

Manuel d'utilisation  
**NetXtreme-E**

**Broadcom<sup>®</sup> NetXtreme-C et NetXtreme-E**  
**MANUEL D'UTILISATION**

## Historique des révisions

<i>Révision</i>	<i>Date</i>	<i>Modification de la description</i>
NetXtreme-E-UG100	26/02/2018	Version initiale pour 20.6.

© 2018 par Broadcom. Tous droits réservés.

Broadcom<sup>®</sup>, le logo Impulsion, Connecting everything<sup>®</sup>, Avago Technologies et le logo A sont des marques de Broadcom et/ou de ses filiales aux Etats-Unis et dans d'autres pays et/ou dans l'UE. Le terme « Broadcom » fait référence à Broadcom Limited et/ou à ses filiales. Pour plus d'informations, rendez-vous sur [www.broadcom.com](http://www.broadcom.com).

Broadcom se réserve le droit de modifier sans préavis ses produits ou les données du présent document en vue d'en améliorer la fiabilité, le fonctionnement ou la conception. Les informations fournies par Broadcom sont censées être précises et fiables. Cependant, Broadcom ne se porte pas garant de l'application et de l'utilisation desdites informations ni de l'application et de l'utilisation de tout produit ou circuit décrit dans les présentes. Broadcom ne cède pas non plus de licence au titre de ses droits de propriété industrielle ni au titre des droits de tout autre tiers.

## Table des matières

<b>Autorisations réglementaires et homologations de sécurité</b> .....	7
Réglementation .....	7
Sécurité .....	7
Compatibilité électromagnétique (EMC).....	8
Conformité aux normes relatives aux décharges électrostatiques (ESD) .....	8
Déclaration de la FCC .....	8
<b>Description du fonctionnement</b> .....	9
<b>Indication de liaison au réseau et d'activité du réseau</b> .....	14
BCM957402AXXXX/BCM957412AXXXX .....	14
BCM957404AXXXX/BCM957414AXXXX .....	15
BCM957406AXXXX/BCM957416AXXXX .....	16
BCM957414M4140D.....	17
BCM957412M4120D.....	18
BCM957416M4160 .....	19
<b>Spécifications</b> .....	20
Fonctions logicielles et matérielles.....	20
Fonctions de virtualisation.....	21
VXLAN.....	22
NVGRE/GRE/IP-in-IP/Geneve .....	22
Déchargements sans état .....	22
RSS.....	22
TPA .....	22
Répartition En-tête/Charge utile (Header-Payload Split) .....	23
Déchargement de la fragmentation UDP .....	23
Déchargement du tunnel de transport sans état .....	23
Prise en charge de multiples files d'attente pour SE .....	23
NDIS VMQ .....	23
VMWare NetQueue.....	24
KVM/Xen Multiqueue .....	24
Matrice de prise en charge des configurations SR-IOV .....	24
SR-IOV .....	24
Partitionnement réseau (NPAR).....	24
RDMA over Converge Ethernet – RoCE .....	25
Combinaisons prises en charge .....	25
NPAR, SR-IOV et RoCE .....	25
NPAR, SR-IOV et DPDK.....	26
Combinaisons non prises en charge.....	26

---

<b>Installation du matériel</b> .....	27
Mesures de sécurité .....	27
Configuration système.....	27
Configuration matérielle requise .....	27
Liste de vérification avant l'installation .....	27
Installation de la carte .....	28
Connexion des câbles réseau .....	28
Câbles et modules pris en charge.....	28
Cuivre.....	29
SFP+ .....	29
SFP28 .....	29
<b>Progiciels et installation</b> .....	30
Systèmes d'exploitation pris en charge .....	30
Installation des pilotes .....	30
Windows.....	30
<i>Dell DUP</i> .....	30
<i>Installation de l'interface graphique utilisateur</i> .....	30
<i>Installation silencieuse</i> .....	30
<i>Installation des fichiers INF</i> .....	30
Linux.....	31
<i>Installation du module</i> .....	31
<i>Commandes Ethtool de Linux</i> .....	32
VMware .....	33
Mise à jour du micrologiciel.....	34
Dell Update Package .....	34
Windows.....	34
Linux.....	34
<b>Propriétés avancées du pilote Windows et messages du journal d'événements</b> .....	35
Propriétés avancées du pilote .....	35
Messages du journal des événements.....	36
<b>Regroupement</b> .....	38
Windows.....	38
Linux.....	38
<b>Configuration au niveau du système</b> .....	39
Menu UEFI HII.....	39
Page de configuration principale.....	39
Propriétés de l'image du microprogramme .....	39
Configuration du niveau du périphérique .....	39
Configuration de la carte NIC.....	39

---

---

Configuration iSCSI.....	40
CCM (Comprehensive Configuration Management) .....	40
Configuration matérielle du périphérique .....	40
Menu de configuration MBA.....	40
Menu principal de l'initialisation iSCSI .....	40
Configuration de la négociation automatique .....	41
<i>Vitesse de liaison opérationnelle</i> .....	44
<i>Vitesse de liaison du microprogramme</i> .....	44
<i>Protocole de négociation automatique</i> .....	44
<i>Paramètres du pilote Windows</i> .....	44
<i>Paramètres du pilote Linux</i> .....	44
<i>Paramètres du pilote ESXi</i> .....	45
Négociation automatique FEC .....	46
Link Training.....	47
Détection automatique des supports.....	48
<b>Initialisation iSCSI</b> .....	50
Systèmes d'exploitation pris en charge pour l'initialisation iSCSI .....	50
Configuration de l'initialisation iSCSI.....	50
Configuration de la cible iSCSI .....	50
Configuration des paramètres d'initialisation iSCSI .....	51
Configuration du protocole de démarrage MBA.....	52
Configuration de l'initialisation iSCSI .....	52
<i>Configuration de l'initialisation iSCSI statique</i> .....	52
<i>Configuration de l'initialisation iSCSI dynamique</i> .....	53
Activation de l'authentification CHAP .....	55
Configuration du serveur DHCP pour la prise en charge de l'initialisation iSCSI.....	55
Configurations DHCP pour l'initialisation iSCSI pour IPv4.....	55
<i>Option DHCP 17, chemin d'accès</i> .....	55
<i>Option DHCP 43, informations concernant le fournisseur</i> .....	56
<i>Configuration du serveur DHCP</i> .....	56
Configurations DHCP pour l'initialisation iSCSI pour IPv6.....	57
<i>Option DHCPv6 16, option de classe fournisseur</i> .....	57
<i>Option DHCPv6 17, informations concernant le fournisseur</i> .....	57
<i>Configuration du serveur DHCP</i> .....	58
<b>VXLAN : Configuration et exemples d'utilisation</b> .....	58
<b>SR-IOV : Configuration et exemples d'utilisation</b> .....	59
Cas d'utilisation Linux.....	59
Cas de Windows .....	60
Cas de VMWare SRIOV .....	61

---

---

<b>NPAR : Configuration et exemples d'utilisation</b> .....	64
Fonctionnalités et configuration requise .....	64
Limitations .....	64
Configuration .....	64
Notes sur la réduction de consommation de mémoire de la carte réseau .....	67
<b>RoCE : Configuration et exemples d'utilisation</b> .....	68
Configuration sous Linux .....	68
Configuration requise .....	68
Dépendances du pilote BNXT_RE .....	68
Installation .....	69
Limitations .....	70
Problèmes connus .....	70
Windows .....	70
Mode noyau .....	70
Vérification de RDMA .....	71
Mode utilisateur .....	72
VMware ESX .....	72
Limitations .....	72
Configuration requise du pilote RoCE de la carte réseau BNXT .....	72
Installation .....	72
Configuration des cartes réseau PVRDMA .....	73
<i>Configuration d'un centre virtuel pour PVRDMA</i> .....	73
<i>Balisage d'un vmknic pour PVRDMA sur les hôtes ESX</i> .....	73
<i>Définition de la règle de pare-feu pour PVRDMA</i> .....	73
<i>Ajout d'un périphérique PVRDMA à la machine virtuelle</i> .....	73
<i>Configuration de la machine virtuelle sur le système d'exploitation invité Linux</i> .....	74
<b>DCBX – Data Center Bridging</b> .....	75
Profil QoS : Profil de file d'attente QoS par défaut .....	75
Mode DCBX = Activer (IEEE uniquement) .....	76
Bit Willing DCBX .....	76
<b>Foire aux questions</b> .....	79

---

## Autorisations réglementaires et homologations de sécurité

Les sections suivantes fournissent des informations détaillées sur le respect des normes en matière de réglementation, sécurité, compatibilité électromagnétique (EMC) et décharges électrostatiques (ESD) pour la NetXtreme-Ecarte d'interface réseau.

### Réglementation

**Tableau 1 : Autorisations réglementaires**

<b>Élément</b>	<b>Normes applicables</b>	<b>Autorisation/Certificat</b>
CE/Union européenne	EN 62368-1:2014	Rapport et certificat CB
UL/Etats-Unis	IEC 62368-1 édition 2	Rapport et certificat CB
CSA/Canada	CSA 22.2 N° 950	Rapport et certificat CSA.
Taiwan	CNS14336 Classe B	–

### Sécurité

**Tableau 2 : Homologations de sécurité**

<b>Pays</b>	<b>Type de certification/Norme</b>	<b>Conformité</b>
International	Schéma CB ICES 003 - Dispositif numérique UL 1977 (sécurité du connecteur) UL 796 (sécurité du câblage PCB) UL 94 (inflammabilité de pièces)	Oui

## Compatibilité électromagnétique (EMC)

**Tableau 3 : Compatibilité électromagnétique**

<b>Norme/Pays</b>	<b>Type de certification</b>	<b>Conformité</b>
CE/EU	EN 55032:2012/AC:2013 Classe B EN 55024:2010 EN 61000-3-2:2014 EN 61000-3-3:2013	Rapport CE et CE DoC
FCC/Etats-Unis	CFR47, Partie 15 Classe B	FCC/IC DoC et rapport EMC faisant référence aux normes FCC et IC
IC/Canada	ICES-003 Classe B	FCC/IC DoC et rapport faisant référence aux normes FCC et IC
ACA/Australie, Nouvelle-Zélande	AS/NZS CISPR 22:2009 +A1:2010 / AS/NZS CISPR 32:2015	Certificat ACA Marquage RCM
BSMI/Taiïwan	CNS13438 Classe B	Certificat BSMI
BSMI/Taiïwan	CNS15663	Certificat BSMI
MIC/S. Corée	KN32 Classe B KN35	Certificat coréen Marquage MSIP
VCCI/Japon	V-3/2014/04 (entrée en vigueur le 31/03/2015)	Copie du certificat VCCI en ligne

## Conformité aux normes relatives aux décharges électrostatiques (ESD)

**Tableau 4 : Synthèse de la conformité ESD**

<b>Normes</b>	<b>Type de certification</b>	<b>Conformité</b>
EN55024:2010 (EN 61000-4-2)	Décharge directe/d'air	Oui

## Déclaration de la FCC

L'équipement a été testé et déclaré conforme aux limites des dispositifs numériques de la Classe B définies par l'alinéa 15 du règlement de la FCC. Ces limites permettent d'assurer une protection raisonnable contre les interférences nuisibles dans le cadre d'une installation résidentielle. Cet équipement génère, utilise et dégage une énergie H.F. S'il n'est pas installé et utilisé en conformité avec les instructions, il peut créer une interférence nuisible aux communications radio. Nous ne pouvons cependant pas garantir qu'aucune interférence ne se produira dans le cadre d'une installation particulière. Si cet équipement cause des interférences nuisibles à la réception radio ou télévisuelle, ce qui peut être déterminé en éteignant et rallumant l'équipement, nous vous conseillons d'essayer de corriger ces interférences en prenant les mesures suivantes :

- Réorientez l'antenne de réception.
- Eloignez davantage l'équipement du récepteur.
- Consultez votre revendeur ou un technicien de radio/télévision expérimenté pour obtenir de l'aide.



**Remarque :** Tout changement ou toute modification n'ayant pas été expressément approuvé(e) par le fabricant responsable de la conformité peut annuler l'autorisation accordée à l'utilisateur d'utiliser cet équipement.



## Description du fonctionnement

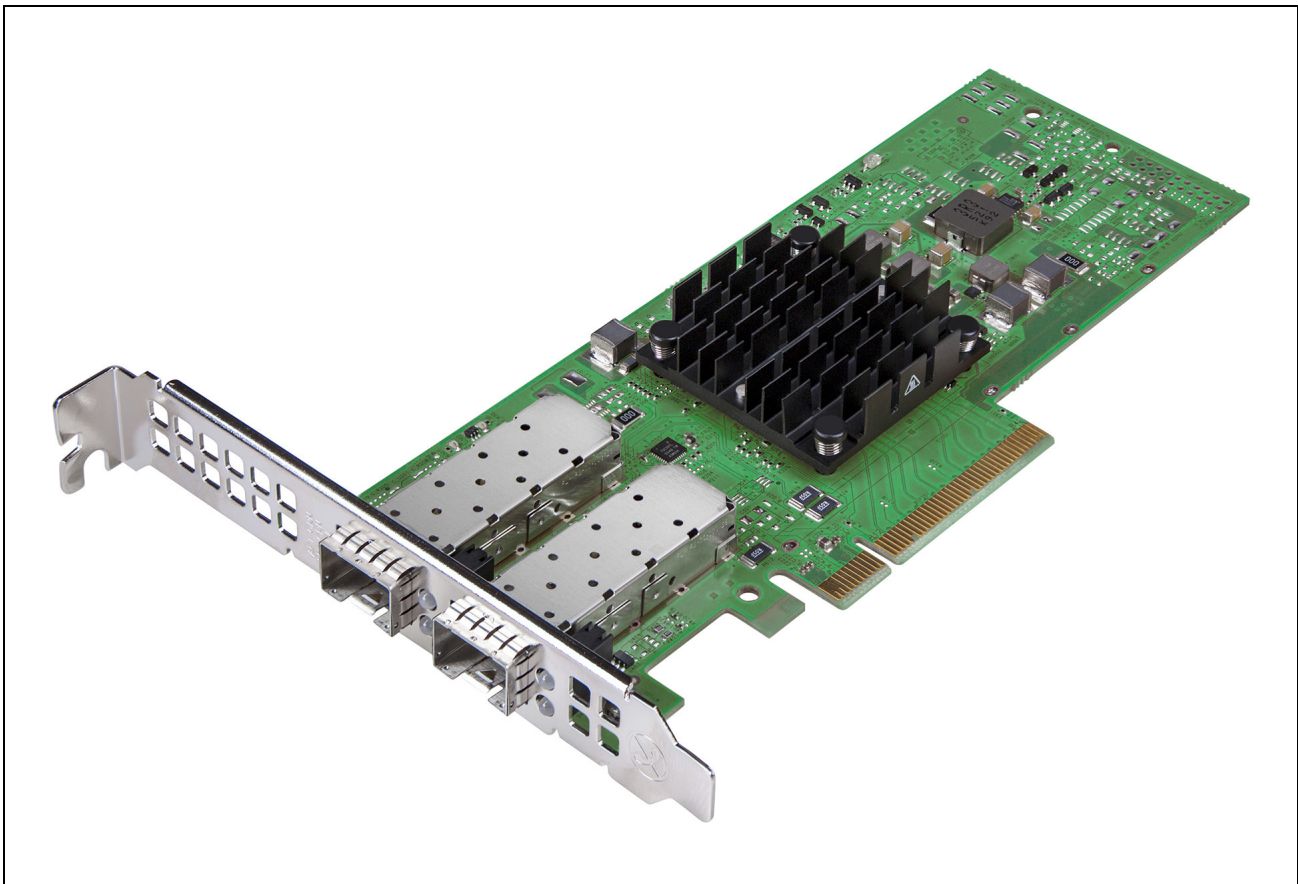
Dell prend en charge les cartes d'interface réseau (NIC) 10GBase-T, 10G SFP+ et 25G SFP28. Ces cartes NIC sont présentées dans le [Tableau 5](#).

**Tableau 5 : Description du fonctionnement**

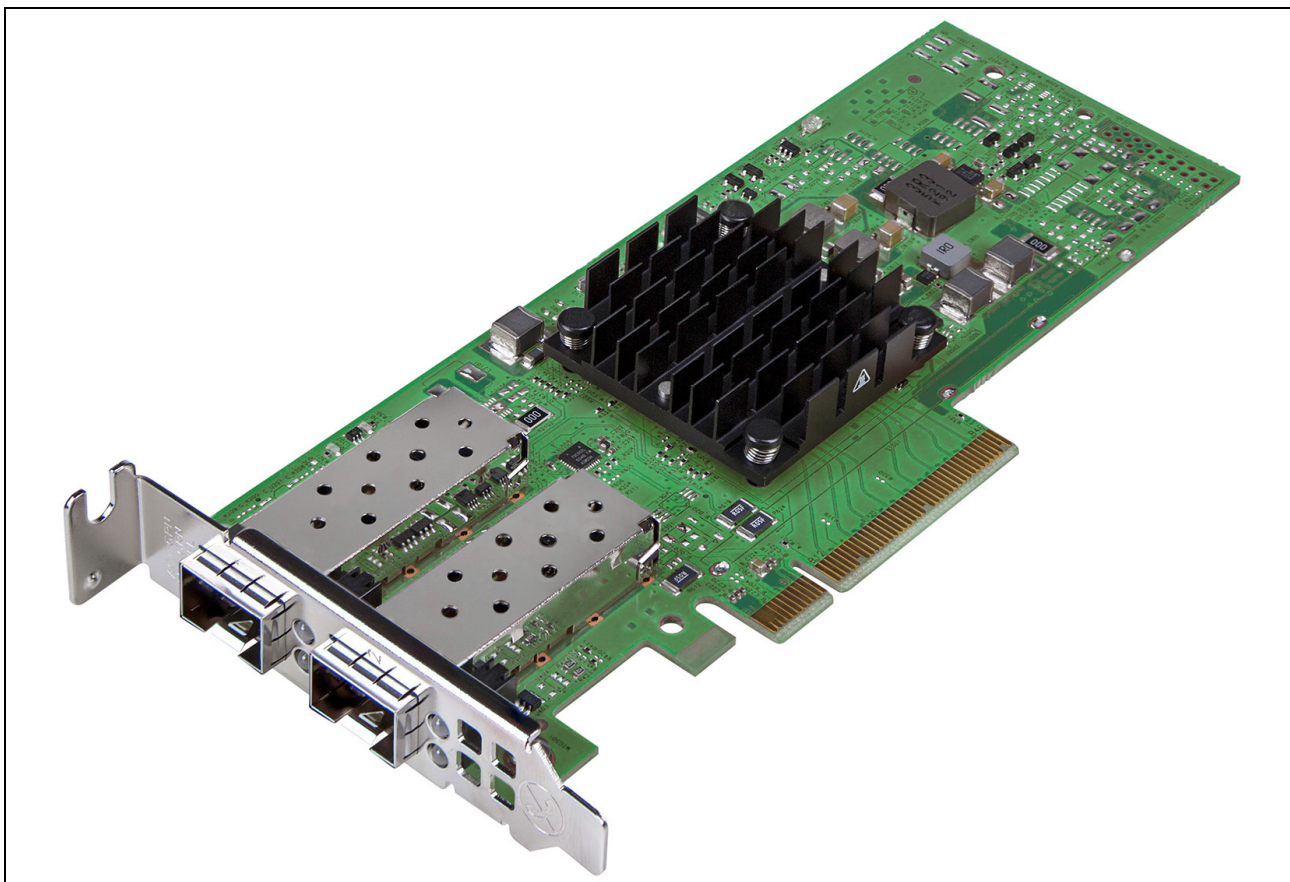
<b>Carte d'interface réseau</b>	<b>Description</b>
<b>BCM957402A4020DLPC/BCM957402A4020DC/BCM957412A4120D/BCM957412M4120D</b>	
Vitesse	Deux ports 10 Gbit/s Ethernet
PCI-E	Gen 3 x8 <sup>a</sup>
Interface	SFP+ pour 10 Gbit/s
Périphérique	Contrôleur Broadcom BCM57402/BCM57412 10 Gbit/s MAC avec émetteur/récepteur SFI bicanal 10 Gbit/s intégré.
Nom NDIS	Carte Broadcom NetXtreme E-Series double port 10 Gbits SFP+ Ethernet PCIe
Nom UEFI	SFP double Broadcom 10 Gbits + Ethernet
<b>BCM57404A4041DLPC/BCM57404A4041DC/BCM957414A4141D/BCM957414M4140D</b>	
Vitesse	Deux ports 25 Gbit/s ou 10 Gbit/s Ethernet
PCI-E	Gen 3 x8 <sup>a</sup>
Interface	SFP28 pour 25 Gbit/s et SFP+ pour 10 Gbit/s
Périphérique	Contrôleur Broadcom BCM57404/BCM57414 25 Gbit/s MAC avec émetteur-récepteur SFI bicanal 25 Gbit/s intégré.
Nom NDIS	Carte Broadcom NetXtreme E-Series double port 25 Gbits SFP28 Ethernet PCI-e
Nom UEFI	SFP double Broadcom 25 Gbits 28 Ethernet
<b>BCM957406A4060DLPC/BCM957406A4060DC/BCM957416A4160D/BCM957416M4160</b>	
Vitesse	Double port 10GBase-T Ethernet
PCI-E	Gen 3 x8 <sup>a</sup>
Interface	RJ45 pour 10 Gbit/s et 1 Gbit/s
Périphérique	Contrôleur Broadcom BCM57406/BCM57416 10 Gbit/s MAC avec émetteur-récepteur 10GBase-T bicanal intégré.
Nom NDIS	Carte Broadcom NetXtreme E-Series double port 10GBASE-T Ethernet PCIe
Nom UEFI	Broadcom Double 10GBASE-T Ethernet

- a. La carte NIC prend en charge les PCI-E Gen 1, Gen 2 et Gen 3. Le PCI-E Gen 3 est toutefois recommandé pour obtenir le débit nominal lorsque deux ports 25G Liens transmettent et reçoivent simultanément le trafic.

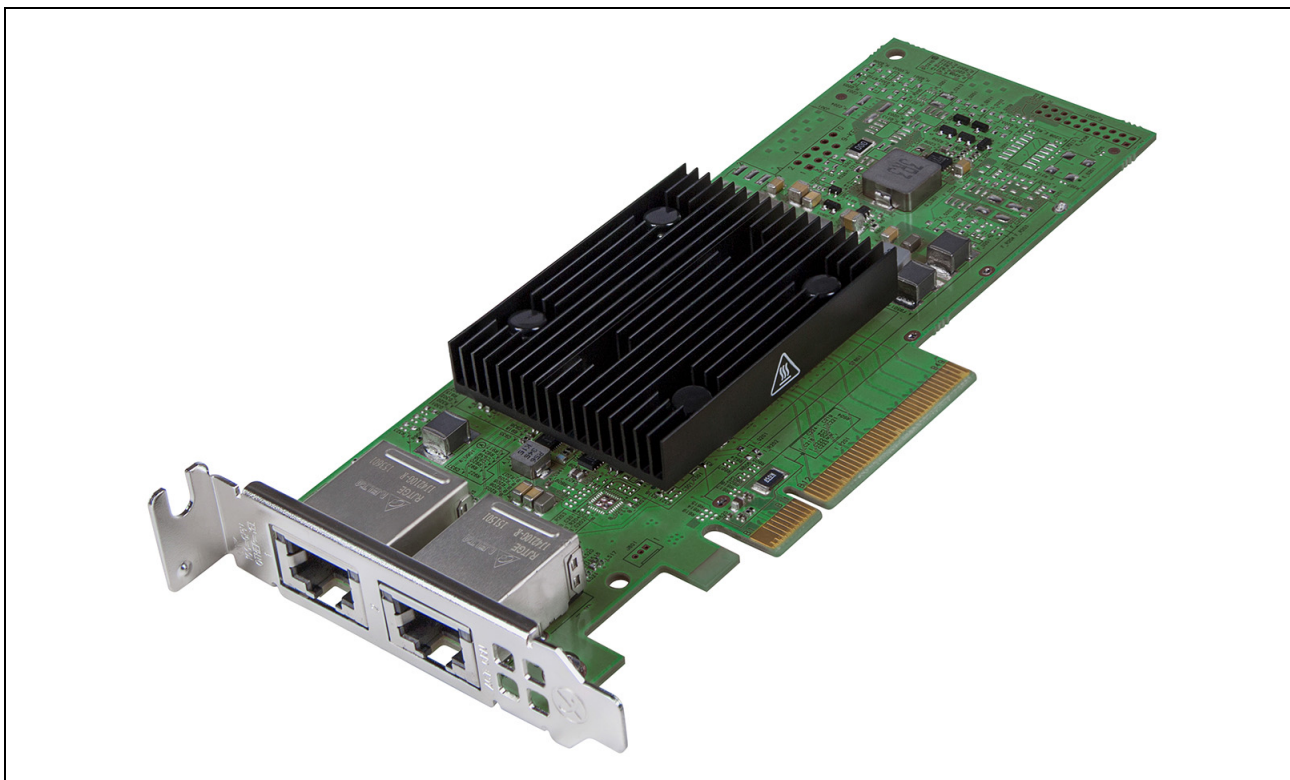
**Figure 1 : Carte d'interface réseau BCM957402A4020DC, BCM957412A4120D**



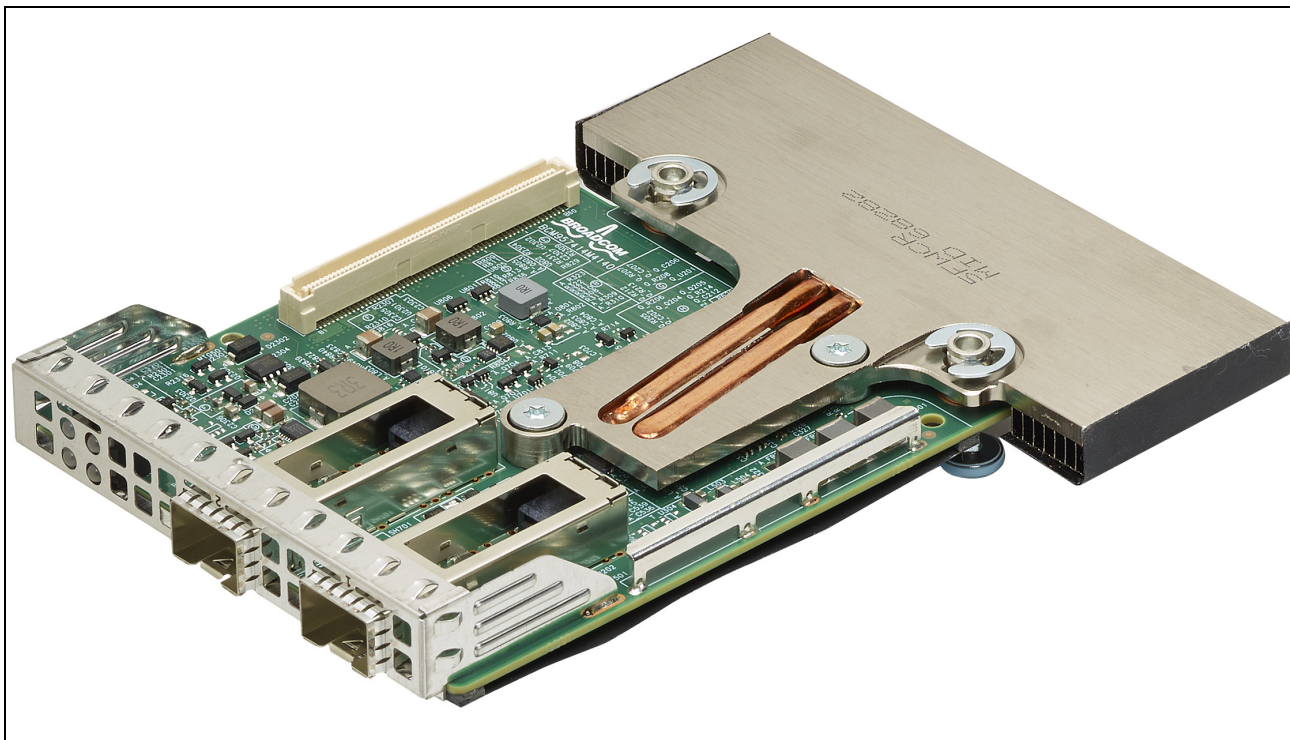
**Figure 2 : Carte d'interface réseau BCM957404A4041DLPC, BCM957414A4141D**



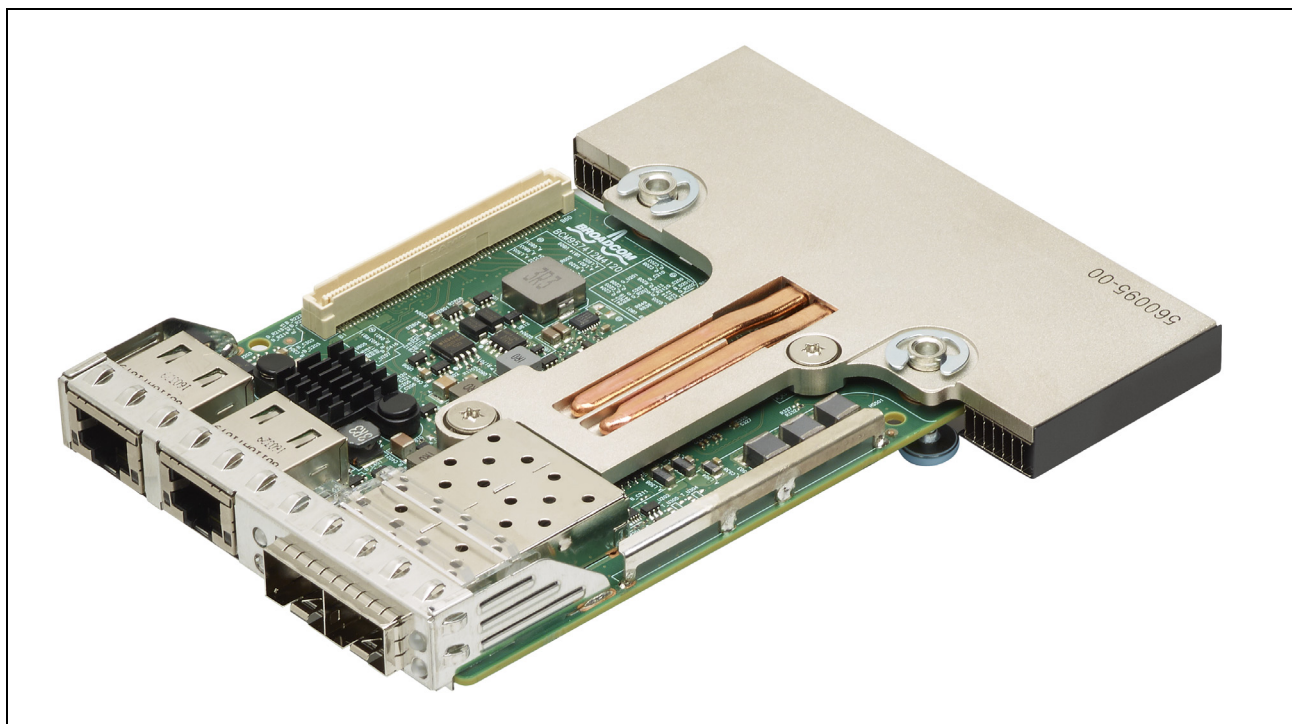
**Figure 3 : Carte d'interface réseau BCM957406A4060DLPC, BCM957416A4160D**



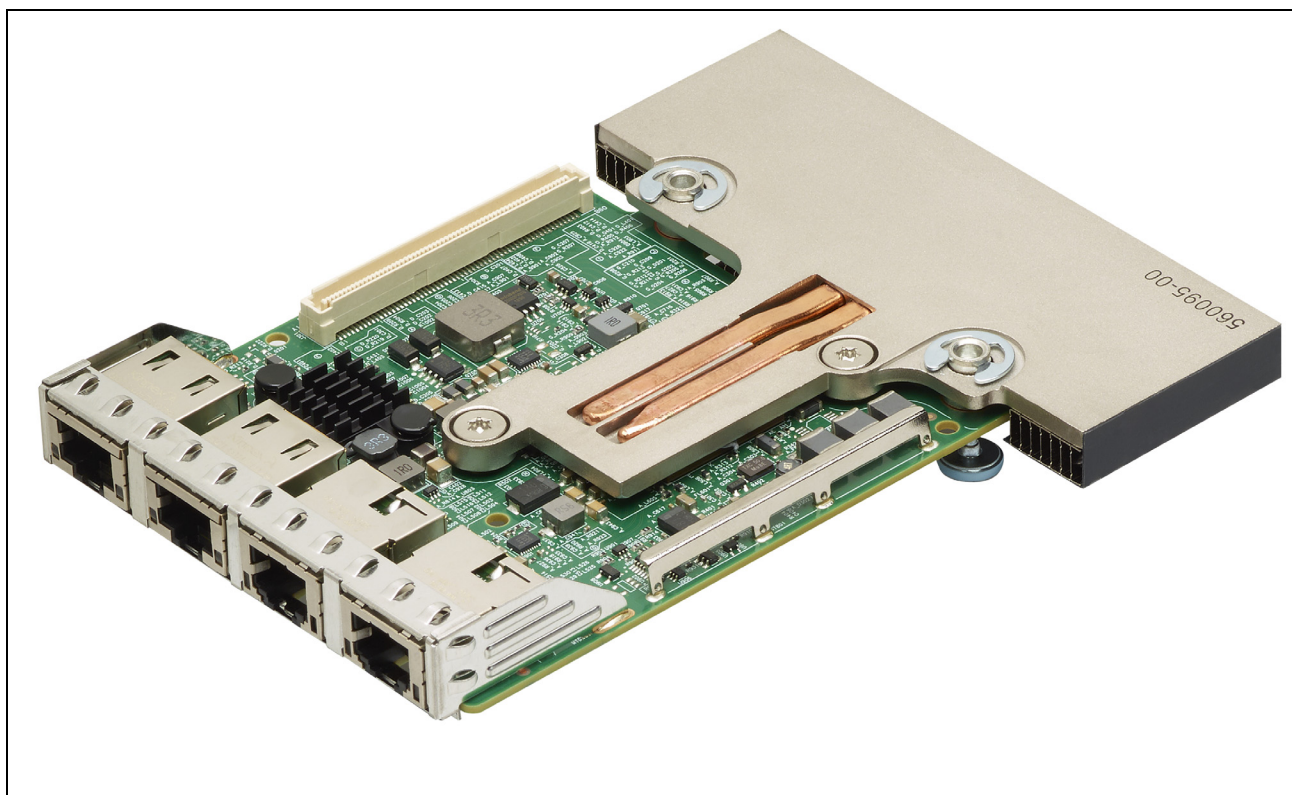
**Figure 4 : Carte fille réseau (rNDC) BCM957414M4140D**



**Figure 5 : Carte fille réseau (rNDC) BCM957412M4120D**



**Figure 6 : Carte fille réseau (rNDC) BCM957416M4160**

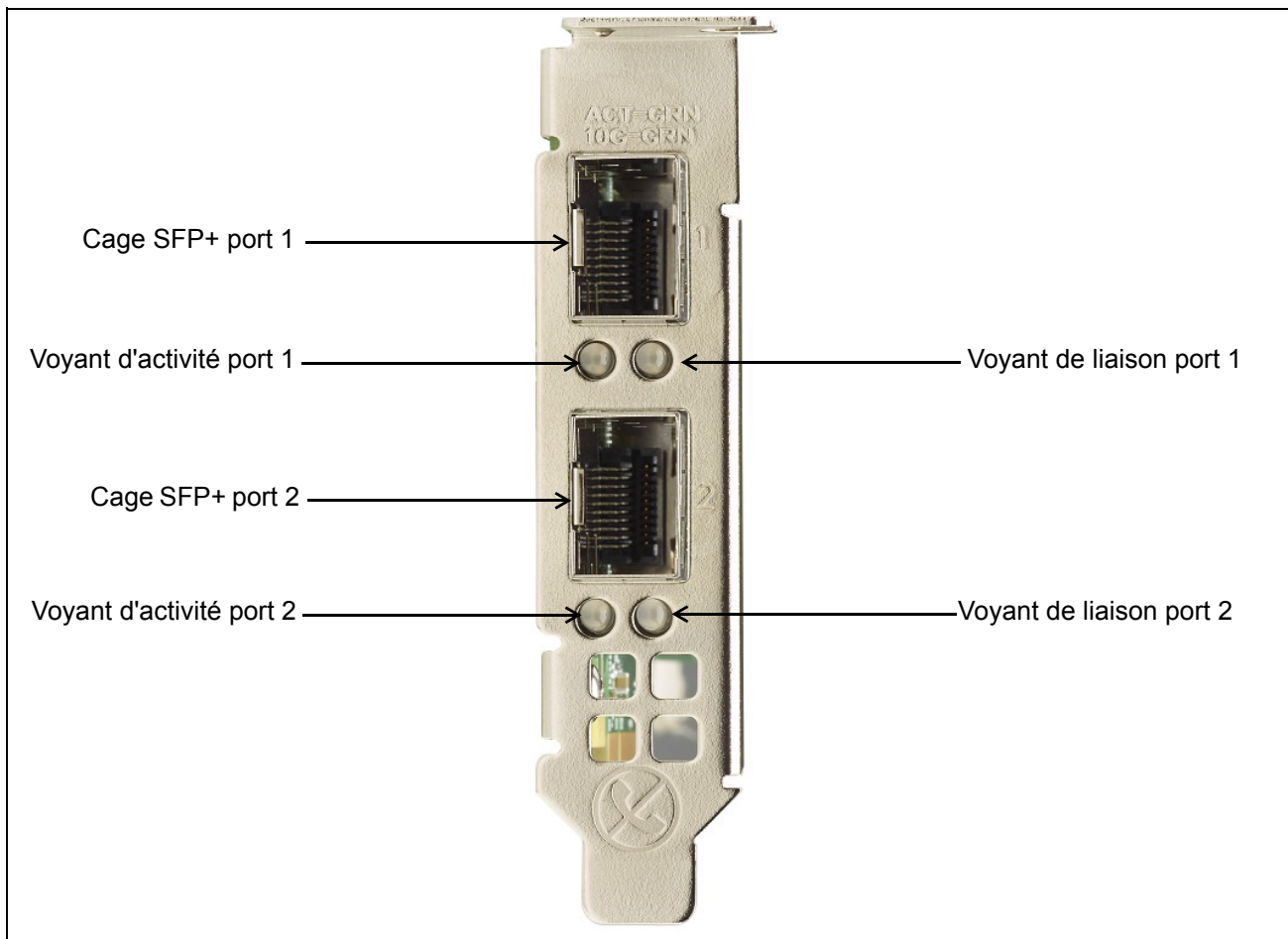


## Indication de liaison au réseau et d'activité du réseau

### BCM957402AXXXX/BCM957412AXXXX

Le port SFP+ possède deux voyants pour indiquer le volume du trafic et la vitesse de liaison. Les voyants sont visibles à travers la découpe, sur le support, comme illustré sur la [Figure 7](#). Les fonctions des voyants sont décrites dans le [Tableau 6](#).

**Figure 7 : Emplacements des voyants de liaison et d'activité des modèles BCM957402AXXXX/BCM957412AXXXX**



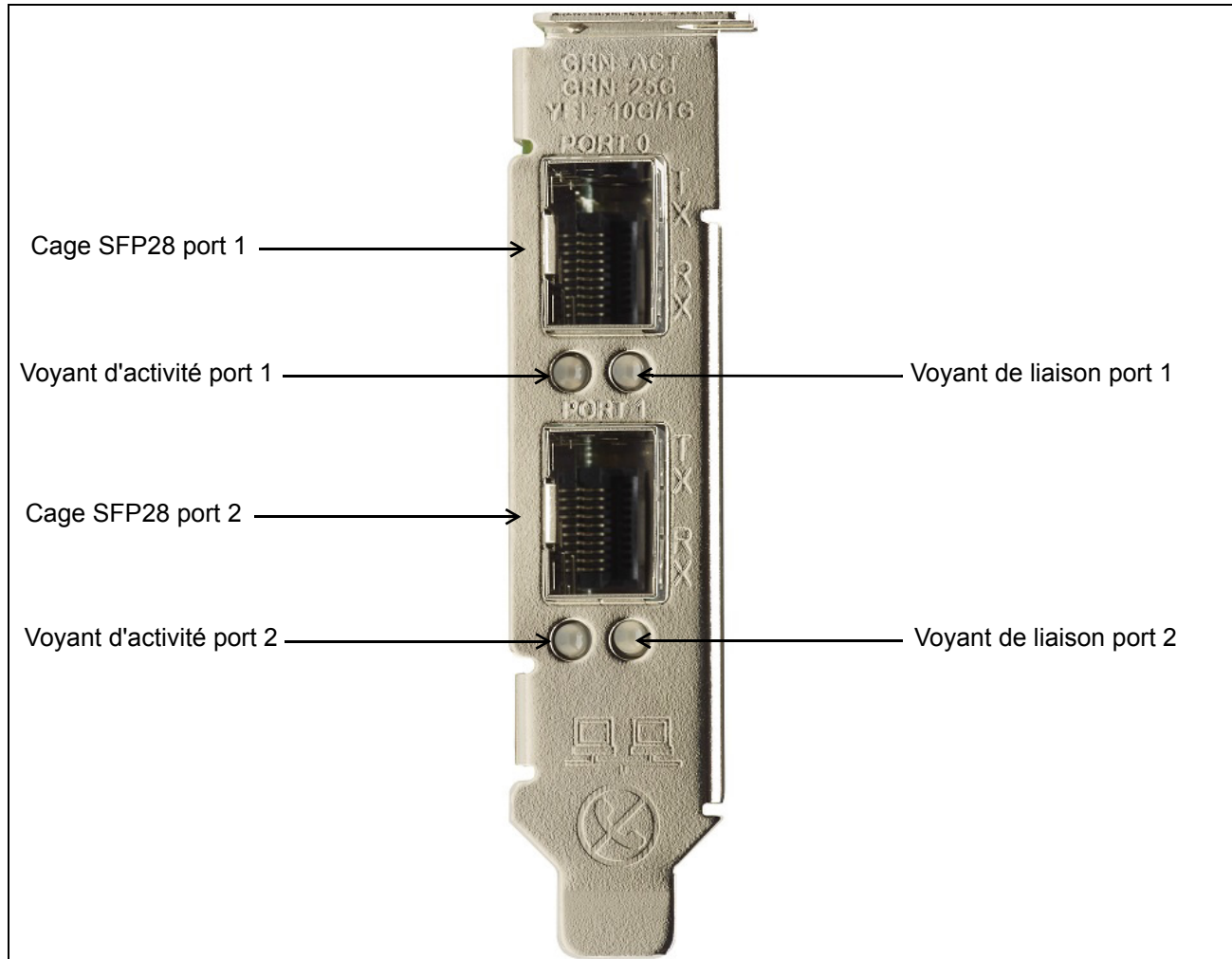
**Tableau 6 : Emplacement des voyants de liaison et d'activité des modèles BCM957402AXXXX/BCM957412AXXXX**

Type de voyant	Couleur/Comportement	Note
Activité	Désactivé	Aucune activité
	Vert clignotant	Activité de flux de trafic
Liaison	Désactivé	Pas de liaison
	Vert	Liaison à 10 Gbit/s

## BCM957404AXXXX/BCM957414AXXXX

Le port SFP28 possède deux voyants pour indiquer le volume du trafic et la vitesse de liaison. Les voyants sont visibles à travers la découpe, sur le support, comme illustré sur la [Figure 8](#). Les fonctions des voyants sont décrites dans le [Tableau 7](#).

**Figure 8 : Emplacements des voyants de liaison et d'activité BCM957404AXXXX/BCM957414AXXXX**



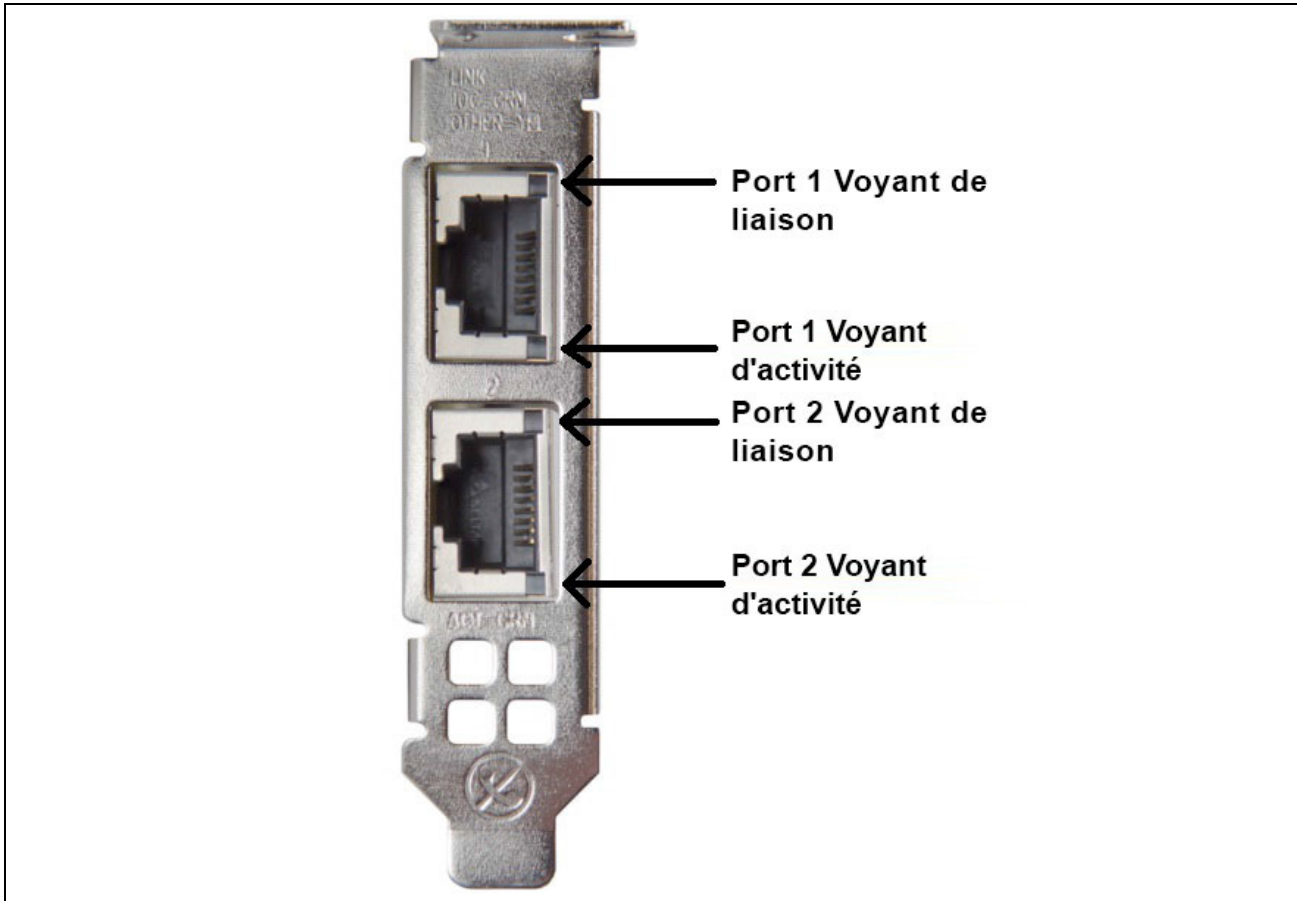
**Tableau 7 : Emplacement des voyants de liaison et d'activité des modèles BCM957404AXXXX/BCM957414AXXXX**

Type de voyant	Couleur/Comportement	Note
Activité	Désactivé	Aucune activité
	Vert clignotant	Activité de flux de trafic
Liaison	Désactivé	Pas de liaison
	Vert	Liaison à 25 Gbit/s
	Jaune	Liaison à 10 Gbit/s

## BCM957406AXXX/BCM957416AXXX

Le port RJ-45 possède deux voyants pour indiquer le volume du trafic et la vitesse de liaison. Les voyants sont visibles à travers la découpe, sur le support, comme illustré sur la [Figure 9](#). Les fonctions des voyants sont décrites dans le [Tableau 8](#).

**Figure 9 : Emplacements des voyants de liaison et d'activité BCM957406AXXX/BCM957416AXXX**



**Tableau 8 : Emplacement des voyants de liaison et d'activité BCM957406AXXX/BCM957416AXXX**

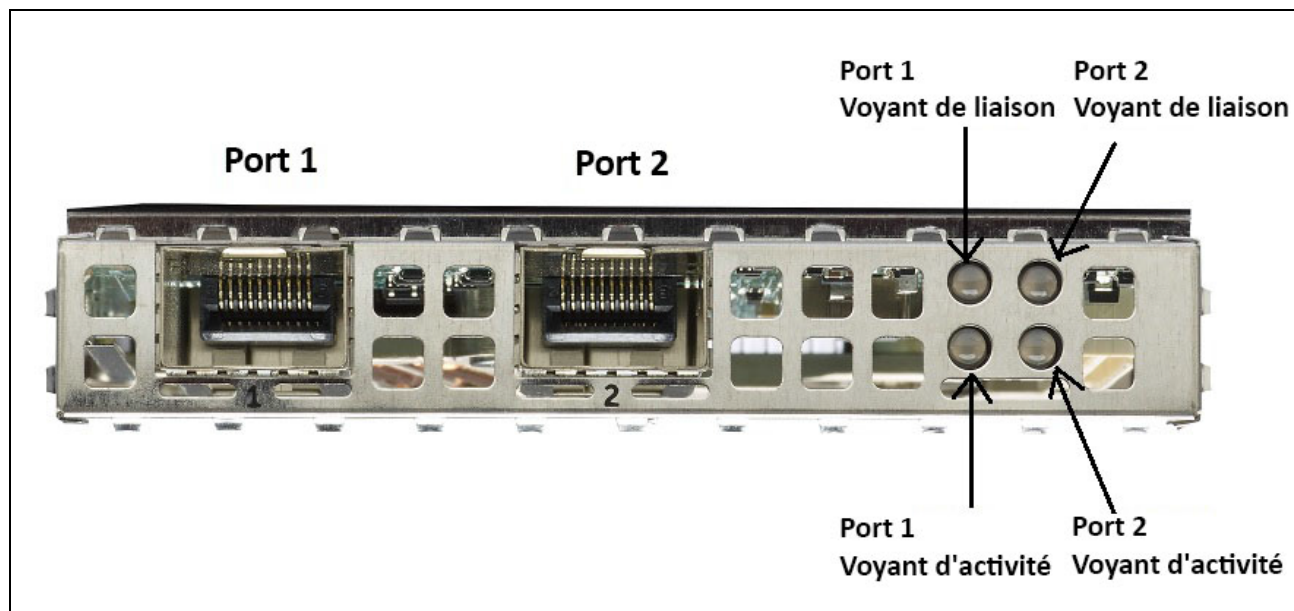
Type de voyant	Couleur/Comportement	Remarques
Activité	Désactivé	Aucune activité
	Vert clignotant	Activité de flux de trafic
Liaison	Désactivé	Pas de liaison
	Vert	Liaison à 10 Gbit/s
	Orange	Liaison à 1 Gbit/s



## BCM957414M4140D

Le port SFP28 possède deux voyants pour indiquer le volume du trafic et la vitesse de liaison. Les voyants sont visibles à travers la découpe, sur le support, comme illustré sur la [Figure 10](#).

**Figure 10 : Emplacements des voyants de liaison et d'activité de la carte fille réseau (rNDC) BCM957414M4140D**



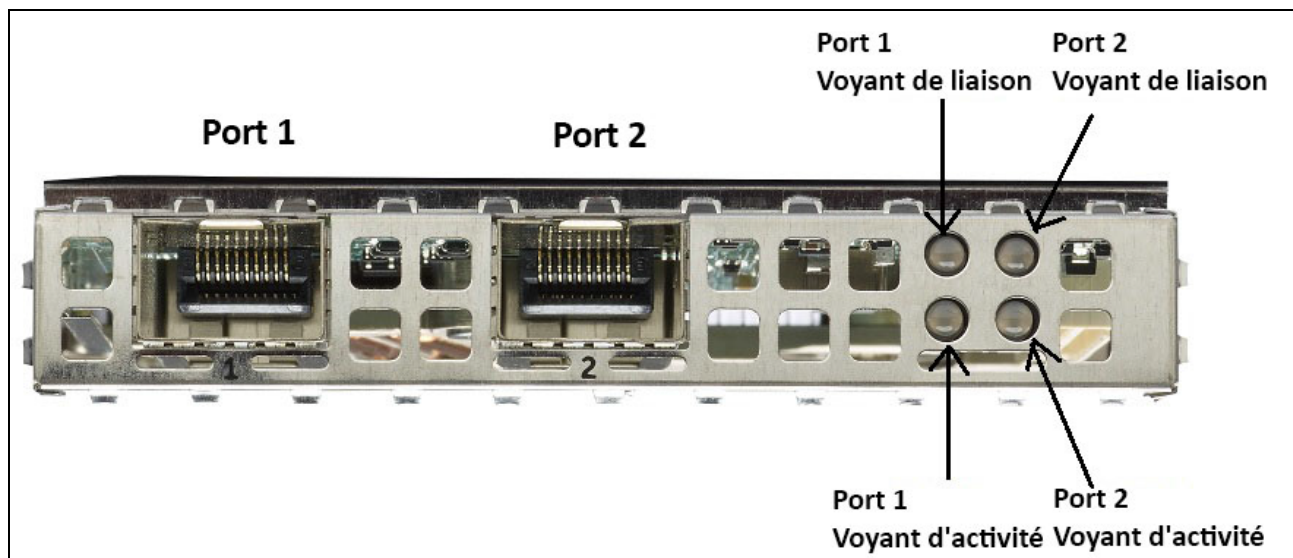
**Tableau 9 : Emplacements des voyants de liaison et d'activité de la carte fille réseau (rNDC) BCM957414M4140D**

Type de voyant	Couleur/Comportement	Remarques
Activité	Désactivé	Aucune activité
	Vert clignotant	Activité de flux de trafic
Liaison	Désactivé	Pas de liaison
	Vert	Liaison à 25 Gbit/s
	Jaune	Liaison à 10 Gbit/s

## BCM957412M4120D

Cette carte fille réseau (rNDC) possède des ports SFP+ et RJ-45, chacun associé à deux voyants indiquant les activités du trafic et la vitesse de liaison. Les voyants sont visibles comme illustré sur la [Figure 11](#).

**Figure 11 : Emplacements des voyants de liaison et d'activité de la carte fille réseau (rNDC) BCM957412M4120D**



**Tableau 10 : Emplacements des voyants de liaison et d'activité de la carte fille réseau (rNDC) BCM957412M4120D ports SFP+ 1 et 2**

Type de voyant	Couleur/Comportement	Remarques
Activité	Désactivé	Aucune activité
	Vert clignotant	Activité de flux de trafic
Liaison	Désactivé	Pas de liaison
	Vert	Liaison à 10 Gbit/s

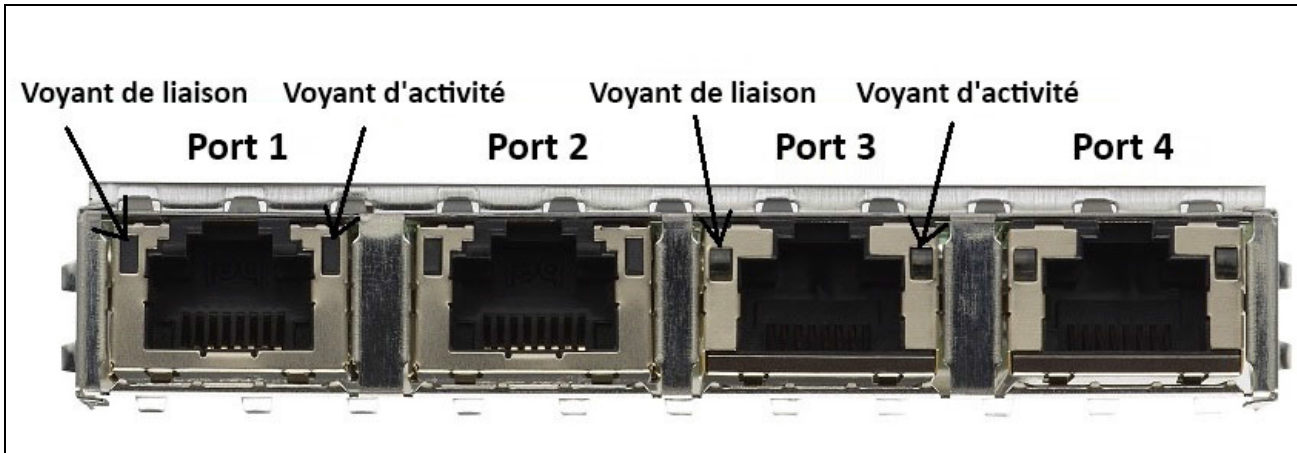
**Tableau 11 : Ports 1000BaseT 3 et 4**

Type de voyant	Couleur/Comportement	Remarques
Activité	Désactivé	Aucune activité
	Vert clignotant	Activité de flux de trafic
Liaison	Désactivé	Pas de liaison
	Vert	Liaison à 1 Gbit/s
	Orange	Liaison à 10/100 Mbit/s

## BCM957416M4160

Cette carte fille réseau (rNDC) possède des ports RJ-45 10GBaseT et 1000BaseT, chacun associé à deux voyants indiquant les activités du trafic et la vitesse de liaison. Les voyants sont visibles comme illustré sur la [Figure 12](#).

**Figure 12 : Emplacements des voyants de liaison et d'activité de la carte fille réseau (rNDC) BCM957416M4160**



**Tableau 12 : Emplacements des voyants de liaison et d'activité de la carte fille réseau (rNDC) BCM957416M4160 Ports 10GBaseT 1 et 2**

Type de voyant	Couleur/Comportement	Remarques
Activité	Désactivé	Aucune activité
	Vert clignotant	Activité de flux de trafic
Liaison	Désactivé	Pas de liaison
	Vert	Liaison à 10 Gbit/s
	Orange	Liaison à 1 Gbit/s

## Spécifications

Les sections suivantes présentent les fonctions du périphérique.

### Fonctions logicielles et matérielles

Le [Tableau 13](#) fournit la liste des fonctions de l'interface hôte.

**Tableau 13 : Fonctions de l'interface hôte**

<b>Fonction</b>	<b>Données</b>
Interface hôte	PCIe v3.0 (Gen 3 : 8 GT/s ; Gen 2 : 5 GT/s ; Gen 1 : 2,5 GT/s).
Nombre de voies PCIe	Connecteur PCI-E Edge : x8.
Données VPD (Vital Product Data)	Prise en charge.
ID de routage alternatif (ARI, Alternate Routing ID)	Prise en charge.
Réinitialisation du niveau de fonction (FLR)	Prise en charge.
Rapport d'erreur avancé	Prise en charge.
ECN PCIe	Prise en charge de TLH (TLP Processing Hint), LTR (Latency Tolerance Reporting) et OBFF (Optimized Buffer Flush/Fill).
Vecteur d'interruption MSI-X par file d'attente	1 par file d'attente RSS, 1 par NetQueue, 1 par file d'attente VM (Virtual Machine Queue, VMQ).
Déchargement de la somme de contrôle (CO, Checksum Offload) IP	Prise en charge pour les côtés transmission/réception.
Déchargement de la somme de contrôle (CO, Checksum Offload) TCP	Prise en charge pour les côtés transmission/réception.
Déchargement de la somme de contrôle (CO, Checksum Offload) UDP	Prise en charge pour les côtés transmission/réception.
Déchargement important à l'émission (LSO, Large Send Offload) TCP NDIS	Prise en charge pour LSOV1 et LSOV2.
Réception de la fusion de segment (RSC, Receive Segment Coalescing) NDIS	Prise en charge pour les environnements Windows.
TSO (Déchargement de la segmentation TCP)	Prise en charge pour les environnements Linux et VMware.
Déchargement important à la réception (LRO, Large Receive Offload)	Prise en charge pour les environnements Linux et VMware.
Déchargement générique à la réception (GRO, Generic Receive Offload)	Prise en charge pour les environnements Linux et VMware.
Mise à l'échelle côté réception (RSS, Receive Side Scaling)	Prise en charge pour les environnements Windows, Linux et VMware. Jusqu'à 8 files d'attente/ports pris(es) en charge pour RSS.

**Tableau 13 : Fonctions de l'interface hôte (suite)**

<b>Fonction</b>	<b>Données</b>
Répartition En-tête/Charge utile (Header-Payload Split)	Permet à la pile logicielle TCP/IP de recevoir des paquets TCP/IP avec un en-tête et des données de charge utile répartis sur plusieurs mémoires tampons distinctes. Prend en charge les environnements Windows, Linux et VMware.
Jumbo Frames (Trames Jumbo)	Prise en charge.
Initialisation iSCSI	Prise en charge.
Partitionnement de la carte réseau (NPAR)	Prend en charge jusqu'à huit fonctions physiques (PF) par port, ou jusqu'à 16 PF par mémoire morte. Cette option est configurable dans NVRAM.
RDMA over Converged Ethernet (RoCE)	La BCM5741X prend en charge le protocole RoCE v1/v2 pour Windows, Linux et VMware.
Data Center Bridging (DCB)	La BCM5741X prend en charge les protocoles DCBX (IEEE et spécification CEE), PFC et AVB.
NCSI (Network Controller Sideband Interface)	Prise en charge.
Réseau local de réveil (WOL)	Pris en charge sur la rNDC avec interfaces 10GBase-T, SFP+ et SFP28.
Initialisation PXE	Prise en charge.
Démarrage UEFI	Prise en charge.
Contrôle de flux (Pause)	Prise en charge.
Négociation automatique	Prise en charge.
VLAN 802.1q	Prise en charge.
Modération de l'interruption	Prise en charge.
Filtres MAC/VLAN	Prise en charge.

## Fonctions de virtualisation

Le [Tableau 14](#) répertorie les fonctions de virtualisation du système NetXtreme-E.

**Tableau 14 : Fonctions de virtualisation**

<b>Fonction</b>	<b>Données</b>
Linux KVM Multiqueue	Prise en charge.
VMware NetQueue	Prise en charge.
NDIS VMQ (Virtual Machine Queue, File d'attente de machines virtuelles)	Prise en charge.
VXLAN (Virtual eXtensible LAN) - Conscience des déchargements sans état (déchargements de la somme de contrôle IP/UDP/TCP)	Prise en charge.
Encapsulation générique de routage (GRE, Generic Routing Encapsulation) - Conscience des déchargements sans état (déchargements de la somme de contrôle IP/UDP/TCP)	Prise en charge.

**Tableau 14 : Fonctions de virtualisation (suite)**

<b>Fonction</b>	<b>Données</b>
Virtualisation de réseaux via l'encapsulation générique de routage (NVGRE, Network Virtualization using Generic Routing Encapsulation) - Conscience des déchargements sans état	Prise en charge.
IP-in-IP - Conscience des déchargements sans état (déchargements de la somme de contrôle IP/UDP/TCP)	Prise en charge
SR-IOV v1.0	128 fonctions virtuelles (VF) pour les systèmes d'exploitation invités (SE invité) par périphérique. Le vecteur MSI-X par VF est défini sur 16.
Port de vecteur MSI-X	Valeur par défaut : 74 par port (configuration à deux ports). 16 par VF ; configurable dans HII et CCM.

## VXLAN

Un réseau local virtuel eXtensible (VXLAN), défini dans IETF RFC 7348, est utilisé pour répondre aux besoins en réseaux superposés des centres de données virtualisés hébergeant plusieurs locataires. VXLAN est un schéma de tunnelage ou une superposition de couche 2 via un réseau de couche 3. Seules les machines virtuelles placées dans le même segment VXLAN peuvent communiquer les unes avec les autres.

## NVGRE/GRE/IP-in-IP/Geneve

La virtualisation réseau via GRE (NVGRE), définie dans IETF RFC 7637, est similaire à un réseau VXLAN.

## Déchargements sans état

### RSS

La mise à l'échelle côté réception (RSS, Receive Side Scaling) utilise un algorithme Toeplitz à 4 correspondances de tuple sur les trames reçues, puis les transmet à une UC déterminée pour le traitement des trames. Cela permet un traitement simplifié des trames et équilibre l'utilisation de l'UC. Une table d'indirection est utilisée pour mapper le flux à une UC.

Une mise à l'échelle côté réception (RSS) symétrique permet le mappage des paquets d'un flux TCP ou UDP donné à la même file d'attente de réception.

### TPA

L'agrégation transparente des paquets (TPA, Transparent Packet Aggregation) est une technique où les trames reçues des 4 mêmes trames correspondantes aux tuples sont agrégées, puis indiquées à la pile réseau. Chaque entrée du contexte TPA est identifiée par les 4 tuples : IP source, IP de destination, port TCP source et port TCP de destination. La TPA améliore les performances du système en réduisant les interruptions du trafic réseau et en réduisant la charge du processeur.

## Répartition En-tête/Charge utile (Header-Payload Split)

La répartition En-tête/Charge utile est une fonction qui permet à la pile logicielle TCP/IP de recevoir des paquets TCP/IP avec un en-tête et des données de charge utile répartis sur plusieurs mémoires tampons distinctes. Cette fonction est prise en charge dans les environnements Windows et Linux. Cette fonction présente les avantages suivants :

- La fonction de répartition En-tête/Charge utile permet une mise en cache compacte et efficace des en-têtes de paquet dans les caches de l'UC hôte. Cela peut entraîner une amélioration des performances TCP/IP côté réception.
- La répartition En-tête/Charge utile permet à la pile TCP/IP hôte d'effectuer des opérations de basculement de pages et zéro copie. Cela peut améliorer encore davantage les performances du chemin de réception.

## Déchargement de la fragmentation UDP

Le déchargement de la fragmentation UDP (UFO) est une fonction qui permet à la pile logicielle de décharger la fragmentation des datagrammes UDP/IP dans des paquets UDP/IP. Cette fonction est uniquement prise en charge dans un environnement Linux. La fonction UFO peut présenter l'avantage suivant :

- La fonction UFO permet à la carte NIC de traiter la fragmentation d'un datagramme UDP dans des paquets UDP/IP. Résultat : la durée UC du traitement UDP/IP est réduite côté transmission.

## Déchargement du tunnel de transport sans état

Le déchargement du tunnel de transport sans état (STT, Stateless Transport Tunnel Offload) est une encapsulation de tunnel qui permet d'activer des réseaux superposés dans des centres de données virtualisés. Le déchargement STT utilise une encapsulation IP avec un en-tête de type TCP. Aucun état de connexion TCP n'est associé au tunnel, c'est pourquoi le déchargement STT est sans état. Le commutateur virtuel ouvert (OVS, Open Virtual Switch) utilise le déchargement STT.

Une trame STT contient l'en-tête et la charge utile de la trame STT. La charge utile de la trame STT est une trame Ethernet non identifiée. La charge utile encapsulée et l'en-tête de la trame STT sont traités comme la charge utile TCP et un en-tête de type TCP. L'en-tête IP (IPv4 ou IPv6) et l'en-tête Ethernet sont créés pour chaque segment STT transmis.

## Prise en charge de multiples files d'attente pour SE

### NDIS VMQ

La fonction NDIS VMQ (File d'attente de machines virtuelles) est prise en charge par Microsoft pour améliorer les performances du réseau Hyper-V. La fonction VMQ prend en charge la classification des paquets selon l'adresse MAC de destination afin de renvoyer les paquets reçus vers différentes files d'attente d'exécution. Cette classification des paquets associée à la possibilité d'envoyer des paquets DMA directement vers la mémoire d'une machine virtuelle permet la mise à l'échelle des machines virtuelles sur plusieurs processeurs.

Reportez-vous à la section « [Propriétés avancées du pilote Windows et messages du journal d'événements](#) » [page 35](#) pour plus d'informations sur la fonction VMQ.

## VMWare NetQueue

La fonction VMwareNetQueue est similaire à la fonction Microsoft NDIS VMQ de Microsoft. La fonction NetQueue prend en charge la classification des paquets selon l'adresse MAC de destination et le réseau VLAN afin de renvoyer les paquets reçus vers différentes files NetQueue. Cette classification des paquets associée à la possibilité d'envoyer des paquets DMA directement vers la mémoire d'une machine virtuelle permet la mise à l'échelle des machines virtuelles sur plusieurs processeurs.

## KVM/Xen Multiqueue

La fonction KVM/Multiqueue renvoie les trames vers différentes files d'attente de la pile hôte en classant la trame entrante via le traitement de l'adresse MAC de destination et/ou la balise 802.1Q VLAN du paquet reçu. La classification associée à la possibilité d'envoyer des trames DMA directement vers la mémoire d'une machine virtuelle permet la mise à l'échelle des machines virtuelles sur plusieurs processeurs.

## Matrice de prise en charge des configurations SR-IOV

- VF Windows via l'hyperviseur Windows
- Windows VF Windows et VF Linux via l'hyperviseur VMware
- VF Linux via Linux KVM

## SR-IOV

PCI-SIG offre une prise en charge optionnelle de la spécification SR-IOV (Single-Root IO Virtualization). SR-IOV est conçu pour permettre à la machine virtuelle d'accéder directement au périphérique via les fonctions virtuelles (VF). La fonction physique (PF) de la carte NIC est divisée en plusieurs fonctions virtuelles et chaque VF est présentée comme une PF aux machines virtuelles.

SR-IOV utilise la fonctionnalité IOMMU pour convertir les adresses virtuelles PCI-E en adresses physiques à l'aide d'une table de conversion.

Les nombres de fonctions physiques (PF) et de fonctions virtuelles (VF) sont gérés via le menu HII UEFI, CCM et les configurations NVRAM. SRIOV peut être pris en charge parallèlement au mode NPAR.

## Partitionnement réseau (NPAR)

La fonction de partitionnement réseau (NPAR) permet à un port d'interface réseau physique de s'afficher pour le système comme plusieurs fonctions de périphérique réseau. Lorsque le mode NPAR est activé, le périphérique NetXtreme-E est énuméré en plusieurs fonctions physiques (PF) PCIe. Chaque PF ou « partition » se voit affecter un ID de fonction PCIe distinct lors de la première mise sous tension. La définition PCIe d'origine autorisait huit PF par périphérique. Pour les systèmes compatibles ARI (Alternative Routing-ID), les cartes Broadcom NetXtreme-E prennent en charge jusqu'à 16 PF par périphérique. Chaque partition dispose de son propre espace de configuration et de ses propres adresses BAR et MAC, ce qui lui permet de fonctionner indépendamment. Les partitions prennent en charge l'affectation directe à des machines virtuelles, à des VLAN, etc., comme toute autre interface physique.





**Remarque :** Sur la page **Configuration système > Paramètres du périphérique [Périphérique Broadcom 5741x] > Configuration du niveau du périphérique**, l'utilisateur peut activer **NParEP** pour permettre à la carte NXE de prendre en charge jusqu'à 16 PF par périphérique. Pour des périphériques à 2 ports, jusqu'à huit PF peuvent être prises en charge pour chaque port.

## RDMA over Converge Ethernet – RoCE

Le protocole RoCE (Remote Direct Memory Access (RDMA) over Converged Ethernet) est une fonction complète de déchargement matériel de BCM5741X qui permet d'activer la fonction RDMA sur un réseau Ethernet. La fonctionnalité RoCE est disponible sur les applications en mode utilisateur et en mode noyau. Les fonctions physiques (PF) RoCE et les fonctions virtuelles (VF) SRIOV sont disponibles en mode monofonction et en mode multifonction (mode de partitionnement NIC). Broadcom prend en charge le protocole RoCE sous Windows, Linux et VMWare.

Veuillez vous reporter à l'un des liens suivants pour connaître la prise en charge du protocole RDMA pour chaque système d'exploitation :

### 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](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>

## Combinaisons prises en charge

Les sections suivantes décrivent les combinaisons de fonctionnalités prises en charge pour ce périphérique.

### NPAR, SR-IOV et RoCE

Le **Tableau 15** décrit les combinaisons de fonctionnalités prises en charge pour les protocoles NPAR, SR-IOV et RoCE.

**Tableau 15 : NPAR, SR-IOV et RoCE**

<b>Fonction logicielle</b>	<b>Remarques</b>
NPAR	Jusqu'à 8 PF ou 16 PF
SR-IOV	Jusqu'à 128 VF (au total par puce)
RoCE sur PF	Jusqu'à 4 PF

**Tableau 15 : NPAR, SR-IOV et RoCE (suite)**

<b>Fonction logicielle</b>	<b>Remarques</b>
RoCE sur VF	Valide pour les VF attachées à des PF compatibles RoCE
Système d'exploitation hôte	Linux, Windows, ESXi (pas de prise en charge vRDMA)
Système d'exploitation invité	Linux et Windows
DCB	Jusqu'à deux COS par port avec mémoire réservée non partagée

## **NPAR, SR-IOV et DPDK**

Le [Tableau 16](#) décrit les combinaisons de fonctionnalités prises en charge pour les protocoles NPAR, SR-IOV et DPDK.

**Tableau 16 : NPAR, SR-IOV et DPDK**

<b>Fonction logicielle</b>	<b>Remarques</b>
NPAR	Jusqu'à 8 PF ou 16 PF
SR-IOV	Jusqu'à 128 VF (au total par puce)
DPDK	Uniquement pris en charge en tant que VF
Système d'exploitation hôte	Linux
Système d'exploitation invité	DPDK (Linux)

## **Combinaisons non prises en charge**

La combinaison des protocoles NPAR, SR-IOV, RoCE et DPDK n'est pas prise en charge.

---

## Installation du matériel

### Mesures de sécurité



**Attention** : La carte est installée dans un système fonctionnant à une tension risquant d'être mortelle. Avant de retirer le capot du système, vous devez prendre les mesures suivantes afin de vous protéger et d'éviter tout risque de destruction des composants du système :

- Enlevez les objets métalliques ou les bijoux que vous portez aux mains et aux poignets.
- Veillez à utiliser uniquement des outils isolés ou non conducteurs.
- Vérifiez que le système est hors tension et que la prise est débranchée avant de toucher tout composant interne.
- Installez ou enlevez les cartes dans un environnement exempt d'électricité statique. Le port d'un bracelet antistatique correctement relié à la terre ou de tout autre dispositif antistatique ainsi que l'utilisation d'un tapis antistatique sont vivement conseillés.

### Configuration système

Avant d'installer la carte NetXtreme-E Ethernet de Broadcom, assurez-vous que la configuration du système répond aux conditions requises par votre système d'exploitation.

### Configuration matérielle requise

La liste ci-après répertorie les éléments de la configuration matérielle requise :

- Systèmes Dell 13G qui répondent à la configuration requise pour le système d'exploitation.
- Systèmes Dell 13G qui prennent en charge des cartes NetXtreme-E Ethernet.
- Un logement PCI-E Gen 3 x8 libre ou un logement PCIE Gen3 rNDC libre pour une carte réseau de format rNDC.
- 4 Go de mémoire ou plus (32 Go ou plus sont recommandés pour les performances du débit nominal du réseau et des applications de virtualisation).

### Liste de vérification avant l'installation

Reportez-vous à la liste suivante avant d'installer le périphérique NetXtreme-E.

1. Assurez-vous que le serveur répond bien à la configuration matérielle et logicielle requise décrite dans la section « [Configuration système](#) ».
2. Vérifiez que le serveur utilise le BIOS le plus récent.
3. Si le système est actif, fermez-le.
4. Une fois le système arrêté, coupez l'alimentation secteur et débranchez la prise de l'ordinateur.
5. En tenant la carte par les bords, enlevez-la de son emballage d'expédition et placez-la sur une surface antistatique.
6. Vérifiez que la carte ne présente aucun signe de détérioration, en particulier sur le connecteur de bord. N'essayez jamais d'installer une carte endommagée.

## Installation de la carte

Les instructions suivantes concernent l'installation de la carte NetXtreme-E Ethernet de Broadcom (NIC complémentaire) sur la plupart des serveurs. Reportez-vous aux manuels fournis avec le serveur pour toute précision sur la réalisation de ces tâches sur ce serveur spécifique.

1. Consultez les sections « [Mesures de sécurité](#) » page 27 et « [Liste de vérification avant l'installation](#) » avant d'installer la carte. Assurez-vous que le système est hors tension et débranché. Veillez à ce que les procédures de mise à la terre aient été correctement suivies.
2. Ouvrez le boîtier du système et sélectionnez un logement PCI Express Gen 3 x8 vide.
3. Retirez la lame d'obturation blanche du logement.
4. Alignez le bord du connecteur de la carte sur le logement du connecteur dans le système.
5. Fixez la carte avec son clip ou sa vis.
6. Refermez le boîtier du système et détachez les dispositifs antistatiques.



**Remarque :** Pour les cartes NIC de format rNDC, recherchez un logement RNDC libre ou remplacez le rNDC par défaut existant par le rNDC NetXtreme.

## Connexion des câbles réseau

Les commutateurs Ethernet de Dell sont fournis avec des ports SFP+/SFP28/QSFP28 qui prennent en charge jusqu'à 100 Gbit/s. Ces ports de 100 Gbit/s peuvent être divisés en 4 ports SFP28 de 25 Gbit/s. Les ports QSFP peuvent être connectés à des ports SFP28 avec 4 câbles adaptateurs 25G SFP28.

## Câbles et modules pris en charge

**Tableau 17 : Câbles et modules pris en charge**

<i>Module optique</i>	<i>Référence Dell</i>	<i>Cartes</i>	<i>Description</i>
FTLX8571D3BCL-DL	3G84K	BCM957404A4041DLPC, BCM957404A4041DC, BCM957402A4020DLPC, BCM957402A4020DC	Emetteur-récepteur SFP+ 10 Gbit/s 850 nm multimode
FCLF-8521-3	8T47V	BCM957406A4060DLPC, BCM957406A4060DC	Emetteur-récepteur SFP 1000Base-T en cuivre
FTLF8536P4BCL	P7D7R	BCM957404A4041DLPC, BCM957404A4041DC	Emetteur-récepteur SFP+ à courte longueur d'onde 25 Gbit/s
Méthode DM7051	PGYJT	BCM57402X, BCM57404X, BCM57412X, BCM57414X	Emetteur-récepteur SFP+ vers 10GBASE-T
FTLX1471D3BCL-FC	RN84N	BCM57402X, BCM57404X, BCM57412X, BCM57414X	Emetteur-récepteur SFP28 à 25 Gbit/s
FTLX8574D3BNL	N8TDR	BCM57402X, BCM57404X, BCM57412X, BCM57414X	Plage de température étendue à 85 °C Emetteur-récepteur SFP+ à 10 Gbit/s
FTLF8536P4BNL-FC	HHHHC	BCM57402X, BCM57404X, BCM57412X, BCM57414X	Plage de température étendue à 85 °C Emetteur-récepteur SFP+ à 10 Gbit/s

**Tableau 17 : Câbles et modules pris en charge (suite)**

<b>Module optique</b>	<b>Référence Dell</b>	<b>Cartes</b>	<b>Description</b>
FTLX8574D3BCL-FC ou PLRXPLSCS43811	WTRD1	BCM57402X, BCM57404X, BCM57412X, BCM57414X	Emetteur-récepteur SFP+ SR à 10 Gbit/s

**Remarque :**

1. Les câbles d'attache directe (DAC) conformes aux normes IEEE peuvent être connectés à la carte.
2. Les composants Dell HHHHC et N8TDR sont requis pour l'utilisation du modèle BCM957414M4140D.

**Cuivre**

Les cartes BCM957406AXXXX, BCM957416AXXXX et BCM957416XXXX disposent de deux connecteurs RJ-45 utilisés pour relier le système à un segment de fil cuivré CAT 6E Ethernet.

**SFP+**

Les cartes BCM957402AXXXX, BCM957412AXXXX et BCM957412MXXXX disposent de deux connecteurs SFP+ utilisés pour relier le système à un commutateur Ethernet 10 Gbit/s.

**SFP28**

Les cartes BCM957404AXXXX, BCM957414AXXXX et BCM957414MXXXX disposent de deux connecteurs SFP28 utilisés pour relier le système à un commutateur Ethernet 100 Gbit/s.

---

## Progiciels et installation

Reportez-vous aux sections suivantes pour plus d'informations sur les progiciels et l'installation.

### Systèmes d'exploitation pris en charge

Le [Tableau 18](#) présente la liste des systèmes d'exploitation pris en charge.

**Tableau 18 : Liste des systèmes d'exploitation pris en charge**

<b>SE type</b>	<b>Distribution</b>
Windows	Windows 2012 R2 ou supérieur
Linux	Redhat 6.9, Redhat 7.1 ou supérieur SLES 11 SP 4, SLES 12 SP 2 ou supérieur
VMWare	ESXi 6.0 U3 ou supérieur

### Installation des pilotes

Les sections suivantes présentent l'installation des pilotes.

#### Windows

##### Dell DUP

Les pilotes du contrôleur NetXtreme E-Series de Broadcom peuvent être installés à l'aide du progiciel DUP de pilote. Le programme d'installation est fourni au format exécutable x64.

##### Installation de l'interface graphique utilisateur

Lorsque le fichier est exécuté, une boîte de dialogue s'affiche, invitant l'utilisateur à saisir des données. Le programme d'installation prend en charge l'option Pilote seul.

##### Installation silencieuse

L'exécutable peut être lancé en mode silencieux via la commande indiquée ci-dessous.

Exemples :

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

##### Installation des fichiers INF

Dell DUP est utilisé pour installer les pilotes des contrôleurs Broadcom NetXtreme-E Ethernet. Utilisez la commande suivante pour extraire les fichiers INF du pilote à partir du progiciel Dell DUP :

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

Une fois les fichiers extraits, l'installation INF est exécutée via la fonction de « pilote de mise à niveau » à l'aide du Gestionnaire de périphériques (devmgmt.msc). Ouvrez le Gestionnaire de périphériques, sélectionnez la carte NIC souhaitée, cliquez avec le bouton droit de la souris et sélectionnez le pilote de mise à niveau pour la mettre à jour.

## Linux

Les pilotes Linux sont fournis aux formats RPM, KMP et de code source. Pour générer le pilote de périphérique à partir du code source via Linux, reportez-vous à l'exemple suivant :

1. Connectez-vous au système Linux en tant qu'utilisateur racine.
2. Exécutez une commande scp ou cp sur le fichier tar du pilote, sur le système Linux. Exemple type :

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

3. Exécutez la commande suivante :

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

4. Exécutez la commande suivante :

```
cd bnxt_en-<version>
```

5. Exécutez la commande suivante :

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

Pour la fonctionnalité RDMA, installez les pilotes bnxt\_en et bnxt\_re. Utilisez netxtreme-bnxt\_en-<version>.tar.gz au lieu de bnxt\_en-<version>.tar.gz.

## Installation du module

### RHEL

L'image du pilote peut être installée à l'aide de l'une des options suivantes :

- Montez l'image bnxt\_en-x.x.x-rhelYuZ-x86\_64-dd.iso à l'aide de la console virtuelle iDRAC de Dell.
- Montez l'image bnxt\_en-x.x.x-rhelYuZ-x86\_64-dd.iso à partir d'un CD/DVD.
- Copiez l'image bnxt\_en-x.x.x-rhelYuZ-x86\_64-dd.iso sur un périphérique USB et montez le périphérique.

Démarrez l'installation du SE, puis appuyez sur la touche de tabulation et entrez « Linux dd ». Continuez l'installation jusqu'à ce que le disque du pilote soit demandé, puis sélectionnez le pilote bnxt\_en.

### SLES

L'image du pilote peut être installée à l'aide de l'une des options suivantes :

- Montez l'image bnxt\_en-x.x.x-rhelYuZ-x86\_64-dd.iso à l'aide de la console virtuelle iDRAC de Dell.
- Montez l'image bnxt\_en-x.x.x-rhelYuZ-x86\_64-dd.iso à partir d'un CD/DVD.
- Extrayez l'image bnxt\_en-x.x.x-rhelYuZ-x86\_64-dd.iso, copiez le contenu sur un périphérique USB et montez ce périphérique USB.

## Commandes Ethtool de Linux



**Remarque :** Dans le [Tableau 19](#), ethX doit être remplacé par le véritable nom de l'interface.

**Tableau 19 : Commandes Ethtool de Linux**

Commande	Description
ethtool -s ethX speed 25000 autoneg off	Définit la vitesse. Si la liaison est active sur un port, le pilote ne permettra pas que l'autre port soit défini sur une vitesse incompatible.
ethtool -i ethX	La sortie inclut la version du progiciel, la version du BIOS de la carte NIC (code d'amorçage).
ethtool -k ethX	Affiche les fonctions de déchargement.
ethtool -K ethX tso off	Désactive TSO.
ethtool -K ethX gro off lro off	Désactivez GRO/LRO.
ethtool -g ethX	Affiche les tailles d'anneau.
ethtool -G ethX rx N	Définit les tailles d'anneau.
ethtool -S ethX	Affiche les statistiques.
ethtool -l ethX	Affiche le nombre d'anneaux.
ethtool -L ethX rx 0 tx 0 combined M	Définit le nombre d'anneaux.
ethtool -C ethX rx-frames N	Définit la fusion de l'interruption. Les autres paramètres pris en charge sont les suivants : rx-usecs, rx-frames, rx-usecs-irq, rx-frames-irq, tx-usecs, tx-frames, tx-usecs-irq, tx-frames-irq.
ethtool -x ethX	Affiche la table d'indirection de hachage du flux RSS et la clé RSS.
ethtool -s ethX autoneg on speed 10000 duplex full	Activez la négociation automatique (voir <a href="#">« Configuration de la négociation automatique » page 41</a> pour plus de détails)
ethtool --show-eee ethX	Affiche l'état EEE.
ethtool --set-eee ethX eee off	Désactive EEE.
ethtool --set-eee ethX eee on tx-lpi off	Active EEE, mais désactiver LPI.
ethtool -L ethX combined 1 rx 0 tx 0	Désactive RSS. Définissez les canaux combinés sur 1.
ethtool -K ethX ntuple off	Désactive le RFS accéléré en désactivant les filtres ntuple.
ethtool -K ethX ntuple on	Active le RFS accéléré.
Ethtool -t ethX	Effectue divers autotests de diagnostic.
echo 32768 > /proc/sys/net/core/rps_sock_flow_entries	Active le RFS pour l'anneau X.
echo 2048 > /sys/class/net/ethX/queues/rx-X/rps_flow_cnt	
sysctl -w net.core.busy_read=50	Cette commande définit la durée accordée au périphérique pour lire l'anneau de réception sur 50 usec. Pour les applications de socket en attente de données, l'utilisation de cette méthode peut généralement réduire la latence de 2 ou 3 usecs, au détriment d'une meilleure utilisation de l'UC.



**Tableau 19 : Commandes Etool de Linux (suite)**

Commande	Description
echo 4 > /sys/bus/pci/devices/0000:82:00.0/sriov_numvfs	Active SR-IOV avec quatre VF sur le bus 82, périphérique 0 et fonction 0.
ip link set ethX vf 0 mac 00:12:34:56:78:9a	Définit l'adresse MAC de la VF.
ip link set ethX vf 0 state enable	Définit l'état de la liaison VF pour VF 0.
ip link set ethX vf 0 vlan 100	Définit VF 0 avec ID VLAN 100.

## VMware

Les pilotes ESX sont fournis au format VIB standard VMware.

1. Pour installer les pilotes Ethernet et RDMA, exécutez les commandes suivantes :

```
$ 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. Le système doit être redémarré pour que le nouveau pilote prenne effet.

Le [Tableau 20](#) présente d'autres commandes VMware utiles.



**Remarque :** Dans le [Tableau 20](#), vmnicX doit être remplacé par le véritable nom de l'interface.



**Remarque :** \$ kill -HUP \$(cat /var/run/vmware/vmkdevmgr.pid) Cette commande doit être utilisée après la commande vmkload\_mod bnxtnet pour que le module puisse s'afficher.

**Tableau 20 : Commandes VMware**

Commande	Description
esxcli software vib list  grep bnx	Répertorie les VIB installés pour savoir si le pilote bnxt a été correctement installé.
esxcfg-module -l bnxtnet	Imprime les informations de module qui s'affichent à l'écran.
esxcli network get -n vmnicX	Affiche les propriétés vmnicX.
esxcfg-module -g bnxtnet	Imprime les paramètres du module.
esxcfg-module -s 'multi_rx_filters=2 disable_tap=0 max_vfs=0,0 RSS=0'	Définit les paramètres du module.
vmkload_mod -u bnxtnet	Décharge le module bnxtnet.
vmkload_mod bnxtnet	Charge le module bnxtnet.
esxcli network nic set -n vmnicX -D full -S 25000	Définit la vitesse et le mode duplex de vmnicx.
esxcli network nic down -n vmnicX	Désactive vmnicx.
esxcli network nic up -n vmnic6	Activer vmnicx.
bnxtnetcli -s -n vmnic6 -S "25000"	Définit la vitesse de liaison. La commande bnxtnetcli est nécessaire pour les anciennes versions ESX afin de permettre la prise en charge du paramètre de débit 25G.

## Mise à jour du micrologiciel

Le micrologiciel de la carte NIC peut être mis à jour via l'une des méthodes suivantes :

- Via le progiciel de mise à jour Dell (DUP) lorsque le système est à l'état SE démarré. Cette méthode s'applique uniquement aux systèmes d'exploitation Windows et Linux.
- Via Dell iDRAC - Lifecycle Controller. Cette méthode peut être utilisée quel que soit le système d'exploitation. Si le système exécute VMware, utilisez Lifecycle Controller pour mettre à jour le micrologiciel.

Reportez-vous à la page de support produit, à l'adresse <http://www.dell.com/support>

## Dell Update Package

Les sections suivantes abordent l'utilisation du progiciel Dell Update Package (DUP) :

### Windows

Le micrologiciel du contrôleur Broadcom NetXtreme E-Series peut être mis à niveau à l'aide du progiciel Dell DUP. L'exécutable est fourni au format exécutable Windows X64 standard. Double-cliquez sur le fichier pour l'exécuter.

Les progiciels DUP peuvent être téléchargés à partir du site <http://support.dell.com>

### Linux

Le progiciel Dell Linux DUP est fourni au format exécutable x86\_64. Utilisez la commande Linux `chmod` standard pour mettre à jour l'autorisation d'exécution et lancez l'exécutable. Reportez-vous à l'exemple suivant :

1. Connectez-vous à Linux.
2. Exécutez une commande `scp` ou `cp` sur l'exécutable du progiciel DUP sur le système de fichiers.  
Exemple type :

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

3. Exécutez la commande suivante :

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

4. Exécutez la commande suivante :

```
./Network_Firmware_<version>.BIN
```

Un redémarrage est nécessaire pour activer le nouveau micrologiciel.

## Propriétés avancées du pilote Windows et messages du journal d'événements

### Propriétés avancées du pilote

Les propriétés avancées du pilote Windows sont affichées dans le [Tableau 21](#).

**Tableau 21 : Propriétés avancées du pilote Windows**

<b>Clé pilote</b>	<b>Paramètres</b>	<b>Description</b>
Déchargement de tâches encapsulées	Activer ou Désactiver	Utilisé pour la configuration du déchargement de tâches encapsulées NVGRE.
Energy Efficient Ethernet	Activer ou Désactiver	La fonction EEE est activée pour les ports en cuivre et désactivée pour les ports SFP+ ou SFP28. Elle est uniquement activée pour la carte BCM957406A4060.
Contrôle de flux	TX ou RX ou TX/RX activé	Configure le contrôle de flux sur RX ou Tx ou sur les deux côtés.
Modération de l'interruption	Activer ou Désactiver	La valeur par défaut est Activé. Permet le traitement par lots des trames en réduisant le temps UC.
Paquet Jumbo	1514 ou 4088 ou 9014	Taille de paquet Jumbo.
Déchargement important à l'émission V2 (IPV4)	Activer ou Désactiver	LSO pour IPV4.
Déchargement important à l'émission V2 (IPV6)	Activer ou Désactiver	LSO pour IPV6.
Adresse administrée localement	Adresse MAC saisie par l'utilisateur.	Remplace l'adresse matérielle MAC par défaut après le redémarrage du SE.
Nombre maximal de files d'attente RSS	2, 4 ou 8.	La valeur par défaut est 8. Permet à l'utilisateur de configurer des files d'attente RSS.
Priorité et VLAN	Priorité et VLAN désactives, Priorité activée, VLAN activé, Priorité et VLAN activés.	La valeur par défaut est Activé. Utilisée pour configurer 802.1Q et 802.1P.
Tampons de réception (0=Auto)	Incrément de 500.	Auto est le paramètre par défaut.
Mise à l'échelle côté réception (RSS)	Activer ou Désactiver.	La valeur par défaut est Activé.
Réception de la fusion de segment (IPV4)	Activer ou Désactiver.	La valeur par défaut est Activé.
Réception de la fusion de segment (IPV6)	Activer ou Désactiver.	La valeur par défaut est Activé.

**Tableau 21 : Propriétés avancées du pilote Windows (suite)**

<b>Clé pilote</b>	<b>Paramètres</b>	<b>Description</b>
Profil d'équilibrage de charge RSS	Mise à l'échelle statique NUMA, Processeur le plus proche, Processeur statique le plus proche, Mise à l'échelle conservatrice, Mise à l'échelle NUMA.	La valeur par défaut est Mise à l'échelle statique NUMA.
Vitesse et Mode duplex	1 Gbit/s, 10 Gbit/s, 25 Gbit/s ou Négociation automatique.	Les ports en cuivre 10 Gbit/s peuvent négocier automatiquement des vitesses, tandis que les ports 25 Gbit/s sont définis sur des vitesses imposées.
SR-IOV	Activer ou Désactiver.	La valeur par défaut est Activé. Ce paramètre fonctionne conjointement au paramètre HW configuré SR-IOV et BIOS configuré SR-IOV.
Déchargement de la somme de contrôle TCP/UDP IPV4	TX/RX activé, TX activé ou RX activé ou déchargement désactivé.	La valeur par défaut est RX/TX activé.
Déchargement de la somme de contrôle TCP/UDP IPV6	TX/RX activé, TX activé ou RX activé ou déchargement désactivé.	La valeur par défaut est RX/TX activé.
Transmit Buffers (0=Auto) (Tampons de transmission)	Incrément de 50.	La valeur par défaut est Auto.
File d'attente de machines virtuelles (VMQ)	Activer ou Désactiver.	La valeur par défaut est Activé.
ID VLAN	Nombre configurable par l'utilisateur.	La valeur par défaut est 0.

## Messages du journal des événements

Le [Tableau 22](#) présente les messages consignés par le pilote Windows NDIS dans le journal des événements.

**Tableau 22 : Messages du journal des événements sous Windows**

<b>ID de message</b>	<b>Commentaire</b>
0x0001	Echec de l'allocation de mémoire.
0x0002	Liaison inactive détectée.
0x0003	Liaison active détectée.
0x0009	Liaison 1000 complète.
0x000A	Liaison 2500 complète.
0x000b	Initialisation réussie.
0x000c	Réinitialisation du miniport.
0x000d	Echec de l'initialisation.
0x000E	Liaison 10 Gbits réussie.

**Tableau 22 : Messages du journal des événements sous Windows (suite)**

<b>ID de message</b>	<b>Commentaire</b>
0x000F	Echec de la liaison de la couche du pilote.
0x0011	Echec de la définition des attributs.
0x0012	Echec du DMA par ventilation-regroupement.
0x0013	Echec de l'initialisation de la file d'attente par défaut.
0x0014	Version de micrologiciel incompatible.
0x0015	Interruption unique.

**Tableau 23 : Messages du journal des événements**

0x0016	Le micrologiciel n'a pas répondu dans le temps alloué.
0x0017	Le micrologiciel a renvoyé un état d'échec.
0x0018	Le micrologiciel se trouve dans un état inconnu.
0x0019	Le module optique n'est pas pris en charge.
0x001A	Sélection d'un débit incompatible entre le port 1 et le port 2. Les débits de liaison indiqués sont corrects mais ne correspondent peut-être pas au paramétrage de la vitesse et du mode duplex.
0x001B	Sélection d'un débit incompatible entre le port 1 et le port 2. La configuration de la liaison est interdite.
0x001C	Network controller configured for 25Gb full-duplex link. (Le contrôleur réseau est configuré pour une liaison duplex intégral de 10 Gbit.)
0x0020	L'initialisation de la prise en charge du protocole RDMA a échoué.
0x0021	Le micrologiciel RDMA du périphérique est incompatible avec ce pilote.
0x0022	La taille Doorbell BAR est insuffisante pour RDMA.
0x0023	Echec du redémarrage RDMA lors de réinitialisation du périphérique.
0x0024	Echec du redémarrage RDMA lors de la mise sous tension du système.
0x0025	Echec du redémarrage RDMA. Ressources insuffisantes.
0x0026	RDMA n'est pas activé dans le micrologiciel.
0x0027	Echec du démarrage, aucune adresse MAC n'est définie.
0x0028	Arrêt de transmission détecté. Le contrôle de flux TX est maintenant désactivé.

---

# Regroupement

## Windows

Les périphériques Broadcom NetXtreme-E installés sur les plates-formes Dell peuvent participer à la fonction de regroupement des cartes réseau via la solution d'association de Microsoft. Reportez-vous à la documentation mise à disposition par Microsoft sur la page Web suivante :

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

Microsoft LBFO est un pilote d'association natif qui peut être utilisé dans le SE Windows. Le pilote d'association fournit également des fonctions de balisage VLAN.

## Linux

La fonction de création de liaisons Linux est utilisée pour le regroupement sous Linux. Le principe est de charger le pilote de liaison, puis d'ajouter des membres d'équipe à la liaison responsable du chargement/de l'équilibrage du trafic.

Pour configurer la fonction de création de liaisons Linux, procédez comme suit :

1. Exécutez la commande suivante :

```
modprobe bonding mode="balance-alb". Une interface de liaison est créée.
```

2. Ajoutez des clients de liaison à l'interface de liaison. Exemple :

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

3. Affectez une adresse IPV4 pour lier l'interface via `ifconfig bond0 IPV4Address netmask NetMask up`. IPV4Address et NetMask représentent une adresse IPV4 et le masque de réseau associé.



**Remarque :** IPV4Address doit être remplacé par la véritable adresse IPV4 du réseau. NetMask doit être remplacé par le véritable masque de réseau IPV4.

4. Affectez une adresse IPV6 pour lier l'interface via `ifconfig bond0 IPV6Address netmask NetMask up`. IPV6Address et NetMask représentent une adresse IPV6 et le masque de réseau associé.



**Remarque :** IPV6Address doit être remplacé par la véritable adresse IPV6 du réseau. NetMask doit être remplacé par le véritable masque de réseau IPV6.

Reportez-vous à la documentation Linux relative à la création de liaisons pour les configurations avancées.

## Configuration au niveau du système

Reportez-vous aux sections suivantes pour plus d'informations sur la configuration des cartes réseau au niveau du système.

### Menu UEFI HII

Les contrôleurs Broadcom NetXtreme E-Series peuvent être définis pour une configuration de pré-amorçage, iSCSI et avancée telle que SR-IOV via le menu HII (Human Interface, Interface humaine).

Pour configurer les paramètres, sélectionnez F2 -> Configuration système -> Paramètres du périphérique pendant le démarrage du système. Sélectionnez la carte réseau souhaitée pour afficher et modifier la configuration.

### Page de configuration principale

Cette page affiche l'état de la liaison réseau en cours, PCI-E Bus:Device:Function, l'adresse MAC de la carte et du périphérique Ethernet.

Une carte 10GBaseT permet à l'utilisateur d'activer ou de désactiver la fonction EEE (Energy Efficient Ethernet).

### Propriétés de l'image du microprogramme

Page de configuration principale -> Les propriétés de l'image du micrologiciel affichent la version du micrologiciel qui se compose des numéros de version du contrôleur BIOS, de l'agent MBA (Multi Boot Agent), de l'interface UEFI, de la norme iSCSI et de l'utilitaire CCM (Comprehensive Configuration Management).

### Configuration du niveau du périphérique

Page de configuration principale -> La configuration du niveau du périphérique permet à l'utilisateur d'activer le mode SR-IOV, le nombre de fonctions virtuelles par fonction physique, les vecteurs MSI-X par fonction virtuelle et le nombre maximum de vecteurs MSI-X par fonction physique.

### Configuration de la carte NIC

Configuration de la carte réseau -> Le protocole d'amorçage hérité est utilisé pour sélectionner et configurer PXE, iSCSI ou désactiver le mode d'amorçage hérité. Le type d'amorçage peut être Auto, int18h (interruption 18h), int19h (interruption 19h) ou BBS.

MBA et iSCSI peuvent également être configurés via CCM. Le mode BIOS hérité utilise CCM pour la configuration. L'invite de définition du masquage peut être utilisée pour activer ou désactiver l'affichage de la bannière.

VLAN pour PXE peut être activé ou désactivé et l'ID VLAN peut être configuré par l'utilisateur. Reportez-vous à la section « [Configuration de la négociation automatique](#) » page 41 pour plus d'informations sur les options de configuration de la vitesse de liaison.

## Configuration iSCSI

La configuration de l'amorçage iSCSI peut être définie via la Page de configuration principale -> Configuration iSCSI. Les paramètres tels que IPV4 ou IPV6, initiateur iSCSI ou cible iSCSI peuvent être définis sur cette page.

Reportez-vous à la section « [Initialisation iSCSI](#) » [page 50](#) pour obtenir des informations détaillées sur cette configuration.

## CCM (Comprehensive Configuration Management)

La configuration de pré-amorçage peut être définie à l'aide de l'option de menu Comprehensive Configuration Management (CCM). Lors du POST du BIOS du système, le message de la bannière Broadcom s'affiche avec une option pour modifier les paramètres à l'aide des touches Ctrl+S. Lorsque les touches Ctrl+S sont enfoncées, une liste de périphériques est renseignée avec toutes les cartes réseau Broadcom trouvées sur le système. Sélectionnez la carte réseau de votre choix pour la configuration.

## Configuration matérielle du périphérique

Les paramètres pouvant être configurés dans cette section sont les mêmes que ceux configurés via l'option « Configuration du niveau du périphérique » du menu HII.

## Menu de configuration MBA

Les paramètres pouvant être configurés dans cette section sont les mêmes que ceux configurés via l'option « Configuration de la carte NIC » du menu HII.

## Menu principal de l'initialisation iSCSI

Les paramètres pouvant être configurés dans cette section sont les mêmes que ceux configurés via l'option « Configuration iSCSI » du menu HII.



## Configuration de la négociation automatique



**Remarque :** Dans les périphériques NPAR (partitionnement de la carte réseau) où un port est partagé par plusieurs fonctions PCI, la vitesse du port est préconfigurée et ne peut pas être modifiée par le pilote.

Le contrôleur Broadcom NetXtreme-E prend en charge les fonctions de négociation automatique suivantes :

- Négociation automatique de la vitesse de liaison
- Négociation automatique de la mise en pause/du contrôle de flux
- Négociation automatique FEC (Forward Error Correction, correction d'erreurs sans voie de retour)



**Remarque :** En ce qui concerne la négociation automatique de la vitesse de liaison, utilisez des émetteurs-récepteurs DAC ou optiques multimodes capables de prendre en charge la négociation automatique si vous utilisez des connecteurs SFP+ ou SFP28. Vérifiez que le port partenaire de liaison a été défini sur le protocole de négociation automatique correspondant. Par exemple, si le port Broadcom local est défini sur le protocole de négociation automatique IEEE 802.3by, le partenaire de liaison doit prendre en charge la négociation automatique et doit être défini sur le protocole de négociation automatique IEEE 802.3by.



**Remarque :** Pour les contrôleurs réseau NetXtreme-E double port, la combinaison de vitesses de liaison 10 Gbit/s et 25 Gbit/s n'est pas prise en charge.

Les combinaisons de vitesse de liaison prises en charge pour le contrôle NetXtreme-E à deux ports sont présentées sur le [Tableau 24 page 42](#).

Tableau 24 : Combinaisons de vitesses de liaison prises en charge

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}



**Remarque :** Utilisez un émetteur-récepteur optique pris en charge ou un câble en cuivre d'attache directe pour atteindre la vitesse de liaison 1 Gbit/s.

- P1 : configuration port 1
- P2 : configuration port 2
- AN : négociation automatique
- No AN : vitesse forcée

- {vitesse de liaison} : vitesse de liaison attendue
- AN {vitesses de liaison} : vitesses de liaison annoncées avec prise en charge de la négociation automatique.

Les vitesses de liaison attendues sur les paramètres locaux et du partenaire de liaison sont indiquées sur le [Tableau 25](#).

**Tableau 25 : Vitesses de liaison attendues**

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



**Remarque :** La vitesse de liaison 1 Gbit/s pour SFP+/SFP28 n'est actuellement pas prise en charge dans cette version.

Pour activer la négociation automatique de la vitesse de liaison, les options suivantes peuvent être activées dans le BIOS du système, dans le menu HII ou dans le CCM :



**Remarque :** La détection automatique des supports doit être activée lors de l'activation de la négociation automatique de la vitesse de liaison en cas de connexion via un module/câble d'émetteur-récepteur optique.

**BIOS système-> Paramètres du périphérique -> NetXtreme-E-> Configuration du niveau du périphérique**

## Vitesse de liaison opérationnelle

Cette option configure la vitesse de liaison utilisée par les pilotes de pré-démarrage (MBA et UEFI) (Linux, ESX), le pilote du système d'exploitation et le micrologiciel. Ce paramètre est remplacé par le paramètre de pilote dans l'état actuel du système d'exploitation. Le pilote Windows (bnxtnet.sys) utilise les paramètres de vitesse de liaison du fichier driver .inf.

## Vitesse de liaison du microprogramme

Cette option permet de configurer la vitesse de liaison utilisée par le micrologiciel lorsque le périphérique est en D3.

## Protocole de négociation automatique

Il s'agit du protocole de négociation automatique pris en charge utilisé pour négocier la vitesse de liaison avec le partenaire de liaison. Cette option doit correspondre au paramètre du protocole de négociation automatique dans le port du partenaire de liaison. La carte réseau Broadcom NetXtreme-E prend en charge les protocoles de négociation automatique suivants : IEEE 802.3by, consortiums 25G/50G et BAM 25G/50G. Par défaut, cette option est définie sur IEEE 802.3by et se rapporte aux consortiums 25G/50G.

Les paramètres de vitesse de liaison et de contrôle de flux/pause doivent être configurés dans le pilote du système d'exploitation hôte.

## Paramètres du pilote Windows

Pour accéder aux paramètres du pilote Windows :

Ouvrez le menu **Gestionnaire de périphériques Windows -> Carte Broadcom NetXtreme E Series -> Propriétés avancées -> onglet Avancé.**

### Contrôle de flux = négociation automatique

Cette fonction permet d'activer la négociation automatique des trames de contrôle de flux/pause.

### Vitesse et duplex = négociation automatique

Cette fonction permet d'activer la négociation automatique de la vitesse de liaison.

## Paramètres du pilote Linux



**Remarque :** Pour les cartes réseau NetXtreme-E 10GBase-T, la négociation automatique doit être activée.



La vitesse 25G annoncée correspond à une norme plus récente, d'abord définie dans l'interface ethtool du noyau 4.7. Pour prendre intégralement en charge ces nouvelles vitesses annoncées pour la négociation automatique, vous devez utiliser un noyau 4.7 (ou plus récent) et une version plus récente de l'utilitaire ethtool (version 4.8).

**ethtool -s eth0 speed 25000 autoneg off**

Cette commande désactive la négociation automatique et impose la vitesse de liaison de 25 Gbit/s.

**ethtool -s eth0 autoneg on advertise 0x0**

Cette commande active la négociation automatique et annonce que le périphérique prend en charge toutes les vitesses : 1G, 10G, 25G.

Les vitesses annoncées ci-dessous sont prises en charge.

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

**ethtool -A eth0 autoneg on|off**

Utilisez cette commande pour activer/désactiver la négociation automatique des trames Pause.

**ethtool -a eth0**

Utilisez cette commande pour afficher le paramètre actuel de négociation automatique du contrôle de flux.

**Paramètres du pilote ESXi**

**Remarque :** Pour les cartes réseau NetXtreme-E 10GBase-T, la négociation automatique doit être activée. L'utilisation d'une vitesse forcée sur une carte 10GBase-T entraîne l'échec de la commande esxcli.



**Remarque :** VMWare ne prend pas en charge les vitesses 25G dans ESX6.0. Dans ce cas, utilisez le second utilitaire (BNXTNETCLI) pour définir la vitesse 25G. Pour ESX6.0U2, la vitesse 25G est prise en charge.

**\$ esxcli network nic get -n <iface>**

Cette commande affiche la vitesse actuelle, le mode duplex, la version du pilote, la version du micrologiciel et l'état de la liaison.

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

Cette commande définit la vitesse imposée sur 10 Gbit/s.

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

Cette commande permet d'activer la négociation automatique de la vitesse de liaison sur l'interface <iface>.

**\$ esxcli network nic pauseParams list**

Utilisez cette commande pour obtenir la liste des paramètres de pause.

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

Utilisez cette commande pour définir les paramètres de pause.



**Remarque :** La négociation automatique du contrôle de flux/pause peut être définie uniquement lorsque l'interface est configurée en mode de négociation automatique de la vitesse de liaison.

## Négociation automatique FEC

Pour activer/désactiver la négociation automatique de liaison FEC, les options suivantes peuvent être activées dans le BIOS du système, dans le menu HII ou dans le CCM :

- **BIOS système-> Paramètres du périphérique -> NetXtreme-E-> Configuration du niveau du périphérique**

La négociation automatique FEC utilise deux paramètres pendant l'échange : FEC capable et FEC request.

Si la carte réseau annonce la compatibilité avec la négociation automatique FEC, les paramètres FEC sont contrôlés par le commutateur. Elle peut alors s'associer à un commutateur qui active ou désactive FEC.

### Exemples:

- switch – capable=1, request = 1 then, the link is a FEC link.
- switch – capable= N/A, request= 0 = then the FEC is disabled.

Pour les contrôleurs NetXtreme-E Ethernet, seul Base-R FEC (CL74) est pris en charge. Le [Tableau 26](#) affiche toutes les configurations prises en charge avec un partenaire de liaison.

**Tableau 26 : Configurations FEC prises en charge pour le modèle BCM5730X/BCM5740X**

Paramètre SFEC local	Paramètre FEC du partenaire de liaison			
	Force Speed No FEC	Force Speed Base-R FEC CL74	AN (None)	AN (None, Base-R)
Force Speed No FEC	Liaison avec/sans FEC	Pas de liaison	Pas de liaison	Pas de liaison
Force Speed Base-R FEC CL74	Pas de liaison	Base-R FEC CL74	Pas de liaison	Pas de liaison
AN (None)	Pas de liaison	Pas de liaison	Liaison avec/sans FEC	Base-R FEC CL74
AN (None, Base-R)	Pas de liaison	Pas de liaison	Base-R FEC CL74	Base-R FEC CL74



**Remarque :** Pour imposer la vitesse, le paramètre de vitesse doit être identique des deux côtés.



**Remarque :** AN {None} signifie qu'AN annonce le bit compatible Base-R. Définissez le bit F0 sur IEEE802.3by et le bit F2 sur Consortium.



**Remarque :** AN {None, Base-R} signifie qu'AN annonce le bit compatible et le bit requis Base-R. Définissez les bits F0 et F1 sur IEEE802.3by, et les bits F2 et F4 sur Consortium.

Pour le contrôleur NetXtreme-E Ethernet, FEC prend en charge Base-R FEC (CL74) et RS-FEC (CL91/CL108). Le [Tableau 27](#) affiche toutes les configurations prises en charge avec un partenaire de liaison.

**Tableau 27 : Configurations FEC prises en charge pour le modèle BCM5741X**

Paramètre FEC local	Paramètre FEC du partenaire de liaison					
	Force Speed No FEC	Force Speed Base-R FEC CL74	Force Speed RS-FEC CL91/CL108	AN (None)	An (None, Base-R)	AN (None, Base-R, RS)
Imposer sans FEC	Liaison avec/sans FEC	Pas de liaison	Pas de liaison	Pas de liaison	Pas de liaison	Pas de liaison
Force Speed Base-R FEC CL74	Pas de liaison	Base-R FEC CL74	Pas de liaison	Pas de liaison	Pas de liaison	Pas de liaison
Imposer RS-FEC CL91/CL108	Pas de liaison	Pas de liaison	RS-FEC CL91/CL108	Pas de liaison	Pas de liaison	Pas de liaison
AN (None)	Pas de liaison	Pas de liaison	Pas de liaison	Liaison avec/sans FEC	Base-R FEC CL74	RS-FEC CL91/CL108
AN (None, Base-R)	Pas de liaison	Pas de liaison	Pas de liaison	Base-R FEC CL74	Base-R FEC CL74	RS-FEC CL91/CL108
AN (None, Base-R, RS)	Pas de liaison	Pas de liaison	Pas de liaison	RS-FEC CL91/CL108	RS-FEC CL91/CL108	RS-FEC CL91/CL108



**Remarque :** Pour imposer la vitesse, le paramètre de vitesse doit être identique des deux côtés.



**Remarque :** AN {None} signifie qu'AN annonce le bit compatible Base-R. Définissez le bit F0 sur IEEE802.3by et le bit F2 sur Consortium.



**Remarque :** AN {None, Base-R} signifie qu'AN annonce le bit compatible et le bit requis Base-R. Définissez les bits F0 et F1 sur IEEE802.3by, et les bits F2 et F4 sur Consortium.

## Link Training

Link Training permet aux deux points de terminaison, à savoir la carte Broadcom et l'autre côté, d'ajuster les paramètres de puissance et autres paramètres de réglage afin d'optimiser la fiabilité et l'efficacité du canal de communication entre les deux périphériques. L'objectif consiste à ne plus avoir besoin de réglage spécifique au canal entre les différents types et longueurs de câbles. Link Training est exécuté pour les vitesses CR/KR et s'effectue avant la négociation automatique. Link Training est opérationnel lorsque la négociation automatique est activée. La politique de liaison désactive automatiquement l'option Link Training si cette dernière n'aboutit à aucune liaison avec le partenaire de liaison. Cette politique garantit la compatibilité avec un partenaire de liaison ne prenant pas en charge Link Training.

Le [Tableau 28](#) affiche la relation entre le type de support et la vitesse pour les contrôleurs Ethernet BCM5730X, BCM5740X et BCM5741X.

**Tableau 28 : Relation Link Training entre le type de support et la vitesse**

Type de câble local et du support	Paramètre Link Training du partenaire de liaison		
	Force Speed Link Training Disabled	Force Speed Link Training Enabled	AN (Link Training automatique)
Force Speed DAC (SFP+/SFP28/QSFP28)	Liaison sans Link Training	Liaison avec Link Training	Liaison avec Link Training
Force Speed optical (Transceiver/AOC)	Liaison sans Link Training	Liaison sans Link Training	S/O
AN DAC (SFP+SFP28/QSFP28)	Pas de liaison	Pas de liaison	Liaison avec Link Training

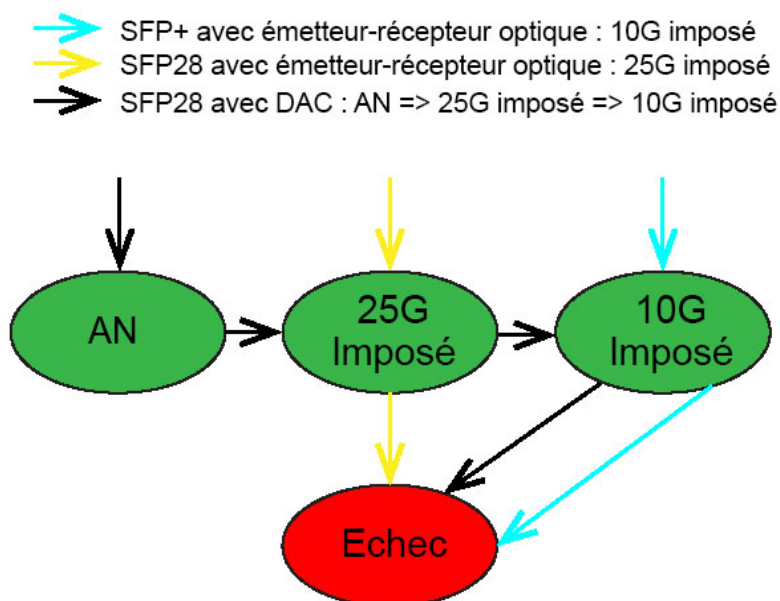
## Détection automatique des supports

Puisque la détection parallèle n'est pas prise en charge dans SerDes, le micrologiciel a mis en œuvre une méthode pour améliorer la détection des liaisons, appelée **détection automatique des supports**. Cette fonctionnalité est contrôlée par CCM/HII, comme indiqué dans la [Figure 13](#).

Lorsque la **détection automatique des supports** est activée, la politique de liaison suit l'ordinateur contrôlant les états (voir la [Figure 13](#)) pour établir une liaison avec un partenaire de liaison. Ce comportement dépend du type de support. Pour les câbles DAC, la méthode renvoie à différents modes d'imposition et s'interrompt en cas de liaison.

**REMARQUE:** La détection parallèle n'est pas prise en charge en cas de négociation automatique à 25G ou 10GbE. Cela signifie que si seul un côté effectue une négociation automatique, la liaison ne pourra pas s'établir.

**Figure 13 : Ordinateur contrôlant les états pour la fonction de détection automatique des supports**





Le [Tableau 29](#) et le [Tableau 30](#) affichent les résultats de la liaison avec la fonction **Détection automatique des supports** activée.

**Tableau 29 : Détection automatique des supports pour les modèles BCM5730X et le BCM5740X**

		Paramètre Link Training du partenaire de liaison	
Paramètres du partenaire de liaison		Détection automatique des supports	
Vitesse	FEC	Sans FEC	Base-R FEC
10G	Sans FEC	Liaison	Liaison avec Base-R
	Base-R	Pas de liaison	Liaison avec Base-R
25G	Sans FEC	Liaison	Liaison
	Base-R	Pas de liaison	Liaison avec Base-R
AN	Désactivé	Liaison	Liaison
	FEC automatique	Liaison	Liaison avec Base-R
	Base-R	Liaison avec Base-R	Liaison avec Base-R

**Tableau 30 : Détection automatique des supports pour le modèle BCM5741X**

		Paramètre Link Training du partenaire de liaison		
Paramètres du partenaire de liaison		Détection automatique des supports		
Vitesse	FEC	Sans FEC	Base-R FEC	RS-FEC
10G	Sans FEC	Liaison	Liaison avec Base-R	Pas de liaison
	Base-R	Pas de liaison	Liaison avec Base-R	Pas de liaison
	RS	Pas de liaison	Pas de liaison	Liaison avec RS-FEC
25G	Sans FEC	Liaison	Pas de liaison	Pas de liaison
	Base-R	Pas de liaison	Liaison avec Base-R	Pas de liaison
	RS	Pas de liaison	Pas de liaison	Liaison avec RS-FEC
AN	Désactivé	Liaison	Liaison	Liaison
	FEC automatique	Liaison	Liaison avec Base-R	Liaison avec RS-FEC
	Base-R	Liaison avec Base-R	Liaison avec Base-R	Liaison avec RS-FEC
	RS	Liaison avec RS	Liaison avec RS-FEC	Liaison avec RS-FEC

---

## Initialisation iSCSI

Les cartes NetXtreme-E Ethernet de Broadcom prennent en charge l'initialisation iSCSI afin de permettre le démarrage réseau de systèmes d'exploitation sur des systèmes sans disque. L'initialisation iSCSI permet le démarrage d'un système d'exploitation Windows, Linux ou VMware à partir d'un ordinateur cible iSCSI se trouvant sur un réseau IP standard distant.

## Systemes d'exploitation pris en charge pour l'initialisation iSCSI

Les cartes NetXtreme-E Gigabit Ethernet de Broadcom prennent en charge l'initialisation iSCSI sur les systèmes d'exploitation suivants :

- Windows Server 2012 et versions ultérieures 64 bits
- Linux RHEL 7.1 et versions ultérieures, SLES11 SP4 et versions ultérieures
- VMware 6.0 U2

## Configuration de l'initialisation iSCSI

Reportez-vous aux sections suivantes pour plus d'informations sur la configuration de l'initialisation iSCSI.

## Configuration de la cible iSCSI



**Remarque :** Windows 2016 (ou version antérieure) peut disposer d'un pilote intégré obsolète qui prend uniquement en charge les vitesses de liaison imposées. Pour éviter tout problème d'installation tel qu'une incapacité à établir une liaison avec un partenaire de liaison incompatible, les utilisateurs sont fortement invités à personnaliser l'image/le support d'installation Windows avec le dernier pilote NIC disponible qui prend en charge la négociation automatique des vitesses de liaison.

La configuration de la cible iSCSI varie en fonction des fournisseurs cibles. Pour plus d'informations sur la configuration de la cible iSCSI, reportez-vous à la documentation du fournisseur. Les étapes générales sont les suivantes :

1. Créez une cible iSCSI.
2. Créez un disque virtuel.
3. Mappez le disque virtuel à la cible iSCSI créée à l'[Etape 1 page 50](#).
4. Associez un initiateur iSCSI à la cible iSCSI.
5. Enregistrez le nom de la cible iSCSI, le numéro du port TCP, le numéro d'unité logique (LUN) iSCSI, le nom IQN de l'initiateur et les détails d'authentification CHAP.
6. Une fois la cible iSCSI configurée, notez les éléments suivants :
  - Nom IQN de la cible
  - Adresse IP cible
  - Numéro de port TCP de la cible

- Numéro d'unité logique (LUN) de la cible
- Nom IQN de l'initiateur
- Réf. et clé secrète CHAP

## Configuration des paramètres d'initialisation iSCSI

Configurez le logiciel d'initialisation iSCSI de Broadcom de manière statique ou dynamique. Reportez-vous au [Tableau 31](#) pour connaître les options de configuration disponibles depuis le menu Paramètres généraux. Le [Tableau 31](#) répertorie les paramètres pour IPv4 et IPv6. Les paramètres spécifiques à IPv4 ou IPv6 sont notés.

**Tableau 31 : Options de configuration**

<b>Option</b>	<b>Description</b>
Paramètres TCP/IP via DHCP	Cette option est spécifique à IPv4. Contrôle si le logiciel hôte d'initialisation iSCSI obtient les informations d'adresse IP via DHCP (Activé) ou par le biais d'une configuration IP statique (Désactivé).
Configuration auto de l'IP	Cette option est spécifique à IPv6. Contrôle si le logiciel hôte d'initialisation iSCSI configure une adresse lien-local sans état et/ou une adresse avec état si DHCPv6 est présent et utilisé (Activé). Des paquets de sollicitation du routeur sont envoyés jusqu'à trois fois à un intervalle de 4 secondes entre chaque tentative. Vous pouvez aussi utiliser une configuration IP statique (Désactivé).
Paramètres iSCSI via DHCP	Contrôle si le logiciel hôte d'initialisation iSCSI obtient ses paramètres cibles iSCSI via DHCP (Activé) ou par le biais d'une configuration statique (Désactivé). Les informations statiques sont à indiquer dans l'écran Configuration des paramètres de l'initiateur iSCSI.
Authentification CHAP	Contrôle si le logiciel hôte d'initialisation iSCSI utilise une authentification CHAP lors de sa connexion à la cible iSCSI. Si l'authentification CHAP est activée, la réf. CHAP et la clé secrète CHAP sont à indiquer dans l'écran Configuration des paramètres de l'initiateur iSCSI.
Réf. fournisseur DHCP	Contrôle de quelle façon le logiciel hôte d'initialisation iSCSI interprète le champ ID de la classe du fournisseur utilisé avec DHCP. Si le champ ID de la classe du fournisseur du paquet de l'Offre DHCP correspond à la valeur du champ, le logiciel hôte d'initialisation iSCSI recherche les extensions d'initialisation iSCSI requises dans le champ Option DHCP 43. Si DHCP est désactivé, il n'est pas nécessaire de désactiver cette valeur.
Durée d'établissement de la liaison	Contrôle la durée d'attente du logiciel hôte d'initialisation iSCSI, en secondes, une fois la liaison Ethernet établie, avant de transmettre des données sur le réseau. Les valeurs valides sont comprises entre 0 et 255. Par exemple, un utilisateur devra définir une valeur pour cette option si un protocole réseau, tel que STP, est activé sur l'interface de commutateur du système client.
Utiliser l'horodatage TCP	Contrôle si l'option d'horodatage TCP est activée ou désactivée.
Cibler comme premier disque dur	Permet de définir le lecteur cible iSCSI comme premier lecteur de disque dur du système.
Nombre d'essais de connexion à une unité logique occupée	Contrôle le nombre de tentatives de reconnexion effectué par l'initiateur de démarrage iSCSI lorsque le numéro d'unité logique de la cible iSCSI est occupé.
Version IP	Cette option est spécifique à IPv6. Permet de passer du protocole IPv4 au protocole IPv6, et inversement. Tous les paramètres IP sont perdus lorsque vous basculez d'une version de protocole à une autre.

## Configuration du protocole de démarrage MBA

Pour configurer le protocole de démarrage, procédez comme suit :

1. Redémarrez le système.
2. A partir de la bannière PXE, sélectionnez **CTRL+S**. Le **menu de configuration MBA** s'affiche.
3. Dans le **menu de configuration MBA**, utilisez les **flèches de direction HAUT** et **BAS** pour sélectionner l'option **Protocole de démarrage**. Utilisez les **flèches de direction GAUCHE** et **DROITE** pour définir l'option **Protocole de démarrage** sur iSCSI.
4. Sélectionnez **Configuration de l'initialisation iSCSI** dans le **menu principal**.

## Configuration de l'initialisation iSCSI

Il existe deux façons de configurer l'initialisation iSCSI :

- Configuration de l'initialisation iSCSI statique
- Configuration de l'initialisation iSCSI dynamique

### Configuration de l'initialisation iSCSI statique

Dans le cadre d'une configuration statique, vous devez indiquer l'adresse IP du système, le nom IQN de l'initiateur du système et les paramètres cibles obtenus dans la section « [Configuration de la cible iSCSI](#) » [page 50](#). Pour plus d'informations sur la configuration des options, reportez-vous au [Tableau 31 page 51](#).

Pour configurer les paramètres d'initialisation iSCSI par le biais d'une configuration statique, procédez comme suit :

1. Dans le menu **Paramètres généraux**, définissez les éléments suivants :
  - Paramètres TCP/IP via DHCP : Désactivé. (Pour IPv4.)
  - Configuration auto de l'IP : Désactivé. (Pour IPv6, sans déchargement.)
  - Paramètres iSCSI via DHCP : Désactivé
  - Authentification CHAP : Désactivé
  - ID du fournisseur DHCP : BRCM ISAN
  - Délai de liaison active : 0
  - Utiliser l'horodateur TCP : Activé (pour certaines cibles telles que Dell/EMC AX100i, il est nécessaire d'activer Utiliser l'horodateur TCP)
  - Cibler comme premier lecteur de disque dur : Désactivé.
  - Nombre de tentatives si le numéro d'unité logique est occupé : 0
  - Version IP : IPv6. (Pour IPv6, sans déchargement.)
2. Appuyez sur **ECHAP** pour revenir au menu **principal**.
3. Dans le menu **principal**, sélectionnez **Paramètres de l'initiateur**.

4. Dans l'écran **Paramètres de l'initiateur**, entrez des valeurs pour les éléments suivants :
  - Adresse IP (les adresses IPv4 et IPv6 non spécifiées doivent être « 0.0.0.0 » et « :: », respectivement)
  - Préfixe de masque de sous-réseau
  - Passerelle par défaut
  - DNS principal
  - DNS secondaire
  - Nom iSCSI (correspond au nom de l'initiateur iSCSI devant être utilisé par le système client)



**Remarque :** Entrez l'adresse IP. L'adresse IP ne fait l'objet d'aucune vérification en matière de duplication ou de segment/d'attribution réseau incorrects.

5. Appuyez sur **ECHAP** pour revenir au menu **principal**.
6. Dans le menu **principal**, sélectionnez **Paramètres de la 1re cible**.



**Remarque :** Lors de la configuration initiale, la configuration d'une deuxième cible n'est pas prise en charge.

7. Dans l'écran **Paramètres de la 1re cible**, activez **Connexion** pour vous connecter à la cible iSCSI. Pour les éléments suivants, entrez les valeurs utilisées lors de la configuration de la cible iSCSI :
  - Adresse IP
  - Port TCP
  - Numéro d'unité logique d'initialisation
  - Nom iSCSI
8. Appuyez sur **ECHAP** pour revenir au menu **principal**.
9. Appuyez sur **ECHAP** et sélectionnez **Quitter** et **enregistrer la configuration**.
10. Appuyez sur **F4** pour enregistrer la configuration MBA.

## Configuration de l'initialisation iSCSI dynamique

Dans le cadre d'une configuration dynamique, spécifiez que l'adresse IP et les informations concernant la cible ou l'initiateur du système sont fournies par un serveur DHCP (voir les configurations IPv4 et IPv6 dans [« Configuration du serveur DHCP pour la prise en charge de l'initialisation iSCSI » page 55](#)). Pour IPv4, à l'exception du nom iSCSI de l'initiateur, tous les paramètres des écrans Paramètres de l'initiateur, Paramètres de la 1re cible ou Paramètres de la 2e cible sont ignorés et il n'est pas nécessaire de les effacer. Pour IPv6, à l'exception de la réf. et de la clé secrète CHAP, tous les paramètres des écrans Paramètres de l'initiateur, Paramètres de la 1re cible ou Paramètres de la 2e cible sont ignorés et il n'est pas nécessaire de les effacer. Pour plus d'informations sur la configuration des options, reportez-vous au [Tableau 31 page 51](#).



**Remarque :** Lors de l'utilisation d'un serveur DHCP, les entrées du serveur DNS sont écrasées par les valeurs provenant du serveur DHCP. Ce problème se pose même si les valeurs fournies localement sont valides et si le serveur DHCP ne transmet aucune information concernant le serveur DNS. Lorsque le serveur DHCP ne transmet aucune information concernant le serveur DNS, les valeurs des serveurs DNS principal et secondaire sont définies sur 0.0.0.0. Lorsque le système d'exploitation Windows prend le contrôle, l'initiateur iSCSI Microsoft récupère les paramètres de l'initiateur iSCSI et configure les registres appropriés de manière statique. Il écrase alors tout élément configuré. Etant donné que le démon DHCP s'exécute dans l'environnement Windows en tant que processus utilisateur, tous les paramètres TCP/IP doivent être configurés de manière statique avant que la pile apparaisse dans l'environnement d'initialisation iSCSI.

- Si l'option DHCP 17 est utilisée, les informations sur la cible sont fournies par le serveur DHCP et le nom iSCSI de l'initiateur est récupéré à partir de la valeur programmée dans l'écran Paramètres de l'initiateur. Si aucune valeur n'a été sélectionnée, le contrôleur utilise le nom par défaut :

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

dans lequel la chaîne **11.22.33.44.55.66** correspond à l'adresse MAC du contrôleur.

- Si l'option DHCP 43 (IPv4 uniquement) est utilisée, tous les paramètres des écrans Paramètres de l'initiateur, Paramètres de la 1re cible ou Paramètres de la 2e cible sont ignorés et il n'est pas nécessaire de les effacer.

Pour configurer les paramètres d'initialisation iSCSI par le biais d'une configuration dynamique, procédez comme suit :

1. Dans l'écran de menu **Paramètres généraux**, définissez les paramètres suivants :

- Paramètres TCP/IP via DHCP : Activé. (Pour IPv4.)
- Configuration auto de l'IP : Activé. (Pour IPv6, sans déchargement.)
- Paramètres iSCSI via DHCP : Activé
- Authentification CHAP : Désactivé
- ID du fournisseur DHCP : BRCM ISAN
- Délai de liaison active : 0
- Utiliser l'horodateur TCP : Activé (pour certaines cibles telles que Dell/EMC AX100i, il est nécessaire d'activer Utiliser l'horodateur TCP)
- Cibler comme premier lecteur de disque dur : Désactivé.
- Nombre de tentatives si le numéro d'unité logique est occupé : 0
- Version IP : IPv6. (Pour IPv6, sans déchargement.)

2. Appuyez sur **ECHAP** pour revenir au menu **principal**.



**Remarque :** Les informations des écrans Paramètres de l'initiateur et Paramètres de la 1re cible sont ignorées et il n'est pas nécessaire de les effacer.

3. Sélectionner **Quitter** et **enregistrer la configuration**.

## Activation de l'authentification CHAP

Assurez-vous que l'authentification CHAP est activée sur la cible.

Pour activer l'authentification CHAP, procédez comme suit :

1. Dans l'écran **Paramètres généraux**, définissez l'option **Authentification CHAP** sur **Activé**.
2. Dans l'écran **Paramètres de l'initiateur**, entrez des valeurs pour les paramètres suivants :
  - Réf. CHAP (jusqu'à 128 octets)
  - Clé secrète CHAP (si l'authentification est nécessaire ; doit contenir au minimum 12 caractères)
3. Appuyez sur **ECHAP** pour revenir au menu **principal**.
4. Dans le menu **principal**, sélectionnez **Paramètres de la 1re cible**.
5. Dans l'écran **Paramètres de la 1re cible**, entrez les valeurs utilisées lors de la configuration de la cible iSCSI pour les éléments suivants :
  - Réf. CHAP (facultatif si CHAP bidirectionnel)
  - Clé secrète CHAP (facultatif si CHAP bidirectionnel ; doit contenir au minimum 12 caractères)
6. Appuyez sur **ECHAP** pour revenir au menu **principal**.
7. Appuyez sur **ECHAP** et sélectionnez **Quitter et enregistrer la configuration**.

## Configuration du serveur DHCP pour la prise en charge de l'initialisation iSCSI

Le serveur DHCP est un composant facultatif. Il est nécessaire uniquement si vous effectuez une configuration dynamique de l'initialisation iSCSI (voir « [Configuration de l'initialisation iSCSI dynamique](#) » page 53).

La configuration du serveur DHCP pour la prise en charge de l'initialisation iSCSI est différente pour IPv4 et IPv6. Reportez-vous aux sections suivantes :

### Configurations DHCP pour l'initialisation iSCSI pour IPv4

Le protocole DHCP englobe plusieurs options apportant des informations de configuration au client DHCP. Pour une initialisation iSCSI, les cartes Broadcom prennent en charge les configurations DHCP suivantes :

#### Option DHCP 17, chemin d'accès

L'option 17 permet de transmettre les informations concernant la cible iSCSI au client iSCSI. Le format du chemin d'accès défini dans IETF RFC 4173 se présente de la manière suivante :

```
"iscsi:"<nomserveur>":"<protocole>":"<port>":"<LUN>":"<nomcible>
```

Les paramètres sont définis dans le [Tableau 32](#).

**Tableau 32 : Définition des paramètres de l'option DHCP 17**

<b>Paramètre</b>	<b>Définition</b>
"iscsi:"	Une chaîne littérale.
<nomserveur>	L'adresse IP ou le nom de domaine complet de la cible iSCSI
":"	Séparateur.
<protocole>	Le protocole IP permettant d'accéder à la cible iSCSI. Seul TCP étant pris en charge actuellement, le protocole est 6.
<port>	Le numéro de port associé au protocole. Le numéro de port standard pour iSCSI est 3260.
<LUN>	Le numéro d'unité logique à utiliser sur la cible iSCSI. La valeur du LUN doit être représentée à l'aide de caractères hexadécimaux. Un LUN dont l'identifiant est 64 doit être défini sur 40, dans l'option 17 du serveur DHCP.
<nomcible>	Le nom de la cible au format IQN ou EUI (voir RFC 3720 pour des détails sur les formats IQN et EUI). Exemple de nom IQN : « iqn.1995-05.com.broadcom:iscsi-target ».

### Option DHCP 43, informations concernant le fournisseur

L'option DHCP 43 (informations concernant le fournisseur) offre davantage d'options de configuration au client iSCSI que l'option DHCP 17. Dans le cadre de cette configuration, trois sous-options supplémentaires permettent d'attribuer le nom IQN de l'initiateur au client d'initialisation iSCSI, ainsi que deux noms IQN de cible iSCSI pouvant être utilisés pour l'initialisation. Le format du nom IQN de la cible iSCSI est identique à celui de l'option DHCP 17, tandis que le nom IQN de l'initiateur iSCSI est tout simplement le nom IQN de l'initiateur.



**Remarque :** L'option DHCP 43 est uniquement prise en charge pour IPv4.

Les sous-options sont définies ci-dessous.

**Tableau 33 : Définition des sous-options de l'option DHCP 43**

<b>Sous-option</b>	<b>Définition</b>
201	Informations sur la première cible iSCSI sous la forme du chemin d'accès standard "iscsi:"<nomserveur>":"<protocole>":"<port>":"<LUN>":"<nomcible>
203	Nom IQN de l'initiateur iSCSI

L'option DHCP 43 requiert une configuration plus importante que l'option DHCP 17, mais elle offre un environnement plus riche et comporte davantage d'options de configuration. Broadcom recommande à ses clients d'utiliser l'option DHCP 43 lors de la réalisation de configurations dynamiques de l'initialisation iSCSI.

### Configuration du serveur DHCP

Configurez le serveur DHCP de sorte qu'il prenne en charge l'option 17 ou l'option 43.





**Remarque :** Si l'option 43 est utilisée, configurez l'option 60. La valeur de l'option 60 doit correspondre à celle de la Réf. fournisseur DHCP. La valeur de la Réf. fournisseur DHCP est BRCM ISAN, comme indiqué dans les Paramètres généraux du menu de configuration de l'initialisation iSCSI.

## Configurations DHCP pour l'initialisation iSCSI pour IPv6

Le serveur DHCPv6 peut fournir plusieurs options, y compris la configuration IP sans état ou avec état, ainsi que des informations au client DHCPv6. Pour une initialisation iSCSI, les cartes Broadcom prennent en charge les configurations DHCP suivantes :



**Remarque :** L'option de chemin d'accès standard de DHCPv6 n'est pas encore disponible. Broadcom suggère d'utiliser l'option 16 ou 17 pour la prise en charge dynamique de IPv6 pour l'initialisation iSCSI.

### Option DHCPv6 16, option de classe fournisseur

L'option DHCPv6 16 (option de classe fournisseur) doit être présente et contenir une chaîne correspondant au paramètre configuré pour la Réf. fournisseur DHCP. La valeur de la Réf. fournisseur DHCP est BRCM ISAN, comme indiqué dans les Paramètres généraux du menu de configuration de l'initialisation iSCSI.

Le contenu de l'option 16 doit être <longueur de 2 octets> <Réf. fournisseur DHCP>.

### Option DHCPv6 17, informations concernant le fournisseur

L'option DHCPv6 17 (informations concernant le fournisseur) fournit d'autres options de configuration au client iSCSI. Dans le cadre de cette configuration, trois sous-options supplémentaires permettent d'attribuer le nom IQN de l'initiateur au client d'initialisation iSCSI, ainsi que deux noms IQN de cible iSCSI pouvant être utilisés pour l'initialisation.

Les sous-options sont répertoriées dans le [Tableau 34 page 57](#).

**Tableau 34 : Définition des sous-options de l'option DHCP 17**

Sous-option	Définition
201	Informations sur la première cible iSCSI sous la forme du chemin d'accès standard "iscsi:[<nomserveur>]":"<protocole>":"<port>":"<LUN>":"<nomcible>"
203	Nom IQN de l'initiateur iSCSI



**Remarque :** Dans le [Tableau 34](#), les crochets [ ] sont requis pour les adresses IPv6.

Le contenu de l'option 17 doit être <numéro d'option de 2 octets 201|202|203> <longueur de 2 octets> <données>.

## Configuration du serveur DHCP

Configurez le serveur DHCP de sorte qu'il prenne en charge l'option 16 et l'option 17.



**Remarque :** Le format des options DHCPv6 16 et 17 est entièrement défini dans RFC 3315.

## VXLAN : Configuration et exemples d'utilisation

L'encapsulation VXLAN permet à de nombreux hôtes de couche 3 résidant sur un serveur d'envoyer et de recevoir des trames en les encapsulant en une seule adresse IP associée à la carte NIC installée sur le même serveur.

L'exemple suivant présente la connectivité VXLAN de base entre deux serveurs RHEL. Chaque serveur possède une carte réseau physique activée avec une adresse IP définie sur 1.1.1.4 et 1.1.1.2.

Une interface VXLAN10 avec ID VXLAN 10 est créée avec un groupe de multidiffusion 239.0.0.10 et est associée à un port réseau physique pxx1 sur chaque serveur.

Une adresse IP pour l'hôte est créée sur chaque serveur et associée à l'interface VXLAN. Une fois l'interface VXLAN mise en service, l'hôte présent sur le système 1 peut communiquer avec l'hôte présent sur le système 2. Le format VLXAN est indiqué dans le [Tableau 35](#).

**Tableau 35 : Format de trame VXLAN**

En-tête MAC	En-tête IP externe avec proto = UDP	En-tête UDP avec port de destination = VXLAN	En-tête VXLAN (indicateurs, VNI)	Trame L2 d'origine	FCS
-------------	-------------------------------------	--	----------------------------------	--------------------	-----

Le [Tableau 36](#) fournit des exemples de configuration et de commande VXLAN.

**Tableau 36 : Exemples de configuration et de commande VXLAN**

Système 1	Système 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 et RHEL)

**Remarque :** x représente le numéro de bus PCIe de la carte physique détectée dans le système. y représente le numéro de port sur la carte physique.

## SR-IOV : Configuration et exemples d'utilisation

SR-IOV peut être configuré, activé et utilisé sur des cartes réseau Broadcom NetExtreme-E de 10 Gbits et 25 Gbits.

### Cas d'utilisation Linux

1. Activez SR-IOV dans les cartes NIC :
  - a. SR-IOV peut être activé dans la carte réseau à l'aide du menu **HII**. Lors du démarrage du système, accédez au **BIOS** système -> **Paramètres du périphérique** -> **Cartes NIC NetXtreme-E** -> **Configuration du niveau du périphérique**.
  - b. Définissez le mode de virtualisation sur SR-IOV.
  - c. Définissez le nombre de fonctions virtuelles par fonction physique.
  - d. Définissez le nombre de vecteurs MSI-X par VF et le nombre maximum de vecteurs MSI-X par fonction physique. Si la VF manque de ressources, équilibrez le nombre de vecteurs MSI-X par machine virtuelle à l'aide de CCM.
2. Activez la virtualisation dans le BIOS :
  - a. Lors du démarrage du système, entrez dans le **BIOS** système -> **Paramètres du processeur** -> **Technologies de virtualisation**, puis configurez-le afin qu'il soit **activé**.
  - b. Lors du démarrage du système, entrez dans le **BIOS** système -> **Périphériques intégrés** -> **SR-IOV Global**, puis configurez-le afin qu'il soit **activé**.
3. Installez la version Linux de votre choix avec la virtualisation activée (libvirt et Qemu).
4. Activez le paramètre de noyau iommu.
  - a. Le paramètre de noyau IOMMU est activé en modifiant le fichier `/etc/default/grub.cfg` et en exécutant `grub2-mkconfig -o /boot/grub2/grub.cfg` pour le mode hérité. Pour le mode UEFI, modifiez le fichier `/etc/default/grub.cfg` et exécutez `grub2-mkconfig -o /etc/grub2-efi.cfg`. Reportez-vous à l'exemple suivant :
 

```
linuxefi /vmlinuz-3.10.0-229.el7.x86_64 root=/dev/mapper/rhel-root no rd.lvm.lv=rhel/swap
          crashkernel=auto rd.lvm.lv=rhel/root rhgb intel_iommu=on quiet LANG=en_US.UTF.8
```
5. Installez le pilote `bnxt_en` :
  - a. Copiez le pilote `bnxt_en` sur le SE, puis exécutez la commande `make; make install; modprobe bnxt_en`.



**Remarque :** Utilisez `netxtreme-bnxt_en<version>.tar.gz` pour installer `bnxt_re` et `bnxt_en` pour la fonctionnalité RDMA sur les VF SRIOV.

6. Activez les fonctions virtuelles via les paramètres du noyau :
  - a. Une fois le pilote installé, `lspci` affiche les cartes NIC NetXtreme-E présentes sur le système. Les bus, périphériques et fonctions sont nécessaires pour activer les fonctions virtuelles.
  - b. Pour activer les fonctions virtuelles, entrez la commande ci-dessous :

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



**Remarque :** Vérifiez que les interfaces PF sont actives. Les VF sont uniquement créées si les PF sont actives. X est le nombre de VF qui seront exportées vers le SE.

Exemple type :

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

7. Vérifiez les fonctions virtuelles PCI-E :

- a. La commande `lspci` affiche les fonctions virtuelles avec DID défini sur 16D3 pour les cartes BCM57402/BCM57404/BCM57406, sur 16DC pour les cartes BCM57412/BCM57414/BCM57416 non RDMA, et sur 16C1 pour les cartes BCM57412/BCM57414/BCM57416 compatibles RDMA.

8. Utilisez Virtual Manager pour installer un système client virtualisé (VM).

Reportez-vous à la documentation Linux pour installer Virtual Manager. Assurez-vous que le pilote intégré à l'hyperviseur est supprimé. Exemple : `NIC:d7:73:a7 rt18139`. Supprimez ce pilote.

9. Affectez une fonction virtuelle aux VM invitées.

- a. Affectez cette carte à une VM invitée en tant que périphérique PCI physique. Reportez-vous à la documentation Linux pour plus d'informations sur l'affectation des fonctions virtuelles à une VM invitée.

10. Installez les pilotes `bnxt_en` sur des machines virtuelles :

- a. Sur les machines virtuelles invitées, copiez le fichier source `netxtreme-bnxt_en-<version>.tar.gz` et extrayez le fichier `tar.gz`. Changez de répertoire pour chaque pilote et exécutez `make; make install; modprobe bnxt_en` (et `bnxt_re` si vous activez RDMA). Assurez-vous que le pilote se charge correctement en vérifiant l'interface à l'aide de la commande `modinfo`. L'utilisateur peut avoir besoin d'exécuter `modprobe -r bnxt_en` pour décharger le module existant ou de réception `bnxt_en` avant de charger la dernière version du module intégré.

11. Testez la connectivité de la VM invitée au monde extérieur :

- a. Affectez une adresse IP correcte à la carte et testez la connectivité au réseau.

## Cas de Windows

1. Activez SR-IOV dans les cartes NIC :

- a. SR-IOV peut être activé dans la carte réseau à l'aide du menu HII. Lors du démarrage du système, accédez à **BIOS -> Paramètres du périphérique -> Cartes NIC NetXtreme-E -> Configuration du niveau du périphérique**.
- b. Définissez le mode de virtualisation sur SR-IOV.
- c. Définissez le nombre de fonctions virtuelles par fonction physique.
- d. Définissez le nombre de vecteurs MSI-X par VF et le nombre maximum de vecteurs MSI-X par fonction physique. Si la VF manque de ressources, équilibrez le nombre de vecteurs MSI-X par machine virtuelle à l'aide de CCM.

2. Activez la virtualisation dans le BIOS :

- a. Lors du démarrage du système, entrez dans le **BIOS système -> Paramètres du processeur -> Technologies de virtualisation**, puis configurez-le afin qu'il soit **activé**.
- b. Lors du démarrage du système, entrez dans le **BIOS système -> Périphériques intégrés -> SR-IOV Global**, puis configurez-le afin qu'il soit **activé**.

3. Installez la dernière mise à jour KB pour votre système d'exploitation Windows 2012 R2 ou Windows 2016.
4. Installez les options de virtualisation (Hyper-V) appropriées. Pour plus d'informations sur la configuration requise et sur la configuration de l'hyperviseur Hyper-V, du commutateur virtuel et des machines virtuelles, rendez-vous sur le site microsoft.com :  
<https://technet.microsoft.com/fr-fr/windows-server-docs/compute/hyper-v/system-requirements-for-hyper-v-on-windows>  
<https://technet.microsoft.com/fr-fr/windows-server-docs/compute/hyper-v/get-started/install-the-hyper-v-role-on-windows-server>
5. Installez le dernier pilote NetXtreme-E sur l'hyperviseur Hyper-V.
6. Activez SR-IOV dans les propriétés avancées du pilote miniport NDIS.
7. Dans Hyper-V Manager, créez votre commutateur virtuel avec l'interface NetXtreme-E sélectionnée.
8. Cochez la case **Activer la virtualisation d'E/S d'une racine unique (SR-IOV)** lors de la création de la carte virtuelle Hyper-V.
9. Créez une machine virtuelle (VM) et ajoutez le nombre souhaité de cartes virtuelles.
10. Pour chaque carte virtuelle, dans les paramètres de carte réseau de la machine virtuelle, cochez **Activer SR-IOV** sous la section **Accélération matérielle**.
11. Démarrez votre machine virtuelle et installez le système d'exploitation invité.
12. Installez le pilote NetXtreme-E correspondant pour chaque système d'exploitation invité.



**Remarque :** Le pilote de fonction virtuelle (VF) pour la carte NetXtreme-E est le même que le pilote de base. Par exemple, si le système d'exploitation invité est Windows 2012 R2, l'utilisateur doit installer Bnxtnd64.sys dans la VM. Pour ce faire, l'utilisateur peut lancer le fichier exécutable du programme d'installation du pilote NetXtreme-E. Une fois le pilote installé dans le système d'exploitation invité, l'utilisateur peut voir la ou les interface(s) du pilote VF dans le Gestionnaire de périphériques du système d'exploitation invité sur la machine virtuelle.

## Cas de VMWare SRIOV

1. Activez SR-IOV dans les cartes NIC :
  - a. SR-IOV peut être activé dans la carte réseau à l'aide du menu HII. Lors du démarrage du système, accédez à **BIOS -> Paramètres du périphérique -> Cartes NIC NetXtreme-E -> Configuration du niveau du périphérique**.
  - b. Définissez le mode de virtualisation sur SR-IOV.
  - c. Définissez le nombre de fonctions virtuelles par fonction physique.
  - d. Définissez le nombre de vecteurs MSI-X par VF et le nombre maximum de vecteurs MSI-X par fonction physique. Si la VF manque de ressources, équilibrez le nombre de vecteurs MSI-X par machine virtuelle à l'aide de CCM.

## 2. Activez la virtualisation dans le BIOS :

- Lors du démarrage du système, entrez dans le **BIOS système -> Paramètres du processeur -> Technologies de virtualisation**, puis configurez-le afin qu'il soit **activé**.
- Lors du démarrage du système, entrez dans le **BIOS système -> Périphériques intégrés -> SR-IOV Global**, puis configurez-le afin qu'il soit **activé**.

## 3. Sur ESXi, installez le pilote Bnxtnet en procédant comme suit :

- Copiez le fichier <bnxtnet>-<version du pilote> dans /var/log/vmware.

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

- Redémarrez l'ordinateur.
- Vérifiez que les pilotes sont correctement installés :

```
$ esxcli software vib list | grep bnxtnet
```

## 4. Installez l'utilitaire BNXTNETCLI (esxcli bnxtnet) fourni par Broadcom pour définir/afficher les divers paramètres de pilote qui ne sont pas nativement pris en charge dans esxcli, par exemple : vitesse de liaison à 25G, afficher les informations sur le pilote/le micrologiciel/la puce, afficher la configuration de la carte réseau (NPAR, SRIOV). Pour plus d'informations, veuillez consulter le fichier README.txt du pilote bnxtnet.

Pour installer cet utilitaire :

- Copiez BCM-ESX-bnxtnetcli-<version>.vib dans /var/log/vmware.

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

- Redémarrez le système.
- Vérifiez si le fichier vib est installé correctement :

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

- Définir la vitesse 10/20/25

```
$ esxcli bnxtnet link set -S <speed> -D <full> -n <iface>  
Le message OK est renvoyé si la vitesse est correctement définie.
```

Exemples :

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

- Afficher les statistiques de liaison

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

- Afficher des informations sur le pilote/le micrologiciel/la puce

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

- Afficher les informations sur la carte réseau (par exemple : BDF ; NPAR ; configuration SRIOV)

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

## 5. Activation des VF SRIOV :

Seules les PF sont automatiquement activées. Si une PF prend en charge SR-IOV, le paramètre PF(vmknix) fait partie de la sortie de la commande illustrée ci-dessous.

```
esxcli network sriovnic list
```

Pour activer une ou plusieurs VF, le pilote utilise le paramètre de module « max\_vfs » pour activer le nombre souhaité de VF pour les PF. Par exemple, pour activer quatre VF sur la PF1 :

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

Pour activer des VF sur un ensemble de PF, utilisez le format de commande illustré ci-dessous. Par exemple, pour activer quatre VF sur la PF 0 et 2 VF sur la PF 2 :

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

Les VF requises de chaque PF prise en charge PF sont activées dans l'ordre d'affichage de la PF. Reportez-vous à la documentation VMware pour plus d'informations sur le mappage d'une VF à une machine virtuelle :

<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 : Configuration et exemples d'utilisation

### Fonctionnalités et configuration requise

- Indépendance vis-à-vis du système d'exploitation/BIOS : les partitions sont présentées au système d'exploitation en tant qu'interfaces réseau « réelles », ce qui signifie qu'aucune prise en charge du BIOS ou du système d'exploitation particulière n'est nécessaire comme pour SR-IOV.
- Cartes réseau supplémentaires sans ports de commutateur, câblage ou logements d'extension PCIe supplémentaires.
- Régulation de flux : l'allocation de bande passante par partition peut être contrôlée de manière à appliquer des limites ou des réserves si besoin.
- Peut être utilisé indépendamment du commutateur : le commutateur n'a pas besoin de configuration particulière ou de connaissances sur l'activation NPAR.
- Peut être utilisé avec RoCE et SR-IOV.
- Prend en charge les déchargements sans état tels que LSO, TPA, RSS/TSS et RoCE (deux PF par port uniquement).
- Prise en charge d'ID de routage alternatif pour plus de huit fonctions par périphérique physique.



**Remarque :** Sur la page **Menu Dell UEFI HII-> Configuration principale -> Configuration du niveau du périphérique**, l'utilisateur peut activer **NParEP** pour permettre à la carte NXE de prendre en charge jusqu'à 16 PF par périphérique sur un système compatible ARI. Pour un périphérique à 2 ports, jusqu'à 8 PF peuvent donc être prises en charge pour chaque port.

### Limitations

- Les paramètres partagés doivent être supprimés pour éviter les conflits d'accès. Par exemple : la vitesse, le mode duplex, le contrôle de flux et les paramètres physiques similaires sont masqués par le pilote de périphérique pour éviter tout conflit.
- Les systèmes non ARI prennent en charge seulement huit partitions par périphérique physique.
- RoCE est uniquement pris en charge sur les deux premières partitions de chaque port physique, soit un total de quatre partitions par périphérique physique. En mode NPAR + SRIOV, seules deux VF de chaque port physique parent peuvent activer la prise en charge du protocole RDMA, soit un total de quatre VF + RDMA par périphérique physique.

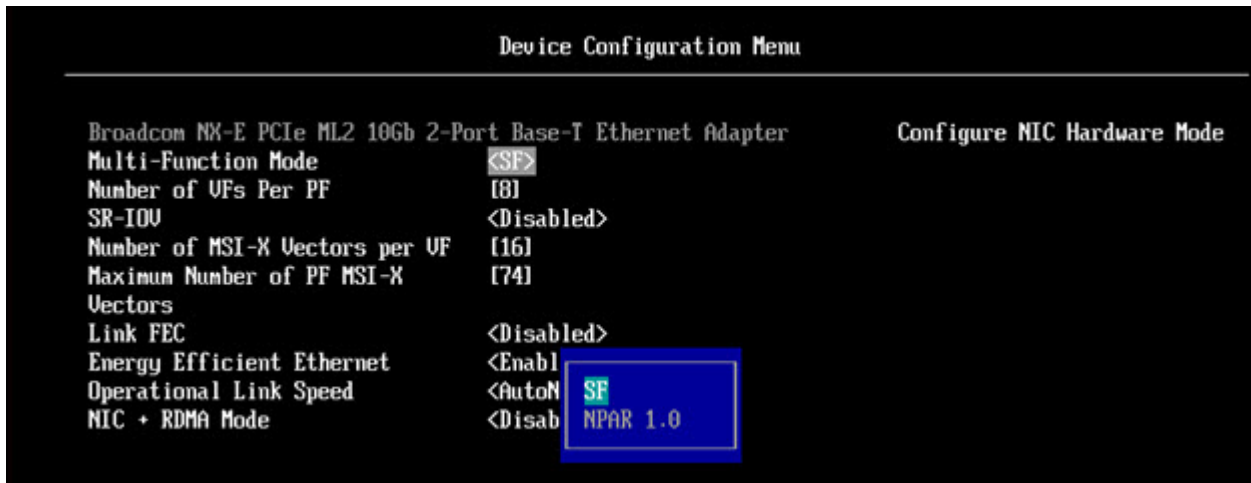
### Configuration

NPAR peut être configuré à l'aide des menus HII de configuration du BIOS ou à l'aide de l'utilitaire Broadcom CCM sur les anciens systèmes de démarrage. Certains fournisseurs exposent également la configuration via d'autres interfaces propriétaires.



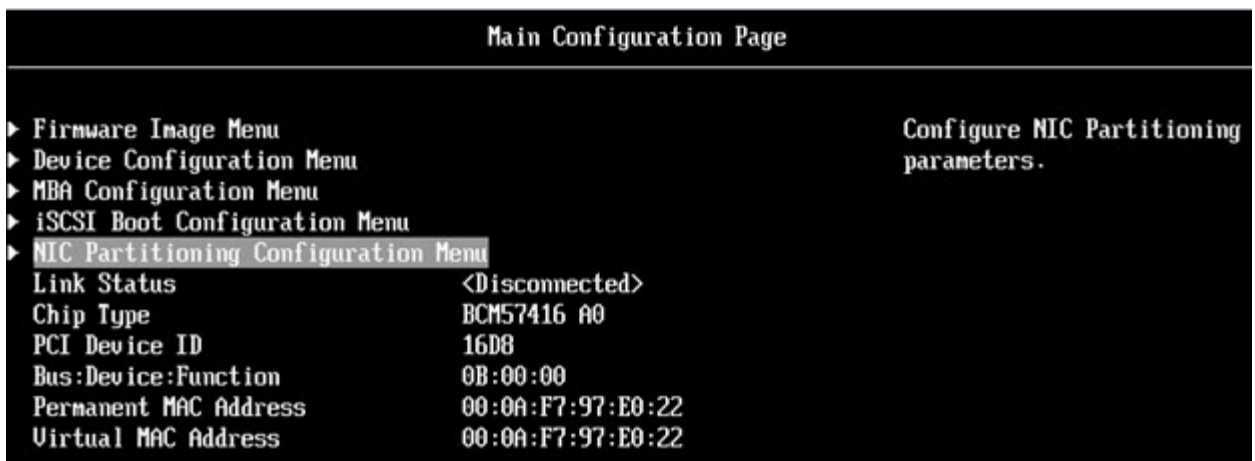
Pour activer NPAR :

1. Sélectionnez la carte réseau cible à partir du menu HII du BIOS ou de l'interface CCM et définissez le mode multifonction ou le mode virtualisation.



NPAR est activé avec SR-IOV. Pour certains systèmes OEM compatibles ARI, le bouton **NParEP** s'affiche pour autoriser explicitement la BCM5741X à prendre en charge jusqu'à 16 partitions. En passant du mode monofonction au mode multifonction, le périphérique doit être énuméré à nouveau. Par conséquent, les modifications ne seront pas prises en compte tant que le système n'a pas redémarré.

2. Une fois NPAR activé, l'option de menu Configuration principale du partitionnement de la carte réseau s'affiche dans le menu Configuration principale de la carte réseau associé à chaque port physique.



3. Le menu Configuration des partitions de la carte réseau (illustré ci-dessous) permet à l'utilisateur de choisir le nombre de partitions qui doivent être allouées à partir du port physique sélectionné. Chaque carte réseau BCM5741X peut prendre en charge 16 partitions maximum sur un serveur compatible ARI. Par défaut, les cartes à deux ports sont configurées pour huit partitions par port physique. Des options de configuration pour chaque partition sont également accessibles à partir de ce menu. Pour certains systèmes OEM, le menu HII contient également une page Allocation de bande passante globale où vous pouvez configurer simultanément la bande passante TX minimale (réservée) et maximale (limite) pour toutes les partitions.

```

NIC Partitioning Configuration Menu
-----
Broadcom BCM57416 NetXtreme-E 10GBASE-T RDMA Ethernet Controller   Configure Number of
Number of Partitions           [8]                               Partitions Per Port
▶ Partition 1 Configuration
▶ Partition 2 Configuration
▶ Partition 3 Configuration
▶ Partition 4 Configuration
▶ Partition 5 Configuration
▶ Partition 6 Configuration
▶ Partition 7 Configuration
▶ Partition 8 Configuration

```

4. Définissez les paramètres de configuration des partitions de la carte réseau (voir le [Tableau 37 page 66](#)).

```

Partition 1 Configuration
-----
Broadcom NX-E PCIe ML2 10Gb 2-Port Base-T Ethernet Adapter       Configure RDMA Support
BW Reservation              [0]
BW Limit                    [100]
BW Reservation Valid       <False>
BW Limit Valid             <False>
Support RDMA                <Disabled>
MAC Address                 00:10:18:99:98:E0

```

**Tableau 37 : Paramètres NPAR**

Paramètre	Description	Options valides
Réservation de bande passante	Pourcentage de bande passante totale disponible qui doit être réservée pour cette partition. 0 indique une répartition égale de la bande passante entre toutes les partitions.	Valeur 0-100
Limitation de bande passante	Pourcentage maximal de bande passante disponible autorisé pour cette partition.	Valeur 0-100
Réservation de bande passante valide	Fonctionne à la manière d'un commutateur marche/arrêt pour le paramètre de réservation de bande passante.	True/False
Limitation de bande passante valide	Fonctionne à la manière d'un commutateur marche/arrêt pour le paramètre de limitation de bande passante.	True/False
Prendre en charge RDMA	Fonctionne à la manière d'un commutateur marche/arrêt pour la prise en charge du protocole RDMA sur cette partition. Veuillez noter que seules deux partitions par port physique peuvent prendre en charge RDMA. Pour un périphérique double port, jusqu'à 4 partitions NPAR peuvent prendre en charge RDMA.	Activé/désactivé
Adresse MAC	Adresse MAC de cette partition.	–

## Notes sur la réduction de consommation de mémoire de la carte réseau

Puisque cette carte réseau prend en charge des vitesses de liaison plus rapides, les tampons de réception sont plus nombreux par défaut. Un plus grand nombre de paquets peut arriver dans un intervalle de temps donné lorsque la vitesse de liaison est plus élevée, et si le système hôte prend du retard dans le traitement des interruptions de réception, la carte réseau doit ignorer les paquets si tous les tampons de réception sont utilisés.

La valeur par défaut des tampons de réception a été choisie pour fonctionner de manière satisfaisante avec les configurations types. Mais si votre système compte de nombreuses cartes réseau, si vous avez activé NPAR sur plusieurs cartes réseau, ou si vous n'avez qu'une quantité de RAM limitée, vous risquez de voir un code 12 avec un signe d'exclamation jaune dans le Gestionnaire de périphériques pour certaines cartes. Le code 12 signifie que le chargement du pilote a échoué en raison de ressources insuffisantes. Dans ce cas, la ressource est un type spécifique de mémoire du noyau appelé « pool non paginé » (NPP).

Si vous recevez un code 12, ou si vous souhaitez pour d'autres raisons réduire la quantité de mémoire NPP consommée par la carte réseau :

- Réduisez le nombre de files d'attente RSS en remplaçant la valeur par défaut de 8 par 4 ou 2. Chaque file d'attente RSS se voit allouer son propre ensemble de tampons de réception, autrement dit le fait de réduire le nombre de files d'attente RSS a pour effet de réduire la mémoire NPP allouée. La réduction du nombre de files d'attente RSS peut avoir des conséquences sur les performances, car moins de cœurs participent au traitement des paquets de réception provenant de cette carte réseau. L'utilisation CPU par processeur doit être surveillée afin de vérifier l'absence de processeurs « chauds » après cette modification.
- Réduisez l'allocation de mémoire en réduisant le nombre de tampons de réception alloués. La valeur par défaut de 0 signifie que le pilote doit déterminer automatiquement le nombre de tampons de réception. Pour les configurations standard, un paramètre de 0 (=auto) sera mappé à XXXX tampons de réception par file d'attente. Vous pouvez choisir une valeur plus petite, par exemple 1 500, 1 000 ou 500. (La valeur doit être un multiple de 500, dans une plage de 500 à 15 000.) Comme mentionné ci-dessus, un plus petit nombre de tampons de réception augmente le risque de perte de paquets, ce qui affecte les retransmissions de paquets et entraîne une baisse de débit.

Les paramètres « Nombre maximal de files d'attente RSS » et « Tampon de réception (0=Auto) » peuvent être modifiés depuis le **Gestionnaire de périphériques**, sous l'onglet des **Propriétés avancées** de chaque carte réseau. Si vous souhaitez modifier plusieurs cartes réseau simultanément, il est plus rapide d'utiliser *l'applet de commande PowerShell Set-NetAdapterAdvancedProperty*. Par exemple, pour affecter deux files d'attente RSS pour toutes les cartes réseau d'un système dont le nom de la carte commence par « S1 », exécutez la commande suivante :

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

De même, pour définir le nombre de tampons de réception sur 1 500, exécutez la commande suivante :

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

Voir <https://blogs.technet.microsoft.com/wincat/2012/08/27/using-powershell-for-nic-configuration-tasks/> pour savoir comment utiliser PowerShell pour modifier les propriétés des cartes réseau.

---

## RoCE : Configuration et exemples d'utilisation

Cette section présente la configuration et les exemples d'utilisation de RoCE.

Pour activer RoCE pour des PF ou des VF, l'utilisateur doit activer la sélection RDMA dans le menu HII du BIOS pour que l'option RDMA prenne effet dans le système d'exploitation hôte ou invité.

Pour activer la technologie RDMA en mode monofonction (si le **Mode de virtualisation** est défini sur **Aucun** sur **SR-IOV**) :

1. Pendant le démarrage du système, accédez à **Configuration du système -> Paramètres du périphérique -> Carte réseau NetXtreme-E -> Page de configuration principale** et définissez le **mode Carte réseau + RMDA** sur **Activé**.

Pour activer la technologie RDMA si la virtualisation est en mode NPAR ou NPAR+SR-IOV :

1. Pendant le démarrage du système, accédez à **Configuration du système -> Paramètres du périphérique -> Carte réseau NetXtreme-E -> Configuration du partitionnement de la carte réseau-> Configuration de la partition 1 (ou 2)** et définissez le **mode Carte réseau + RMDA** sur **Activé**.



**Remarque** : Si vous utilisez le mode NPAR + SRIOV, seules deux VF de chaque port physique parent peuvent activer la prise en charge du protocole RDMA, soit un total de quatre VF + RDMA par périphérique physique.

## Configuration sous Linux

### Configuration requise

Pour configurer RoCE dans Linux, vous avez besoin des éléments suivants :

- `bnxt_en-roce` RoCE (pilote `bnxt_en` pris en charge par RoCE qui fait partie de l'archive tar publiée et compressée sous gzip)
- `bnxt_re` (pilote RoCE)
- `libbnxtre` (module de bibliothèque RoCE en mode utilisateur)

### Dépendances du pilote `BNXT_RE`

Le pilote `Bnxt_re` requiert une version spéciale compatible RoCE de `bnxt_en`, qui est incluse dans le package `netxtreme-bnxt_en-1.7.9.tar.gz` (ou plus récent). La compilation du pilote `bnxt_re` intervient selon que la pile IB est disponible avec la distribution du système d'exploitation ou qu'un OFED externe est nécessaire.



**Remarque** : Il est nécessaire de charger la bonne version de `bnxt_en` incluse dans le même package `netxtreme-bnxt_en-1.7.x.tar.gz`. `Bnxt_re` et `Bnxt_en` fonctionnent par paire pour activer le trafic RoCE. L'utilisation de versions incompatibles de ces deux pilotes entraîne des résultats peu fiables ou imprévisibles.

- Distributions ayant une pile IB disponible avec la distribution du système d'exploitation :

RH7.1/7.2/7.3/6.7/6.8, SLES12SP2 et Ubuntu 16.04

S'ils ne sont pas déjà installés, vous pouvez installer la pile IB et des utilitaires pratiques dans Redhat en exécutant les commandes suivantes avant la compilation de `bnxt_re` :

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

Pour compiler `bnxt_re` :

```
$make
```

- Distributions pour lesquelles un OFED externe doit être installé :

SLES11SP4

Veillez consulter les notes de version OFED à partir du lien suivant et installer OFED avant de compiler le pilote `bnxt_re`.

[http://downloads.openfabrics.org/downloads/OFED/release\\_notes/OFED\\_3.18-2\\_release\\_notes](http://downloads.openfabrics.org/downloads/OFED/release_notes/OFED_3.18-2_release_notes)

Pour compiler `bnxt_re` :

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

## Installation

Pour installer RoCE sous Linux :

1. Mettez à niveau la NVRAM de la carte réseau en utilisant les packages de micrologiciel pris en charge par RoCE à partir du logiciel version 20.06.04.01 ou antérieure.
2. Dans le système d'exploitation, décompressez, générez et installez les pilotes Linux L2 et RoCE de BCM5741X.
  - a. # tar -xzf netxtreme-bnxt\_en-1.7.9.tar.gz
  - b. # cd netxtreme-bnxt\_en-bnxt\_re
  - c. # make build && make install
3. Décompressez, générez et installez la bibliothèque utilisateur RoCE Linux pour NetXtreme-E.
  - 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

Veillez vous reporter au fichier `README.txt` du pilote `bnxt_re` pour plus de détails sur les options configurables et les recommandations applicables.

## Limitations

Dans les cartes réseau à deux ports, si les deux ports se trouvent sur le même sous-réseau, les commandes rdma perftest peuvent échouer. Cet échec peut être dû à un problème de flux arp dans le système d'exploitation Linux. Pour contourner cette limitation, utilisez plusieurs sous-réseaux de test ou désactivez le deuxième port ou la deuxième interface.

## Problèmes connus

Bnxt\_en et Bnxt\_re sont conçus pour fonctionner par paire. Les pilotes Bnxt\_en antérieurs à la version 1.7.x ne prennent pas en charge RDMA et ne peuvent pas être chargés en même temps que le pilote Bnxt\_re (RDMA). L'utilisateur peut rencontrer un plantage système et devoir redémarrer s'il charge le pilote Bnxt\_re avec des pilotes Bnxt\_en de génération précédente. Il est recommandé que l'utilisateur charge les modules Bnxt\_en et Bnxt\_re à partir du même bundle netxtreme-bnxt\_en-<1.7.x>.tar.gz.

Pour éviter de charger une combinaison incompatible de pilotes bnxt\_en et bnxt\_re, respectez les conditions suivantes :

- Si RedHat/CentOS 7.2 a été installé sur le système cible à l'aide de PXEboot avec un DUD bnxt\_en ou un RPM de module noyau, supprimez le fichier bnxt\_en.ko qui se trouve dans /lib/modules/\$(uname -r)/extra/bnxt\_en/bnxt\_en.ko ou modifiez /etc/depmo.d/.
- Ignorez le fichier bnxt\_en.conf pour utiliser la version à jour. Les utilisateurs peuvent également effacer le pilote du noyau Linux BCM5741X à l'aide de la commande rpm -e kmod-bnxt\_en. RHEL 7.3/SLES 12 SP2 possède un pilote intégré bnxt\_en (antérieur à la version 1.7.x). Ce pilote doit être supprimé et le dernier pilote bnxt\_en être ajouté avant d'appliquer le bnxt\_re (pilotes RoCE).

## Windows

### Mode noyau

Windows Server 2012 (et versions supérieures) appelle la fonction RDMA dans la carte réseau pour le trafic de fichiers SMB si les deux extrémités sont activées pour RDMA. Le pilote miniport Broadcom NDIS bnxtnd.sys v20.6.2 et versions supérieures prend en charge les protocoles RoCEv1 et RoCEv2 via l'interface NDKPI. RoCEv1 est le paramètre par défaut.

Pour activer RDMA :

1. Mettez à niveau la NVRAM de la carte réseau à l'aide des packages appropriés. Dans CCM ou dans le menu HII UEFI, activez la prise en charge de RDMA.
2. Accédez à la page **Propriétés avancées** de l'adaptateur et définissez **Fonctionnalités NetworkDirect** sur **Activé** pour chaque miniport BCM5741X, ou, à l'aide de la fenêtre de commande PowerShell, exécutez la commande suivante :

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

3. Les commandes PowerShell suivantes retournent la valeur true si **NetworkDirect** est activé.
  - a. Get-NetOffloadGlobalSetting
  - b. Get-NetAdapterRDMA

## Vérification de RDMA

Pour vérifier la fonctionnalité RDMA :

1. Créez un partage de fichiers sur le système distant et ouvrez ce partage à l'aide de l'Explorateur Windows ou de « net use... ». Pour éviter tout goulot d'étranglement en lecture/écriture du disque dur, il est recommandé d'utiliser un disque RAM comme partage réseau de test.
2. A partir de PowerShell, exécutez les commandes suivantes :

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

Si le client et le serveur affichent la valeur true, les transferts de fichiers sur cette connexion SMB utilisent SMB.

3. Les commandes suivantes peuvent être utilisées pour activer/désactiver la fonctionnalité SMB Multichannel :

### Côté serveur :

- Activer : Set-SmbServerConfiguration -EnableMultiChannel \$true
- Désactiver : Set-SmbServerConfiguration -EnableMultiChannel \$false

### Côté client :

- Activer : Set-SmbClientConfiguration -EnableMultiChannel \$true
- Désactiver : Set-SmbClientConfiguration -EnableMultiChannel \$false



**Remarque :** Par défaut, le pilote configure deux connexions RDMA pour chaque partage réseau par adresse IP (sur un sous-réseau unique). L'utilisateur peut augmenter le nombre de connexions RDMA en ajoutant plusieurs adresses IP, chacune associée à un sous-réseau différent, pour le même port physique de test. Plusieurs partages réseau peuvent être créés et mappés pour chaque partenaire de liaison à l'aide des adresses IP uniques créées.

Par exemple :

```
Sur le serveur 1, créez les adresses IP suivantes pour le Port réseau1.  
172.1.10.1  
172.2.10.2  
172.3.10.3
```

```
Sur le même serveur 1, créez 3 partages.  
Partage1  
Partage2  
Partage3
```

```
Sur les partenaires de liaison réseau,  
Connectez-vous à \\172.1.10.1\partage1  
Connectez-vous à \\172.2.10.2\partage2  
Connectez-vous à \\172.3.10.3\partage3  
...etc.
```

## Mode utilisateur

Avant d'exécuter une application en mode utilisateur écrite sur NDSPI, copiez et installez le pilote en mode utilisateur `bxndspi.dll`. Pour copier et installer le pilote en mode utilisateur :

1. Copiez `bxndspi.dll` dans `C:\Windows\System32`.
2. Installez le pilote en exécutant la commande suivante :

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

## VMware ESX

### Limitations

La version actuelle du pilote pris en charge par RoCE requiert ESXi-6.5.0 GA version 4564106 ou supérieure.

### Configuration requise du pilote RoCE de la carte réseau BNXT

Le pilote BNXTNET L2 doit être installé avec le paramètre de module `disable_roce=0` avant l'installation du pilote.

Pour définir le paramètre de module, exécutez la commande suivante :

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

Veuillez utiliser le pilote ESX6.5 L2 version 20.6.9.0 (pilote L2 pris en charge par RoCE) ou supérieure.

### Installation

Pour installer le pilote RoCE :

1. Copiez le fichier `<bnxtroce>-<version du pilote>.vib` dans `/var/log/vmware` à l'aide des commandes suivantes :

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

2. Redémarrez l'ordinateur.
3. Vérifiez que les pilotes sont correctement installés à l'aide de la commande suivante :

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

4. Pour désactiver l'ECN (activé par défaut) pour le trafic RoCE, utilisez le paramètre de module « `tos_ecn=0` » pour `bnxtroce`.



## Configuration des cartes réseau PVRDMA

Veillez vous reporter au lien VMware ci-dessous pour plus d'informations sur la configuration et l'utilisation des cartes réseau Paravirtualized RDMA (PVRDMA).

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

### Configuration d'un centre virtuel pour PVRDMA

Pour configurer un centre virtuel pour PVRDMA :

1. Créez un DVS (commutateur virtuel distribué requis pour PVRDMA)
2. Ajoutez l'hôte au DVS.

### Balitage d'un vmknic pour PVRDMA sur les hôtes ESX

Pour baliser un vmknic pour PVRDMA afin de l'utiliser sur des hôtes ESX :

1. Sélectionnez l'hôte et cliquez avec le bouton droit de la souris sur **Paramètres** pour passer à la page des paramètres de l'onglet **Gérer**.
2. Sur la page **Paramètres**, développez **Système** et cliquez sur **Paramètres système avancés** pour afficher la valeur de paire de clés des paramètres système avancés et son résumé.
3. Cliquez sur **Edition** pour afficher l'option **Modifier les paramètres système avancés**.  
Filtrez sur **PVRDMA** pour réduire tous les paramètres à `net.PVRDMAMknics`.
4. Définissez la valeur `Net.PVRDMAMknics` sur vmknic, comme dans l'exemple vmk0

### Définition de la règle de pare-feu pour PVRDMA

Pour définir la règle de pare-feu pour PVRDMA :

1. Sélectionnez l'hôte et cliquez avec le bouton droit de la souris sur **Paramètres** pour passer à la page des paramètres de l'onglet **Gérer**.
2. Sur la page **Paramètres**, développez **Système** et cliquez sur **Profil de sécurité** pour afficher le résumé du pare-feu.
3. Cliquez sur **Edition** pour afficher l'option **Modifier le profil de sécurité**.
4. Faites défiler vers le bas pour rechercher pvr dma et cochez la case pour définir le pare-feu.

### Ajout d'un périphérique PVRDMA à la machine virtuelle

Pour ajouter un périphérique PVRDMA à la machine virtuelle :

1. Sélectionnez la machine virtuelle et cliquez avec le bouton droit de la souris sur **Modifier les paramètres**.
2. Ajoutez une nouvelle carte réseau.
3. Sélectionnez le réseau comme **Commutateur virtuel distribué** et **Groupe de ports**.
4. Pour le **Type de carte**, sélectionnez **PVRDMA** et cliquez sur **OK**.

## Configuration de la machine virtuelle sur le système d'exploitation invité Linux



**Remarque :** L'utilisateur doit installer les outils de développement appropriés, y compris git, avant de passer aux étapes de configuration ci-dessous.

1. Téléchargez le pilote et la bibliothèque PVRDMA en utilisant les commandes suivantes :

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

2. Compilez et installez le pilote et la bibliothèque invité PVRDMA.
3. Pour installer le pilote, exécutez `make && sudo insmod pvrdma.ko` dans le répertoire du pilote.  
Le pilote doit être chargé après le chargement du pilote vmxnet3 apparié.



Les modules de noyau RDMA installés peuvent ne pas être compatibles avec le pilote PVRDMA. Si tel est le cas, supprimez l'installation actuelle et redémarrez. Suivez ensuite les instructions d'installation. Veuillez lire le fichier README figurant dans le répertoire du pilote pour plus d'informations sur les différentes piles RDMA.

4. Pour installer la bibliothèque, exécutez `./autogen.sh && ./configure --sysconfdir=/etc && make && sudo make install` dans le répertoire de la bibliothèque.



**Remarque :** Le chemin d'installation de la bibliothèque doit se trouver dans le cache partagé de la bibliothèque. Suivez les instructions du fichier d'installation contenu dans le répertoire de la bibliothèque.



**Remarque :** Les paramètres de pare-feu peuvent être modifiés pour autoriser le trafic RDMA. Veuillez vérifier que les paramètres du pare-feu sont en place.

5. Ajoutez `/usr/lib` dans le fichier `/etc/ld.so.conf` et rechargez `ldconf` en exécutant `ldconfig`.
6. Chargez les modules `ib` en utilisant `modprobe rdma_ucm`.
7. Chargez le module de noyau PVRDMA à l'aide de `insmod pvrdma.ko`.
8. Affectez une adresse IP à l'interface PVRDMA.
9. Vérifiez si le périphérique IB a été créé en exécutant la commande `ibv_devinfo -v`.

## DCBX – Data Center Bridging

Les contrôleurs Broadcom NetXtreme-E prennent en charge IEEE802.1Qaz DCBX ainsi que l'ancienne spécification CEE DCBX. La configuration de DCB est obtenue en échangeant les paramètres configurés en local avec l'homologue de liaison. Etant donné que les deux extrémités d'une liaison peuvent être configurées différemment, DCBX utilise un concept de « préparation » pour indiquer quelle extrémité de la liaison est prête à accepter les paramètres de l'autre extrémité. Ceci est indiqué dans le protocole DCBX en utilisant un seul bit dans la configuration ETS et dans PFC TLV ; ce bit n'est pas utilisé avec les TLV ETS Recommandation et Application Priority. Par défaut, la carte réseau NetXtreme-E est en mode « prêt » tandis que le commutateur réseau du partenaire de liaison est en mode « non prêt ». Cela permet de propager à l'ensemble du réseau le même paramètre DCBX sur le commutateur.

Les utilisateurs peuvent définir manuellement la carte réseau NetXtreme-E en mode non prêt et effectuer diverses configurations PFC, Strict Priority, ETS et APP à partir de l'hôte. Veuillez vous reporter au fichier readme.txt du pilote pour plus de détails sur les configurations disponibles. Ce document fournit un exemple de configuration possible dans Windows avec Windows PowerShell. Des informations supplémentaires sur DCBX, la QoS et des exemples d'utilisation associés sont fournis plus en détail dans un autre livre blanc, au-delà du périmètre de ce manuel de l'utilisateur.

Les paramètres suivants du menu HII UEFI sont requis pour activer la prise en charge DCBX :

**Configuration du système-> Paramètres du périphérique -> NetXtreme-E-> Configuration du niveau du périphérique**

### Profil QoS : Profil de file d'attente QoS par défaut

Il est nécessaire de configurer les ressources QoS (qualité de service) pour prendre en charge diverses exigences PFC et ETS lorsqu'un affinage du réglage au-delà de l'allocation de bande passante est requis. La carte NetXtreme-E permet à l'administrateur de choisir de dédier des ressources matérielles de la carte réseau pour prendre en charge les trames Jumbo et/ou de combiner des files d'attente CoS (classe de service) avec et sans perte. De nombreuses combinaisons de configuration sont possibles, ce qui peut compliquer le calcul. Cette option permet à un utilisateur d'effectuer sa sélection dans une liste de profils de file d'attente QoS précalculés. Ces profils précalculés sont conçus pour optimiser la prise en charge des exigences PFC et ETS dans les déploiements clients types.

Voici une brève description de chaque profil QoS.

**Tableau 38 : Profils QoS**

<i>N° de profil</i>	<i>Prise en charge des trames Jumbo</i>	<i>Nb de files d'attente CoS/ports avec perte</i>	<i>Nb de files d'attente CoS/ports sans perte</i>	<i>Prise en charge des SKU à 2 ports</i>
<b>Profil n° 1</b>	Oui	0	1 (PFC pris en charge)	Oui (25 Gbit/s)
<b>Profil n° 2</b>	Oui	4	2 (PFC pris en charge)	Non
<b>Profil n° 3</b>	Non. (MTU <= 2 Ko)	6	2 (PFC pris en charge)	Oui (25 Gbit/s)
<b>Profil n° 4</b>	Oui	1	2 (PFC pris en charge)	Oui (25 Gbit/s)

Tableau 38 : Profils QoS (suite)

N° de profil	Prise en charge des trames Jumbo	Nb de files d'attente CoS/ports avec perte	Nb de files d'attente CoS/ports sans perte	Prise en charge des SKU à 2 ports
Profil n° 5	Oui	1	0 (aucune prise en charge PFC)	Oui (25 Gbit/s)
Profil n° 6	Oui	8	0 (aucune prise en charge PFC)	Oui (25 Gbit/s)
Profil n° 7	Cette configuration optimise l'affectation du tampon des paquets sur deux CoS sans perte Files d'attente qui optimisent les performances RoCE tout en offrant plus de flexibilité.			
	Oui	0	2	Oui (25 Gbit/s)
Par défaut	Oui	Identique au profil n° 4		Oui

## Mode DCBX = Activer (IEEE uniquement)

Cette option permet à un utilisateur d'activer/désactiver DCBX avec la spécification indiquée. IEEE indique uniquement que DCBX IEEE802.1Qaz est sélectionné.

Paramétrage du pilote Windows :

Après avoir activé les options indiquées dans le menu HII UEFI pour définir les paramètres au niveau du micrologiciel, effectuez les sélections suivantes dans les propriétés avancée du pilote Windows.

Ouvrez le menu **Gestionnaire de périphériques Windows -> Carte Broadcom NetXtreme E Series -> Propriétés avancées -> onglet Avancé.**

- Qualité de service = Activée
- Priorité et VLAN = Priorité et VLAN activés
- VLAN = <ID>
- Définissez l'ID de VLAN souhaité

Pour exécuter la commande associée à DCB dans Windows PowerShell, installez la fonction DCB Windows appropriée.

1. Dans la **barre des tâches**, cliquez avec le bouton droit sur l'icône Windows PowerShell, puis cliquez sur **Exécuter en tant qu'administrateur**. Windows PowerShell s'ouvre en mode élevé.
2. Dans la console Windows PowerShell, entrez :

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

## Bit Willing DCBX

Le bit Willing DCBX est spécifié dans la spécification DCB. Si le bit Willing d'un périphérique a la valeur true, le périphérique est prêt à accepter les configurations à partir d'un périphérique distant via DCBX. Si le bit Willing d'un périphérique a la valeur false, le périphérique rejette toute configuration à partir d'un périphérique distant et applique uniquement les configurations locales.

Utilisez la commande suivante pour définir le bit Willing sur True ou False. 1 pour activé, 0 pour désactivé.

Par exemple `set-netQoSdcbxSetting -willing 1`

Utilisez la commande suivante pour créer une classe de trafic.

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



**Remarque :** Par défaut, toutes les valeurs 802.1p sont mappées à une classe de trafic par défaut, qui possède 100 % de la bande passante de la liaison physique. La commande ci-dessus crée une nouvelle classe de trafic à laquelle sont mappés tous les paquets marqués avec huit IEEE 802.1p de valeur 4, avec un algorithme de sélection de transmission (TSA) défini sur ETS et une allocation de 30 % de bande passante.

Il est possible de créer jusqu'à sept nouvelles classes de trafic. En plus de la classe de trafic par défaut, il y a au maximum huit classes de trafic dans le système.

Utilisez la commande suivante pour afficher la classe de trafic créée :

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

Utilisez la commande suivante pour modifier la classe de trafic :

```
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
```

Utilisez la commande suivante pour supprimer la classe de trafic :

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

Utilisez la commande suivante pour créer une classe de trafic (Priorité stricte) :

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

Activation de PFC :

```
PS C:\> Enable-NetQoSFlowControl -priority 4
PS C:\> Get-NetQoSFlowControl -priority 4
Priority Enabled (Priorité activée)
-----
4 True

PS C:\> Get-NetQoSFlowControl
```

Désactivation de PFC :

```
PS C:\> disable-NetQosflowControl -priority 4
PS C:\> get-NetQosFlowControl -priority 4
Priority Enabled (Priorité activée)
-----
4 False
```

Utilisez la commande suivante pour créer une règle QoS :

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

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



**Remarque :** La commande ci-dessus crée une nouvelle règle pour SMB. -SMB est un filtre intégré qui correspond au port TCP 445 (réservé à SMB). Si un paquet est envoyé sur le port TCP 445, il est identifié par le système d'exploitation avec la valeur 802.1p de 4 avant que le paquet soit transmis à un pilote miniport réseau.

D'autres filtres par défaut s'ajoutent au filtre SMB : -iSCSI (correspondant au port TCP 3260), -NFS (correspondant au port TCP 2049), -LiveMigration (correspondant au port TCP 6600), -FCOE (correspondant à EtherType 0x8906) et -NetworkDirect.

NetworkDirect est une couche d'abstraction créée en amont de toute implémentation RDMA sur une carte réseau. -NetworkDirect doit être suivi d'un port Network Direct.

En plus des filtres par défaut, un utilisateur peut classer le trafic par nom d'exécutable de l'application (comme dans le premier exemple ci-dessous), ou par adresse IP, par port ou par protocole.

Utilisez la commande suivante pour créer une règle QoS basée sur l'adresse source/destination :

```
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 : All
Precedence : 127
IPProtocol : Both
IPDstPrefix : 10.240.1.0/24
PriorityValue : 7
```

Utilisez la commande suivante pour afficher la règle QoS :

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

```
NetworkProfile : All
Precedence : 127
Template : SMB
PriorityValue : 4
```

Utilisez la commande suivante pour modifier la règle QoS :

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

Utilisez la commande suivante pour supprimer la règle QoS :

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

---

## Foire aux questions

- La négociation automatique est-elle prise en charge à la vitesse 25G ?  
Oui. Veuillez vous reporter à [« Configuration de la négociation automatique » page 41](#) pour plus d'informations.
- Comment dois-je connecter le câble SFP28 aux ports QSFP ?  
Des câbles adaptateurs sont disponibles pour connecter des ports QSFP à 4 ports SFP28.
- Quelles sont les vitesses de port compatibles ?  
Pour les périphériques à double port BCM57404AXXX/BCM57414, les vitesses des deux ports doivent être compatibles entre elles. Les vitesses 10 Gbit/s et 25 Gbit/s ne sont pas compatibles. Si un port est défini sur 10 Gbit/s, l'autre port ne peut pas être défini sur 25 Gbit/s. Si l'utilisateur tente de définir des vitesses de port incompatibles, le deuxième port à activer ne pourra pas effectuer la liaison. Veuillez vous reporter à [« Configuration de la négociation automatique » page 41](#) pour plus d'informations.
- Puis-je utiliser une vitesse de 10 Gbit/s pour la connectivité PXE sur un port à 25 Gbit/s ?  
Seule la vitesse 25 Gbit/s est prise en charge pour la connectivité PXE. Il n'est pas recommandé d'utiliser une connectivité PXE de 10 Gbit/s sur une carte 25 Gbit/s. Ceci est dû à l'absence de prise en charge de la négociation automatique sur les commutateurs 10 Gbit/s existants et aux éventuelles complications liées à des paramètres incompatibles de vitesse de liaison des ports.