

Groupe de travail Réseau
Request for Comments : 4392
 Catégorie : Information

V. Kashyap, IBM
 avril 2006
 Traduction Claude Brière de L'Isle

Architecture IP sur InfiniBand (IPoIB)

Statut de ce mémoire

Le présent mémoire apporte des informations pour la communauté de l'Internet. Le présent mémoire ne spécifie aucune sorte de norme de l'Internet. La distribution du présent mémoire n'est soumise à aucune restriction.

Notice de Copyright

Copyright (C) The Internet Society (2006).

Résumé

InfiniBand est une interconnexion à grande vitesse, fondée sur le canal entre systèmes et appareils.

Le présent document présente une vue d'ensemble de l'architecture InfiniBand. Il décrit de plus les exigences et lignes directrices pour la transmission de IP sur InfiniBand. Les discussions de ce document sont applicables à IPv4 et IPv6 sauf mention explicite contraire. L'encapsulation de IP sur InfiniBand et le mécanisme pour la résolution d'adresse IP sur des tissus IB sont couverts dans d'autres documents.

Table des matières

| | |
|--|----|
| 1. Introduction à InfiniBand..... | 1 |
| 1.1 Spécification de l'architecture InfiniBand..... | 1 |
| 1.2 Vue d'ensemble de l'architecture InfiniBand..... | 2 |
| 1.3 Gestion de groupe de diffusion groupée InfiniBand..... | 5 |
| 2. Gestion de sous réseau InfiniBand..... | 7 |
| 3. IP sur IB..... | 7 |
| 3.1 InfiniBand comme liaison de données..... | 8 |
| 3.2 Prise en charge de la diffusion groupée..... | 8 |
| 3.3 Sous réseaux IP à travers des sous réseaux IB..... | 9 |
| 4. Sous réseaux IP dans un tissu InfiniBand..... | 9 |
| 4.1 VLAN IPoIB..... | 10 |
| 4.2 Diffusion groupée dans des sous réseaux IPoIB..... | 10 |
| 4.3 Transmission des paquets IPoIB..... | 12 |
| 4.4 Protocole de résolution d'adresse inverse (RARP) et entrées ARP statiques..... | 12 |
| 4.5 DHCPv4 et IPoIB..... | 12 |
| 5. QS et questions en rapport..... | 13 |
| 6. Considérations sur la sécurité..... | 13 |
| 7. Remerciements..... | 13 |
| 8. Références..... | 13 |
| 8.1 Références normatives..... | 13 |
| 8.2 Références pour information..... | 13 |
| Adresse de l'auteur..... | 14 |
| Déclaration complète de droits de reproduction..... | 14 |

1. Introduction à InfiniBand

L'association commerciale InfiniBand (IBTA, *InfiniBand Trade Association*) a été formée pour développer une spécification d'une technologie de tissu commuté fondée sur le canal. La norme InfiniBand vise à satisfaire les exigences d'adaptabilité, de fiabilité, de disponibilité, et de performances des serveurs dans des centres de données.

1.1 Spécification de l'architecture InfiniBand

On peut télécharger la spécification de l'association commerciale InfiniBand à <http://www.infinibandta.org>.

1.2 Vue d'ensemble de l'architecture InfiniBand

Pour une vue plus complète, le lecteur se reportera au Chapitre 3 de la spécification InfiniBand.

L'architecture InfiniBand (IBA, *InfiniBand Architecture*) définit un réseau de zone système (SAN, *System Area Network*) pour connecter plusieurs plates-formes indépendantes de processeurs, des plates-formes entrée/sortie, et des appareils entrée/sortie. Le SAN IBA est une infrastructure de communications et de gestion qui prend en charge les communications à la fois entrée/sortie et inter-processeurs pour un ou plusieurs systèmes informatiques.

Un SAN IBA consiste en nœuds processeurs et en unités d'entrée/sortie connectés par un tissu IBA constitué de commutateurs en cascade et de routeurs IB (qui connectent des sous réseaux IB). Les unités d'entrée/sortie peuvent avoir une complexité qui va d'un seul appareil à circuit intégré spécifique d'une application (ASIC, *Application-specific Integrated Circuit*) rattaché à l'IBA comme un adaptateur de LAN) à un grand sous système riche en mémoire de matrice redondante de disques indépendants (RAID, *Redundant Array of Independent Disks*).

Un réseau IBA peut être subdivisé en sous réseaux interconnectés par des routeurs. Ce sont des routeurs IB et des sous réseaux IB et non des routeurs IP ou sous réseaux IP. Le présent document va se référer aux routeurs et sous réseaux InfiniBand comme respectivement des "routeurs IB" et des "sous réseaux IB". Les routeurs et sous réseaux IP sont appelés respectivement "routeurs" et "sous réseaux".

Chaque nœud ou commutateur IB peut se rattacher à un seul commutateur ou à plusieurs ou directement à chaque autre. Chaque unité IB s'interface avec la liaison au moyen d'adaptateurs de canaux (CA, *Channel Adapter*). L'architecture supporte plusieurs CA par unité avec chaque CA fournissant un ou plusieurs accès qui connectent au tissu. Chaque CA apparaît comme un nœud au tissu.

Les accès sont les points d'extrémité auxquels sont envoyées les données. Cependant, chacun des accès peut inclure plusieurs paires de file d'attente (QP, *Queue Pair*) auxquelles un homologue distant peut s'adresser directement. Du point de vue du transfert des données le numéro de paire de file d'attente (QPN, *QP Number*) fait partie de l'adresse.

IBA prend en charge entre les accès les deux services en mode connexion et de datagramme. Les homologues sont identifiés par le QPN et l'identifiant d'accès. Il y a deux exceptions. Les QPN ne sont pas utilisés quand les paquets sont en diffusion groupée. Les QPN ne sont pas non plus utilisés dans le mode datagramme brut.

Un accès, dans un paquet de données, est identifié par un identifiant local (LID, *Local Identifier*) et facultativement par un identifiant mondial (GID, *Global Identifier*). Le GID dans le paquet n'est nécessaire que quand la communication est à travers un sous réseau IB, mais il peut toujours être inclus.

Le GID est long de 128 bits et est formé par l'enchaînement d'un préfixe de sous réseau IB de 64 bits et d'une portion conforme à EUI-64 de 64 bits (EUI, *Extended Unique Identifier*). La portion EUI-64 d'un GID est appelée l'identifiant unique au monde (GUID, *Global Unique Identifier*). Le LID est une valeur de 16 bits qui est allouée quand l'accès devient actif. Le GUID est le seul identifiant persistant d'un accès. Cependant, il ne peut pas être utilisé comme adresse dans un paquet. Si le préfixe est modifié, alors le GID peut changer. Le gestionnaire de sous réseau peut tenter de garder les valeurs de LID constantes à travers les réamorçages, mais ce n'est pas exigé.

L'allocation du GID et du LID est faite par le gestionnaire de sous réseau (SM, *Subnet Manager*). Chaque sous réseau IB a au moins un composant gestionnaire de sous réseau qui contrôle le tissu. Il alloue les LID et GID. Le gestionnaire de sous réseau programme aussi les commutateurs afin qu'ils acheminent les paquets entre les destinations. Le SM et un composant en rapport avec lui, l'administrateur de sous réseau (SA, *Subnet Administrator*) sont les dépositaires centraux de toutes les informations requises pour établir et activer le tissu.

Les routeurs IB sont les composants qui acheminent les paquets entre les sous réseaux IB sur la base des GID. Donc, au sein d'un sous réseau IB, un paquet peut ou non inclure un GID mais quand il passe à travers un sous réseau IB, le GID doit être inclus. Un LID est toujours nécessaire dans un paquet car la destination au sein d'un sous réseau est déterminée par lui.

Un CA et un commutateur peuvent avoir plusieurs accès. Chaque accès de CA reçoit son propre LID ou gamme de LID. Les accès d'un commutateur ne sont pas adressables par LID/GID ou, en d'autres termes, sont transparents aux nœuds de l'autre extrémité. Chaque accès a son propre ensemble de mémoires tampon. La mise en mémoire tampon est transportée par des canaux virtuels (VL, *Virtual Lane*) où chaque VL a son propre contrôle de flux. Il peut y avoir jusqu'à 16 VL.

Les VL fournissent un mécanisme pour créer de multiples liaisons virtuelles au sein d'une seule liaison physique. Tous les accès doivent prendre en charge VL15 qui est réservé exclusivement aux datagrammes de gestion de réseau et donc ne concerne pas les discussions de IP sur InfiniBand (IPoIB). Le VL réel qu'utilise un paquet est configuré par le SM dans les tableaux d'adaptateur de commutateur/canal et est déterminé sur la base du niveau de service (SL, *Service Level*) spécifié dans chaque paquet. Il y a 16 SL possibles.

En plus des caractéristiques décrites ci-dessus avec les QP, SL, et l'adressage (GID/LID) IBA définit aussi ce qui suit :

Partitionnement : chaque paquet, sauf les datagrammes bruts, porte la clé de partition (P_Key). Ces valeurs sont utilisées pour créer un flot dans le tissu. Un commutateur (qui est une caractéristique facultative) peut être programmé par le SM pour éliminer les paquets qui n'ont pas une certaine clé. Les accès de CA vérifient toujours les P_Key. Un accès de CA peut appartenir à plusieurs partitions. La vérification de P_Key est facultative chez les routeurs IB. Une P_Key peut être décrite comme ayant une "appartenance limitée" ou une "appartenance complète". Pour qu'un paquet soit accepté, au moins une des P_Key (c'est-à-dire, la P_Key dans le paquet ou la P_Key dans l'accès) doit être une P_Key à "appartenance complète".

Q_Key : les Q_Key sont utilisées pour appliquer les droits d'accès pour les services de datagrammes IB fiables et non fiables. Les services de datagrammes bruts n'utilisent pas de Q_Key. À l'établissement de la communication, les points d'extrémité échangent les Q_Key et doivent toujours utiliser les Q_Key pertinentes quand ils communiquent les uns avec les autres. Les paquets en diffusion groupée utilisent la Q_Key associée au groupe de diffusion groupée. Les Q_Key avec le bit de poids fort établi sont considérées comme des Q_Key contrôlées (comme la Q_Key d'interface de service général (GSI, *General Service Interface*) [IB_ARCH]) et un adaptateur de canal hôte (HCA, *Host Channel Adapter*) ne permettent pas à un consommateur de spécifier arbitrairement une Q_Key contrôlée. Une tentative d'envoi d'une Q_Key contrôlée résulte en l'utilisation de la Q_Key dans le contexte de QP. Donc, le système d'exploitation conserve le contrôle car il peut configurer le contexte de QP pour la Q_Key contrôlée pour les utilisateurs privilégiés. On notera que bien que la notion d'une Q_Key contrôlée soit suggérée par la spécification IB, elle n'exige pas son utilisation ni sa mise en œuvre.

Prise en charge de la diffusion groupée : un commutateur peut prendre en charge la diffusion groupée, c'est-à-dire, la duplication des paquets sur plusieurs accès de sortie. C'est une caractéristique facultative. De même, la prise en charge de l'envoi/réception de paquets en diffusion groupée est facultative dans les CA. Un groupe de diffusion groupée est identifié par un GID. Le format de GID est défini dans la RFC 2373 sur l'adressage IPv6 [IB_ARCH]. Donc, du point de vue de IPv6 sur InfiniBand, l'adresse de liaison de données de diffusion groupée ressemble à l'adresse réseau. Un accès IB doit explicitement se joindre à un groupe de diffusion groupée en envoyant une demande au SM pour recevoir les paquets en diffusion groupée. Un accès peut envoyer des paquets à tout groupe de diffusion groupée. Dans les deux cas, le LID de diffusion groupée à utiliser dans les paquets est reçu du SM.

Il y a six méthodes pour le transfert de données dans l'architecture IB :

1. **Datagramme non fiable (sans accusé de réception – sans connexion)** : le service de datagramme non fiable (UD, *Unreliable Datagram*) est sans connexion et sans accusé de réception. Il permet à la QP de communiquer avec toute QP de datagramme non fiable dans tout nœud.
Les commutateurs, et donc chaque liaison, peuvent prendre en charge seulement une certaine MTU. Les gammes de MTU sont 256 octets, 512 octets, 1024 octets, 2048 octets, et 4096 octets. Un paquet UD ne peut pas être plus grand que la MTU de liaison entre les deux homologues.
2. **Datagramme fiable (acquitté – multiplexé)** : le service de datagrammes fiables (RD, *Reliable Datagram*) est multiplexé sur les connexions entre des nœuds appelés des contextes de bout en bout (EEC, *End-to-End Context*) ce qui permet à chaque QP RD de communiquer avec toute QP RD sur tout nœud avec un EEC établi. Plusieurs QP peuvent utiliser le même EEC et une seule QP peut utiliser plusieurs EEC (un pour chaque nœud distant par domaine de datagrammes fiables).
3. **Connecté fiable (acquitté – en mode connexion)** : le service connecté fiable (RC, *Reliable Connected*) associe une QP locale à une et une seule QP distante. Les tailles de message peuvent être de jusqu'à 2^{31} octets. La mise en œuvre de CA veille à la segmentation et au réassemblage.
4. **Connecté non fiable (non acquitté- en mode connexion)** : le service connecté non fiable (UC, *Unreliable Connected*) associe une QP locale à une et une seule QP distante. Il n'y a pas d'accusé de réception et donc pas de renvoi de paquets perdus ou corrompus. De tels paquets sont donc simplement éliminés. Il est par ailleurs similaire à RC.
5. **Ethertype brut (non acquitté – sans connexion)** : le paquet de datagramme brut Ethertype contient un en-tête générique

de transport qui n'est pas interprété par le CA mais il spécifie le type de protocole. Les valeurs de Ethertype sont les mêmes que définies par l'IANA [IANA] pour Ethertype.

6. IPv6 brut (non acquitté – sans connexion) : en utilisant le service IPv6 de datagrammes bruts, le CA IBA peut prendre en charge des couches de protocole standard par dessus IPv6 (comme TCP/UDP). Donc, les paquets IPv6 natifs peuvent être pontés dans le SAN IBA et livrés directement à un accès et à sa QP de datagrammes bruts IPv6.

Les quatre premiers types sont appelés des transports IB. Les deux derniers sont classés comme datagrammes bruts. Il n'y a pas d'indication du numéro de QP dans les paquets de datagrammes bruts. Les paquets de datagramme brut sont limités en taille par la MTU de liaison.

Les deux modes connectés et le mode datagramme fiable peuvent aussi prendre en charge la migration automatique de chemin (APM, *Automatic Path Migration*). C'est une facilité facultative qui permet une récupération du chemin sur défaillance fondée sur le matériel. Un chemin de remplacement est associé à la QP quand la connexion/contexte EE est créée. En présence d'erreurs irrécupérables, la connexion passe sur le chemin de remplacement.

1.2.1 Adresses InfiniBand

L'architecture InfiniBand emprunte beaucoup à l'architecture IPv6 en termes de structure de sous réseau InfiniBand et de GID.

L'architecture InfiniBand définit le GID associé à un accès comme un identifiant d'envoi individuel ou de diffusion groupée de 128 bits. IBA déduit le format d'adresse de GID, comme défini dans la RFC 2373 [IB_ARCH], avec des propriétés et restrictions supplémentaires définies pour faciliter une découverte, communication, et acheminement, efficaces.

Note : IBA se réfère explicitement à la RFC 2373, qui est obsolète [RFC3513]. On notera que IBA n'est donc pas affecté par les autres changements qui sont introduits dans l'architecture d'adressage IPv6.

IBA définit deux types de GID : envoi individuel et diffusion groupée.

1.2.1.1 GID en envoi individuel

Les GID en envoi individuel sont définis, comme dans IPv6, avec trois portées. La spécification IB déclare :

- a. liaison locale : FE80/10. Les routeurs IB ne vont pas transmettre de paquets avec une adresse de liaison locale dans une source ou destination au delà du sous réseau IB.
- b. site local : FEC0/10. Un GID en envoi individuel utilisé au sein d'une collection de sous réseaux qui est unique au sein de cette collection (par exemple, un centre de données ou un campus) mais n'est pas nécessairement unique au monde. Les routeurs IB ne doivent pas transmettre de paquets avec un GID source de site local ou un GID de destination de site local en dehors du site.
- c. global : un GID en envoi individuel avec un préfixe mondial ; un routeur IB peut utiliser ce GID pour acheminer les paquets dans toute une entreprise ou internet.

1.2.1.2 GID en diffusion groupée

Les GID en diffusion groupée sont aussi semblables aux adresses de diffusion groupée IPv6. La spécification IB définit les GID en diffusion groupée comme suit :

Ffxy : <112 bits>

Bits fanions :

Le demi octet, noté ci-dessus par x, sont les quatre bits de fanion : 000T.

Les trois premiers bits sont réservés et sont réglés à zéro. Le dernier bit est défini comme suit :

T = 0 : note un GID alloué en permanence, c'est-à-dire, un GID bien connu.

T = 1 : note un groupe transitoire.

Bits de portée :

Les 4 bits, notés par y dans le GID ci-dessus, sont les bits de portée.
Ces valeurs de portée sont décrites dans le Tableau 1.

| Valeur de portée | Valeur d'adresse |
|------------------|---------------------|
| 0 | Réservé |
| 1 | Non allouée |
| 2 | Liaison locale |
| 3 | Non allouée |
| 4 | Non allouée |
| 5 | Site-local |
| 6 | Non allouée |
| 7 | Non allouée |
| 8 | Organisation locale |
| 9 | Non allouée |
| 0xA | Non allouée |
| 0xB | Non allouée |
| 0xC | Non allouée |
| 0xD | Non allouée |
| 0xE | Mondiale |
| 0xF | Réservée |

Tableau 1

La spécification IB se réfère de plus aux RFC 2373 et RFC 2375 pour définir les adresses de diffusion groupée bien connues. Cependant, elle déclare alors que les adresses bien connues s'appliquent seulement aux datagrammes IPv6 bruts sur IB. Il doit être noté cependant qu'un groupe de diffusion groupée peut être associé à seulement un identifiant mondial de diffusion groupée (MGID, *Multicast Global Identifier*). Donc le même MGID ne peut pas être associé au mode UD et au mode datagramme brut.

1.3 Gestion de groupe de diffusion groupée InfiniBand

Les groupes de diffusion groupée IB, identifiés par les MGID, sont gérés par le SM. Le SM programme explicitement les commutateurs IB dans le tissu pour s'assurer que les paquets sont reçus par tous les membres du groupe de diffusion groupée qui demandent la réception des paquets. Le SM doit aussi programmer les commutateurs de façon à ce que les paquets transmis au groupe par tout membre du groupe atteignent tous les receveurs dans le groupe de diffusion groupée.

IBA distingue les envoyeurs et les receveurs de diffusion groupée. Bien que tous les membres d'un groupe de diffusion groupée puissent transmettre au groupe (et s'attendre à ce que leurs paquets soient correctement transmis) tous les membres du groupe ne sont pas des receveurs. Un accès doit demander explicitement que les paquets en diffusion groupée adressés au groupe lui soient transmis.

Un groupe de diffusion groupée est créé en envoyant une demande de jonction au SM. Comme on l'explique plus loin, IBA définit plusieurs modes pour se joindre à un groupe de diffusion groupée. Le gestionnaire de sous réseau enregistre le GID de diffusion groupée du groupe et les caractéristiques qui lui sont associées. Les caractéristiques du groupe sont définies par la MTU de chemin du groupe, si le groupe va être utilisé pour des datagrammes bruts ou des datagrammes non fiables, le niveau de service, la clé de partition associée au groupe, l'identifiant local (LID, *Local Identifier*) associé au groupe, et ainsi de suite. Ces caractéristiques sont définies au moment de la création du groupe. Le lecteur intéressé peut chercher l'attribut 'MCMemberRecord' dans la spécification d'architecture IB [IB_ARCH] pour avoir la liste complète des caractéristiques qui définissent un groupe.

Un LID est associé au groupe de diffusion groupée par le SM au moment de la création du groupe de diffusion groupée. Le SM détermine l'arborescence de diffusion groupée sur la base de tous les membres du groupe et programme les commutateurs pertinents. L'identifiant local de diffusion groupée (MLID, *Multicast LID*) est utilisé par les commutateurs pour acheminer les paquets.

Tout accès IB membre qui veut participer au groupe de diffusion groupée doit se joindre au groupe. Au titre de l'opération de jonction, le nœud reçoit les caractéristiques du groupe du SM. En même temps, le gestionnaire de sous réseau s'assure que le demandeur peut bien participer au groupe en vérifiant qu'il peut supporter la MTU de groupe et son accessibilité au reste des membres du groupe. D'autres caractéristiques du groupe peuvent aussi devoir être vérifiées.

Le SM, pour les groupes qui franchissent les limites de sous réseau IB, doit interagir avec les routeurs IB pour déterminer la présence de ce groupe dans d'autres sous réseaux InfiniBand. Si il est présent, la MTU doit correspondre à travers les sous réseaux InfiniBand.

La P_Key est un autre caractéristique qui doit correspondre à travers les sous réseaux InfiniBand car la P_Key insérée dans un paquet n'est pas modifiée par les commutateurs ou routeurs IB. Donc, si les P_Key ne correspondent pas, le ou les routeurs IB eux-mêmes pourraient éliminer les paquets, ou les destinations sur d'autres sous réseaux pourraient éliminer les paquets.

Une opération de jonction peut causer la reprogrammation du tissu par le SM afin que le nouveau membre puisse participer au groupe de diffusion groupée. Par le même jeton, un départ peut causer la reprogrammation du tissu par le SM pour arrêter de transmettre les paquets au demandeur.

1.3.1 Enregistrement de membre de diffusion groupée

La maintenance du groupe de diffusion groupée est faite par le SM avec chaque membre du groupe représenté par un enregistrement MCMemberRecord [IB_ARCH]. Certains de ses composants sont :

MGID : GID de diffusion groupée pour ce groupe de diffusion groupée
 PortGID : GID valide de l'accès qui se joint à ce groupe de diffusion groupée
 Q_Key : Q_Key à utiliser par ce groupe de diffusion groupée
 MLID : LID de diffusion groupée pour ce groupe de diffusion groupée
 MTU : MTU pour ce groupe de diffusion groupée
 P_Key : clé de partition pour ce groupe de diffusion groupée
 SL : niveau de service pour ce groupe de diffusion groupée
 Scope : la même que la portée de l'adresse MGID
 JoinState : état de jonction/départ demandé par l'accès :
 bit 0 : FullMember (*membre à part entière*)
 bit 1 : NonMember (*non membre*)
 bit 2 : SendOnlyNonMember (*en envoi seul pour non membre*)

1.3.1.1 JoinState

L'état de jonction (*JoinState*) indique la qualité d'adhésion qu'un accès souhaite ajouter lors de la jonction/création d'un groupe ou supprimer lors du départ d'un groupe. La signification des bits de JoinState est la suivante :

FullMember : les messages destinés au groupe sont acheminés du et vers l'accès. Un groupe peut être supprimé par le SM si il n'y a plus de FullMember dans le groupe.

NonMember : les messages destinés au groupe sont acheminés du et vers l'accès. L'accès n'est pas considéré comme membre pour les besoins de la création/suppression du groupe.

SendOnlyNonMember : les messages du groupe sont seulement acheminés à partir de l'accès mais pas vers l'accès. L'accès n'est pas considéré comme membre pour les besoins de la création/suppression du groupe.

Un accès peut avoir plusieurs bits établis dans son enregistrement. Dans ce cas, les qualités de membre sont une union des JoinState. Un accès peut quitter le groupe de diffusion groupée pour chacun des JoinState individuels ou dans toute combinaison des bits de JoinState [IB_ARCH].

1.3.2 Opérations de jonction et de départ

Un accès IB se joint à un groupe de diffusion groupée en envoyant une demande d'adhésion (méthode SubnAdmSet()) et quitte un groupe de diffusion groupée en envoyant un message de départ (méthode SubnAdmDelete()) au SM. La spécification IBA [IB_ARCH] décrit les méthodes et attributs à utiliser dans l'envoi de ces messages.

1.3.2.1 Création d'un groupe de diffusion groupée

Il n'y a pas de commande 'create' pour former un nouveau groupe de diffusion groupée. Le bit FullMember dans le JoinState doit être établi pour créer un groupe de diffusion groupée. En d'autres termes, la première demande d'adhésion

comme membre à part entière (*FullMember*) va causer la création du groupe comme effet collatéral de la demande de jonction. Les demandes suivantes de jonction ou de départ peuvent contenir toutes les combinaisons de bits JoinState.

Le créateur du groupe spécifie les valeurs de Q_Key, MTU, P_Key, SL, FlowLabel, TClass, et Scope. Un créateur peut demander qu'un MGID convenable soit créé pour lui. Autrement, la demande peut spécifier le MGID désiré. Dans les deux cas, le MLID est alloué par le SM.

Donc, un groupe va être créé avec les valeurs spécifiées quand le demandeur établit le bit FullMember et qu'aucun tel groupe n'existe déjà dans le sous réseau.

1.3.2.2 Suppression d'un groupe de diffusion groupée

Quand le dernier membre à part entière quitte le groupe de diffusion groupée, le SM peut supprimer le groupe de diffusion groupée en libérant toutes les ressources, y compris celles qui pourraient exister dans le tissu lui-même, associées au groupe.

Noter qu'il n'existe pas de message "supprimer" particulier. C'est un effet collatéral de la dernière opération "leave" de FullMember.

1.3.2.3 Codes de création/suppression de groupe de diffusion groupée

Les accès peuvent demander au SA de générer un rapport chaque fois qu'un groupe de diffusion groupée est créé ou supprimé. L'accès peut spécifier le ou les groupes de diffusion groupée auxquels il est intéressé en utilisant son MGID ou en soumettant une demande générique. Le SA va rapporter ces événements en utilisant les codes 66 (pour les créations) et 67 (pour les suppressions) [IB_ARCH].

Donc, un accès qui souhaite se joindre à un groupe mais non le créer par lui-même peut demander une notification de création ou un accès pourrait même demander une notification pour tous les groupes qui sont créés (une demande générique). Le SA va l'informer de la création en utilisant les codes sus-mentionnés. Le demandeur peut alors se joindre au groupe de diffusion groupée indiqué. De même, un SendOnlyNonMember ou un NonMember pourrait demander que le SA l'informe des suppressions de groupe. Le nœud d'extrémité, à réception d'un rapport de suppression, peut en toute sécurité libérer les ressources associées au groupe. Le MLID associé n'est plus valide pour le groupe et peut être réalloué à un nouveau groupe de diffusion groupée par le SM.

2. Gestion de sous réseau InfiniBand

Pour aider à la surveillance et la configuration des composants de sous réseau InfiniBand, un ensemble de modules de MIB devra être défini. Les modules de MIB sont nécessaires pour les adaptateurs de canal, les interfaces InfiniBand, le gestionnaire de sous réseau InfiniBand, et les agents de gestion de sous réseau InfiniBand et pour permettre la gestion des propriétés d'appareil spécifiques. On notera que les objets de gestion visés dans les documents IPoIB sont pour tous les composants de sous réseau IB et ne sont pas limités à IP (sur IB). Les modules de MIB pertinents sont décrits dans des documents séparés et ne sont pas traités ici.

3. IP sur IB

Comme décrit à la Section 1, l'architecture InfiniBand fournit un large ensemble de capacités de choix quand on met en œuvre IP sur des réseaux InfiniBand.

La spécification IPoIB ne doit pas exiger, et n'exige pas, de changements à IP et aux protocoles de couches supérieures. Elle n'oblige pas non plus que les piles IP mettent en œuvre des programmes particuliers de niveau utilisateur. Un des buts de la spécification IPoIB est que les changements à IPoIB soient accessibles à la modularisation et l'incorporation dans les mises en œuvre existantes au même niveau que les autres types de supports.

3.1 InfiniBand comme liaison de données

L'architecture InfiniBand donne plusieurs méthodes d'échange de données entre deux points d'extrémité comme on l'a noté précédemment. Ce sont les suivantes :

Connecté fiable (RC, *Reliable Connected*)
Datagramme fiable (RD, *Reliable Datagram*)
Connecté non fiable (UC, *Unreliable Connected*)
Datagramme non fiable (UD, *Unreliable Datagram*)
Datagramme brut : IPv6 brut (R6)
 : Ethertype brut (RE)

IPoIB peut être mis en œuvre sur un, plusieurs, ou tous ces services. Un cas particulier peut être fait pour la prise en charge d'une des méthodes de transport selon les caractéristiques désirées.

La spécification IB exige que le mode Datagramme non fiable soit pris en charge par tous les nœuds IB. Il est spécifiquement exigé des adaptateurs de canal d'hôte (HCA, *host channel adapter*) qu'ils prennent en charge les modes Connecté fiable (RC) et Connecté non fiable (UC) mais ce n'est pas le cas des adaptateurs de canal cible (TCA, *target channel adapter*). La prise en charge des deux modes de datagramme brut est entièrement facultative. Le mode Datagramme brut prend en charge une vérification de redondance cyclique (CRC, *Cyclic Redundancy Check*) de 16 bits à comparer à la meilleure protection fournie par l'utilisation d'un CRC de 32 bits dans les autres modes.

Par souci de simplicité, de facilité de mise en œuvre et d'intégration avec les piles de protocoles existantes, il est souhaitable que le tissu supporte la diffusion groupée. Ceci n'est possible que dans les modes Datagramme non fiable (UD) et Datagramme brut de IB.

Donc, c'est seulement le mode UD qui est universel, prend en charge la diffusion groupée et un CRC robuste. Dans ces conditions, c'est le choix évident pour IP sur InfiniBand [RFC4391].

De futurs documents pourraient considérer les modes connectés. À la différence de la MTU de liaison limitée offerte par le mode UD, les modes connectés peuvent offrir des avantages significatifs en termes de performances en utilisant une plus grande MTU. La fiabilité est aussi améliorée si les caractéristiques sous-jacentes de la migration automatique de chemin des modes connectés sont utilisées.

3.2 Prise en charge de la diffusion groupée

La spécification InfiniBand rend la prise en charge de la diffusion groupée facultative dans les commutateurs. La diffusion groupée est cependant une exigence de base dans les réseaux IP. Donc, IPoIB exige que des tissus InfiniBand capables de diffusion groupée soient utilisés pour mettre en œuvre les sous réseaux IPoIB.

3.2.1 Transposition de diffusion groupée IP en diffusion groupée IB

Des groupes de diffusion groupée IP bien connus sont définis pour IPv4 et IPv6 [IANA], [RFC3513]. Des groupes de diffusion groupée peuvent aussi être créés dynamiquement à tout moment. Pour éviter de créer des duplications inutiles de paquets en diffusion groupée dans le tissu, et pour éviter un traitement inutile de tels paquets chez les hôtes, chaque groupe de diffusion groupée IP doit être associé à un groupe de diffusion groupée IB différent autant que possible. Un processus est défini dans la [RFC4391] pour transposer les adresses de diffusion groupée IP en adresses uniques de diffusion groupée IB.

3.2.2 Fanion transitoire dans les MGID IB

La spécification IB décrit les bits de fanion comme expliqué au paragraphe 1.2. La spécification IB définit aussi des MGID IB bien connus. Les MGID sont réservés au mode Datagramme brut de IB qui est incompatible avec les autres transports de IB. Toute transposition qui est définie à partir d'adresses de diffusion groupée IP ne doit donc pas entrer dans la définition de IB d'une adresse bien connue.

Donc tous les GID de diffusion groupée relatifs à IPoIB établissent toujours le bit transitoire.

3.3 Sous réseaux IP à travers des sous réseaux IB

Certaines mises en œuvre peuvent souhaiter prendre en charge plusieurs grappes de machines dans leurs propres sous réseaux InfiniBand mais par ailleurs faire partie d'un sous réseau IP commun. Pour une telle solution, la spécification IB a besoin de plusieurs mises à niveau. Certaines des améliorations requises sont comme suit :

- 1) Une méthode de création de GID de diffusion groupée IB qui s'étende sur plusieurs sous réseaux InfiniBand. Les clés de partition et autres paramètres doivent être cohérents à travers les sous réseaux InfiniBand.
- 2) Développer un protocole d'acheminement IB pour déterminer la topologie IB à travers les sous réseaux InfiniBand.
- 3 Définir le traitement et les protocoles nécessaires entre les nœuds IB et les routeurs IB.

Tant que les conditions ci-dessus ne sont pas satisfaites, il n'est pas possible de mettre en œuvre des sous réseaux IPoIB qui s'étendent sur des sous réseaux InfiniBand. Les normes IPoIB ont cependant été définis en vue de cette possibilité.

4. Sous réseaux IP dans un tissu InfiniBand

Le sous réseau IPoIB est superposé au sous réseau IB. Le sous réseau IPoIB est établi selon les étapes suivantes :

Note : dans ce document, l'opération jonction/départ au niveau IP est appelé IP_join/IP_leave et les opérations jonction/départ au niveau IB sont appelées IB_join.

1. Le groupe de diffusion groupée IB Tous les nœuds IPoIB est créé.

L'administrateur du tissu crée un groupe de diffusion groupée IB (appelé à partir d'ici "groupe de diffusion") quand le sous réseau IP est établi. Le "groupe de diffusion" est défini dans la [RFC4391]. La méthode d'établissement du groupe de diffusion n'est pas définie par IPoIB. Le groupe peut être établi au SM par l'administrateur ou par le premier IB_join.

Comme noté précédemment, au moment de la création d'un groupe de diffusion groupée IB, plusieurs valeurs comme P_Key, Q_Key, SL, Limite de bonds, Identifiant de flux, TClass, MTU, etc. ont dues être spécifiées. Ces valeurs devraient être telles que tous les membres potentiels du groupe de diffusion groupée IB soient capables de communiquer les uns avec les autres quand ils les utilisent. A l'avenir, quand la spécification IB associera plus de signification aux divers paramètres et définira une qualité de service IB, différentes valeurs seront possibles pour le trafic de diffusion groupée IP. Tous les paquets en envoi individuel ont aussi besoin d'utiliser la P_Key et la Q_Key spécifiée dans le groupe de diffusion [RFC4391]. Il est évident qu'une configuration bien réfléchiée est requise pour un établissement réussi du sous réseau IPoIB.

2. Toutes les interfaces IPoIB se joignent au groupe de diffusion IB.

Le groupe de diffusion définit la portée et les membres de la liaison IPoIB. Cette liaison est établie lorsque les nœuds IPoIB se joignent au groupe de diffusion IB.

La jonction IB au groupe de diffusion a l'avantage supplémentaire de distribuer les paramètres sus-mentionnés de groupe de diffusion groupée à tous les membres du sous réseau.

Noter que ce IB_join au groupe de diffusion est une jonction de membre à part entière (*FullMember*). Si un des accès ou des commutateurs reliant l'accès au reste du sous réseau IPoIB ne peut pas supporter les paramètres (par exemple, la MTU de chemin ou la P_Key) associés au groupe de diffusion, la demande IB_join va échouer et l'accès demandeur ne va pas faire partie du sous réseau IPoIB.

3. Paramètres de configuration.

Comme noté ci-dessus, les paramètres tels que Q_Key et MTU de chemin, qui sont nécessaires pour toutes les communications IPoIB, sont retournés au nœud IPoIB lors de la jonction IB au "groupe de diffusion". La [RFC4391] note aussi que les paramètres utilisés dans le groupe de diffusion sont utilisés lors de la création d'autres groupes de diffusion groupée.

Cependant, la P_Key doit quand même être connue du nœud d'extrémité IPoIB avant qu'il puisse se joindre au groupe de diffusion. La P_Key est incluse dans la transposition du groupe de diffusion [RFC4391]. Un autre paramètre, la portée du groupe de diffusion, doit aussi être connu du nœud d'extrémité avant qu'il puisse se joindre au groupe de diffusion. Il relève d'une mise en œuvre de choisir comment la P_Key et les bits de portée relatifs au sous réseau IPoIB sont déterminés par la mise en œuvre. Ce pourrait être par des paramètres de configuration initialisés par l'administrateur.

Les méthodes employées par une mise en œuvre pour déterminer la P_Key et les bits de portée ne sont pas spécifiés par IPoIB.

4.1 VLAN IPoIB

Les points d'extrémité dans un sous réseau IB doivent avoir des P_Key compatibles pour communiquer l'un avec l'autre. Donc, quand l'administrateur établit un sous réseau IP sur un sous réseau IB, il doit s'assurer que tous les membres ont des P_Key compatibles. Un sous réseau IP ne peut avoir qu'une seule P_Key associée pour assurer que tous les nœuds IP qu'il contient peuvent parler les uns avec les autres. Un point d'extrémité peut cependant avoir plusieurs P_Key.

L'architecture IB spécifie qu'il ne peut y avoir qu'un seul MGID associé à un groupe de diffusion groupée dans le sous réseau IB. La P_Key est incluse dans les transposition de MGID à partir des adresses de diffusion groupée IP [RFC4391]. Comme la P_Key est unique dans le sous réseau IB, l'inclusion de la P_Key dans les MGID IB assure que des transpositions uniques de MGID sont créées. Chaque MGID unique de groupe de diffusion ainsi formé crée une liaison IPoIB abstraite séparée et donc un VLAN IPoIB.

4.2 Diffusion groupée dans des sous réseaux IPoIB

La diffusion groupée IP sur des sous réseaux InfiniBand suit les mêmes concepts et règles que sur tous les autres supports. Cependant, à la différence de la plupart de autres supports, la diffusion groupée sur InfiniBand exige l'interaction avec une autre entité, le gestionnaire de sous réseau IB. Ce paragraphe décrit les grandes lignes du processus et suggère des lignes directrices.

L'architecture IB spécifie le format suivant pour les paquets de diffusion groupée IB quand ils sont utilisés sur le mode Datagramme non fiable (UD) :

```
+-----+-----+-----+-----+-----+-----+-----+
|En-tête|En-tête| En-tête | En-tête |Charge utile|  CRC   |  CRC   |
|d'achem|d'achem|transport|datagram.| du paquet |invariant|variant |
| local |mondial| de base |transport|          |        |        |
|      |      |      | étendu | (IP)    |        |        |
+-----+-----+-----+-----+-----+-----+-----+
```

Pour les détails des divers en-têtes, se reporter à la spécification de l'architecture InfiniBand [IB_ARCH].

L'en-tête d'acheminement mondial (GRH, *Global Routing Header*) inclut le GID de groupe de diffusion groupée IB. L'en-tête d'acheminement local (LRH, *Local Routing Header*) inclut l'identifiant local (LID, *Local Identifier*). Les commutateurs IB dans le tissu acheminement le paquet sur la base du LID.

Le GID est rendu disponible à l'utilisateur IB receveur (le pilote d'interface IPoIB par exemple).Le pilote peut donc déterminer le groupe IB auquel appartient le paquet.

IPv4 définit trois niveaux de conformité de diffusion groupée [RFC1112].

- niveau 0 : pas de prise en charge de la diffusion groupée IP.
- niveau 1 : prise en charge de l'envoi mais pas de la réception de diffusion groupée.
- niveau 2 : plein support de la réception de diffusion groupée.

Dans IPv6, il n'y a pas de telle distinction. Le plein support de la diffusion groupée est obligatoire. De plus, tous les sous réseaux IPv4 prennent en charge la diffusion (255.255.255.255). La diffusion IPv4 peut toujours être envoyée/reçue par toutes les interfaces IPv4.

Chaque sous réseau IPoIB exige que le GID de diffusion soit défini. Donc, un paquet peut toujours être en diffusion.

4.2.1 Envoi de datagrammes en diffusion groupée IP

Un hôte IP peut envoyer un paquet en diffusion groupée à tout moment à toute adresse de diffusion groupée.

La couche IP convoie le paquet en diffusion groupée au pilote/module de l'interface IPoIB. Ce module tente de se joindre (*IB_join*) au groupe de diffusion groupée IB pertinent. Ceci est nécessaire car autrement, l'architecture InfiniBand ne

garantit pas que le paquet atteigne ses destinations.

Un pur expéditeur peut choisir de se joindre au groupe de diffusion groupée comme membre à part entière (*FullMember*). Dans ce cas, l'expéditeur va recevoir tous les paquets en diffusion groupée transmis au groupe IB. De plus, le groupe IB ne sera pas supprimé tant que l'expéditeur restera dans le groupe.

Autrement, un expéditeur pourrait se joindre à IB comme non membre en envoi seul (*SendOnlyNonMember*). Dans ce cas, les paquets ne sont pas acheminés à l'expéditeur bien que les paquets qu'il transmet puissent atteindre les autres membres du groupe. De plus, le groupe peut être supprimé quand tous les membres à part entière ont quitté le groupe. L'expéditeur peut de plus demander les mises à jour de suppression au SM.

Si l'expéditeur ne trouve pas le groupe existant, il est recommandé dans la [RFC4391] que les paquets soient envoyés au MGID correspondant à l'adresse Tous les routeurs IP. Un expéditeur pourrait aussi envoyer les paquets au groupe de diffusion. L'expéditeur pourrait aussi choisir de demander des rapports de "création" au SM.

4.2.2 Réception de paquets en diffusion groupée

L'hôte IP doit se joindre au groupe de diffusion groupée IB correspondant à l'adresse IP. Cela découle de l'exigence IBA que le receveur doit se joindre au groupe de diffusion groupée IB pertinent. Le groupe est automatiquement créé si il n'existe pas [IB_ARCH].

Les receveurs IP doivent quitter (*IB_leave*) le groupe IB quand la couche IP arrête d'écouter sur l'adresse IP correspondante. Le SM peut alors choisir de supprimer le groupe.

4.2.3 Considérations de routeurs pour IPoIB

Les routeurs IP connaissent les nouveaux groupes IP créés dans le sous réseau par l'utilisation de protocoles comme le protocole de gestion de groupe Internet (IGMP, *Internet Group Management Protocol*) / découverte d'écoute de diffusion groupée (MLD, *Multicast Listener Discovery*) [RFC3376], [RFC2710]. Cependant, ce n'est pas assez pour IPoIB car le routeur a besoin de se joindre aux groupes IB pertinents pour être capable de recevoir et transmettre les paquets. Il n'y a pas de mode de promiscuité pour l'écoute de tous les paquets.

Les routeurs IPoIB ont donc besoin de demander au SM de faire rapport de toutes les créations de groupes IB dans le tissu. Le routeur IPoIB peut alors se joindre au groupe rapporté. Il n'est pas souhaitable que la jonction IB d'un routeur à un groupe de diffusion groupée soit considérée de la même façon que la jonction IB d'un receveur – la jonction IB du routeur ne devrait pas interdire la dissolution du groupe quand tous les receveurs l'ont quitté. Pour surmonter ce type de situation, IBA a créé le mode de jonction IB non membre (*NonMember IB_join*).

Le mode de jonction IB non membre peut être utilisé par les routeurs IP quand ils se joignent en réponse à des rapports de création. Un routeur devrait idéalement demander aussi les rapports de suppression afin qu'il puisse libérer toutes les ressources associées au groupe. Le MLID associé à un MGID supprimé peut être réalloué par le SM, et donc il y a une possibilité de transmissions erronées si le MLID est mis en antémémoire. Un routeur qui ne demande pas les rapports de suppression va quand même fonctionner correctement car il va recevoir le MLID correct, et purger toutes les vieilles valeurs en antémémoire, quand il se joint au groupe IB en réponse à un rapport de création.

Il est raisonnable pour un routeur de se joindre comme membre à part entière si il se joint au groupe IB en réponse à la demande d'un automate d'application/acheminement. Dans ce cas, le routeur pourrait finir par contrôler l'existence du groupe IB (car il est un membre à part entière du groupe).

4.2.4 Impact des limites de l'architecture InfiniBand

Un HCA ou TCA peut avoir une limite sur le nombre de MGID qu'il peut prendre en charge. Donc, même si les groupes ne peuvent pas être limités chez le gestionnaire de sous réseau et dans le sous réseau en tant que tel, ils peuvent être limités à une interface particulière. Il est conseillé de choisir un HCA/TCA provisionné de façon adéquate lors de l'établissement d'un sous réseau IPoIB.

4.2.5 Départ/suppression d'un groupe de diffusion groupée

Un expéditeur IPv4 (niveau 1 de conformité) ne se joint au groupe de diffusion groupée IB que parce que c'est la seule façon de garantir la réception des paquets par tous les récepteurs du groupe. L'expéditeur doit cependant quitter (*IB_leave*) le groupe à un certain moment. Un expéditeur pourrait, quand il n'est pas un récepteur du groupe, lancer un temporisateur par groupe de diffusion groupée auquel il envoie. L'expéditeur quitte le groupe IB quand le temporisateur arrive à expiration. Il relance le temporisateur si un autre message est envoyé.

Cette suggestion ne s'applique pas au groupe de diffusion IB. Elle ne s'applique pas non plus au groupe IB correspondant au groupe de diffusion groupée Tous-les-hôtes. Un hôte IPv4 doit toujours rester membre du groupe de diffusion.

Un récepteur de diffusion groupée IP quitte (*IB_leave*) le groupe de diffusion groupée IB correspondant quand il quitte (*IP_leave*) le groupe de diffusion groupée IP. Dans le cas d'une mise en œuvre de IPv4, le récepteur peut choisir de continuer d'être un expéditeur (niveau 1 de conformité) auquel cas il peut choisir de ne pas quitter le groupe IB mais de lancer un temporisateur comme expliqué ci-dessus.

On a noté ailleurs que le SM peut choisir de libérer les ressources (par exemple, les entrées d'acheminement dans les commutateurs) associées au groupe IB quand le dernier membre à part entière quitte le groupe. Le MLID devient donc invalide pour le groupe. Le MLID peut être réalloué quand un nouveau groupe est créé.

La mise en antémémoire du MLID des accès *SendOnlyNonMember/NonMember* doit éviter cette possibilité. La façon de le faire est qu'il demandent des rapports de suppression de groupe. Un routeur IP qui demande des rapports pour tous les groupes n'a pas besoin de demander le rapport de suppression car un *IB_join* en réponse à un rapport de création va retourner le nouveau MLID qui lui est associé.

Un routeur pourrait préférer quitter le groupe de diffusion groupée IB quand il n'y a plus de membres à l'adresse IP de diffusion groupée dans le sous-réseau et qu'il n'a pas de connaissance explicite d'un besoin de transmission de tels paquets.

4.3 Transmission des paquets IPoIB

L'encapsulation des paquets IP dans InfiniBand est décrite dans la [RFC4391].

Elle spécifie l'utilisation d'une valeur de "Ethertype" [IANA] dans tous les paquets de communication IPoIB. L'adresse de couche de liaison est composée du GID et du numéro de paire de file d'attente (QPN, *Queue Pair Number*) [RFC4391].

Pour permettre aux sous-réseaux IPoIB de s'étendre sur plusieurs sous-réseaux IB, la spécification utilise le GID au titre de l'adresse de couche de liaison. Comme tous les paquets dans IB doivent utiliser l'identifiant local (LID, *Local Identifier*) le processus de résolution d'adresse a une étape supplémentaire pour résoudre le GID de destination, retourné en réponse à la demande de découverte de voisin (ND, *Neighbor Discover*) du protocole de résolution d'adresse (ARP, *Address Resolution Protocol*) sur le LID [RFC4391]. Cette phase de la résolution d'adresse pourrait aussi être utilisée pour déterminer d'autres paramètres essentiels (par exemple, le SL, le débit de chemin, etc.) pour une communication IB réussie entre deux homologues.

Comme noté plus tôt, toutes les communications dans le sous-réseau IPoIB déduisent la *Q_Key* à utiliser de la *Q_Key* spécifiée dans le groupe de diffusion.

4.4 Protocole de résolution d'adresse inverse (RARP) et entrées ARP statiques

Les entrées de RARP ou les entrées de ARP statique se fondent sur des adresses de liaison invariantes. Dans le cas de IPoIB, l'adresse de liaison inclut le QPN, qui pourrait n'être pas constant à travers les réamorçages ou même les réinitialisations d'interface réseau. Donc, les entrées d'ARP statiques ou les entrées de serveur RARP ne vont fonctionner que si la ou les mises en œuvre qui utilisent ces options peuvent assurer que le QPN associé à une interface est invariant à travers les réamorçages/réinitialisation de réseau [RFC4391].

4.5 DHCPv4 et IPoIB

DHCPv4 [RFC2131] utilise un champ "identifiant de client" (supposé contenir l'adresse de couche de liaison) de 16 octets. L'adresse dans le cas de IPoIB est de 20 octets. Pour contourner ce problème, IPoIB spécifie [RFC4390] le fanion de diffusion à utiliser par le client quand il demande une adresse IP.

5. QS et questions en rapport

La spécification IB suggère l'utilisation de niveaux de service pour l'équilibrage de charge, la QS, et l'évitement d'impasses avec un sous réseau IB. Mais la spécification IB laisse l'usage et le mode de détermination du SL à la décision de l'application. Le SL et la liste des SL sont disponibles dans le SA, mais il appartient à l'application du nœud d'extrémité de choisir la "bonne" valeur.

Chaque mise en œuvre IPoIB va déterminer la valeur de SL pertinente sur la base de sa propre politique. Aucune méthode ou processus de choix du SL n'a été défini par les normes IPoIB.

6. Considérations sur la sécurité

Le présent document décrit l'architecture IB en tant que pertinente pour IPoIB. Il réaffirme des questions spécifiées dans d'autres documents. Il ne spécifie par lui-même aucune exigence. Aucun problème de sécurité n'est introduit par ce document. Les questions de sécurité relatives à IPoIB sont décrites dans les [RFC4390] et [RFC4391].

7. Remerciements

Le présent document a bénéficié des commentaires et suggestions des membres du groupe de travail IPoIB et des membres de l'association commerciale InfiniBand(SM).

8. Références

8.1 Références normatives

[IB_ARCH] "InfiniBand Architecture Specification, version 1.2". www.infinibandta.org

[RFC2131] R. Droms, "Protocole de [configuration dynamique d'hôte](#)", mars 1997. (DS) (Mà J par RFC3396, RFC4361, RFC5494, et RFC6849)

[RFC2373] R. Hinden, S. Deering, "Architecture d'adressage IP version 6", juillet 1998. (Obsolète, voir [RFC4291](#)) (PS)

[RFC4390] V. Kashyap, "[Protocole de configuration dynamique d'hôte](#) (DHCP) sur InfiniBand", avril 2006. (P.S.)

[RFC4391] J. Chu, V. Kashyap, "[Transmission de IP sur InfiniBand](#) (IPoIB)", avril 2006. (P.S.)

8.2 Références pour information

[IANA] Internet Assigned Numbers Authority, URL <http://www.iana.org>

[RFC1112] S. Deering, "Extensions d'hôte pour [diffusion groupée sur IP](#)", STD 5, août 1989. (Mise à jour par la RFC2236)

[RFC2375] R. Hinden, S. Deering, "[Allocation des adresses de diffusion groupée IPv6](#)", juillet 1998. (Information)

[RFC2710] S. Deering, W. Fenner et B. Haberman, "[Découverte d'écouteur de diffusion groupée](#) (MLD) pour IPv6", octobre 1999.

[RFC3376] B. Cain et autres, "[Protocole Internet de gestion de groupe](#), IGMP version 3", octobre 2002. (P.S.)

[RFC3513] R. Hinden et S. Deering, "[Architecture d'adressage du protocole Internet](#) version 6 (IPv6)", avril 2003. (Obs. voir [RFC4291](#))

Adresse de l'auteur

Vivek Kashyap
15350, SW Koll Parkway
Beaverton
OR 97006
USA
téléphone : +1 503 578 3422
mél : vivk@us.ibm.com

Déclaration complète de droits de reproduction

Copyright (C) The Internet Society (2006)

Le présent document est soumis aux droits, licences et restrictions contenus dans le BCP 78, et sauf pour ce qui est mentionné ci-après, les auteurs conservent tous leurs droits.

Le présent document et les informations contenues sont fournies sur une base "EN L'ÉTAT" et le contributeur, l'organisation qu'il ou elle représente ou qui le/la finance (s'il en est), la INTERNET SOCIETY, le IETF TRUST et la INTERNET ENGINEERING TASK FORCE déclinent toutes garanties, exprimées ou implicites, y compris mais non limitées à toute garantie que l'utilisation des informations encloses ne viole aucun droit ou aucune garantie implicite de commercialisation ou d'aptitude à un objet particulier.

Propriété intellectuelle

L'IETF ne prend pas position sur la validité et la portée de tout droit de propriété intellectuelle ou autres droits qui pourraient être revendiqués au titre de la mise en œuvre ou l'utilisation de la technologie décrite dans le présent document ou sur la mesure dans laquelle toute licence sur de tels droits pourrait être ou n'être pas disponible ; pas plus qu'elle ne prétend avoir accompli aucun effort pour identifier de tels droits. Les informations sur les procédures de l'ISOC au sujet des droits dans les documents de l'ISOC figurent dans les BCP 78 et BCP 79.

Des copies des dépôts d'IPR faites au secrétariat de l'IETF et toutes assurances de disponibilité de licences, ou le résultat de tentatives faites pour obtenir une licence ou permission générale d'utilisation de tels droits de propriété par ceux qui mettent en œuvre ou utilisent la présente spécification peuvent être obtenues sur le répertoire en ligne des IPR de l'IETF à <http://www.ietf.org/ipr>.

L'IETF invite toute partie intéressée à porter son attention sur tous copyrights, licences ou applications de licence, ou autres droits de propriété qui pourraient couvrir les technologies qui peuvent être nécessaires pour mettre en œuvre la présente norme. Prière d'adresser les informations à l'IETF à ietf-ipr@ietf.org.

Remerciement

Le financement de la fonction d'édition des RFC est fourni par l'activité de soutien administratif de l'IETF (IASA).