado.
Moderação de interrupção	Suportado.
Filtros de MAC/VLAN	Suportado.

Recursos de virtualização

Tabela 14 lista os recursos de virtualização do NetXtreme-E.

Tabela 14: Recursos de virtualização

Recurso	Dados
Linux KVM Multiqueue	Suportado.
VMware NetQueue	Suportado.
Fila de máquina virtual (VMQ) NDIS	Suportado.
LAN eXtensível virtual (VXLAN) - descarregamentos sem estado consciente (descarregamentos de soma de verificação IP/UDP/TCP)	Suportado.
Encapsulamento de roteamento genérico (GRE) - descarregamentos sem estado consciente (descarregamentos de soma de verificação IP/UDP/TCP)	Suportado.
Virtualização de rede usando Encapsulamento de roteamento genérico (NVGRE) - descarregamentos sem estado consciente	Suportado.
Descarregamentos sem estado consciente IP em IP (descarregamentos de soma de verificação IP/UDP/TCP)	Suportado
SR-IOV v1.0	128 funções virtuais (VFs) para sistemas operacionais convidados (GOS) por dispositivo. Vetor MSI-X por VF é definido para 16.
Porta do vetor MSI-X	74 por valor de porta padrão (configuração de duas portas). 16 por VF e é configurável no HII e CCM.

VXLAN

Uma rede de área local eXtensível virtual (VXLAN), definida em IETF RFC 7348, é usada para abordar a necessidade de sobrepor redes em centros de dados virtualizados acomodando múltiplas hospedagens. VXLAN é uma sobreposição de Camada 2 ou esquema de tunelamento sobre uma rede de Camada 3. Somente VMs no mesmo segmento de VXLAN podem se comunicar entre si.

NVGRE/GRE/IP-in-IP/Geneve

A virtualização de rede usando GRE (NVGRE), definida em IETF RFC 7637, é semelhante a uma VXLAN.

Descarregamentos sem estado

RSS

O Escalonamento do lado de recepção (RSS) usa um algoritmo Toeplitz com correspondência 4-tupla nos quadros recebidos, que são encaminhados para uma CPU determinística para processamento de quadros. Isso permite um processamento de quadro otimizado e equilibra a utilização da CPU. Uma tabela indireta é usada para mapear a transmissão para uma CPU.

O RSS simétrico permite o mapeamento de pacotes de um fluxo específico de TCP ou UDP para a mesma fila de recepção.

TPA

A Agregação de pacote transparente (TPA) é uma técnica onde os quadros recebidos dos mesmos quadros 4-tupla correspondentes são agregados e, em seguida, indicados para a pilha da rede. Cada entrada no contexto TPA é identificada por 4-tupla: IP de origem, IP de destino, porta TCP de origem e porta TCP de destino. O TPA melhora o desempenho do sistema reduzindo as interrupções do tráfego de rede e diminuindo a sobrecarga da CPU.

Divisão de cabeçalho/carga

A divisão de cabeçalho/carga é um recurso que permite que a pilha de TCP/IP de software receba pacotes TCP/IP com dados de cabeçote e de carga divididos em buffers separados. O suporte para esse recurso está disponível tanto para ambientes Windows como Linux. A seguir estão alguns dos benefícios possíveis da divisão de cabeçalho/carga:

- A divisão de cabeçalho/carga possibilita o armazenamento em cache compacto e eficaz de cabeçalhos de pacotes em caches da CPU host. Isso pode resultar em uma melhoria de desempenho de TCP/IP do lado de recepção.
- Dividir cabeçalhos/cargas permite operações de virar páginas e de cópia zero pela pilha TCP/IP do host. Isso pode melhorar ainda mais o desempenho do caminho de recepção.

Descarregamento da fragmentação de UDP

O descarregamento de fragmentação de UDP (UFO) é um recurso que permite que a pilha de software descarregue a fragmentação de datagramas UDP/IP em pacotes UDP/IP. O suporte a esse recurso está disponível somente em ambiente Linux. A seguir está um possível benefício do UFO:

- O UFO permite que a NIC lide com a desfragmentação de um datagrama UDP em pacotes UDP/IP. Isso pode resultar na redução da sobrecarga da CPU para o processamento de UDP/IP do lado de transmissão.

Descarregamento do túnel de transporte sem estado

O descarregamento do túnel de transporte sem estado (STT) é um encapsulamento de túnel que permite sobrepor redes em centros de dados virtualizados. O STT usa um encapsulamento baseado em IP com um cabeçalho semelhante ao TCP. Não existe nenhum estado de conexão TCP associado ao túnel e é por isso que o STT é "sem estado". O Switch virtual aberto (OVS) usa STT.

Um quadro STT contém o cabeçalho e a carga do quadro STT. A carga do quadro STT é um quadro Ethernet desmarcado. O cabeçalho do quadro STT e a carga encapsulada são tratados como a carga TCP e o cabeçalho semelhante ao TCP. O cabeçalho de IP (IPv4 ou IPv6) e o cabeçalho de Ethernet são criados para cada segmento STT transmitido.

Suporte Multiqueue para OS

NDIS VMQ

A Fila de máquina virtual (VMQ) NDIS é um recurso suportado pela Microsoft para aumentar o desempenho da rede Hyper-V. O recurso VMQ suporta classificação de pacote baseada no endereço MAC de destino para devolver pacotes recebidos em diferentes filas de conclusão. Essa classificação de pacote combinada com a capacidade de pacotes DMA diretamente na memória de uma máquina virtual permite o escalonamento de máquinas virtuais entre múltiplos processadores.

Consulte [“Propriedades avançadas do driver de Windows e mensagens do log de erros”](#) na página 34 para obter informações sobre VMQ.

VMWare NetQueue

VMware NetQueue é um recurso semelhante ao recurso NDIS VMQ da Microsoft. O recurso NetQueue suporta classificação de pacote baseada no endereço MAC de destino e em VLAN para devolver pacotes recebidos em diferentes NetQueues. Essa classificação de pacote combinada com a capacidade de pacotes DMA diretamente na memória de uma máquina virtual permite o escalonamento de máquinas virtuais entre múltiplos processadores.

KVM/Xen Multiqueue

KVM/Multiqueue devolve os quadros para diferentes filas da pilha do host pela classificação do quadro recebido pelo processamento do endereço MAC de destino ou tag 802.1Q VLAN do pacote recebido. A classificação combinada com a capacidade de executar DMA nos quadros diretamente para a memória de uma máquina virtual permite o escalonamento de máquinas virtuais entre múltiplos processadores.

Matriz de suporte à configuração SR-IOV

- Windows VF sobre hipervisor Windows
- Windows VF e Linux VF sobre hipervisor VMware
- Linux VF sobre Linux KVM

SR-IOV

A PCI-SIG define o suporte opcional para virtualização de IO de raiz única (SR-IOV). O SR-IOV é projetado para permitir o acesso da VM diretamente para o dispositivo usando Funções virtuais (VFs). A Função física (PF) da NIC é dividida em múltiplas funções virtuais e cada VF é apresentada como uma PF para as VMs.

SR-IOV usa a funcionalidade IOMMU para traduzir endereços virtuais PCI-E em endereços físicos usando uma tabela de tradução.

O número de funções físicas (PFs) e de funções virtuais (VFs) é gerenciado pelo menu UEFI HII, pelo CCM e pelas configurações de NVRAM. O SRIOV pode ser suportado em combinação com o modo NPAR.

Particionamento da rede (NPAR)

O recurso de Particionamento da rede (NPAR) permite que uma única porta de interface de rede física seja exibida para o sistema como funções múltiplas do dispositivo de rede. Quando o NPAR está ativado, o dispositivo NetXtreme-E é enumerado como múltiplas funções físicas (PF) de PCIE. Cada PF ou "partição" é atribuída a um ID de função PCIe separado na primeira inicialização. A definição do PCIe original permitia oito PFs por dispositivo. Para sistemas compatíveis com ID de roteamento alternativo (ARI), os adaptadores Broadcom NetXtreme-E suportam até 16 PFs por dispositivo. Cada partição é atribuída ao próprio espaço de configuração, endereço BAR e endereço MAC, permitindo que ela funcione independentemente. As partições suportam a atribuição direta a VMs, VLANs, etc., assim como qualquer outra interface física.



Nota: Na página **Configuração do sistema > Configurações do dispositivo > [Dispositivo Broadcom 5741x] > Configuração do nível do dispositivo**, o usuário pode ativar o **NParEP** para permitir que o adaptador NXE suporte até 16 PFs por dispositivo. Para dispositivos de duas portas, isso significa até oito PFs para cada porta.

RDMA over Converge Ethernet – RoCE

Remote Direct Memory Access (RDMA) via Converge Ethernet (RoCE) é um recurso de offload completo de hardware no BCM5741X que permite a funcionalidade RDMA por uma rede Ethernet. A funcionalidade RoCE está disponível no modo de usuário e no aplicativo do modo kernel. As funções físicas (PF) e funções virtuais SRIOV (VF) do RoCE estão disponíveis no modo de função única e no modo de multifunções (modo de Particionamento NIC). O Broadcom suporta o RoCE no Windows, Linux e VMWare.

Consulte os links a seguir para suporte de RDMA para cada sistema operacional:

Windows

[https://technet.microsoft.com/en-us/library/jj134210\(v=ws.11\).aspx](https://technet.microsoft.com/en-us/library/jj134210(v=ws.11).aspx)

Redhat Linux

https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/7/html/Networking_Guide/ch-Configure_InfiniBand_and_RDMA_Networks.html

VMware

<https://pubs.vmware.com/vsphere-65/index.jsp?topic=%2Fcom.vmware.vsphere.networking.doc%2FGUID-4A5EBD44-FB1E-4A83-BB47-BBC65181E1C2.html>

Combinações suportadas

As seções a seguir descrevem as combinações de recursos suportadas para este dispositivo.

NPAR, SR-IOV e RoCE

A [Tabela 15](#) mostra as combinações de recursos suportadas de NPAR, SR-IOV e RoCE.

Tabela 15: NPAR, SR-IOV e RoCE

Recurso de SW	Observações
NPAR	Até 8 PFs ou 16 PFs
SR-IOV	Até 128 VFs (total por chip)
RoCE em PFs	Até 4 PFs
RoCE em VFs	Válido para VFs conectados a PFs ativados para RoCE
SO host	Linux, Windows, ESXi (sem suporte para vRDMA)
SO convidado	Linux e Windows
DCB	Até dois COS por porta com memória reservada não-compartilhada

NPAR, SR-IOV e DPDK

A [Tabela 16](#) mostra as combinações de recursos suportadas de NPAR, SR-IOV e DPDK.

Tabela 16: NPAR, SR-IOV e DPDK

Recurso de SW	Observações
NPAR	Até 8 PFs ou 16 PFs
SR-IOV	Até 128 VFs (total por chip)
DPDK	Suportado somente como VF
SO host	Linux
SO convidado	DPDK (Linux)

Combinações não suportadas

A combinação de NPAR, SR-IOV, RoCE e DPDK não é suportada.

Instalação do hardware

Precauções de segurança



Cuidado! O adaptador está sendo instalado em um sistema que opera com tensões que podem ser fatais. Antes de remover a tampa do sistema, observe as precauções a seguir para sua proteção e para evitar danos aos componentes do sistema:

- Remova qualquer objeto metálico ou joia das mãos e dos pulsos.
- Utilize apenas ferramentas isoladas ou não-condutoras de eletricidade.
- Verifique se o sistema está DESLIGADO e desconectado antes de tocar qualquer um dos componentes internos.
- Instale ou remova adaptadores em um ambiente livre de energia estática. É recomendável o uso de uma pulseira antiestática devidamente aterrada, ou outros dispositivos pessoais antiestáticos, e uma esteira antiestática.

Requisitos do sistema

Antes de instalar o adaptador Broadcom NetXtreme-E Ethernet, verifique se o sistema atende aos requisitos descritos para o sistema operacional.

Requisitos de hardware

Consulte a seguinte lista de requisitos de hardware:

- Sistemas Dell 13G que atendem aos requisitos de sistema operacional.
- Sistemas Dell 13G que suportam placas NetXtreme-E Ethernet.
- Um slot PCI-E Gen 3 x 8 ou slots abertos PCIE Gen3 rNDC para um Adaptador NIC com fator de forma rNDC.
- 4 GB de memória ou mais (32 GB ou mais é recomendado para aplicativos de virtualização e desempenho de produtividade de rede nominal).

Lista de verificação de pré-instalação

Consulte a lista a seguir antes de instalar o dispositivo NetXtreme-E.

1. Verifique se o servidor atende aos requisitos de hardware e software listados em [“Requisitos do sistema”](#).
2. Verifique se o servidor está utilizando o BIOS mais recente.
3. Se o sistema estiver ativo, encerre-o.
4. Quando o encerramento do sistema estiver concluído, desligue a alimentação do sistema e desconecte o cabo de alimentação.
5. Segurando a placa do adaptador pelas bordas, remova-a da sua embalagem de envio e coloque-a sobre uma superfície antiestática.
6. Verifique se há sinais visíveis de danos, particularmente no conector da borda da placa. Nunca tente instalar um adaptador danificado.

Instalação do adaptador

As instruções a seguir aplicam-se à instalação do adaptador Broadcom NetXtreme-E Ethernet (NIC integrada) na maioria dos servidores. Consulte os manuais fornecidos com o servidor para obter detalhes sobre como executar essas tarefas nesse servidor específico.

1. Consulte “[Precauções de segurança](#)” na página 26 e “[Lista de verificação de pré-instalação](#)” antes de instalar o adaptador. Verifique se a alimentação do sistema está DESLIGADA e desconectada da tomada de energia e se os procedimentos corretos de aterramento elétrico foram seguidos.
2. Abra o sistema e selecione qualquer slot PCI Express Gen 3 x8 vazio.
3. Remova a tampa branca do slot.
4. Alinhe a borda do conector do adaptador com o slot do conector no sistema.
5. Prenda o adaptador com o prendedor do adaptador ou com o parafuso.
6. Feche o sistema e desconecte todos os dispositivos antiestáticos pessoais.



Nota: Para adaptadores NIC com fator de forma rNDC, localize um slot aberto RNDC ou remova e substitua o rNDC padrão de saída pelo NetXtreme rNDC.

Conexão dos cabos de rede

Os switches Dell Ethernet são produzidos com portas SFP+/SFP28/QSFP28 compatíveis com até 100 Gbps. Essas portas de 100 Gbps podem ser divididas entre 4 portas SFP28 de 25 Gbps. Portas QSFP podem ser conectadas a portas SFP28 usando 4 cabos breakout SFP28 de 25G.

Cabos e módulos suportados

Tabela 17: Cabos e módulos suportados

Módulo óptico	Número de peça da Dell	Adaptadores	Descrição
FTLX8571D3BCL-DL	3G84K	BCM957404A4041DLPC, BCM957404A4041DC, BCM957402A4020DLPC, BCM957402A4020DC	Transceptor multimodo SFP+ de 10 Gbps 850 nm
FCLF-8521-3	8T47V	BCM957406A4060DLPC, BCM957406A4060DC	Transceptor SFP de cobre de 1000Base-T
FTLF8536P4BCL	P7D7R	BCM957404A4041DLPC, BCM957404A4041DC	Transceptor de comprimento de onda curta SFP+ de 25 Gbps
Método DM7051	PGYJT	BCM57402X, BCM57404X, BCM57412X, BCM57414X	SFP+ para transceptor 10GBASE-T
FTLX1471D3BCL-FC	RN84N	BCM57402X, BCM57404X, BCM57412X, BCM57414X	Transceptor SFP28 de 25 Gbps
FTLX8574D3BNL	N8TDR	BCM57402X, BCM57404X, BCM57412X, BCM57414X	Faixa de temperatura estendida de 85C Transceptor SFP+ de 10 Gbps

Tabela 17: Cabos e módulos suportados (Continuação)

Módulo óptico	Número de peça da Dell	Adaptadores	Descrição
FTLF8536P4BNL-FC	HHHHC	BCM57402X, BCM57404X, BCM57412X, BCM57414X	Faixa de temperatura estendida de 85C Transceptor SFP+ de 10 Gbps
FTLX8574D3BCL-FC ou PLRXPLSCS43811	WTRD1	BCM57402X, BCM57404X, BCM57412X, BCM57414X	Transceptor SFP+ de 10 Gbps-SR

Nota:

1. Cabos de conexão direta (DAC) que seguem os padrões IEEE podem ser conectados ao adaptador.
2. O HHHHC de peça da Dell e o N8TDR são necessários para o BCM957414M4140D.

Cobre

Os adaptadores BCM957406AXXXX, BCM957416AXXXX e BCM957416XXXX possuem dois conectores RJ-45 usados para conectar o sistema a um segmento de fio de cobre Ethernet CAT 6E.

SFP+

Os adaptadores BCM957402AXXXX, BCM957412AXXXX e BCM957412MXXXX possuem dois conectores SFP+ usados para conectar o sistema a um switch Ethernet de 10 Gbps.

SFP28

Os adaptadores BCM957404AXXXX, BCM957414AXXXX e BCM957414MXXXX possuem dois conectores SFP28 usados para conectar o sistema a um switch Ethernet de 100 Gbps.

Pacotes de software e instalação

Consulte as seções a seguir para obter informações sobre os pacotes de software e a instalação.

Sistemas operacionais suportados

Tabela 18 fornece uma lista de sistemas operacionais suportados.

Tabela 18: Lista de sistemas operacionais suportados

OS Flavor	Distribuição
Windows	Windows 2012 R2 ou superior
Linux	Redhat 6.9, Redhat 7.1 ou superior SLES 11 SP 4, SLES 12 SP 2 ou superior
VMWare	ESXi 6.0 U3 ou superior

Instalação de drivers

Consulte as seguintes seções para saber mais sobre a instalação de drivers.

Windows

Dell DUP

Os drivers do controlador Broadcom NetXtreme série E podem ser instalados usando o driver DUP. O instalador é fornecido no formato executável x64.

Instalação da GUI

Quando o arquivo é executado, uma caixa de diálogo aparece, solicitando entrada do usuário. O instalador suporta somente a opção de driver.

Instalação silenciosa

O executável pode ser instalado silenciosamente usando o comando mostrado abaixo.

Por exemplo:

```
Network_Driver_<versão>.EXE /s /driveronly
```

Instalação de INF

O Dell DUP é usado para instalar os drivers de controladores Ethernet Broadcom NetXtreme-E. Use o seguinte comando para extrair os arquivos INF do driver do Dell DUP:

```
Network_Driver_<versão>.EXE /s /v"EXTRACTDRIVERS=c:\dell\drivers\network"
```

Após a extração dos arquivos, a instalação do arquivo INF é executada pela funcionalidade "atualizar driver" usando o Gerenciador de dispositivos (devmgmt.msc). Abra o Gerenciador de dispositivos, selecione a NIC desejada, clique com o botão direito e selecione a atualização de driver para atualizá-lo.

Linux

Os drivers do The são fornecidos em RPM, KMP e no formato de código fonte. Para compilar o driver do dispositivo pelo código fonte usando Linux, consulte o exemplo a seguir:

1. Como um usuário-raiz, faça login no sistema Linux.
2. scp ou cp no pacote tar do driver no sistema Linux. Um exemplo típico é:

```
cp /var/run/media/usb/bnxt_en-<versão>.tar.gz /root/
```

3. Execute o seguinte comando:

```
tar -zxvf /root/bnxt_en-<versão>.tar.gz
```

4. Execute o seguinte comando:

```
cd bnxt_en-<versão>
```

5. Execute o seguinte comando:

```
make; make install; modprobe -r bnxt_en; modprobe bnxt_en
```

Para a funcionalidade RDMA, instale o driver `bnxt_en` e `bnxt_re`. Use `netxtreme-bnxt_en-<version>.tar.gz` em vez de `bnxt_en-<version>.tar.gz`.

Instalação de módulo

RHEL

A imagem do driver será instalada com uma das seguintes opções:

- Monte a imagem `bnxt_en-x.x.x-rhelYuZ-x86_64-dd.iso` usando o console Dell iDRAC Virtual.
- Monte a imagem `bnxt_en-x.x.x-rhelYuZ-x86_64-dd.iso` em um CD/DVD.
- Copie a imagem `bnxt_en-x.x.x-rhelYuZ-x86_64-dd.iso` em um dispositivo USB e monte o dispositivo.

Inicie a instalação do SO, pressione a tecla `tab` e insira "linux dd". Continue a instalação até que o disco do driver seja solicitado e selecione o driver `bnxt_en`.

SLES

A imagem do driver será instalada com uma das seguintes opções:

- Monte a imagem `bnxt_en-x.x.x-rhelYuZ-x86_64-dd.iso` usando o console Dell iDRAC Virtual.
- Monte a imagem `bnxt_en-x.x.x-rhelYuZ-x86_64-dd.iso` em um CD/DVD.
- Extraia a imagem `bnxt_en-x.x.x-rhelYuZ-x86_64-dd.iso`, copie o conteúdo para um dispositivo USB e monte o dispositivo USB.

Comandos Ethtool do Linux



Nota: Na [Tabela 19](#), `ethX` deve ser substituído pelo nome real da interface.

Tabela 19: Comandos Ethtool do Linux

Comando	Descrição
<code>ethtool -s ethX speed 25000 autoneg off</code>	Define a velocidade. Se o link estiver ativo em uma porta, o driver não permitirá que outra porta seja definida para uma velocidade incompatível.
<code>ethtool -i ethX</code>	A saída inclui a versão do pacote, versão de NIC do BIOS (código de inicialização).
<code>ethtool -k ethX</code>	Mostra recursos de descarregamento.
<code>ethtool -K ethX tso off</code>	Desativa o TSO.
<code>ethtool -K ethX gro off lro off</code>	Desativa o GRO/LRO.
<code>ethtool -g ethX</code>	Mostra tamanhos do anel.
<code>ethtool -G ethX rx N</code>	Define tamanhos do anel.
<code>ethtool -S ethX</code>	Obtém estatísticas.
<code>ethtool -l ethX</code>	Mostra o número de anéis.
<code>ethtool -L ethX rx 0 tx 0 combined M</code>	Define o número de anéis.
<code>ethtool -C ethX rx-frames N</code>	Define a coalescência de interrupção. Outros parâmetros suportados são: rx-usecs, rx-frames, rx-usecs-irq, rx-frames-irq, tx-usecs, tx-frames, tx-usecs-irq, tx-frames-irq.
<code>ethtool -x ethX</code>	Mostra a tabela indireta de fluxo de RSS e a chave de RSS.
<code>ethtool -s ethX autoneg on speed 10000 duplex full</code>	Ative o Autoneg (consulte para obter “Configuração de negociação automática” na página 40 mais detalhes)
<code>ethtool --show-eee ethX</code>	Mostra o estado de EEE.
<code>ethtool --set-eee ethX eee off</code>	Desativa EEE.
<code>ethtool --set-eee ethX eee on tx-lpi off</code>	Ativa EEE, mas desativa LPI.
<code>ethtool -L ethX combined 1 rx 0 tx 0</code>	Desativa RSS. Define os canais combinados para 1.
<code>ethtool -K ethX ntuple off</code>	Desativa o RFS acelerado pela desativação de filtros ntuple.
<code>ethtool -K ethX ntuple on</code>	Ativa RFS acelerado.
<code>Ethtool -t ethX</code>	Executa vários autotestes diagnósticos.
<code>echo 32768 > /proc/sys/net/core/rps_sock_flow_entries</code>	Ativa RFS ara o anel X.
<code>echo 2048 > /sys/class/net/ethX/queues/rx-X/rps_flow_cnt</code>	
<code>sysctl -w net.core.busy_read=50</code>	Isso define o tempo para leitura ocupada do anel de recepção do dispositivo para 50 microssegundos. Para aplicativos de soquete esperando pela chegada de dados, usar esse método geralmente pode diminuir a latência em 2 ou 3 microssegundos à custa de uma maior utilização da CPU.
<code>echo 4 > /sys/bus/pci/devices/0000:82:00.0/sriov_numvfs</code>	Ativa SR-IOV com quatro VFs no barramento 82, Dispositivo 0 e Função 0.
<code>ip link set ethX vf 0 mac 00:12:34:56:78:9a</code>	Define o endereço MAC da FV.
<code>ip link set ethX vf 0 state enable</code>	Define o estado do link VF para VF 0.
<code>ip link set ethX vf 0 vlan 100</code>	Define VF 0 com ID de VLAN 100.

VMware

Os drivers ESX são fornecidos no formato VIB padrão para VMware.

1. Para instalar o driver Ethernet e RDMA, emita os comandos:

```
$ esxcli software vib install --no-sig-check -v <bnxtnet>-<driver version>.vib
```

```
$ esxcli software vib install --no-sig-check -v <bnxtroce>-<driver version>.vib
```

2. Uma reinicialização do sistema é necessária para que o novo driver tenha efeito.

Outros comandos úteis para VMware mostrados na [Tabela 20](#).



Nota: Na [Tabela 20](#), vmnicX deve ser substituído pelo nome real da interface.



Nota: \$ kill -HUP \$(cat /var/run/vmware/vmkdevmgr.pid) Esse comando é exigido após vmkload_mod bnxtnet para exibir o módulo com sucesso.

Tabela 20: Comandos de VMware

Comando	Descrição
esxcli software vib list grep bnx	Liste os VIBs instalados para ver se o driver bnxt foi instalado com êxito.
esxcfg-module -l bnxtnet	Imprime informações do módulo na tela.
esxcli network get -n vmnicX	Obtém as propriedades de vmnicX.
esxcfg-module -g bnxtnet	Imprime parâmetros do módulo.
esxcfg-module -s 'multi_rx_filters=2 disable_tap=0 max_vfs=0,0 RSS=0'	Define os parâmetros do módulo.
vmkload_mod -u bnxtnet	Descarrega o módulo bnxtnet.
vmkload_mod bnxtnet	Carrega o módulo bnxtnet.
esxcli network nic set -n vmnicX -D full -S 25000	Define a velocidade e a duplexação de vmnicX.
esxcli network nic down -n vmnicX	Desativa vmnicX.
esxcli network nic up -n vmnic6	Ativa vmnicX.
bnxtnetcli -s -n vmnic6 -S "25000"	Define a velocidade do link. O Bnxtnetcli é necessário para as versões mais antigas do ESX, para suportar a configuração de velocidade de 25G.

Atualização de firmware

O firmware da NIC pode ser atualizado usando um dos seguintes métodos:

- Usando o Pacote de atualização da Dell (DUP) quando o sistema está em modo de SO inicializado. Esse método se aplica somente aos sistemas operacionais Windows e Linux.
- Usando o Dell iDRAC – controlador de ciclo de vida. Esse método pode ser usado independentemente do sistema operacional. Se o sistema estiver executando VMware, use o Controlador de ciclo de vida para atualizar o firmware.

Consulte a página de suporte do produto em <http://www.dell.com/support>

Pacote de atualização da Dell

Consulte as seções a seguir para usar o Pacote de atualização da Dell (DUP):

Windows

O firmware do controlador Broadcom NetXtreme série E pode ser atualizado usando o pacote Dell DUP. O executável é fornecido em formato executável padrão do Windows x64. Clique duas vezes no arquivo para executá-lo.

Pacotes DUP podem ser baixados em <http://support.dell.com>

Linux

O Dell Linux DUP é fornecido no formato executável x86_64. Use o `chmod` padrão do Linux para atualizar a permissão de execução e executar o executável. Consulte o seguinte exemplo:

1. Faça login no Linux.
2. `scp` ou `cp` no executável DUP para o sistema de arquivo. Um exemplo típico é:

```
cp /var/run/media/usb/Network_Firmware_<versão>.BIN /root/
```

3. Execute o seguinte comando:

```
chmod 755 Network_Firmware_<versão>.BIN
```

4. Execute o seguinte comando:

```
./Network_Firmware_<versão>.BIN
```

Uma reinicialização é necessária para ativar o novo firmware.

Propriedades avançadas do driver de Windows e mensagens do log de erros

Propriedades avançadas do driver

As propriedades avançadas do driver de Windows são mostradas na [Tabela 21](#).

Tabela 21: Propriedades avançadas do driver de Windows

Chave do driver	Parâmetros	Descrição
Descarregamento de tarefa encapsulada	Ativar ou desativar	Usado para configuração do descarregamento de tarefa encapsulada de NVGRE.
Ethernet eficiente em energia	Ativar ou desativar	EEE ativado para portas de cobre e desativado para portas SFP+ ou SFP28. Esse recurso está ativado somente para o adaptador BCM957406A4060.
Controle de fluxo	TX, RX ou TX/RX ativado	Configure o controle de fluxo em RX, TX ou nos dois lados.
Moderação de interrupção	Ativar ou desativar	Ativado por padrão. Permite que quadros sejam processados em lote, poupando tempo da CPU.
Pacote Jumbo	1514, 4088 ou 9014	Tamanho do pacote Jumbo.
Descarregamento de envio de grande porte V2 (IPV4)	Ativar ou desativar	LSO para IPV4.
Descarregamento de envio de grande porte V2 (IPV6)	Ativar ou desativar	LSO para IPV6.
Locally Administered Address	Endereço MAC inserido pelo usuário.	Substitua o endereço MAC de hardware padrão após a inicialização do SO.
Número de filas RSS	2, 4 ou 8.	O padrão é 8. Permite que o usuário configure as filas de escalonamento do lado de recepção.
Prioridade e VLAN	Prioridade e VLAN desativadas, Prioridade ativada, VLAN ativada, Prioridade e VLAN ativadas.	Ativado por padrão. Usado para configuração de 802.1Q e 802.1P.
Buffer de recebimento (0=Auto)	Incrementos de 500.	O padrão é Auto.
Receive Side Scaling (RSS)	Ative ou desative.	Ativado por padrão
Receive Segment Coalescing (IPV4)	Ative ou desative.	Ativado por padrão
Receive Segment Coalescing (IPV6)	Ative ou desative.	Ativado por padrão
Perfil de equilíbrio de carga de RSS	Estática de escalonamento NUMA, Processador mais próximo, Estática do processador mais próximo, escalonamento conservador, escalonamento NUMA.	Estática de escalonamento NUMA padrão.

Tabela 21: Propriedades avançadas do driver de Windows (Continuação)

Chave do driver	Parâmetros	Descrição
Velocidade e duplexação	1 Gbps, 10 Gbps, 25 Gbps ou negociação automática.	Portas de cobre de 10 Gbps podem negociar velocidades automaticamente, enquanto portas de 25 Gbps estão definidas para velocidades forçadas.
SR-IOV	Ative ou desative.	Ativado por padrão. Esse parâmetro funciona com configuração SR-IOV configurado por HW e SR-IOV configurado por BIOS.
Descarregamento da soma de verificação de TCP/UDP IPV4	TX/RX ativado, TX ativado ou RX ativado ou descarregamento desativado.	RX e TX ativados por padrão.
Descarregamento da soma de verificação de TCP/UDP IPV6	TX/RX ativado, TX ativado ou RX ativado ou descarregamento desativado.	RX e TX ativados por padrão.
Buffers de transmissão (0=Auto)	Incremento de 50.	O padrão é Auto.
Fila de máquina virtual	Ative ou desative.	Ativado por padrão.
ID de VLAN	Número configurado por usuário.	O padrão é 0.

Mensagens de log de evento

[Tabela 22](#) fornece as mensagens do Log de eventos registradas pelo driver NDIS do Windows nos logs de eventos.

Tabela 22: Mensagens de log de eventos do Windows

ID da mensagem	Comentário
0x0001	Falha na alocação da memória.
0x0002	Indicação de link inoperante.
0x0003	Indicação de link operante.
0x0009	Link 1000 cheio.
0x000A	Link 2500 cheio.
0x000b	Inicialização bem-sucedida.
0x000c	Reinicialização de miniporta.
0x000d	Falha na inicialização.
0x000E	Link de 10Gb bem-sucedido.
0x000F	Falha na vinculação de camada de driver.
0x0011	Falha ao definir atributos.
0x0012	Falha de DMA de dispersão/reunião.
0x0013	Falha na inicialização de fila padrão.
0x0014	Versão de firmware incompatível.
0x0015	Interrupção única.

Tabela 23: Mensagens de log de evento

0x0016	O firmware falhou ao responder dentro do tempo alocado.
0x0017	O firmware retornou o status de falha.
0x0018	O firmware está em um estado desconhecido.
0x0019	O módulo óptico não é suportado.
0x001A	Seleção de velocidade incompatível entre a Porta 1 e a Porta 2. As velocidades de link relatadas estão corretas e podem não coincidir com a configuração de Velocidade e Duplex.
0x001B	Seleção de velocidade incompatível entre a Porta 1 e a Porta 2. A configuração do link tornou-se ilegal.
0x001C	Controlador de rede configurado para link full-duplex de 25Gb.
0x0020	Falha na inicialização do suporte RDMA.
0x0021	O firmware RDMA do dispositivo é incompatível com este driver.
0x0022	O tamanho de BAR do Doorbell é muito pequeno para o RDMA.
0x0023	Falha ao reiniciar o RDMA após a reinicialização do dispositivo.
0x0024	Falha ao reiniciar o RDMA após a inicialização do sistema
0x0025	Falha na inicialização do RDMA. Não há recursos suficientes.
0x0026	O RDMA não está ativado no firmware.
0x0027	Falha ao iniciar, um endereço MAC não está configurado.
0x0028	Parada de transmissão detectada. Controle de fluxo TX será desativado a partir de agora.

Agrupamentos

Windows

Os dispositivos Broadcom NetXtreme-E instalados em plataformas Dell podem participar da funcionalidade de agrupamento NIC usando a solução de agrupamento da Microsoft. Consulte a documentação pública da Microsoft descrita no seguinte link:

<https://www.microsoft.com/en-us/download/details.aspx?id=40319>

Microsoft LBFO é um driver de agrupamento nativo que pode ser usado no SO Windows. O driver de agrupamento também fornece recursos de marcação de VLAN.

Linux

A vinculação do Linux é usada para agrupamento em Linux. O conceito se refere a carregar o driver de vinculação e a adicionar membros do agrupamento ao vínculo, o que poderia equilibrar a carga do tráfego.

Use as etapas a seguir para configurar a vinculação do Linux:

1. Execute o seguinte comando:

```
modprobe bonding mode="balance-alb". Isso criará uma interface de vínculo.
```

2. Adicione clientes vinculados à interface de vínculo. Um exemplo é mostrado abaixo:

```
ifenslave bond0 ethX; ifenslave bond0 ethY
```

3. Atribua um endereço IPV4 para vincular a interface usando `ifconfig bond0 IPV4Address netmask NetMask up`. IPV4Address e NetMask são o endereço IPV4 e a máscara de rede associada.



Nota: O endereço IPV4 deve ser substituído pelo endereço IPV4 da rede real. NetMask deve ser substituído pela máscara de rede IPV4 real.

4. Atribua um endereço IPV6 para vincular a interface usando `ifconfig bond0 IPV6Address netmask NetMask up`. IPV6Address e NetMask são o endereço IPV6 e a máscara de rede associada.



Nota: O endereço IPV6 deve ser substituído pelo endereço IPV6 da rede real. NetMask deve ser substituído pela máscara de rede IPV6 real.

Consulte a documentação de Vinculação do Linux para conhecer as configurações avançadas.

Configuração de nível de sistema

Consulte as seguintes seções para obter informações sobre a configuração de NIC ao nível do sistema.

Menu UEFI HII

Os controladores Broadcom NetXtreme série E podem ser configurados para pré-inicialização, iscsi e configuração avançada, como SR-IOV usando menu HII (Interface humana).

Para definir as configurações, durante a inicialização do sistema selecione F2 -> Configuração do sistema -> Configurações do dispositivo. Selecione o adaptador de rede desejado para visualizar e alterar a configuração.

Página de configuração principal

Essa página exibe o status do link de rede atual, Barramento PCI-E:Dispositivo:Função, endereço MAC do adaptador e do dispositivo Ethernet.

A placa 10GBaseT permitirá que o usuário ative ou desative a função Ethernet eficiente em energia (EEE).

Propriedades da imagem do firmware

Página de configuração principal -> Exibições de propriedades de imagem de firmware, Versão da família, que consiste dos números de versão do BIOS do controlador, e dos números de versão de Agente de inicialização múltipla (MBA), UEFI, iSCSI e Gerenciamento de configuração abrangente (CCM).

Configuração de nível de dispositivo

Página de configuração principal -> Configuração de nível de dispositivo permite que o usuário ative o modo SR-IOV, o número de funções virtuais por função física, os vetores MSI-X por função virtual e o número máximo de vetores MSI-X da função física.

Configuração de NIC

Configuração de NIC -> Protocolo de inicialização legada é usado para selecionar e configurar PXE e iSCSI ou para desativar o modo de inicialização legada. O tipo de inicialização automática pode ser Auto, int18h (interrupção 18h), int19h (interrupção 19h) ou BBS.

MBA e iSCSI também podem ser configurados usando CCM. O modo de BIOS legado usa o CCM para configuração. O prompt de configuração oculto pode ser usado para desativar ou ativar a exibição da faixa.

A VLAN para PXE pode ser ativada ou desativada e o ID da VLAN pode ser configurado pelo usuário. Consulte a [“Configuração de negociação automática” na página 40](#) para obter detalhes sobre as opções de configuração de velocidade do link.

Configuração iSCSI

A configuração de inicialização iSCSI pode ser definida pela página Configuração principal -> Configuração iSCSI. Parâmetros como IPV4 ou IPV6, programa de inicialização iSCSI ou destino iSCSI podem ser definidos nessa página.

Consulte ["Inicialização iSCSI" na página 49](#) para obter informações detalhadas sobre a configuração.

Gerenciamento de configuração abrangente

A configuração pré-inicialização pode ser configurada usando a opção de menu Gerenciamento de configuração abrangente (CCM). Durante o PÓS-BIOS do sistema, a mensagem da faixa da Broadcom será exibida com a opção de alterar os parâmetros pelo menu Control-S. Quando Control-S é pressionado, uma lista de dispositivos é preenchida com todos os adaptadores de rede Broadcom encontrados no sistema. Selecione a NIC desejada para configuração.

Configuração de nível de hardware

Os parâmetros que podem ser configurados usando essa seção são os mesmos do menu HII "Configuração de nível de dispositivo".

Menu de configuração MBA

Os parâmetros que podem ser configurados usando essa seção são os mesmos do menu HII "Configuração de NIC".

Menu principal de inicialização iSCSI

Os parâmetros que podem ser configurados usando essa seção são os mesmos do menu HII "Configuração de iSCSI".

Configuração de negociação automática



Nota: Nos dispositivos NPAR (particionamento NIC) em que uma porta é compartilhada por várias funções PCI, a velocidade da porta é pré-configurada e não pode ser alterada pelo driver.

O controlador Broadcom NetXtreme-E suporta os seguintes recursos de negociação automática:

- Negociação automática de velocidade do link
- Negociação automática de Pausa/Controle de fluxo
- Negociação automática da FEC (Forward Error Correction, correção de erro de encaminhamento)



Nota: Em relação à velocidade do link AN, ao usar conectores SFP+, SFP28, use os transceptores ópticos DAC ou Multimode capazes de suportar AN. Certifique-se de que a porta do parceiro do link tenha sido configurada para o protocolo de negociação automática correspondente. Por exemplo, se a porta Broadcom local for configurada como protocolo AN do IEEE 802.3by, o parceiro de link deve suportar AN e ser configurado como protocolo AN do IEEE 802.3by.



Nota: Para controladores de rede NetXtreme-E de portas duplas, 10 Gbps e 25 Gbps não são uma combinação suportada de velocidade do link.

A combinação suportada de configurações de velocidade do link para o controlador de rede NetXtreme-E de duas portas é mostrada na [Tabela 24 na página 41](#).

Tabela 24: Combinação suportada de configurações de velocidade do link

Port1 Link Speed Setting	Port 2 Link Setting									
	Forced 1G	Forced 10G	Forced 25G	AN Enabled {1G}	AN Enabled {10G}	AN Enabled {25G}	AN Enabled {1/10G}	AN Enabled {1/25G}	AN Enabled {10/25G}	AN Enabled {1/10/25G}
Forced 1G	P1: no AN	P1: no AN	P1: no AN	P1: no AN	P1: no AN	P1: no AN	P1: no AN	P1: no AN	P1: no AN	P1: no AN
	P2: no AN	P2: no AN	P2: no AN	P2: {1G}	P2: AN {10G}	P2: AN {25G}	P2: AN {1/10G}	P2: AN {1/25G}	P2: AN {10/25G}	P2: AN {1/10/25G}
Forced 10G	P1: no AN	P1: no AN	Not supported	P1: no AN	P1: no AN	Not supported	P1: no AN	P1: no AN	P1: no AN	P1: no AN
	P2: no AN	P2: no AN		P2: {1G}	P2: {10G}		P2: AN {1/10G}	P2: AN {1G}	P2: AN {10G}	P2: AN {1/10G}
Forced 25G	P1: no AN	Not supported	P1: no AN	P1: no AN	P1: no AN	P1: no AN	P1: no AN	P1: no AN	P1: no AN	P1: no AN
	P2: no AN		P2: no AN	P2: no AN	P2: no AN	P2: AN {1G}	P2: AN {1/25G}	P2: AN {25G}	P2: AN {1/25G}	
AN Enabled {1G}	P1: {1G}	P1: {1G}	P1: {1G}	P1: AN {1G}	P1: AN {1G}	P1: AN {1G}	P1: AN {1G}	P1: AN {1G}	P1: AN {1G}	P1: AN {1G}
	P2: no AN	P2: no AN	P2: no AN	P2: AN {1G}	P2: AN {10G}	P2: AN {25G}	P2: AN {1/10G}	P2: AN {1/25G}	P2: AN {10/25G}	P2: AN {1/10/25G}
AN Enabled {10G}	P1: AN {10G}	P1: AN {25G}	Not supported	P1: AN {10G}	P1: AN {10G}	Not supported	P1: AN {25G}	P1: AN {10G}	P1: AN {10G}	P1: AN {10G}
	P2: no AN	P2: no AN		P2: AN {1G}	P2: AN {10G}		P2: AN {1G}	P2: AN {10G}	P2: AN {1/10G}	
AN Enabled {25G}	P1: AN {25G}	Not supported	P1: AN {25G}	P1: AN {25G}	Not supported	P1: AN {25G}	P1: AN {1/10G}	P1: AN {25G}	P1: AN {25G}	P1: AN {25G}
	P2: no AN		P2: no AN	P2: AN {1G}		P2: AN {25G}	P2: AN {1/10G}	P2: AN {1/25G}	P2: AN {25G}	P2: AN {1/25G}
AN Enabled {1/10G}	P1: AN {1/10G}	P1: AN {1/10G}	P1: AN {1G}	P1: AN {1/10G}	P1: AN {1/10G}	P1: AN {1/10G}	P1: AN {1/25G}	P1: AN {1G}	P1: AN {1/10G}	P1: AN {1/10G}
	P2: no AN	P2: no AN	P2: no AN	P2: AN {1G}	P2: AN {10G}	P2: AN {25G}	P2: AN {1/10G}	P2: AN {1G}	P2: AN {10G}	P2: AN {1/10G}
AN Enabled {1/25G}	P1: AN {1/25G}	P1: {1G}	P1: AN {1/25G}	P1: AN {1/25G}	P1: AN {1G}	P1: AN {1/25G}	P1: AN {10/25G}	P1: AN {1/25G}	P1: AN {1/25G}	P1: AN {1/25G}
	P2: no AN	P2: no AN	P2: no AN	P2: AN {1G}	P2: AN {10G}	P2: AN {25G}	P2: AN {1/10G}	P2: AN {1/25G}	P2: AN {25G}	P2: AN {1/25G}
AN Enabled {10/25G}	P1: AN {10/25G}	P1: {10G}	P1: AN {25G}	P1: AN {10/25G}	P1: AN {10G}	P1: AN {25G}	P1: AN {1/10/25G}	P1: AN {25G}	P1: AN {10/25G}	P1: AN {10/25G}
	P2: no AN	P2: no AN	P2: no AN	P2: AN {1G}	P2: AN {10G}	P2: AN {25G}	P2: AN {1/10G}	P2: AN {25G}	P2: AN {10/25G}	P2: AN {1/10/25G}
AN Enabled {1/10/25G}	P1: AN {1/10/25G}	P1: {1/10G}	P1: AN {1/25G}	P1: AN {1/10/25G}	P1: AN {1/10G}	P1: AN {1/25G}	P1: AN {1/10/25G}	P1: AN {1/25G}	P1: AN {1/10/25G}	P1: AN {1/10/25G}
	P2: no AN	P2: no AN	P2: no AN	P2: AN {1G}	P2: AN {10G}	P2: AN {25G}	P2: AN {1/10G}	P2: AN {1/25G}	P2: AN {10/25G}	P2: AN {1/10/25G}



Nota: Use um transceptor óptico compatível ou um cabo de cobre de conexão direta para alcançar 1 Gbps de velocidade do link.

- P1 – configuração da porta 1
- P2 – configuração da porta 2
- AN – negociação automática
- Sem AN – velocidade forçada
- {link speed} – velocidade esperada do link

- AN {link speeds} – velocidades de link de negociação automática suportadas anunciadas.

As velocidades de link esperadas, com base nas configurações do local e do parceiro de link, são mostradas na [Tabela 25](#).

Tabela 25: Velocidades de link esperadas

Local Speed Settings	Link Partner Speed Settings									
	Forced 1G	Forced 10G	Forced 25G	AN Enabled {1G}	AN Enabled {10G}	AN Enabled {25G}	AN Enabled {1/10G}	AN Enabled {1/25G}	AN Enabled {10/25G}	AN Enabled {1/10/25G}
Forced 1G	1G	No link	No link	No link	No link	No link	No link	No link	No link	No link
Forced 10G	No link	10G	No link	No link	No link	No link	No link	No link	No link	No link
Forced 25G	No link	No link	25G	No link	No link	No link	No link	No link	No link	No link
AN {1G}	No link	No link	No link	1G	No link	No link	1G	1G	No link	1G
AN {10G}	No link	No link	No link	No link	10G	No link	10G	No link	10G	10G
AN {25G}	No link	No link	No link	No link	No link	25G	No link	25G	25G	25G
AN {1/10G}	No link	No link	No link	1G	10G	No link	10G	1G	10G	10G
AN {1/25G}	No link	No link	No link	1G	No link	25G	1G	25G	25G	25G
AN {10/25G}	No link	No link	No link	No link	10G	25G	10G	25G	25G	25G
AN {1/10/25G}	No link	No link	No link	1G	10G	25G	10G	25G	25G	25G



Nota: A velocidade de 1 Gbps do link para o SFP+/SFP28 não é atualmente suportada nesta versão.

Para ativar a negociação automática da velocidade do link, as seguintes opções podem ser ativadas no menu BIOS HII do sistema ou no CCM:



Nota: A Detecção Automática de Mídia deve estar ativada ao habilitar a negociação automática da velocidade do link quando estiver conectado por um módulo do transceptor ótico/cabo.

BIOS do sistema -> Configurações do dispositivo -> NetXtreme-E NIC -> Configuração do nível do dispositivo

Velocidade do link operacional

Esta opção configura a velocidade de link usada pelos drivers de pré-inicialização (MBA e UEFI), pelos drivers de sistema operacional (Linux, ESX) e pelo firmware. Essa configuração é substituída pela configuração do driver no estado atual do SO. O driver do Windows (bnxnd_.sys) usa configurações de velocidade de link no arquivo .inf.

Velocidade do link do firmware

Essa opção configura a velocidade do link usada pelo firmware quando o dispositivo está em D3.

Protocolo de negociação automática

Protocolo de negociação automática suportado, usado para negociar a velocidade do link com o parceiro de link. Essa opção deve corresponder à configuração do protocolo AN na porta do parceiro de link. O Broadcom NetXtreme-E NIC suporta os seguintes protocolos de negociação automática: IEEE 802.3by, consórcios 25G/50G e 25G/50G BAM. Por padrão, esta opção é configurada como IEEE 802.3by e retorna aos consórcios de 25 G/50 G.

A velocidade do link e Controle de fluxo/Pausa devem ser configurados no driver no SO host.

Configurações do driver do Windows

Para acessar as configurações do driver do Windows:

Abra **Gerenciador do Dispositivo Windows -> Adaptador do Broadcom NetXtreme Série E -> Propriedades avançadas -> guia Avançado**

Controle de fluxo = Negociação automática

Permite o AN do quadro de Controle de fluxo/Pausa.

Velocidade e duplex = negociação automática

Permite a velocidade do link AN.

Configurações do driver do Linux



Nota: Para adaptadores de rede 10GBase-T NetXtreme-E, a negociação automática deve ser ativada.



Os anúncios de 25 G são padrões mais novos, configurados pela primeira vez na interface ethtool do kernel 4.7. Para suportar totalmente essas novas velocidades de anúncios para a negociação automática, um kernel 4.7 (ou mais recente) e um utilitário ethtool mais recente (versão 4.8) são necessários.

ethtool -s eth0 speed 25000 autoneg off

Esse comando desativa a negociação automática e força a velocidade do link para 25 Gbps.

ethtool -s eth0 autoneg on advertise 0x0

Este comando permite a negociação automática e anuncia que o dispositivo suporta todas as velocidades: 1 G, 10 G, 25 G.

As velocidades anunciadas a seguir são suportadas.

- 0x020 – 1000baseT Full
- 0x1000 – 10000baseT Full
- 0x80000000 – 25000baseCR Full

ethtool -A eth0 autoneg on|off

Use esse comando para ativar/desativar a negociação automática no quadro Pausa.

ethtool -a eth0

Use esse comando para exibir a atual configuração de negociação automática do controle de fluxo.

Configurações do driver ESXi



Nota: Para adaptadores de rede 10GBase-T NetXtreme-E, a negociação automática deve ser ativada. O uso da velocidade forçada no adaptador 10GBase-T resulta na falha do comando esxcli.



Nota: O VMWare não suporta velocidades de 25 G no ESX6.0. Neste caso, use o segundo utilitário (BNXTNETCLI) para configurar a velocidade de 25 G. Para o ESX6.0U2, a velocidade de 25 G é suportada.

\$ esxcli network nic get -n <iface>

Esse comando mostra a velocidade atual, o duplex, a versão do driver, a versão do firmware e o status do link.

\$ esxcli network nic set -S 10000 -D full -n <iface>

Esse comando configura a velocidade forçada para 10 Gbps.

\$ esxcli network nic set -a -n <iface>

Permite a negociação automática da velocidade do link na interface <iface>.

\$ esxcli network nic pauseParams list

Use esse comando para obter a lista de Parâmetros de pausa.

\$ esxcli network nic pauseParams set --auto <1/0> --rx <1/0> --tx <1/0> -n <iface>

Use esse comando para configurar os Parâmetros de pausa.



Nota: A negociação automática do controle de fluxo/pausa pode ser configurada somente quando a interface for configurada no modo de negociação automática da velocidade do link.

Negociação automática da FEC

Para ativar ou desativar a negociação automática da FEC, as seguintes opções podem ser ativadas no menu BIOS HII do sistema ou no CCM:

- **BIOS do sistema -> Configurações do dispositivo -> NetXtreme-E NIC -> Configuração do nível do dispositivo**

A negociação automática da FEC usa dois parâmetros durante a troca de negociações: a FEC compatível e a FEC solicitada.

Se uma NIC anuncia que é capaz de fazer negociação automática da FEC, então as configurações da FEC são acionadas pelo switch. Então, ela poderá se vincular a um switch que ative a FEC ou a um switch de desative a FEC.

Por exemplo:

- switch – compatível = 1, compatível = 1. Então, este será um link de FEC.
- switch – compatível = N/D, solicitado = 0. Então a FEC estará desativada.

Para os controladores Ethernet NetXtreme-E, há compatibilidade somente com Base-R FEC (CL74). A [Tabela 26](#) mostra todas as configurações compatíveis com um parceiro de link.

Tabela 26: Configurações de FEC compatíveis para BCM5730X/BCM5740X

Configuração de SFEC local	Configuração de FEC do parceiro de link			
	Velocidade forçada sem FEC	Velocidade forçada Base-R FEC CL74	AN (nenhum)	AN (nenhum, Base-R)
Velocidade forçada sem FEC	Link sem FEC	Sem link	Sem link	Sem link
Velocidade forçada Base-R FEC CL74	Sem link	Base-R FEC CL74	Sem link	Sem link
AN (nenhum)	Sem link	Sem link	Link sem FEC	Base-R FEC CL74
AN (nenhum, Base-R)	Sem link	Sem link	Base-R FEC CL74	Base-R FEC CL74



Nota: Para a velocidade forçada, a configuração de velocidade deve ser a mesma nos dois lados.



Nota: AN {nenhum} significa que o AN anuncia o bit compatível com Base-R. Configure o bit F0 no IEEE802.3by e o bit F2 no consórcio.



Nota: AN {nenhum, Base-R} significa que o AN anuncia o bit compatível com Base-R e o bit solicitado. Configure os bits F0 e F1 no IEEE802.3by, e os bits F2 e F4 no consórcio.

Para o controlador Ethernet NetXtreme-E, a FEC é compatível com Base-R FEC (CL74) e com RS-FEC (CL91/CL108). A [Tabela 27](#) mostra todas as configurações compatíveis com um parceiro de link.

Tabela 27: Configurações de FEC compatíveis para o BCM5741X

Configuração de FEC local	Configuração de FEC do parceiro de link					
	Velocidade forçada sem FEC	Velocidade forçada Base-R FEC CL74	Velocidade forçada RS-FEC CL91/CL108	AN (nenhum)	An (nenhum, Base-R)	AN (nenhum, Base-R, RS)
Forçada sem FEC	Link sem FEC	Sem link	Sem link	Sem link	Sem link	Sem link
Velocidade forçada Base-R FEC CL74	Sem link	Base-R FEC CL74	Sem link	Sem link	Sem link	Sem link
Forçar RS-FEC CL91/CL108	Sem link	Sem link	RS-FEC CL91/CL108	Sem link	Sem link	Sem link
AN (nenhum)	Sem link	Sem link	Sem link	Link sem FEC	Base-R FEC CL74	RS-FEC CL91/CL108
AN (nenhum, Base-R)	Sem link	Sem link	Sem link	Base-R FEC CL74	Base-R FEC CL74	RS-FEC CL91/CL108
AN (nenhum, Base-R, RS)	Sem link	Sem link	Sem link	RS-FEC CL91/CL108	RS-FEC CL91/CL108	RS-FEC CL91/CL108



Nota: Para a velocidade forçada, a configuração de velocidade deve ser a mesma nos dois lados.



Nota: AN {nenhum} significa que o AN anuncia o bit compatível com Base-R. Configure o bit F0 no IEEE802.3by e o bit F2 no consórcio.



Nota: AN {nenhum, Base-R} significa que o AN anuncia o bit compatível com Base-R e o bit solicitado. Configure os bits F0 e F1 no IEEE802.3by, e os bits F2 e F4 no consórcio.

Treinamento do link

O treinamento do link permite que ambas as extremidades, o adaptador Broadcom e o outro lado ajustem as configurações de potência e demais parâmetros para maximizar a confiabilidade e a eficiência do canal de comunicação entre os dois dispositivos. O objetivo é eliminar a necessidade de se realizar um ajuste específico para o canal entre diferentes tipos e comprimentos de cabos. O treinamento do link é realizado para velocidades CR/KR, precedendo a negociação automática. O treinamento do link é operacional quando a negociação automática está ativada. A política de link desativa automaticamente o treinamento do link se ele não resultar em um link ativo com o parceiro de link. Essa política de link garante a compatibilidade com um parceiro de link que não seja compatível com treinamento de link.

A [Tabela 28](#) mostra a relação entre o tipo e a velocidade da mídia para os controladores Ethernet BCM5730X, BCM5740X e BCM5741X.

Tabela 28: Relação entre o tipo e a velocidade da mídia no treinamento do link

Tipo de cabo de mídia e local	Configuração do treinamento do link do parceiro de link		
	Treinamento do link de velocidade forçada desativado	Treinamento do link de velocidade forçada ativado	AN (Treinamento Automático do Link)
DAC de velocidade forçada (SFP+/SFP28/QSFP28)	Link sem treinamento do link	Link com treinamento do link	Link com treinamento do link
Óptico de velocidade forçada (transceptor/AOC)	Link sem treinamento do link	Link sem treinamento do link	Ausente
DAC de AN (SFP+SFP28/QSFP28)	Sem link	Sem link	Link com treinamento do link

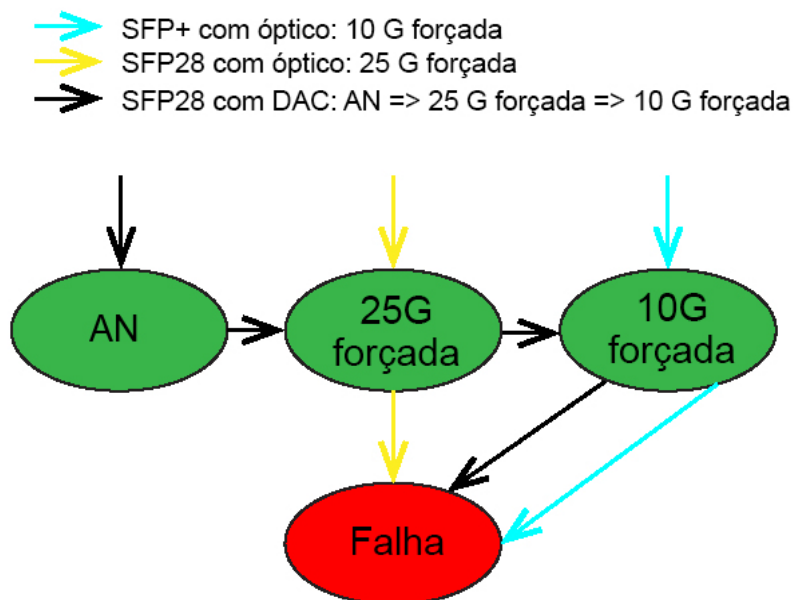
Detecção Automática de Mídia

Como não há compatibilidade com a detecção paralela no SerDes, o firmware implementou um método para aprimorar a detecção do link, chamado de **Detecção Automática de Mídia**. Este recurso é controlado pelo CCm/HII, conforme mostrado na [Figura 13](#).

Quando a **Detecção Automática de Mídia** está ativada, a política do link segue a máquina de estados (consulte a [Figura 13](#)) para estabelecer um link com um parceiro de link. Este comportamento depende do tipo de mídia. Para os cabos DAC, o método retorna a diferentes modos de força e para quando há um link ativo.

NOTA: Não há compatibilidade com detecção paralela quando a negociação automática é realizada a 25 G ou 10 GbE. Isso significa que, se um lado estiver realizando a negociação automática e que o outro lado não estiver, o link não ficará ativo.

Figura 13: Máquina de estados para o recurso de Detecção Automática de Mídia



A [Tabela 29](#) e a [Tabela 30](#) mostram os resultados do link com o recurso de **Deteção Automática de Mídia** ativado.

Tabela 29: Deteção Automática de Mídia para o BCM5730X e o BCM5740X

		Configuração do treinamento do link do parceiro de link	
Configurações do Parceiro de Link		Deteção Automática de Mídia	
Velocidade	FEC	Sem FEC	FEC de Base-R
10G	Sem FEC	Link	Link com Base-R
	Base-R	Sem link	Link com Base-R
25 G	Sem FEC	Link	Link
	Base-R	Sem link	Link com Base-R
AN	Desativada	Link	Link
	FEC automática	Link	Link com Base-R
	Base-R	Link com Base-R	Link com Base-R

Tabela 30: Deteção Automática de Mídia para o BCM5741X

		Configuração do treinamento do link do parceiro de link		
Configurações do Parceiro de Link		Deteção Automática de Mídia		
Velocidade	FEC	Sem FEC	FEC de Base-R	RS-FEC
10G	Sem FEC	Link	Link com Base-R	Sem link
	Base-R	Sem link	Link com Base-R	Sem link
	RS	Sem link	Sem link	Link com RS-FEC
25 G	Sem FEC	Link	Sem link	Sem link
	Base-R	Sem link	Link com Base-R	Sem link
	RS	Sem link	Sem link	Link com RS-FEC
AN	Desativada	Link	Link	Link
	FEC automática	Link	Link com Base-R	Link com RS-FEC
	Base-R	Link com Base-R	Link com Base-R	Link com RS-FEC
	RS	Link com RS	Link com RS-FEC	Link com RS-FEC

Inicialização iSCSI

Os adaptadores Broadcom NetXtreme-E Ethernet suportam a inicialização iSCSI para permitir que sistemas sem disco realizem a inicialização em rede dos sistemas operacionais. A inicialização iSCSI permite que um sistema operacional Windows, Linux ou VMware seja inicializado a partir de uma máquina iSCSI alvo, localizada remotamente em uma rede IP padrão.

Sistemas operacionais suportados para inicialização iSCSI

Os adaptadores Broadcom NetXtreme-E Gigabit Ethernet suportam a inicialização iSCSI nos seguintes sistemas operacionais:

- Windows Server 2012 e 64 bits posterior
- Linux RHEL 7.1 e posterior, SLES11 SP4 ou posterior
- VMware 6.0 U2

Configuração da inicialização iSCSI

Consulte as seguintes seções para obter informações sobre a configuração da inicialização iSCSI.

Configuração do destino iSCSI



Nota: Windows 2016 (ou mais antigo) podem ter um driver da caixa de entrada desatualizado, suportando somente velocidade forçada do link. Para evitar problemas de instalação, como a incapacidade de realizar link ativo com um parceiro de link incompatível, é altamente recomendável que o usuário personalize a imagem/mídia de instalação do Windows com o driver de NIC mais recente disponível que seja compatível com negociação automática de velocidade do link.

A configuração do destino iSCSI varia de acordo com o fornecedor visado. Para obter mais informações sobre como configurar o destino iSCSI, consulte a documentação fornecida pelo fornecedor. O processo geral inclui as seguintes etapas:

1. Crie um destino iSCSI.
2. Crie um disco virtual.
3. Mapeie o disco virtual para o destino iSCSI criado em [Step 1 na página 49](#).
4. Associe um programa de inicialização de iSCSI ao destino iSCSI.
5. Grave o nome do destino iSCSI, o número da porta TCP, o número da unidade lógica (LUN) do iSCSI, o nome de qualificação para a Internet (IQN) do programa de inicialização e as informações sobre a autenticação CHAP.
6. Após configurar o destino iSCSI, obtenha as seguintes informações:
 - IQN de destino
 - Endereço IP do alvo
 - Número da porta TCP do alvo

- LUN do alvo
- IQN do programa de inicialização
- ID e segredo CHAP

Configuração dos parâmetros da inicialização iSCSI

Defina a configuração do software de inicialização iSCSI Broadcom como estática ou dinâmica. Consulte a [Tabela 31](#) para conhecer as opções de configuração disponíveis no menu Parâmetros gerais. [Tabela 31](#) lista os parâmetros para IPv4 e IPv6. Os parâmetros específicos de IPv4 ou IPv6 são anotados.

Tabela 31: Opções de configuração

Opção	Descrição
Parâmetros TCP/IP via DHCP	Essa opção é específica ao IPv4. Controla se o software host de inicialização iSCSI irá obter o endereço IP por meio do DHCP (opção ativada) ou de uma configuração de IP estática (opção desativada).
Autoconfiguração de IP	Essa opção é específica ao IPv6. Controla se o software host da inicialização iSCSI irá configurar um endereço de link local em estado e/ou um endereço com informações de estado se o DHCPv6 estiver presente e em uso (Ativado). Pacotes de solicitação do roteador são enviados por três vezes em intervalos de quatro segundos entre cada tentativa. Opcionalmente, use uma configuração de IP estática (opção desativada).
Parâmetros iSCSI via DHCP	Controla se o software host de inicialização iSCSI irá obter os parâmetros da sua iSCSI-alvo por meio do DHCP (opção ativada) ou de uma configuração estática (opção desativada). As informações estáticas são inseridas na tela Configuração dos parâmetros do programa de inicialização iSCSI.
Autenticação CHAP	Controla se o software host de inicialização iSCSI irá utilizar a autenticação CHAP ao se conectar com a iSCSI-alvo. Se Autenticação CHAP estiver ativada, o ID de CHAP e o Segredo de CHAP são inseridos pela tela Configuração de parâmetros do programa de inicialização iSCSI.
ID do fornecedor de DHCP	Controla o modo como o software host de inicialização iSCSI interpreta o campo ID de classe do fornecedor usado durante o DHCP. Se o campo ID de classe do fornecedor do pacote da Oferta DHCP tiver o mesmo valor, o software host de inicialização iSCSI tenta encontrar os 43 campos de Opções de DHCP para as extensões de inicialização iSCSI necessárias. Quando a função DHCP está desativada, esse valor não precisa ser definido.
Link Up Delay Time (Tempo de espera do link ativo)	Controla quanto tempo, em segundos, o software host de inicialização iSCSI deve aguardar para enviar algum dado pela rede depois que um link Ethernet é estabelecido. Os valores válidos são de 0 a 255. Por exemplo, um usuário pode ter que definir um valor para esta opção se um protocolo de rede, como o Spanning Tree, estiver ativado na interface do switch com o sistema cliente.
Usar carimbo de data/hora do TCP	Controla se a opção Carimbo de data/hora do TCP está ativada ou desativada.
Alvo como o primeiro HDD	Permite que a iSCSI-alvo seja exibida como a primeira unidade de disco rígido do sistema.

Tabela 31: Opções de configuração (Continuação)

Opção	Descrição
Contagem de tentativas para LUN ocupado	Controla o número de novas tentativas de conexão que o programa de inicialização iSCSI poderá fazer se o LUN da iSCSI-alvo estiver ocupado.
Versão IP	Esta opção é específica do IPv6. Alterna entre o protocolo IPv4 ou IPv6. Todas as configurações IP serão perdidas ao alternar de uma versão de protocolo para outra.

Configuração do protocolo de inicialização do MBA

Para configurar o protocolo de inicialização:

1. Reinicialize o sistema.
2. Na faixa PXE, selecione **CTRL+S**. É exibido o **menu Configuração de MBA**.
3. No menu **Configuração de MBA**, use a **SETA PARA CIMA** ou a **SETA PARA BAIXO** para se mover para a opção Protocolo de inicialização. Use a **SETA PARA A ESQUERDA** ou a **SETA PARA A DIREITA** para alterar a opção Protocolo de inicialização para iSCSI.
4. Selecione **Configuração de inicialização iSCSI** no **Menu principal**.

Configuração de inicialização iSCSI

Existem duas maneiras de configurar a inicialização iSCSI:

- Configuração de inicialização iSCSI estática.
- Configuração de inicialização iSCSI dinâmica.

Configuração de inicialização iSCSI estática

Em uma configuração estática, você deve fornecer o endereço IP do sistema, o IQN do programa de inicialização do sistema e os parâmetros do alvo obtidos em [“Configuração do destino iSCSI” na página 49](#). Para obter informações sobre as opções de configuração, consulte [Tabela 31 na página 50](#).

Para configurar os parâmetros da inicialização iSCSI por meio de configuração estática:

1. No menu **Parâmetros gerais**, defina:
 - Parâmetros TCP/IP via DHCP – Desativados. (Para IPv4.)
 - Autoconfiguração de IP – Desativada. (Para IPv6, sem descarregamento.)
 - Parâmetros iSCSI via DHCP – Desativados
 - Autenticação CHAP – Desativada
 - ID do fornecedor de DHCP – BRCM ISAN
 - Tempo de espera do link ativo – 0
 - Usar carimbo de data/hora de TCP – Ativado (para alguns destinos, como Dell/EMC AX100i, é necessário ativar Usar carimbo de data/hora de TCP)
 - Alvo como o primeiro HDD – Desativado
 - Contagem de tentativas para LUN ocupado – 0
 - Versão de IP – IPv6. (Para IPv6, sem descarregamento.)

2. Pressione **ESC** para retornar ao menu **Principal**.
3. No menu **Principal**, selecione a opção **Initiator Parameters** (Parâmetros da inicialização).
4. Na tela **Parâmetros do programa de inicialização**, insira os valores para:
 - Endereço IP (endereços IPv4 e IPv6 não especificados devem ser "0.0.0.0" e ":::", respectivamente)
 - Prefixo da máscara de sub-rede
 - Gateway padrão
 - DNS primário
 - DNS secundário
 - Nome iSCSI (corresponde ao nome do programa de inicialização iSCSI que será usado pelo sistema cliente)



Nota: Insira o endereço IP. Não há verificação de erro referente ao endereço IP para confirmar se não há duplicatas ou atribuições incorretas de segmento/rede.

5. Pressione **ESC** para retornar ao menu **Principal**.
6. No menu **Principal**, selecione a opção **1st Target Parameters** (Parâmetros do primeiro alvo).



Nota: Não é possível configurar um segundo destino para a configuração inicial.

7. Na tela **Parâmetros do 1º alvo**, ative **Conectar** para conectar a iSCSI-alvo. Digite os valores para os campos descritos abaixo, usando os mesmos valores inseridos durante a configuração da iSCSI-alvo:
 - Endereço IP
 - Porta TCP
 - LUN de inicialização
 - Nome iSCSI
8. Pressione **ESC** para retornar ao menu **Principal**.
9. Selecione **ESC**, em seguida, selecione **Sair e Salvar configuração**.
10. Selecione **F4** para salvar a configuração de MBA.

Configuração de inicialização iSCSI dinâmica

Em uma configuração dinâmica, você precisa especificar apenas que as informações do endereço IP do sistema e do alvo/programa de inicialização são fornecidas pelo servidor DHCP (consulte as configurações de IPv4 e IPv6 em [“Configuração do servidor DHCP para suportar a inicialização iSCSI” na página 54](#). No caso do IPv4, com exceção do nome da inicialização iSCSI, todas as configurações contidas nas telas Parâmetros do programa de inicialização, Parâmetros do 1º alvo ou Parâmetros do 2º alvo são ignoradas e não precisam ser apagadas. No caso do IPv6, com exceção do ID e segredo CHAP, todas as configurações contidas nas telas Parâmetros do programa de inicialização, Parâmetros do 1º alvo ou Parâmetros do 2º alvo são ignoradas e não precisam ser apagadas. Para obter informações sobre as opções de configuração, consulte [Tabela 31 na página 50](#).



Nota: Quando um servidor DHCP é utilizado, as entradas do servidor DNS são substituídas pelos valores fornecidos pelo servidor DHCP. Isso ocorre mesmo quando os valores fornecidos localmente são válidos e o servidor DHCP não fornece nenhuma informação sobre o servidor DNS. Quando o servidor DHCP não fornece nenhuma informação sobre o servidor DNS, os valores dos servidores DNS primário e secundário são definidos como 0.0.0.0. Quando o SO Windows assume o controle, o software de inicialização iSCSI da Microsoft recupera os parâmetros do programa de inicialização iSCSI e configura estaticamente os registros apropriados. Isso faz com que todas as configurações sejam substituídas. Como o daemon de DHCP é executado no ambiente do Windows na forma de um processo do usuário, todos os parâmetros TCP/IP precisam ser estaticamente configurados antes que a pilha seja restabelecida no ambiente da inicialização iSCSI.

- Quando a opção 17 de DHCP é utilizada, as informações do alvo são fornecidas pelo servidor DHCP, e o nome da inicialização iSCSI é recuperado a partir do valor programado na tela Parâmetros do programa de inicialização. Se nenhum valor tiver sido selecionado, o controlador define como padrão o nome:

```
iqn.1995-05.com.broadcom.<11.22.33.44.55.66>.iscsiboot
```

onde a sequência **11.22.33.44.55.66** corresponde ao endereço MAC do controlador.

- Quando a opção 43 de DHCP (somente IPv4) é utilizada, todas as configurações das telas Parâmetros do programa de inicialização, Parâmetros do 1º alvo ou Parâmetros do 2º alvo são ignoradas e não precisam ser apagadas.

Para configurar os parâmetros da inicialização iSCSI usando uma configuração dinâmica:

1. Na tela do menu **Parâmetros gerais**, defina os seguintes parâmetros:

- Parâmetros TCP/IP via DHCP – Ativados. (Para IPv4.)
- Autoconfiguração de IP – Ativada. (Para IPv6, sem descarregamento.)
- Parâmetros iSCSI via DHCP – Ativados
- Autenticação CHAP – Desativada
- ID do fornecedor de DHCP – BRCM ISAN
- Tempo de espera do link ativo – 0
- Usar carimbo de data/hora de TCP – Ativado (para alguns destinos, como Dell/EMC AX100i, é necessário ativar Usar carimbo de data/hora de TCP)
- Alvo como o primeiro HDD – Desativado
- Contagem de tentativas para LUN ocupado – 0
- Versão de IP – IPv6. (Para IPv6, sem descarregamento.)

2. Pressione **ESC** para retornar ao menu **Principal**.



Nota: As informações das telas Initiator Parameters e 1st Target Parameters são ignoradas e não precisam ser apagadas.

3. Selecione **Sair** e **Salvar configurações**.

Ativação da autenticação CHAP

Certifique-se de que a autenticação CHAP está ativada no destino.

Para ativar a autenticação CHAP:

1. Na tela **Parâmetros gerais**, defina **Autenticação CHAP** para **Ativada**.
2. Na tela **Parâmetros do programa de inicialização**, insira os parâmetros para:
 - ID de CHAP (até 128 bytes)
 - Segredo CHAP (usado quando a autenticação é obrigatória; deve ter no mínimo 12 caracteres)
3. Pressione **ESC** para retornar ao menu **Principal**.
4. No menu **Principal**, selecione os **Parâmetros do 1º alvo**.
5. Na tela **Parâmetros do 1º alvo**, digite os valores para os itens a seguir usando os valores usados na configuração do destino iSCSI:
 - ID de CHAP (opcional se CHAP for de duas vias)
 - Segredo CHAP (opcional se CHAP for de duas vias; deve ter no mínimo 12 caracteres)
6. Pressione **ESC** para retornar ao menu **Principal**.
7. Selecione **ESC**, em seguida, selecione **Sair e Salvar configuração**.

Configuração do servidor DHCP para suportar a inicialização iSCSI

O servidor DHCP é um componente opcional e é necessário apenas se você for realizar uma configuração de inicialização iSCSI dinâmica (consulte "[Configuração de inicialização iSCSI dinâmica](#)" na página 52).

A configuração do servidor DHCP para oferecer suporte à inicialização iSCSI é diferente para os protocolos IPv4 e IPv6. Consulte as seções a seguir:

Configurações de inicialização iSCSI do DHCP para IPv4

O protocolo DHCP contém várias opções que fornecem informações de configuração ao DHCP cliente. Para a inicialização iSCSI, os adaptadores Broadcom suportam as seguintes configurações de DHCP:

Opção 17 de DHCP, Caminho raiz

A opção 17 é utilizada para o transporte de informações da iSCSI-alvo à iSCSI cliente. O formato do caminho raiz, conforme definido pelo padrão IETF RFC 4173, é:

```
iscsi:"<nome_do_servidor>": "<protocolo>": "<porta>": "<LUN>": "<nome_do_destino>
```

Os parâmetros estão definidos na [Tabela 32](#).

Tabela 32: Definição de parâmetros para a opção 17 de DHCP

Parâmetro	Definição
"iscsi:"	Uma sequência literal.
<servername>	O endereço IP ou o FQDN da iSCSI-alvo
":"	Separador.

Tabela 32: Definição de parâmetros para a opção 17 de DHCP (Continuação)

Parâmetro	Definição
<protocol>	O protocolo IP usado para acessar a iSCSI-alvo. Atualmente, apenas o TCP é suportado, portanto, o valor que define o protocolo é 6.
<port>	O número de porta associado ao protocolo. O número de porta padrão para a iSCSI é 3260.
<LUN>	O Número da unidade lógica que será usado na iSCSI-alvo. É necessário que o valor do LUN seja representado no formato hexadecimal. Um LUN com uma ID OF 64 deveria ser configurado como 40 na opção do parâmetro 17 no servidor DHCP.
<targetname>	O nome do alvo em formato IQN ou EUI (consulte o padrão RFC 3720 para obter informações detalhadas sobre os formatos IQN e EUI). Um exemplo de nome IQN poderia ser "iqn.1995-05.com.broadcom:iscsi-target".

Opção 43 de DHCP, informações específicas do fornecedor

A opção 43 de DHCP (informações específicas do fornecedor) fornece mais opções de configuração para o cliente iSCSI do que a opção 17 de DHCP. Nessa configuração, são fornecidas três subopções adicionais que atribuem o IQN do programa de inicialização ao cliente de inicialização iSCSI, juntamente com dois IQNs da iSCSI-alvo que podem ser usados para a inicialização. O formato do IQN da iSCSI-alvo é igual ao formato da opção 17 de DHCP, enquanto o IQN do programa de inicialização iSCSI é composto apenas pelo IQN do programa de inicialização.



Nota: A opção 43 de DHCP tem suporte apenas para o protocolo IPv4.

Veja abaixo as subopções.

Tabela 33: Definição das subopções da opção 43 de DHCP

Subopções	Definição
201	Informações sobre a primeira iSCSI-alvo no formato padrão do caminho raiz iscsi:"<nome_do_servidor>":"<protocolo>":"<porta>":"<LUN>":"<nome_do_destino>
203	IQN do programa de inicialização iSCSI

A opção 43 de DHCP requer mais configurações do que a opção 17 de DHCP, mas ela oferece um ambiente mais completo e mais opções de configuração. A Broadcom recomenda que os clientes usem a opção 43 de DHCP ao realizarem configurações dinâmicas de inicialização iSCSI.

Configuração do servidor DHCP

Configure o servidor DHCP para suportar a opção 17 ou a opção 43.



Nota: Se a opção 43 for usada, configure a opção 60. O valor da opção 60 deve coincidir com o valor de ID do fornecedor de DHCP. O valor do ID do fornecedor de DHCP é BRCM ISAN, conforme descrito na opção Parâmetros gerais do menu Configuração de inicialização iSCSI.

Configuração de inicialização iSCSI do DHCP para IPv6

O servidor DHCPv6 é capaz de fornecer uma série de opções, incluindo a configuração IP com ou sem informações de estado, além de informações para o cliente DHCPv6. Para a inicialização iSCSI, os adaptadores Broadcom suportam as seguintes configurações de DHCP:



Nota: A opção de caminho raiz padrão do DHCPv6 ainda não está disponível. A Broadcom recomenda usar a opção 16 ou 17 para obter suporte dinâmico à inicialização iSCSI com IPv6.

Opção 16 de DHCPv6, opção de classe do fornecedor

A opção 16 de DHCPv6 (opção de classe do fornecedor) deve estar presente e conter uma string que coincida com o parâmetro de ID do fornecedor de DHCP configurado por você. O valor do ID do fornecedor de DHCP é BRCM ISAN, conforme descrito na opção Parâmetros gerais do menu Configuração de inicialização iSCSI.

O conteúdo da opção 16 deve ser <comprimento de 2 bytes> <ID do fornecedor de DHCP>.

Opção 17 de DHCPv6, informações específicas do fornecedor

A opção 17 de DHCPv6 (informações específicas do fornecedor) fornece mais opções de configuração para o cliente iSCSI. Nessa configuração, são fornecidas três subopções adicionais que atribuem o IQN do programa de inicialização ao cliente de inicialização iSCSI, juntamente com dois IQNs da iSCSI-alvo que podem ser usados para a inicialização.

As subopções estão listadas na [Tabela 34 na página 56](#).

Tabela 34: Definição das subopções da opção 17 de DHCP

Subopções	Definição
201	Informações sobre a primeira iSCSI-alvo no formato padrão do caminho raiz "iscsi:[<nome_do_servidor>]":"<protocolo>":"<porta>":"<LUN>":"<nome_do_destino>"
203	IQN do programa de inicialização iSCSI



Nota: No [Tabela 34](#), os colchetes [] são necessários para endereços IPv6.

O conteúdo da opção 17 deve ser <2-byte Option Number 201|202|203> <2-byte length> <data>.

Configuração do servidor DHCP

Configure o servidor DHCP para suportar a opção 16 ou a opção 17.



Nota: O formato das opções 16 e 17 de DHCPv6 são completamente definidas em RFC 3315.

VXLAN: Exemplos de configuração e de casos de uso

O encapsulamento de VXLAN permite que muitos hosts de 3 camadas residam em um servidor para enviar e receber quadros pelo encapsulamento no endereço IP único associado à placa NIC instalada no mesmo servidor.

Esse exemplo discute a conectividade VxLan básica entre dois servidores RHEL. Cada servidor tem uma NIC física ativada com um endereço IP externo definido para 1.1.1.4 e 1.1.1.2.

Uma interface VXLAN10 com ID de VXLAN 10 é criada com grupo multicast 239.0.0.10 e associada à porta da rede física pxp1 em cada servidor.

Um endereço IP para o host é criado em cada servidor e associado à interface VXLAN. Após a criação da interface VXLAN, o host presente no sistema 1 pode se comunicar com o host presente no sistema 2. O formato da VLXAN é mostrado na [Tabela 35](#).

Tabela 35: Formato do quadro de VXLAN

Cabeçalho MAC	Cabeçalho do IP externo com proto = UDP	Cabeçalho UDP com porta de destino = VXLAN	Cabeçalho de VXLAN (sinalizadores, VNI)	Quadro L2 original	FCS
---------------	---	--	---	--------------------	-----

[Tabela 36](#) fornece exemplos de comandos e configurações de VXLAN.

Tabela 36: Exemplos de comandos e configurações de VXLAN

Sistema 1	Sistema 2
PxPy: ifconfig PxPy 1.1.1.4/24	PxPy: ifconfig PxPy 1.1.1.2/24
ip link add vxlan10 type vxlan id 10 group 239.0.0.10 dev PxPy dstport 4789	ip link add vxlan10 type vxlan id 10 group 239.0.0.10 dev PxPy dstport 4789
ip addr add 192.168.1.5/24 broadcast 192.168.1.255 dev vxlan10	ip addr add 192.168.1.10/24 broadcast 192.168.1.255 dev vxlan10
ip link set vxlan10 up	ip link set vxlan10 up
ip -d link show vxlan10	
Ping 192.168.1.10	ifconfig vxlan10 (MTU 1450) (SUSE e RHEL)

Nota: x representa o número do barramento PCIe do adaptador físico encontrado no sistema. y representa o número da porta do adaptador físico.

SR-IOV: Exemplos de configuração e de casos de uso

SR-IOV pode ser configurado, ativado e usado em NICs de 10Gb e 25Gb Broadcom NetExtreme-E.

Caso de uso Linux

1. Ative SR-IOV nas placas NIC:
 - a. SR-IOV na placa NIC pode ser ativada usando o menu **HII**. Durante uma inicialização do sistema, acesse o **BIOS do sistema -> Configurações do dispositivo -> NetXtreme-E NIC -> Configuração de nível de dispositivo**.
 - b. Defina o modo Virtualização para SR-IOV.
 - c. Defina o número de funções virtuais por função física.
 - d. Defina o número de vetores MSI-X por VF e o Número máximo de vetores MSI-X de função física. Se a VF estiver sendo executada sem recursos, equilibre o número de vetores MSI-X por VM usando CCM.
2. Ative a virtualização no BIOS:
 - a. Durante a inicialização do sistema, entre no **BIOS do sistema -> Configurações do processador -> Tecnologias de virtualização** e configure como **Ativado**.
 - b. Durante a inicialização do sistema, entre no **BIOS do sistema -> Dispositivos integrados -> SR-IOV Global** e configure como **Ativado**.
3. Instale a versão de Linux desejada com Virtualização ativada (libvirt e Qemu).
4. Ative o parâmetro de kernel iommu.
 - a. O parâmetro kernel IOMMU é ativado editando `/etc/default/grub.cfg` e executando `grub2-mkconfig -o /boot/grub2/grub.cfg` para o modo legado. Para o modo UEFI, edite `/etc/default/grub.cfg` e execute `grub2-mkconfig -o /etc/grub2-efi.cfg`. Consulte o seguinte exemplo:
`Linuxefi /vmlinuz-3.10.0-229.el7.x86_64 root=/dev/mapper/rhel-root ro rd.lvm.lv=rhel/swap crashkernel=auto rd.lvm.lv=rhel/root rhgb intel_iommu=on quiet LANG=en_US.UTF.8`
5. Instale o driver `bnxt_en`:
 - a. Copie o driver `bnxt_en` para o SO e execute `make`; `make install`; `modprobe bnxt_en`.



Nota: Use `netxtreme-bnxt_en<version>.tar.gz` para instalar `bnxt_re` e `bnxt_en` para a funcionalidade RDMA nos VFs SRIOV.

6. Ativa funções virtuais pelos parâmetros de kernel:
 - a. Após a instalação do driver, `lspci` exibirá as NICs NetXtreme-E presentes no sistema. O barramento, o dispositivo e a função são necessários para ativar funções virtuais.
 - b. Para ativar as funções virtuais, insira o comando mostrado abaixo:

```
echo X >/sys/bus/pci/device/0000\:Bus\:Dev.Function/sriov_numvfs
```



Nota: Certifique-se de que as interfaces PF estejam ativas. Os VFs são criados somente se os PFs estiverem ativos. X é o número de VFs que serão exportadas para o SO.

Um exemplo típico seria:

```
echo 4 > /sys/bus/pci/devices/0000\:04\:00.0/sriov_numvfs
```

7. Verifique as funções virtuais de PCI-E:
 - a. O comando `lspci` exibirá as funções virtuais com DID configurado como 16D3 para BCM57402/BCM57404/BCM57406, 16DC para o BCM57412/BCM57414/BCM57416 sem RDMA e como 16C1 ou BCM57412/BCM57414/BCM57416 ativado para RDMA.
8. Use o Gerenciador virtual para instalar um sistema cliente virtualizado (VMs).

Consulte a documentação do Linux sobre a instalação do Gerenciador virtual. Verifique se o driver integrado ao hipervisor foi removido. Um exemplo seria NIC:d7:73:a7 rt18139. Remova o driver.
9. Atribua uma função virtual às VMs convidadas.
 - a. Atribua o adaptador a uma VM convidada como um dispositivo PCI físico. Consulte a documentação do Linux para obter informações sobre a atribuição de funções virtuais a uma VM convidada.
10. Instale os drivers `bnxt_en` nas VMs:
 - a. Nas VMs convidadas, copie o arquivo de origem `netxtreme-bnxt_en-<version>.tar.gz` e extraia o arquivo `tar.gz`. Altere o diretório para cada driver e execute `make; make install; modprobe bnxt_en` (e `bnxt_re` se estiver ativando o RDMA). Verifique se o driver é carregado corretamente, verificando a interface pelo comando `modinfo`. O usuário pode precisar executar `modprobe -r bnxt_en` para descarregar o módulo `bnxt_en` existente ou da `inbox` antes de carregar o último módulo criado.
11. Teste a conectividade da VM convidada com o mundo exterior:
 - a. Atribua o endereço IP adequado ao adaptador e teste a conectividade da rede.

Caso do Windows

1. Ative SR-IOV nas placas NIC:
 - a. SR-IOV na placa NIC pode ser ativada usando o menu HII. Durante uma inicialização do sistema, acesse o **BIOS do sistema -> Configurações do dispositivo -> NetXtreme-E NIC -> Configuração do nível do dispositivo**.
 - b. Defina o modo Virtualização para SR-IOV.
 - c. Defina o número de funções virtuais por função física.
 - d. Defina o número de vetores MSI-X por VF e o Número máximo de vetores MSI-X de função física. Se a VF estiver sendo executada sem recursos, equilibre o número de vetores MSI-X por VM usando CCM.
2. Ative a virtualização no BIOS:
 - a. Durante a inicialização do sistema, entre no **BIOS do sistema -> Configurações do processador -> Tecnologias de virtualização** e configure como **Ativado**.
 - b. Durante a inicialização do sistema, entre no **BIOS do sistema -> Dispositivos integrados -> SR-IOV Global** e configure como **Ativado**.
3. Instale a última atualização da KB para o SO Windows 2012 R2 ou Windows 2016.
4. Instale as opções apropriadas de virtualização (Hyper-V). Para obter detalhes adicionais e as etapas da configuração do Hyper-V, switch virtual e máquinas virtuais, visite o site Microsoft.com:
<https://technet.microsoft.com/en-us/windows-server-docs/compute/hyper-v/system-requirements-for-hyper-v-on-windows>

<https://technet.microsoft.com/en-us/windows-server-docs/compute/hyper-v/get-started/install-the-hyper-v-role-on-windows-server>

5. Instale o driver mais recente do NetXtreme-E no Hyper-V.
6. Ative o SR-IOV nas propriedades avançadas do driver da miniporta NDIS.
7. No Gerenciador Hyper-V, crie o seu switch virtual com a interface do NetXtreme-E selecionado.
8. Marque a caixa **Ativar virtualização de E/S de raiz única (SR-IOV)** ao criar o Adaptador virtual do Hyper-V.
9. Crie uma máquina virtual (VM) e adicione o número desejado de Adaptadores virtuais
10. Nas Configurações do adaptador de rede da máquina virtual para cada adaptador virtual, marque **Ativar SR-IOV** na seção **Aceleração do hardware**.
11. Inicie sua VM e instale o SO convidado desejado.
12. Instale o driver do NetXtreme-E correspondente para cada SO convidado.



Nota: O driver de Função virtual (VF) para o NetXtreme-E é mesmo que o driver de base. Por exemplo, se o SO convidado for o Windows 2012 R2, o usuário precisará instalar o Bnxtnd64.sys na VM. O usuário pode fazer isso usando o executável do Instalador de driver do NetXtreme-E. Uma vez que o driver foi instalado no SO convidado, o usuário pode ver a(s) interface(s) do driver do VF no Gerenciador de Dispositivos do SO convidado na VM.

Caso do VMWare SRIOV

1. Ative SR-IOV nas placas NIC:
 - a. SR-IOV na placa NIC pode ser ativada usando o menu HII. Durante uma inicialização do sistema, acesse o **BIOS do sistema -> Configurações do dispositivo -> NetXtreme-E NIC -> Configuração do nível do dispositivo**.
 - b. Defina o modo Virtualização para SR-IOV.
 - c. Defina o número de funções virtuais por função física.
 - d. Defina o número de vetores MSI-X por VF e o Número máximo de vetores MSI-X de função física. Se a VF estiver sendo executada sem recursos, equilibre o número de vetores MSI-X por VM usando CCM.
2. Ative a virtualização no BIOS:
 - a. Durante a inicialização do sistema, entre no **BIOS do sistema -> Configurações do processador -> Tecnologias de virtualização** e configure como **Ativado**.
 - b. Durante a inicialização do sistema, entre no **BIOS do sistema -> Dispositivos integrados -> SR-IOV Global** e configure como **Ativado**.
3. No ESXi, instale o driver Bnxtnet executando as seguintes etapas:
 - a. Copie o arquivo <bnxtnet>-<driver version>.vib no /var/log/vmware.


```
$ cd /var/log/vmware.
$ esxcli software vib install --no-sig-check -v <bnxtnet>-<driver version>.vib.
```
 - b. Reinicialize a máquina.
 - c. Verifique se os drivers foram instalados corretamente:

```
software vib list |grep bnx
```

4. Instale o utilitário BNXTNETCLI (esxcli bnxtnet) fornecido pela Broadcom para configurar/exibir os diversos parâmetros do driver que não são suportados nativamente no esxcli, como: velocidade de 25 G do link, mostrar informações sobre o driver/chip do firmware, mostrar a configuração do NIC (NPAR, SRIOV). Para obter mais informações, consulte o README.txt do driver do bnxtnet.

Para instalar esse utilitário:

- a. Copie o BCM-ESX-bnxtnetcli-<version>.vib in /var/log/vmware.

```
$ cd /var/log/vmware
$ esxcli software vib install --no-sig-check -v /BCM-ESX-bnxtnetcli-<version>.vib
```

- b. Reinicialize o sistema.

- c. Verifique se o vib está instalado corretamente:

```
$ esxcli software vib list | grep bcm-esx-bnxtnetcli
```

- d. Velocidade definida 10/20/25

```
$ esxcli bnxtnet link set -S <speed> -D <full> -n <iface>
```

Isso retornará a mensagem OK se a velocidade estiver configurada corretamente.

Por exemplo:

```
$ esxcli bnxtnet link set -S 25000 -D full -n vmnic5
```

- e. Mostrar as estatísticas do link

```
$ esxcli bnxtnet link get -n vmnic6
```

- f. Mostrar as informações do driver/firmware/chip

```
$ esxcli bnxtnet drvinfo get -n vmnic4
```

- g. Mostrar informações da NIC (por exemplo, BDF, NPAR, configuração SRIOV)

```
$ esxcli bnxtnet nic get -n vmnic4
```

5. Ativação de VFs SRIOV:

Somente os PFs são ativados automaticamente. Se um PF suportar SR-IOV, o PF (vmknicX) fará parte da saída do comando mostrado abaixo.

```
esxcli network sriovnic list
```

Para ativar um ou mais VFs, o driver usa o parâmetro de módulo "max_vfs" para ativar o número desejado de VFs para PFs. Por exemplo, para ativar quatro VFs no PF1:

```
esxcfg-module -s 'max_vfs=4' bnxtnet (reboot required)
3
```

Para ativar VFs em um conjunto do PFs, use o formato do comando mostrado abaixo. Por exemplo, para ativar 4 VFs no PF 0 e 2 VFs no PF 2:

```
esxcfg-module -s 'max_vfs=4,2' bnxtnet (reboot required)
```

Os VFs exigidos de cada PF suportado são ativados na ordem durante a exibição do PF. Consulte a documentação da VMware para obter mais informações sobre como mapear uma VF para uma VM:

<https://pubs.vmware.com/vsphere-60/index.jsp#com.vmware.vsphere.networking.doc/GUID-EE03DC6F-32CA-42EF-98FC-12FDE06C0BE0.html>

<https://pubs.vmware.com/vsphere-60/index.jsp#com.vmware.vsphere.networking.doc/GUID-CC021803-30EA-444D-BCBE-618E0D836B9F.html>

NPAR – Exemplo de configuração e de caso de uso

Recursos e requisitos

- SO/BIOS desconhecido – As partições são apresentadas ao sistema operacional como interfaces de rede "reais", portanto nenhum suporte de BIOS ou SO especial é exigido, como o SR-IOV.
- NICs adicionais sem exigir portas de switch adicionais, cabeamento, slots de expansão PCIe.
- Modelagem do tráfego – A alocação de largura de banda por partição pode ser controlada para limitar ou reservar conforme necessário.
- Pode ser usado de uma maneira independente do switch. O switch não precisa de nenhuma configuração especial ou conhecimento sobre a ativação de NPAR.
- Pode ser usado em conjunto com RoCE e SR-IOV.
- Suporta descarregamentos sem estado como LSO, TPA, RSS/TSS e RoCE (apenas dois PFs por porta).
- Suporte de ID de Roteamento Alternativo para mais de oito funções por dispositivo físico.



Nota: Na página **Menu Dell UEFI HII > Configuração principal > Configuração do nível do dispositivo**, o usuário pode ativar o **NParEP** para permitir que o adaptador NXE suporte até 16 PFs por dispositivo, em um sistema compatível com ARI. Para um dispositivo de 2 portas, isso significa até 8 PFs para cada porta.

Limitações

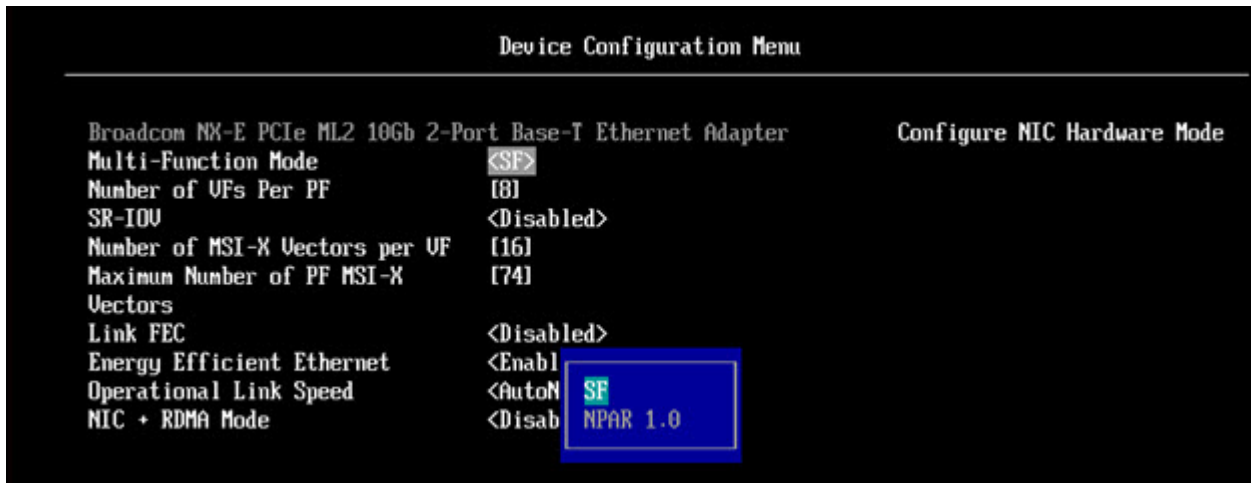
- As configurações compartilhadas devem ser suprimidas para evitar conflitos. Por exemplo: As configurações de velocidade, duplex, controle de fluxo e outras físicas semelhantes são ocultadas pelo driver de dispositivo para evitar a contenção.
- Os sistemas não ARI permitem apenas oito partições por dispositivo físico.
- O RoCE é suportado apenas nas duas primeiras partições de cada porta física, ou um total de quatro partições por dispositivo físico. No modo NPAR + SRIOV, apenas duas VFs de cada porta física principal podem ativar o suporte RDMA, ou o total de quatro VFs + RDMA por dispositivo físico.

Configuração

O NPAR pode ser configurado usando os menus HII de configuração do BIOS ou o utilitário Broadcom CCM nos sistemas de inicialização legados. Alguns fornecedores também expõem a configuração em interfaces proprietárias adicionais.

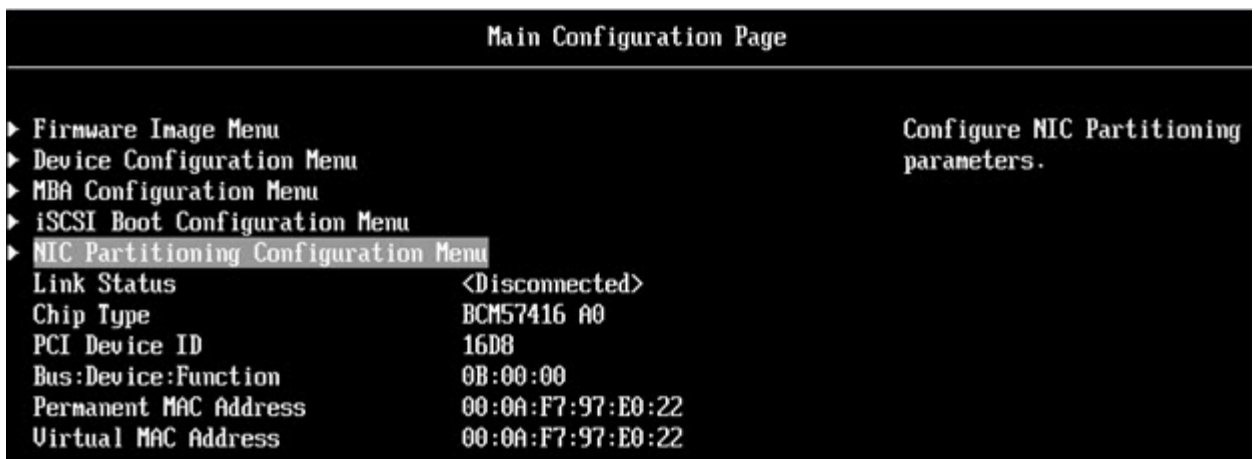
Para ativar o NPAR:

1. Selecione a NIC de destino no menu BIOS HII ou da interface de CCM e configure a opção Modo multifunções ou Modo de virtualização.

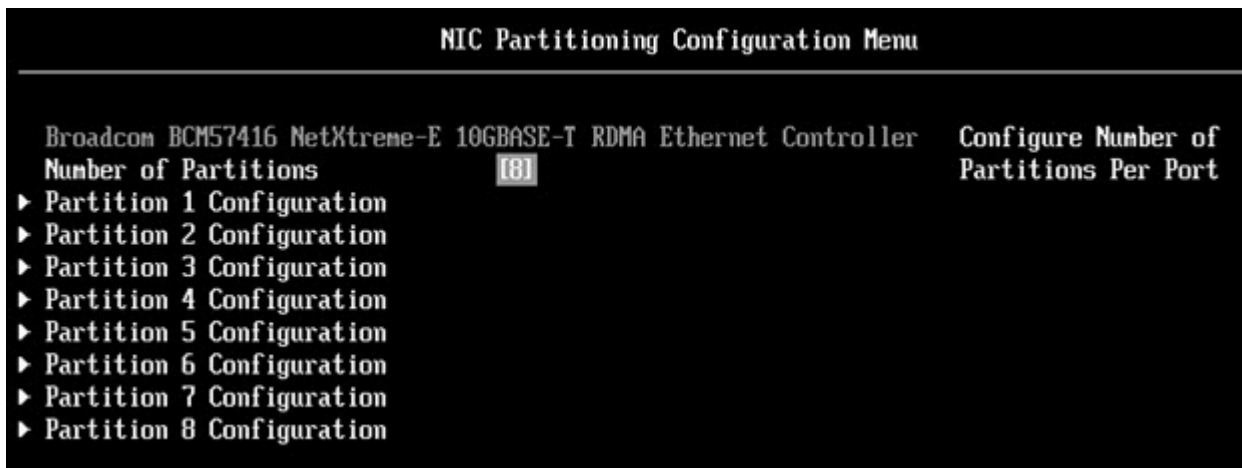


O NPAR é ativado em combinação com o SR-IOV. Para alguns sistemas OEM compatíveis com ARI, o botão **NParEP** está disponível para permitir explicitamente que o BCM5741X suporte até 16 partições. Ao trocar o modo Função única para o modo Multifunções, o dispositivo precisará ser enumerado novamente, portanto, as alterações não terão efeito até que o sistema seja reiniciado.

2. Uma vez ativado o NPAR, a opção Menu de configuração principal da partição NIC estará disponível no Menu de configuração da NIC associada a cada porta física.



3. O Menu de configuração principal da partição NIC (mostrado abaixo) permite que o usuário escolha o número de partições que devem ser alocadas a partir da porta física selecionada. Cada NIC BCM5741X pode suportar um máximo de 16 partições em um servidor compatível com ARI. Por padrão, os adaptadores de porta dupla são configurados para oito partições por porta física. As opções de configuração para cada partição também são acessíveis nesse menu. Para alguns sistemas OEM, o menu HII também inclui uma página Alocação de largura de banda global, em que a largura de banda TX mínima (reservada) e máxima (limite) para todas as partições pode ser configurada simultaneamente.



4. Configure os parâmetros de Configuração da partição NIC (consulte [Tabela 37 na página 64](#)).

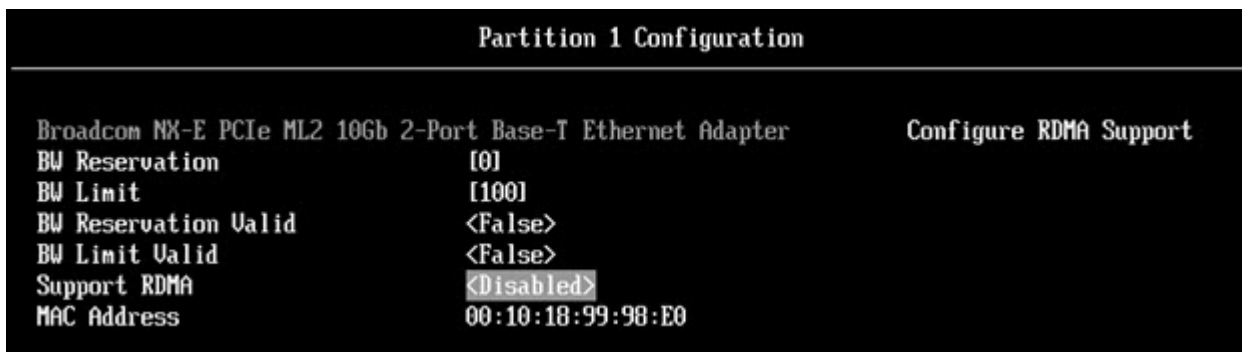


Tabela 37: Parâmetros NPAR

Parâmetro	Descrição	Opções válidas
Reserva de BW	Porcentagem do total de largura de banda disponível que deve ser reservada para esta partição. 0 indica uma divisão igual da largura de banda entre todas as partições.	Valor 0-100
Limite de BW	Porcentagem máxima da largura de banda disponível permitida para essa partição.	Valor 0-100
Reserva de BW válida	Funciona como um botão liga/desliga para a configuração da Reserva de BW.	Verdadeiro/ Falso
Limite de BW válida	Funciona como um botão liga/desliga para a configuração do Limite de BW.	Verdadeiro/ Falso

Tabela 37: Parâmetros NPAR (Continuação)

Parâmetro	Descrição	Opções válidas
Suporte ao RDMA	Funciona como um botão liga/desliga para o suporte ao RDMA nessa partição. Observe que somente as duas partições por porta física podem suportar o RDMA. Para um dispositivo de porta dupla, até 4 partições NPAR podem suportar o RDMA.	Ativado/ Desativado
Endereço MAC	Endereço MAC para essa partição.	–

Observações sobre a redução do consumo de memória NIC

Devido às velocidades de link mais rápidas suportadas neste NIC, o número padrão de buffers de recebimento é maior. Mais pacotes podem chegar dentro de um determinado intervalo de tempo quando a velocidade do link é mais alta e, se o sistema host estiver atrasado no processamento das interrupções de recebimento, o NIC deverá descartar pacotes se todos os buffers de recebimento disponíveis estiverem em uso.

O valor padrão de buffers de recebimento foi selecionado para funcionar bem nas configurações típicas. Porém, se você tem muitas NICs em um sistema, tiver habilitado o NPAR em várias NICs ou tiver apenas uma pequena quantidade de RAM, você poderá ver um Código 12 de "explosão em amarelo" no Gerenciador de Dispositivos para algumas NICs. O Código 12 significa que o driver falhou ao carregar porque não havia recursos suficientes. Nesse caso, o recurso é um tipo específico de memória do kernel chamada Reserva de memória não paginável (NPP).

Se você estiver recebendo um Código 12 ou, por outras razões, deseja reduzir a quantidade de memória NPP consumida pelo NIC:

- Reduza o número de filas RSS do padrão de 8 para 4 ou 2. Cada fila RSS tem o seu próprio conjunto de buffers de recebimento alocado, portanto a redução do número de filas RSS reduz a memória NPP alocada. Pode haver implicações no desempenho com a redução do número de filas RSS, pois menos núcleos participam do processamento dos pacotes de recebimento do NIC. O uso de CPU por processador deve ser monitorado para garantir que não haja processadores "quentes" após essa alteração.
- Reduza a alocação de memória, reduzindo o número de buffers de recebimento alocados. O valor padrão de 0 significa que o driver deve determinar automaticamente o número de buffers de recebimento. Para configurações típicas, uma configuração 0 (= auto) mapeará para XXXX buffers de recebimento por fila. Você pode escolher um valor menor, como 1500, 1000 ou 500. (O valor deve ser múltiplos de 500 entre o intervalo de 500 e 15000.) Conforme mencionado acima, um menor número de buffers de recebimento aumenta o risco da exclusão de pacote, um impacto correspondente nas retransmissões de pacotes e diminuição da produtividade.

Os parâmetros "Número máximo de filas RSS" e "Buffers de recebimento (0 = Automático)" podem ser modificados usando a guia de propriedades **Avançadas** para cada NIC no **Gerenciador de dispositivos**. Se você quiser modificar múltiplos NICs ao mesmo tempo, é mais rápido usar o *cmdlet Set-NetAdapterAdvancedProperty PowerShell*. Por exemplo, para atribuir duas filas RSS para todos os NICs em um sistema cujo nome do NIC comece com "SI", execute o seguinte comando:

```
Set-NetAdapterAdvancedProperty S1* -RegistryKeyword *NumRSSQueues -RegistryValue 2
```

Da mesma forma, para configurar o número de buffers de recebimento para 1500, execute o seguinte comando:

```
Set-NetAdapterAdvancedProperty S1* -RegistryKeyword *ReceiveBuffers -RegistryValue 1500
```

Consulte <https://blogs.technet.microsoft.com/wincat/2012/08/27/using-powershell-for-nic-configuration-tasks/> para obter uma visão geral do uso do PowerShell para modificar as propriedades do NIC.

RoCE – Exemplos de configuração e de casos de uso

Esta seção fornece exemplos de configuração e casos de uso para o RoCE.

Para ativar o RoCE para PFs ou VFs, o usuário deve ativar a seleção de RDMA no menu HII no BIOS antes que a opção RDMA entre em vigor no SO host ou convidado.

Para ativar o RDMA no modo de função única (se o **Modo de virtualização** for **Nenhum** ou **SR-IOV**):

1. Durante a inicialização do sistema, acesse **Configuração do sistema -> Configurações do dispositivo -> NetXtreme-E NIC -> Página de configuração principal** e configure o **Modo NIC+ RMDA** como **Ativado**.

Para ativar o RDMA se o Modo de virtualização for NPAR ou NPAR+SR-IOV:

1. Durante a inicialização do sistema, acesse **Configuração do sistema -> Configurações do dispositivo -> NetXtreme-E NIC -> Configuração de particionamento da NIC -> Configuração da partição 1 (ou 2)** e configure o **Modo NIC+ RMDA** como **Ativado**.



Nota: Se estiver usando o modo NPAR+SRIOV, apenas duas VFs de cada porta física principal podem ativar o suporte RDMA, ou um total de quatro VFs+RDMA por dispositivo físico.

Configuração do Linux

Requisitos

Para configurar o RoCE no Linux, os seguintes itens são exigidos:

- `bnxt_en-roce` (driver de `bnxt_en` suportado pelo RoCE que faz parte do arquivo tar compactado `gzip`)
- `bnxt_re` (driver do RoCE)
- `libbnxtre` (módulo de biblioteca do RoCE no modo de usuário)

Dependências do driver `BNXT_RE`

O driver `Bnxt_re` requer uma versão especial do `bnxt_en` ativada para RoCE, que está incluída no pacote `netxtreme-bnxt_en-1.7.9.tar.gz` (ou mais recente). A compilação do driver `bnxt_re` depende de a pilha IB estar disponível junto com a distribuição do SO ou se um OFED externo é necessário.



Nota: É necessário carregar a versão correta do `bnxt_en`, incluída no mesmo pacote `netxtreme-bnxt_en-1.7.x.tar.gz`. `Bnxt_re` e `Bnxt_en` funcionam como um par para ativar o tráfego RoCE. O uso de versões incompatíveis desses dois drivers produz resultados não confiáveis ou imprevisíveis.

- Para distribuições que tiverem uma pilha IB disponível junto com a distribuição do SO:
RH7.1/7.2/7.3/6.7/6.8, SLES12SP2 e Ubuntu 16.04

Se ainda não estiver instalada, a pilha IB e os utilitários úteis podem ser instalados no Redhat com os seguintes comandos antes de compilar o `bnxt_re`:

```
yum -y install libibverbs* infiniband-diag perftest qperf librdmacm utils
```

Para compilar o `bnxt_re`:

```
$make
```

- Para distribuições que exigem que o OFED externo seja instalado:

SLES11SP4

Consulte as notas de versão do OFED no link a seguir e instale o OFED antes de compilar o driver `bnxt_re`.

http://downloads.openfabrics.org/downloads/OFED/release_notes/OFED_3.18-2_release_notes

Para compilar o `bnxt_re`:

```
$export OFED_VERSION=OFED-3.18-2  
$make
```

Instalação

Para instalar o RoCE no Linux:

1. Atualize o NIC NVRAM usando os pacotes de firmware suportados pelo RoCE a partir da versão do Software 20.06.04.01 ou posterior.
2. No SO, descompacte, compile e instale os drivers BCM5741X Linux L2 e RoCE.
 - a. # tar -xzf netxtreme-bnxt_en-1.7.9.tar.gz
 - b. # cd netxtreme-bnxt_en-bnxt_re
 - c. # make build && make install
3. Descompacte, compile e instale o NetXtreme-E Linux RoCE User Library.
 - a. # tar xzf libbnxtre-0.0.18.tar.gz
 - b. #cd libbnxtre-0.0.18
 - c. # configure && make && make install.
 - d. # cp bnxtre.driver /etc/libibverbs.d/
 - e. # echo "/usr/local/lib" >> /etc/ld.so.conf
 - f. # ldconfig -v

Consulte o README.txt do `bnxt_re` para obter mais detalhes sobre as opções configuráveis e recomendações.

Limitações

Em NICs de porta dupla, se ambas as portas estiverem na mesma sub-rede, os comandos `rdma perftest` poderão falhar. A possível causa é decorrente de um problema de fluxo arp no SO Linux. Para solucionar essa limitação, use várias sub-redes para testar ou desativar a segunda porta/interface.

Problemas conhecidos

Bnxt_en e Bnxt_re foram projetados para funcionar em par. Os drivers Bnxt_en mais antigos, antes da versão 1.7.x, não suportam RDMA e não podem ser carregados ao mesmo tempo que o driver Bnxt_re (RDMA). O usuário poderá enfrentar uma falha do sistema e uma reinicialização se o Bnxt_re for carregado com drivers Bnxt_en mais antigos. Recomenda-se que o usuário carregue o módulo Bnxt_en e Bnxt_re do mesmo pacote netxtreme-bnxt_en-<1.7.x>.tar.gz.

Para evitar que uma combinação incompatível de bnxt_en e bnxt_re seja carregada, o seguinte é exigido:

- Se o SO RedHat/CentOS 7.2 foi instalado no sistema de destino usando o PXEboot com bnxt_en DUD ou um RPM de módulo do kernel, exclua o arquivo bnxt_en.ko encontrado em /lib/modules/\$(uname -r)/extra/bnxt_en/bnxt_en.ko ou edite /etc/depmo.d/.
- bnxt_en.conf para substituir e usar a versão atualizada. Os usuários também podem apagar o atual driver do kernel Linux BCM5741X usando o comando rpm -e kmod-bnxt_en. RHEL 7.3/SLES 12 Sp2 tem o driver da inbox bnxt_en (anterior ao v1.7.x). Esse driver deve ser removido, e o bnxt_en mais recente, adicionado antes de aplicar o bnxt_re (drivers RoCE).

Windows

Modo do kernel

O Windows Server 2012 e posterior invocará o recurso RDMA no NIC o para tráfego de arquivos SMB se ambas as extremidades forem ativadas para o RDMA. A miniporta NDIS bnxnd.sys v20.6.2 do Broadcom e posterior suporta o RoCEv1 e RoCEv2 via interface NDKPI. A configuração padrão é RoCEv1.

Para ativar o RDMA:

1. Atualize o NIC NVRAM usando os pacotes de placa apropriados. No CCM ou UEFI HII, ative o suporte para RDMA.
2. Vá para a página **Propriedades avançadas** do adaptador e configure **Funcionalidade do NetworkDirect** como **Ativada** para cada miniporta BCM5741X ou, usando a janela do PowerShell, execute o seguinte comando:

```
Set-NetAdapterAdvancedProperty -RegistryKeyword *NetworkDirect -RegistryValue 1
```

3. Os seguintes comandos do PowerShell retornarão true se **NetworkDirect** for ativado.
 - a. Get-NetOffLoadGlobalSetting
 - b. Get-NetAdapterRDMA

Verificação do RDMA

Para verificar o RDMA:

1. Crie um compartilhamento de arquivos no sistema remoto e abra esse compartilhamento usando o Windows Explorer ou "uso de rede...". Para evitar o gargalo de velocidade de leitura/gravação do disco rígido, um disco RAM é recomendado como o compartilhamento de rede em teste.

2. No PowerShell, execute os seguintes comandos:

```
Get-SmbMultichannelConnection | fl *RDMA*  
ClientRdmaCapable : Verdadeiro  
ServerRdmaCapable : Verdadeiro
```

Se o Cliente e o Servidor mostrarem True, então qualquer transferência de arquivo por essa conexão SMB usará o SMB.

3. Os seguintes comandos podem ser usados para ativar/desativar o SMB Multicanais:

Lado do servidor:

- Ative: Set-SmbServerConfiguration -EnableMultiChannel \$true
- Desative: Set-SmbServerConfiguration -EnableMultiChannel \$false

Lado do cliente:

- Ative: Set-SmbClientConfiguration -EnableMultiChannel \$true
- Desative: Set-SmbClientConfiguration -EnableMultiChannel \$false



Nota: Por padrão, o driver configura duas conexões RDMA para cada compartilhamento de rede por endereço IP (em uma sub-rede exclusiva). O usuário pode escalonar o número de conexões RDMA adicionando vários endereços IP, cada qual com uma sub-rede diferente, para a mesma porta física em teste. Vários compartilhamentos de rede podem ser criados e mapeados para cada parceiro de link usando os endereços IP exclusivos criados.

Por exemplo:

```
On Server 1, create the following IP addresses for Network Port1.  
172.1.10.1  
172.2.10.2  
172.3.10.3
```

```
On the same Server 1, create 3 shares.  
Share1  
Share2  
Share3
```

```
On the network link partners,  
Connect to \\172.1.10.1\share1  
Connect to \\172.2.10.2\share2  
Connect to \\172.3.10.3\share3  
...etc
```

Modo do usuário

Antes de você executar um aplicativo gravado no modo do usuário NDSPI, copie e instale o driver do modo do usuário bxndspi.dll. Para copiar e instalar o driver do modo do usuário:

1. Copie bxndspi.dll para C:\Windows\System32.
2. Instale o driver executando o seguinte comando:

```
rundll32.exe .\bxndspi.dll,Config install|more
```

VMware ESX

Limitações

A versão atual do driver suportado pelo RoCE requer o ESXi-6.5.0 GA build 4564106 ou acima.

Requisitos do driver BNXT RoCE

O driver BNXTNET L2 deve ser instalado com o parâmetro de módulo `disable_roce=0` antes de instalar o driver.

Para configurar o parâmetro do módulo, execute o seguinte comando:

```
esxcfg-module -s "disable_roce=0" bnxtnet
```

Use a versão do driver ESX6.5 20.6.9.0 L2 (driver L2 suportado pelo RoCE) ou acima.

Instalação

Para instalar o driver RoCE:

1. Copie o arquivo `<bnxtroce>-<driver version>.vib` no `/var/log/vmware` usando os seguintes comandos:

```
$ cd /var/log/vmware
$ esxcli software vib install --no-sig-check -v <bnxtroce>-<driver version>.vib
```

2. Reinicialize a máquina.
3. Verifique se os drivers estão instalados corretamente, usando o seguinte comando:

```
esxcli software vib list | grep bnxtroce
```

4. Para desativar o ECN (ativado por padrão) para o tráfego RoCE, use parâmetro de módulo `"tos_ecn=0"` para o `bnxtroce`.

Configuração de adaptadores de rede RDMA paravirtualizados

Consulte o link do vmware abaixo para obter informações adicionais sobre como configurar adaptadores de rede RDMA paravirtualizados (PVRDMA).

<https://pubs.vmware.com/vsphere-65/index.jsp#com.vmware.vsphere.networking.doc/GUID-4A5EBD44-FB1E-4A83-BB47-BBC65181E1C2.html>

Configuração de um Virtual Center para o PVRDMA

Para configurar um Virtual Center para o PVRDMA:

1. Crie o DVS (requer um switch virtual distribuído para o PVRDMA)
2. Adicione o host ao DVS.

Marcação do vmknic para PVRDMA em hosts ESX

Para marcar um vmknic para PVRDMA para usar em hosts ESX:

1. Selecione o host e clique com o botão direito em **Configurações** para mudar para a página de configurações das guias **Gerenciar**.
2. Na página **Configurações**, expanda **Sistema** e clique em **Configurações avançadas do sistema** para exibir o valor de par de chaves das Configurações avançadas do sistema e seu resumo.
3. Clique em **Editar** para exibir **Editar configurações avançadas do sistema**.
Filtre o **PVRDMA** para reduzir todas as configurações para apenas `Net.PVRDMAvmknic`.
4. Configure o valor `Net.PVRDMAvmknic` como `vmknic`, como no exemplo `vmk0`

Configuração da regra de firewall para o PVRDMA

Para configurar a regra de firewall para o PVRDMA:

1. Selecione o host e clique com o botão direito em **Configurações** para mudar para a página de configurações das guias **Gerenciar**.
2. Na página **Configurações**, expanda **Sistema** e clique em **Perfil de segurança** para exibir o resumo do firewall.
3. Clique em **Editar** para exibir **Editar perfil de segurança**.
4. Role para baixo até encontrar o `pvrDMA` e marque a caixa para configurar o firewall.

Adição de um dispositivo PVRDMA ao VM

Para adicionar um dispositivo PVRDMA ao VM:

1. Selecione o VM e clique em **Editar configurações**.
2. Adicione um novo Adaptador de rede.
3. Selecione a rede como um **Switch virtual distribuído** e o **Grupo de portas**.
4. Em **Tipo de adaptador**, selecione **PVRDMA** e clique em **OK**.

Configuração do VM no SO convidado do Linux



Nota: O usuário deve instalar as ferramentas de desenvolvimento apropriadas, incluindo o git, antes de prosseguir com as etapas de configuração abaixo.

1. Faça o download do driver PVRDMA e da biblioteca usando os seguintes comandos:

```
git clone git://git.openfabrics.org/~aditr/pvrdma_driver.git
git clone git://git.openfabrics.org/~aditr/libpvrdma.git
```

2. Compile e instale o driver convidado PVRDMA e a biblioteca.
3. Para instalar o driver, execute `make && sudo insmod pvrdma.ko` no diretório do driver.
O driver deve ser carregado depois que o driver vmxnet3 emparelhado for carregado.



Os módulos de kernel RDMA instalados podem não ser compatíveis com o driver PVRDMA. Nesse caso, remova a instalação atual e reinicie. Em seguida, siga as instruções de instalação. Leia o README no diretório do driver para obter mais informações sobre as diferentes pilhas RDMA.

4. Para instalar a biblioteca, execute `./autogen.sh && ./configure --sysconfdir=/etc && make && sudo make install` no diretório da biblioteca.



Nota: O caminho de instalação da biblioteca precisa estar no cache da biblioteca compartilhada. Siga as instruções no arquivo INSTALL no diretório da biblioteca.



Nota: Pode ser necessário modificar as configurações do firewall para permitir o tráfego RDMA. Verifique se as configurações do firewall estão adequadas

5. Adicione o arquivo `/usr/lib` in the `/etc/ld.so.conf` e recarregue o `ldconf` by executando `ldconfig`
6. Carregue os módulos `ib` usando `modprobe rdma_ucm`.
7. Carregue o módulo kernel PVRDMA usando `insmod pvrdma.ko`.
8. Atribua um endereço IP à interface PVRDMA.
9. Verifique se o dispositivo é criado pelo IB executando o comando `ibv_devinfo -v`.

DCBX – Data Center Bridging

Os controladores Broadcom NetXtreme-E suportam o IEEE802.1Qaz DCBX, bem como a especificação CEE DCBX mais antiga. A configuração DCB é obtida trocando as configurações efetuadas localmente pelo peer do link. Como as duas extremidades de um link podem ser configuradas de maneira diferente, o DCBX usa um conceito de "disposição" para indicar qual extremidade do link está pronta para aceitar parâmetros da outra extremidade. Isso é indicado no protocolo DCBX usando um único bit na Configuração ETS e PFC TLV; esse bit não é usado com a Recomendação ETS e a Prioridade do aplicativo TLV. Por padrão, o NetXtreme-E NIC está no modo "disposto", enquanto o switch de rede do parceiro de link está no modo "não disposto". Isso garante que a mesma configuração de DCBX no Switch se propague por toda a rede.

Os usuários podem configurar manualmente o NetXtreme-E NIC para o modo não disposto e executar diversas configurações de PFC, Prioridade estrita, ETS e APP do lado do host. Consulte o readme.txt do driver para obter mais detalhes sobre as configurações disponíveis. Esse documento fornecerá um exemplo de como essa configuração pode ser feita no Windows com o Windows PowerShell. Informações adicionais sobre DCBX, QoS e casos de uso associados são descritas em mais detalhes em um documento oficial separado, além do escopo deste manual do usuário.

As configurações a seguir no menu UEFI HII são necessárias para ativar o suporte DCBX:

Configuração do sistema -> Configurações do dispositivo -> NetXtreme-E NIC -> Configuração do nível do dispositivo

Perfil QoS – Perfil da fila QoS padrão

A configuração dos recursos de Qualidade do servidor (QoS) é necessária para suportar vários requisitos de PFC e ETS em que é necessário um ajuste mais fino, além da alocação de largura de banda. O NetXtreme-E permite que o administrador escolha entre dedicar os recursos de hardware do NIC para dar suporte aos quadros Jumbo e/ou combinações de filas de Classe de serviço (CoS) com e sem perda. Muitas combinações de configuração são possíveis e, portanto, podem ser complicadas de computar. Essa opção permite que o usuário selecione em uma lista de Perfis de fila QoS pré-computados. Esses perfis pré-computados foram projetados para otimizar o suporte para requisitos de PFC e ETS nas implantações do cliente.

A seguir está uma descrição resumida para cada Perfil QoS.

Tabela 38: Perfis QoS

Nº do perfil	Suporte ao quadro Jumbo	Nº de filas/porta CoS com perda	Nº de filas/porta CoS sem perda	Suporte para SKU de 2 portas
Perfil nº 1	Sim	0	1 (PFC suportado)	Sim (25 Gbps)
Perfil nº 2	Sim	4	2 (PFC suportado)	Não
Perfil nº 3	Não. (MTU <= 2 KB)	6	2 (PFC suportado)	Sim (25 Gbps)
Perfil nº 4	Sim	1	2 (PFC suportado)	Sim (25 Gbps)
Perfil nº 5	Sim	1	0 (Não há suporte a PFC)	Sim (25 Gbps)
Perfil nº 6	Sim	8	0 (Não há suporte a PFC)	Sim (25 Gbps)

Tabela 38: Perfis QoS (Continuação)

Nº do perfil	Suporte ao quadro Jumbo	Nº de filas/porta CoS com perda	Nº de filas/porta CoS sem perda	Suporte para SKU de 2 portas
Perfil nº 7	Essa configuração maximiza alocações de buffer de pacote para duas filas CoS sem perda para maximizar o desempenho do RoCE ao negociar a flexibilidade.			
	Sim	0	2	Sim (25 Gbps)
Padrão	Sim	Mesmo que o Perfil nº 4		Sim

Modo DCBX = Ativar (IEEE apenas)

Essa opção permite que um usuário ative/desative o DCBX com a especificação indicada. O IEEE indica apenas que o IEEE802.1Qaz DCBX está selecionado.

Configuração do driver do Windows:

Depois de ativar as opções indicadas no menu UEFI HII para fazer as configurações do nível do firmware, execute a seguinte seleção nas propriedades avançadas do driver do Windows.

Abra **Gerenciador do Dispositivo Windows -> Adaptador do Broadcom NetXtreme Série E -> Propriedades avançadas -> guia Avançado**

Qualidade do serviço = Ativada

Prioridade e VLAN = Prioridade e VLAN ativadas

VLAN = <ID>

Configure o ID desejado da VLAN

Para exercer o comando relacionado ao DCB no Windows PowerShell, instale o recurso DCB Windows apropriado.

1. Na **Barra de tarefas**, clique com o botão direito no ícone do Windows PowerShell e clique em **Executar como administrador**. O Windows PowerShell é aberto no modo elevado.
2. No console do Windows PowerShell, digite:

```
Install-WindowsFeature "data-center-bridging"
```

Bit de disposição do DCBX

O bit de disposição do DCBX é mencionado na especificação do DCB. Se o bit de Disposição em um dispositivo for true, o dispositivo estará disposto a aceitar configurações de um dispositivo remoto via DCBX. Se o bit de Disposição em um dispositivo for false, o dispositivo rejeitará qualquer configuração de um dispositivo remoto e imporá apenas as configurações locais.

Use o seguinte para configurar o bit de Disposição como True ou False. 1 para ativado, 0 para desativado.

```
Exemplo set-netQoSdcbxSetting -Willing 1
```

Use o seguinte para criar uma Classe de tráfego.

```
C:\> New-NetQoSTrafficClass -name "SMB class" -priority 4 -bandwidthPercentage 30 -Algorithm ETS
```



Nota: Por padrão, todos os valores 802.1p são mapeados para uma classe de tráfego padrão, que tem 100% da largura de banda do link físico. O comando mostrado acima cria uma nova classe de tráfego, para a qual qualquer pacote marcado com o valor IEEE 802.1p valor 4 é mapeado, e seu Algoritmo de seleção de transmissão (TSA) é ETS e tem 30% da largura de banda.

É possível criar até sete novas classes de tráfego. Além da classe de tráfego padrão, existem no máximo oito classes de tráfego no sistema.

Use o seguinte na exibição da Classe de tráfego criada:

```
C:\> Get-NetQoSTrafficClass
Name           Algorithm Bandwidth(%) Priority
-----
[Default]      ETS        70          0-3,5-7
SMB class      ETS        30          4
```

Use o seguinte na modificação da Classe de tráfego:

```
PS C:\> Set-NetQoSTrafficClass -Name "SMB class" -BandwidthPercentage 40
PS C:\> get-NetQoSTrafficClass
Name Algorithm Bandwidth(%) Priority
-----
[Default] ETS        60 0-3,5-7
SMB class ETS        40    4
```

Use o seguinte para remover a Classe de tráfego:

```
PS C:\> Remove-NetQoSTrafficClass -Name "SMB class"
PS C:\> Get-NetQoSTrafficClass
Name Algorithm Bandwidth(%) Priority
-----
[Default] ETS        100         0-7
```

Use o seguinte para criar a Classe de tráfego (Prioridade estrita):

```
C:\> New-NetQoSTrafficClass -name "SMB class" -priority 4 -bandwidthPercentage 30-Algorithm
Strict
```

Ativação do PFC:

```
PS C:\> Enable-NetQoSFlowControl -priority 4
PS C:\> Get-NetQoSFlowControl -priority 4
Priority Ativado
-----
4 True

PS C:\> Get-NetQoSFlowControl
```

Desativação do PFC:

```
PS C:\> disable-NetQosFlowControl -priority 4
PS C:\> get-NetQosFlowControl -priority 4
Priority Ativado
-----
4 False
```

Use o seguinte para criar uma Política QoS:

```
PS C:\> New-NetQosPolicy -Name "SMB policy" -SMB -PriorityValue8021Action 4

Name : SMB policy
Owner : Group Policy (Machine)
NetworkProfile : Todos
Precedence : 127
```



Nota: O comando acima criará uma nova política para SMB. –SMB é um filtro de inbox que corresponde à porta TCP 445 (reservada para SMB). Se um pacote for enviado para a porta TCP 445, ele será marcado pelo sistema operacional com o valor 802.1p de 4 antes que o pacote seja transmitido para um driver de miniporta de rede.

Além do –SMB, outros filtros padrão incluem –iSCSI (porta TCP correspondente 3260), –NFS (porta TCP correspondente 2049), –LiveMigration (porta TCP correspondente 6600), –FCOE (EtherType correspondente 0x8906) e –NetworkDirect.

NetworkDirect é uma camada abstrata que criamos sobre qualquer implementação RDMA em um adaptador de rede. O –NetworkDirect deve ser seguido por uma porta Direta de rede.

Além dos filtros padrão, um usuário pode classificar o tráfego pelo nome do executável do aplicativo (como no primeiro exemplo abaixo) ou pelo endereço IP, porta ou protocolo.

Use o seguinte para criar a Política QoS com base no Endereço de origem/destino:

```
PS C:\> New-NetQosPolicy "Network Management" -IPDstPrefixMatchCondition 10.240.1.0/24 -
IPProtocolMatchCondition both -NetworkProfile all -PriorityValue8021Action 7
Name : Network Management
Owner : Group Policy (Machine)
Network Profile : Todos
Precedence : 127
IPProtocol : Both
IPDstPrefix : 10.240.1.0/24
PriorityValue : 7
```

Use o seguinte para exibir a Política QoS:

```
PS C:\> Get-NetQosPolicy
Name : Network Management
Owner : (382ACFAD-1E73-46BD-A0A-6-4EE0E587B95)
NetworkProfile : Todos
Precedence : 127
IPProtocol : Both
IPDstPrefix : 10.240.1.0/24
PriorityValue : 7
Name : SMB policy
Owner : (382AFAD-1E73-46BD-A0A-6-4EE0E587B95)
NetworkProfile : Todos
Precedence : 127
```

Template : SMB
PriorityValue : 4

Use o seguinte para modificar a Política QoS:

```
PS C:\> Set-NetqosPolicy -Name "Network Management" -IPSrcPrefixMatchCondition 10.235.2.0/24 -  
IPProtocolMatchCondition both -PriorityValue802.1Action 7  
PS C:\> Get-NetQosPolicy -name "network management"  
Name : Network Management  
Owner : {382ACFD-1E73-46BD-A0A0-4EE0E587B95}  
NetworkProfile : Todos  
Precedence : 127  
IPProtocol : Both  
IPSrcPrefix : 10.235.2.0/24  
IPDstPrefix : 10.240.1.0/24  
PriorityValue : 7
```

Use o seguinte para remover a Política QoS:

```
PS C:\> Remove-NetQosPolicy -Name "Network Management"
```

Perguntas frequentes

- Existe suporte para AutoNeg com velocidade de 25G?

Sim. Consulte [“Configuração de negociação automática” na página 40](#) para obter mais detalhes.

- Como posso conectar o cabo SFP28 a portas QSFP?

Cabos breakout estão disponíveis de portas QSFP para 4xSFP28.

- Quais são as velocidades de portas compatíveis?

Para dispositivos de porta dupla BCM57404AXXX/BCM57414, a velocidade de porta de cada porta deve ser compatível com a velocidade de porta da outra porta. 10 Gbps e 25 Gbps não são velocidades compatíveis. Se uma porta for configurada para 10 Gbps, a outra porta não poderá ser configurada para 25 Gbps. Se o usuário tentar configurar velocidades de portas incompatíveis, a segunda porta que deveria ser exibida não será vinculada. Consulte [“Configuração de negociação automática” na página 40](#) para obter mais detalhes.

- Posso usar 10 Gbps para conectividade PXE em uma porta de 25 Gbps?

Atualmente, há compatibilidade somente com a velocidade de 25 Gbps PXE. Não é recomendável usar a conectividade 10 Gbps PXE em um adaptador de 25 Gbps. Isso se deve à falta de suporte para negociação automática em switches de 10 Gbps existentes e às possíveis complicações causadas pelas configurações incompatíveis de velocidade do link de porta.