Groupe de travail Réseau
Request for Comments : 3347
Catégorie : Information

Traduction Claude Brière de L'Isle

M. Krueger & R. Haagens
Hewlett-Packard Corporation
C. Sapuntzakis, Stanford
M. Bakke, Cisco Systems
juillet 2002

Protocole d'interface de systèmes de petits ordinateurs sur l'Internet (iSCSI) – exigences et considérations sur leur conception

Statut de ce mémoire
Le présent mémoire fournit des informations pour la communauté de l'Internet. Il ne spécifié aucune norme de l'Internet. La distribution de ce mémoire n'est soumise à aucune restriction.

Résumé
Le présent document spécifie les exigences auxquelles iSCSI et l'infrastructure qui s'y rapporte devraient satisfaire ainsi que les considérations de conception qui guident les efforts de développement du protocole iSCSI. Dans l'intérêt de l'accélération de l'adoption du protocole iSCSI, le groupe IPS a choisi de concentrer la première version du protocole sur l'architecture et les commandes SCSI existantes, et sur la couche transport TCP/IP existante. Ces deux protocoles sont largement déployés et bien compris. L'idée est que l'utilisation de ces protocoles arrivés à maturité va entraîner un minimum d'inventions, l'adoption la plus rapide possible, et la plus grande compatibilité avec l'architecture, les protocoles, et l'équipement de l'Internet.

Conventions utilisées dans ce document
Le présent document décrit les exigences d'un concept de protocole, mais ne définit pas une norme de protocole. Néanmoins, les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAIT PAS", "RECOMMANDE", "PEUT", et "FACULTATIF" dans ce document sont à interpréter comme décrit dans la RFC2119 [2].

Table des Matières

1. Introduction

Le groupe de travail IP Storage *(mémorisation IP)* est chargé de développer une technologie complète pour le transport de blocs de données de mémorisation sur les protocoles IP. Cet effort inclut un protocole pour transporter le protocole d'interface de système de petit ordinateur (SCSI, *Small Computer Systems Interface*) sur l'Internet (iSCSI). La version initiale du protocole iSCSI va définir une transposition du protocole de transport SCSI sur TCP/IP afin que le contrôleur de mémorisation SCSI (principalement les dispositifs de disque et bandes et les bibliothèques) puisse être rattaché aux réseaux IP, notamment Gigabit Ethernet (GbE) et 10 Gigabit Ethernet (10 GbE).

Le protocole iSCSI est une transposition de SCSI à TCP, et constitue un "transport SCSI" comme défini par le groupe T10 de l'ANSI, à la page 3 du document SCSI SAM-2 [3], "Transport Protocols".

2. Résumé des exigences

La norme iSCSI :
À partir du paragraphe 3.2 "Performance/coût" :
DOIT permettre que les mises en œuvre égalent ou améliorent l'état de l'art actuel de l'interconnexion SCSI.
DOIT permettre que les mises en œuvre soient d'un prix compétitif.
DEVRAIT minimiser la redondance du contrôle pour permettre des communications à faible délai.
DOIT fournir une forte bande passante et l'agrégation de bande passante.
DOIT avoir de faibles utilisations de CPU d'hôte, égales ou meilleures que la technologie actuelle.
DOIT rendre possible de construire des adaptateurs entrée/sortie qui traitent la totalité de la tâche de SCSI.
DEVRAIT permettre des architectures de placement direct des données.
NE DOIT PAS imposer des opérations complexes au logiciel d'hôte.
DOIT assurer une pleine utilisation de la bande passante disponible de la liaison.
DOIT permettre qu'une mise en œuvre exploite le parallélisme (connexions multiples) aux interfaces d'appareil et au sein de la matrice d'interconnexion.

À partir du paragraphe 3.4 "Large bande passante/Agrégation de bande passante" :
DOIT fonctionner sur une seule connexion TCP.
DEVRAIT prendre en charge le lien de connexion, et elle DOIT être de mise en œuvre facultative.
À partir de la section 4 "Facilité de mise en œuvre/complexité du protocole" :
DEVRAIT garder le protocole simple.
DEVRAIT minimiser les caractéristiques facultatives.
DOIT spécifier la négociation de caractéristiques à l'établissement de session *(login)*.
DOIT fonctionner correctement lorsque aucune caractéristique facultative n'est négociée, ainsi que lorsque la négociation d'option individuelle ne réussit pas.

À partir du paragraphe 5.1 "Détection de la corruption des données" :
DOIT prendre en charge le format de vérification d'intégrité des données à utiliser dans la génération de résumés.
PEUT utiliser des résumés séparés pour les données et les en-têtes.
Le format d'en-tête iSCSI DEVRAIT être extensible pour inclure d'autres méthodes de calcul de résumé d'intégrité des données.

À partir du paragraphe 5.2 "Récupération" :
DOIT spécifier des mécanismes pour récupérer à temps des défaillances chez l'initiateur, la cible, ou l'infrastructure de connexion.
DOIT spécifier des méthodes de récupération pour les demandes non idempotentes.
DEVRAIT prendre en compte les schémas de récupération pour les cibles miroir ou les configurations de mémorisation à haute disponibilité.
DEVRAIT fournir une méthode pour que les sessions se terminent et redémarrent en douceur, qui puisse être initiée par l'initiateur ou la cible.

À partir de la section 6 "Interopérabilité" :

Le document du protocole iSCSI DOIT être clair et sans ambiguïté.

À partir du paragraphe 6.1 "Infrastructure Internet" :
DOIT :
-- être compatible avec IPv4 et IPv6,
-- utiliser avec prudence les connexions TCP, en se souvenant qu'il peut y avoir d'autres utilisateurs de TCP sur une même machine.
NE DOIT PAS exiger de changement aux protocoles Internet existants.
DEVRAIT minimiser les changements exigés des mises en œuvre TCP/IP existantes.
DOIT être conçu de façon à permettre une future substitution de SCTP (à TCP) comme protocole transport IP avec un minimum de changements au fonctionnement du protocole iSCSI, à la structure et aux formats d'unités de données de protocole (PDU).

À partir du paragraphe 6.2 "SCSI" :
Toute caractéristique SAM2 exigée dans une transposition de transport valide DOIT être spécifiée par iSCSI.
DOIT spécifier une livraison strictement ordonnée des commandes SCSI sur une session iSCSI entre une paire initiateur/cible.
Le mécanisme de rangement des commandes DEVRAIT chercher à minimiser la quantité de communication nécessaire à travers les divers adaptateurs qui font le transport des charges.
DOIT spécifier pour chaque caractéristique si elle est FACULTATIVE, RECOMMANDÉE ou EXIGÉE à la mise en œuvre et/ou l'utilisation.
NE DOIT PAS exiger de changements aux ensembles de commandes SCSI-3 et au code client SCSI sauf lorsque les spécifications SCSI pointent sur les champs et comportements qui "dépendent du transport".
DEVRAIT relever les changements à SCSI et au modèle d'architecture SCSI.
DOIT être capable de prendre en charge tous les ensembles de commandes SCSI-3 et de types d'appareils.
DEVRAIT prendre en charge une mise en œuvre à obéissance auto-limitée (ACA, *Auto Contingent Allegiance*).
DOIT permettre la construction de passerelles vers d'autres transports SCSI.
DOIT transporter de façon fiable les commandes SCSI de l'initiateur à la cible.
DOIT traiter correctement les abandons de paquets iSCSI, la duplication, la corruption, les paquets périmés, et leur réarrangement.

À partir du paragraphe 7.1 "Sécurité extensible" :
DEVRAIT exiger une configuration et redondance minimales en fonctionnement non sécurisé.
DOIT assurer une authentification forte lorsque une sécurité accrue est exigée.
DEVRAIT permettre l'intégration de nouveaux mécanismes de sécurité sans casser le fonctionnement rétro compatible.

À partir du paragraphe 7.2 "Authentification" :
PEUT prendre en charge divers niveaux de sécurité d'authentification.
DOIT prendre en charge la connexion privée authentifiée.
La connexion iSCSI authentifiée DOIT être résiliente aux attaques.
DOIT prendre en charge l'authentification d'origine des données de ses communications ; l'authentification d'origine des données PEUT être d'utilisation facultative.

À partir du paragraphe 7.3 "Intégrité des données" :
NE DEVRAIT PAS empêcher l'utilisation de protocoles supplémentaires de protection de l'intégrité des données (IPsec, TLS).

À partir du paragraphe 7.4 "Confidentialité des données" :
DOIT fournir l'utilisation d'un protocole de chiffrement des données tel que TLS ou IPsec ESP pour assurer la confidentialité des données entre les points d'extrémité iSCSI.

À partir de la section 8 "Gestion" :
DEVRAIT être gérable en utilisant des protocoles de gestion standard fondés sur IP.
Le document de protocole iSCSI NE DOIT PAS définir l'architecture de gestion pour iSCSI, ou faire des références explicites aux objets de gestion tels que les variables de MIB.

À partir du paragraphe 8.1 "Dénomination" :
DOIT prendre en charge l'architecture de dénomination de SAM-2. Les moyens par lesquels une ressource iSCSI est localisée DOIVENT utiliser ou étendre les méthodes standard existantes de localisation de ressource Internet.
DOIT fournir un moyen d'identifier les cibles iSCSI par un identifiant unique qui est indépendant du chemin sur lequel il se trouve.
Le format des noms iSCSI DOIT utiliser les autorités de dénomination existantes.
Un nom iSCSI DEVRAIT être une chaîne lisible par l'homme dans un codage de jeu de caractères international.

Les services standard de recherche Internet DEVRAIENT être utilisés pour résoudre les noms iSCSI.
DEVRAIT composer avec les complications de la nouvelle architecture de sécurité SCSI.
L'architecture de dénomination iSCSI DOIT traiter la prise en charge des opérations SCSI de tiers tels que COPIE ÉTENDUE.

À partir du paragraphe 8.2 "Découverte" :
DOIT n'avoir pas d'impact sur l'utilisation des techniques actuelles de découverte de réseau IP.
DOIT fournir des moyens de déterminer si un service iSCSI est disponible par une adresse IP.
Les techniques dépendantes du protocole SCSI DEVRAIENT être utilisées pour d'autres découvertes au delà de la couche iSCSI.
DOIT fournir une méthode de découverte, étant donné un point d'extrémité IP sur son accès bien connu, de la liste des cibles SCSI disponibles pour le demandeur. L'utilisation de ce service de découverte DOIT être facultative.

À partir de la section 9 "Accessibilité Internet" .
Les questions de déni de service DEVRAIENT être examinée en détail et elles devraient être réglées.

À partir du paragraphe 9.2 "Pare-feu et serveurs mandataires" :
DEVRAIT permettre le déploiement lorsque sont présents des boîtiers de médiation tels que des pare-feu, des serveurs mandataires et des NAT.
L'utilisation des adresses IP et des accès TCP DEVRAIT être neutre à l'égard des pare-feu.

À partir du paragraphe 9.3 "Contrôle d'encombrement et choix du transport" :
DOIT être un bon citoyen du réseau avec le contrôle d'encombrement compatible avec TCP (comme défini dans la RFC2914 [13]).
Les mises en œuvre iSCSI NE DOIVENT PAS utiliser plusieurs connexions comme moyen d'éviter le contrôle d'encombrement de couche transport.

3. Considérations sur le concept de iSCSI

3.1 Discussion générale

Traditionnellement, les contrôleurs de stockage (par exemple, les contrôleurs de disque, les contrôleurs de bibliothèque de bandes) ont pris en charge le protocole SCSI-3 et ont été rattachés aux ordinateurs par un bus SCSI parallèle ou par canal fibre.

L'infrastructure IP offre des avantages irrésistibles pour le rattachement de mémorisations en volume ou par blocs. Elle offre l'opportunité de tirer parti des avantages en termes de performances/coûts fournis par la compétition sur le marché de l'Internet. Cela pourrait réduire les coûts des infrastructures de mémorisation réseau en effectuant des économies sur le besoin d'installer et faire fonctionner un seul type de réseau.

De plus, la suite des protocoles IP offre l'opportunité d'un riche éventail de solutions de gestion, de sécurité et de qualité de service. Les organisations peuvent initialement choisir de faire fonctionner les réseaux de mémorisation sur la base de iSCSI indépendamment (isolés) de leur réseau de données actuel sauf pour l'acheminement sécurisé du trafic de gestion des mémorisations. Ces organisations attendent des bénéfices du fort rapport performances/coûts des équipements IP et de l'opportunité d'une architecture de gestion unifiée. Avec l'évolution de la sécurité et de la qualité de service (QS) il devient raisonnable de construire des réseaux combinés avec une infrastructure partagée; néanmoins, il est vraisemblable que des utilisateurs sophistiqués vont choisir de garder leurs sous-réseaux de stockage isolés pour en assurer un meilleur contrôle de la sécurité et de la QS pour assurer un environnement à hautes performances pour le trafic de mémorisations.

La transposition de SCSI sur IP donne aussi :
-- des gammes de distance étendues,
-- la connexité aux services de "classe de transporteur" qui prennent en charge IP.

Les applications suivantes pour iSCSI sont envisagées :
-- Accès à la mémorisation locale, consolidation, groupement et mise en commun (comme dans le centre de données).
-- Accès du client réseau à la mémorisation distante (par exemple, un "fournisseur de service de mémorisation").
-- Réplication synchrone et asynchrone locale et distante entre contrôleurs de mémorisation.
-- Sauvegarde et récupération locale et distante..

iSCSI va accepter les topologies suivantes :
-- Connexions directes en point à point.

-- LAN de mémorisation dédié, consistant en un ou plusieurs segments de LAN.
-- LAN partagé, portant un mélange de trafic de LAN traditionnel plus du trafic de mémorisation.
-- Une extension de LAN à WAN utilisant des routeurs IP ou des "Datatone IP" fournis par le transporteur.
-- Des réseaux privés et l'Internet public.

Des routeurs LAN-WAN IP peuvent être utilisés pour étendre le réseau de mémorisation IP à la large zone, permettant l'accès au disque distant (comme pour un utilitaire de mémorisation) la réplication synchrone et asynchrone, et la sauvegarde et la récupération à distance (comme pour l'archivage de bandes). Dans le WAN, utiliser TCP de bout en bout évite d'avoir besoin d'équipements spécialisés pour la conversion de protocole, assure la fiabilité des données, surmonte l'encombrement du réseau, et fournit des stratégies de retransmission adaptées au délais de WAN.

Le déploiement de la technologie iSCSI va impliquer les éléments suivants :
1. la réalisation d'un protocole standard complet et des mises en œuvre qui le prennent en charge ;
2. le développement de cartes d'interface réseau (NIC, *Network Interface Card*) de mémorisation Ethernet et du logiciel de pilote et de protocole qui s'y rapporte ; [noter que des applications haut débit de iSCSI vont demander des portions significatives de la mise en œuvre iSCSI/TCP/IP dans le matériel pour réaliser le débit nécessaire] ;
3. le développement de contrôleurs de mémorisation compatibles ;
4. le développement vraisemblable de passerelles de traduction pour fournir la connexité entre le réseau de mémorisation Ethernet et les domaines SCSI de canal fibre et/ou bus parallèle ;
5. le développement de spécifications pour la gestion d'appareils iSCSI tels que des MIB, des schémas LDAP ou XML, etc. ;
6. le développement d'applications de service de gestion et de répertoire pour soutenir une infrastructure SAN robuste.

Les produits pourraient initialement être offerts pour le rattachement sur Gigabit Ethernet, avec une migration rapide sur 10 GbE. Pour des performances compétitives avec les autres transports SCSI, il sera nécessaire de mettre en œuvre dans le matériel le chemin de réalisation de la pile de protocole complète. Ces nouveaux NIC de mémorisation pourraient effectuer un traitement de pleine pile d'une tâche SCSI complète, de façon analogue aux HBA SCSI et canal Fibre d'aujourd'hui et pourraient aussi prendre en charge tous les protocoles d'hôte qui utilisent TCP (NFS, CIFS, HTTP, etc.).

Le mandat du groupe de travail Mémorisation IP de l'IETF (IPSWG) décrit les grandes lignes de l'objectif de transposition de SCSI en IP en utilisant un transport qui a démontré son comportement d'évitement d'encombrement et sa large mise en œuvre sur des plateformes diverses. Dans ce mandat, plusieurs solutions de remplacement de transport peuvent être examinées. Les travail initial d'IPS se concentre sur TCP, et le présent document de description des exigences se restreint à ce domaine d'intérêt.

3.2 Performances/coût

En général, iSCSI DOIT permettre aux mises en œuvre d'égaler ou d'améliorer l'état actuel de l'art pour les interconnexions iSCSI. Cet objectif se subdivise en plusieurs types d'exigences :

- Coût compétitif avec les autres technologies de réseau de stockage :
  Pour être adopté par les fabricants et la communauté des utilisateurs, le protocole iSCSI DOIT permettre des mises en œuvre d'un coût compétitif comparé aux autres transports SCSI (canal Fibre).

- Faible délai de communication :
  L'accès conventionnel à la mémorisation est un type de procédure d'appel à distance à arrêt-attente. Les applications emploient normalement très peu le traitement en parallèle de leurs accès de mémorisation, et donc les délais d'accès aux mémorisations impactent directement les performances. Les délais imposés par les interconnexions courantes de mémorisations, y compris le traitement du protocole, est généralement de l'ordre des 100 microsecondes. L'utilisation d'antémémoires dans les contrôleurs de mémorisation signifie que de nombreux accès de mémorisations s'achèvent presque instantanément, et que donc le délai de l'interconnexion peut avoir un fort impact relatif sur les performances globales. Lorsque une entrée/sortie à arrêt-attente est utilisé, le délai de l'interconnexion va affecter les performances. Le protocole iSCSI DEVRAIT minimiser la redondance du contrôle, qui s'ajoute au délai.

- Faible utilisation de la CPU de l'hôte, égale ou meilleure que celle de la technologie actuelle :
  Pour des performances compétitives, le protocole iSCSI DOIT permettre la réalisation de trois objectifs clés de mise en œuvre :
  (1) iSCSI DOIT rendre possible la construction d'adaptateurs entrée/sortie qui traitent une tâche SCSI entière, comme le font les autres mises en œuvre de transport SCSI.
  (2) Le protocole DEVRAIT permettre un placement direct des données (architectures de mémoire "copie zéro", où les adaptateurs entrée/sortie lisent ou écrivent la mémoire de l'hôte exactement une fois par transaction de disque.

(3) Le protocole NE DEVRAIT PAS imposer d'opérations complexes au logiciel d'hôte, ce qui augmenterait la longueur du chemin des instructions chez l'hôte par rapport aux solutions de remplacement.

- Placement direct des données (iSCSI à copie zéro) :
  Le placement direct des données se réfère aux données iSCSI qui sont placées directement "hors réseau" dans la localisation allouée dans la mémoire sans copie intermédiaire. Le placement direct des données réduit significativement la charge du bus de mémoire et du bus d'entrée/sortie dans les systèmes de point d'extrémité, permettant des performances améliorées. Il réduit la mémoire requise pour les NIC, réduisant éventuellement le coût de ces solutions.

  C'est un important objectif de mise en œuvre. Dans un système iSCSI, chacun des nœuds d'extrémité (par exemple l'ordinateur hôte et le contrôleur de mémorisation) devrait avoir une mémoire ample, mais les nœuds intermédiaires (NIC, commutateurs) n'en ont normalement pas.

- Forte bande passante, agrégation de bande passante :
  La bande passante (taux de transfert, Mbit/s) prise en charge par les contrôleurs de mémorisation augmente rapidement, à cause de plusieurs facteurs :
  1. augmentation de la vitesse de rotation des disques et des performances des contrôleurs ;
  2. utilisation d'antémémoires toujours plus grandes, et amélioration des algorithmes de mise en antémémoire ;
  3. taille accrue des contrôleurs de mémorisation (nombre de fuseaux pris en charge, vitesse d'interconnexion).

Le protocole iSCSI DOIT fournir une pleine utilisation de la bande passante de liaison disponible. Le protocole DOIT aussi permettre à une mise en œuvre d'exploiter le parallélisme (connexions multiples) aux interfaces de l'appareil et au sein de la machine d'interconnexion.

Les deux paragraphes qui suivent exposent plus en détails le besoin du placement direct des données et d'une forte bande passante.

3.3 Tramage

Le tramage se réfère à l'ajout d'informations dans un en-tête, ou au flux de données pour permettre aux mises en œuvre de localiser les limites d'une unité de données de protocole (PDU, *protocol data unit*) iSCSI au sein du flux d'octets TCP. Il y a deux exigences techniques qui pilotent le tramage : les besoins de l'interface, et les besoins d'accélération du traitement.

Une solution de tramage qui s'adresse au "besoins de l'interface" du protocole iSCSI va faciliter la mise en œuvre d'un protocole (iSCSI) de couche supérieure fondé sur le message par dessus un protocole sous-jacent de flux d'octets (TCP). Comme TCP est un transport fiable, cela peut se réaliser en incluant un champ de longueur dans l'en-tête iSCSI. Trouver la trame du protocole suppose que le receveur va analyser depuis le début du flux de données TCP, et ne jamais faire une faute (perdre l'alignement sur les en-têtes de paquet).

L'autre exigence technique de tramage, "l'accélération du traitement", découle du besoin de traiter des débits de données de plus en plus élevés dans l'interface physique du support. Deux besoins découlent des débits de données plus élevés :
(1) environnement de LAN – les fabricants de NIC cherchent des moyens pour fournir des méthodes "zéro copie" de déplacement des données directement du réseau dans les mémoires tampon de l'application.
(2) environnement de WAN – l'émergence de supports physiques à large bande passante, forte latence, faible taux d'erreur binaire fait peser des exigences énormes de mémoire tampon sur les solutions d'interface physique.

D'abord, les fabricants produisent des matériels de traitement de réseau qui transfèrent les charges des protocoles réseau aux matériels pour réaliser des taux plus élevés de transfert de données. Le concept de "zéro copie" cherche à mémoriser les blocs de données dans des mémoires situées aux endroits appropriés (alignés) directement en dehors du réseau, même lorsque les données sont réarrangées suite à une perte de paquet. Ceci est nécessaire pour piloter les taux réels de données de 10 Gigabit/s et au-delà.

Ensuite, afin que iSCSI réussisse dans le domaine du WAN, il doit être possible d'opérer efficacement dans les réseaux à forte bande passante et fort délai. L'émergence de réseaux IP multi-gigabit avec des latences de l'ordre du dixième au centième de milliseconde présente un défi. Pour remplir d'aussi gros tuyaux, il est nécessaire d'avoir des dizaines de méga octets de demandes en instance de la part de l'application. De plus, certains protocoles exigent potentiellement des dizaines de méga octets à la couche transport pour assurer la mise en mémoire tampon pour le réassemblage des données lorsque les paquets sont reçus dans le désordre.

Dans les deux cas, le problème est le désir de minimiser la quantité de mémoire et la bande passante de mémoire requise pour les solutions de matériel iSCSI.

Considérons qu'un canal réseau à 10 Gbit/s x 200 ms contient 250 MB. [On suppose une communication terrestre sur la moitié de la circonférence terrestre à l'équateur. On ignore la distance supplémentaire due à l'acheminement par câble. On ignore les délais de répéteur et de commutateur ; on considère seulement un délai dû à la vitesse de la lumière de 5 micros/km. La circonférence du globe à l'équateur est d'environ 40 000 km (le délai d'aller retour doit être considéré comme gardant le tuyau plein). 10 Gbit/s x 40 000 km x 5 micros/km x B / 8b = 250 MB]. Dans une mise en œuvre TCP conventionnelle, la perte d'un segment TCP signifie que le traitement du flux DOIT s'arrêter jusqu'à ce que le segment soit récupéré, ce qui prend au moins un temps de <aller-retour réseau> à accomplir. Suivant l'exemple ci-dessus, une mise en œuvre serait obligée de capter 250 MB de données dans une mémoire tampon anonyme avant de reprendre le traitement du flux ; plus tard, ces données devront être déplacées à la localisation appropriée. Certains défenseurs de iSCSI cherchent des moyens pour mettre les données directement là où elles doivent aller, et à éviter des mouvements de données supplémentaires dans le cas d'un abandon de segment. Ceci est un concept clé pour la compréhension du débat sur les méthodologies de tramage.

Le tramage du protocole iSCSI a un impact aussi bien sur les "besoins d'interface" que sur les "besoins d'accélération du traitement", cependant, bien que l'inclusion d'une longueur dans un en-tête puisse suffire pour les "besoins d'interface", elle ne servira pas au besoin de placement direct des données. Le mécanisme de tramage développé devrait permettre la resynchronisation des limites de paquet même dans le cas où un paquet est temporairement manquant dans le flux de données entrant.

3.4 Haut débit, agrégation de bande passante

Au débit de transport de mémorisation de bloc d'aujourd'hui, toute liaison peut être saturée par le volume de trafic de mémorisation. Les applications de données scientifiques et de duplication de données sont des exemples d'applications de mémorisation qui touchent les limites du débit.

Certaines applications, comme les mises à jour d'enregistrement de données, de bandes de direct, et de duplication, exigent le rangement des mises à jour et donc le rangement des commandes SCSI. Un initiateur peut conserver l'ordre de rangement en attendant que chaque mise à jour s'achève avant de produire la suivante (autrement dit, des mises à jour synchrones). Cependant, le débit de mises à jour synchrones diminue à l'inverse des distances dans le réseau.

Pour un plus grand débit, le mécanisme de mise en file d'attente des tâches SCSI permet à un initiateur d'avoir plusieurs commandes en instance simultanément à la cible et d'exprimer les contraintes de rangement sur l'exécution de ces commandes. Le mécanisme de mise en file d'attente des tâches n'est efficace que si les commandes arrivent à la cible dans l'ordre dans lequel elles étaient présentées à l'initiateur (ordre FIFO, premier entré premier sorti). La norme iSCSI doit fournir un transport ordonné des commandes SCSI, même lorsque les commandes sont envoyées sur des chemins réseau différents (voir le paragraphe 6.2 "SCSI"). On appelle cela le "rangement des commandes".

Le protocole iSCSI DOIT fonctionner sur une seule connexion TCP pour s'accommoder des mises en œuvre au moindre coût. Pour permettre des appareils de mémorisation avec de meilleures performances, le protocole devrait spécifier un moyen pour permettre de fonctionner sur plusieurs connexions tout en conservant le comportement d'un seul accès SCSI. Cela permettrait à l'initiateur et à la cible d'utiliser plusieurs interfaces réseau et plusieurs chemins à travers le réseau pour un débit accru. Il y a quelques moyens potentiels de satisfaire les exigences de chemins multiples et de rangement.

Un moyen populaire de satisfaire l'exigence de chemins multiples est d'avoir un pilote au dessus de la couche SCSI qui instancie plusieurs copies du transport SCSI, chacune communiquant avec la cible sur un chemin différent. Les pilotes "de coin" utilisent aujourd'hui cette technique pour obtenir de hautes performances. Malheureusement, les pilotes de coin doivent attendre un accusé de réception d'achèvement de chaque demande (arrêt et attente) pour s'assurer de mises à jour dans l'ordre.

Une autre approche pourrait être que le protocole iSCSI utilise plusieurs instances de son transport sous-jacent (par exemple, TCP). La couche iSCSI ferait apparaître ces instances de transport indépendantes comme une seule instance de transport SCSI et conserverait la capacité de faire une mise en file d'attente ordonnée des commandes SCSI. Le présent document se réfèrera à cette technique sous le nom de "lien de connexion" pour des raisons pratiques.

Le protocole iSCSI DEVRAIT prendre en charge le lien de connexion, et il DOIT être de mise en œuvre facultative.

En présence d'un lien de connexion, il y a deux façons d'allouer des caractéristiques aux connexions. Dans l'approche symétrique, toutes les connexions sont identiques du point de vue des caractéristiques. Dans le modèle asymétrique, les connexions ont des caractéristiques différentes. Par exemple, certaines connexions peuvent être utilisées principalement pour des transferts de données alors que d'autres sont utilisés principalement pour des commandes SCSI.

Comme le protocole iSCSI doit prendre en charge le cas où il y a seulement une connexion de transport, il doit avoir la commande, les données, et l'état qui voyagent sur la même connexion.

Dans le cas de connexions multiples, le protocole iSCSI doit garder la commande et ses données et son état associés sur la même connexion (allégeance de connexion). L'envoi des données et de l'état sur la même connexion est souhaitable parce que cela garantit que l'état est reçu après les données (TCP fournit une livraison ordonnée). Dans le cas où chaque connexion est gérée par un processeur séparé, l'allégeance diminue le besoin de communication entre les processeurs. Cette approche symétrique est une extension naturelle de l'approche de la connexion unique.

Une autre approche qui a été largement discutée impliquait l'envoi de toutes les commandes sur une seule connexion et des données et de l'état associé sur une connexion différente (approche asymétrique). Dans ce schéma, le transport assure que les commandes arrivent dans l'ordre. Le protocole sur les connexions de données et d'état est plus simple, se prêtant peut-être lui-même à une réalisation plus simple dans le matériel. Un désavantage de cette approche est que la procédure de récupération est différente si une connexion de commande échoue par rapport à une connexion de données. Certains avancent que cette approche exigerait plus de communications inter-processeur lorsque les connexions sont éparpillées entre les processeurs.

Le lecteur peut se référer aux archives de la messagerie de la liste de diffusion du groupe de travail IPS entre juin et septembre 2000 pour voir les discussions intenses sur les modèles de connexion symétrique et asymétrique .

4. Facilité de mise en œuvre/complexité du protocole

L'expérience a montré que l'adoption d'un protocole par la communauté de l'Internet est inversement proportionnelle à sa complexité. De plus, plus simple est le protocole, plus il est facile de diagnostiquer les problèmes. Les concepteurs de iSCSI DEVRAIENT s'efforcer de satisfaire aux exigences de la création d'un transport SCSI sur IP, tout en gardant le protocole aussi simple que possible.

Dans l'intérêt de la simplicité, iSCSI DEVRAIT minimiser les caractéristiques optionnelles. Lorsque des caractéristiques sont réputées nécessaires, le protocole DOIT spécifier la négociation des caractéristiques à l'établissement de la session (à l'amorçage). Le transport iSCSI DOIT fonctionner correctement lorsque aucune caractéristique facultative n'est négociée ainsi que lorsque les négociations d'options individuelles ne réussissent pas.

5. Fiabilité et disponibilité

5.1 Détection des données corrompues

Il y a eu plusieurs papiers de recherche qui suggéraient que le calcul de la somme de contrôle TCP permet à un certain nombre d'erreurs binaires de passer indétectées [10], [11].

Pour protéger contre la corruption des données, le protocole iSCSI DOIT prendre en charge un format de vérification de l'intégrité des données à utiliser à la génération d'un résumé.

Le protocole iSCSI PEUT utiliser des résumés séparés pour les données et les en-têtes. Dans une situation de mandataire iSCSI ou de passerelle, les en-têtes iSCSI sont retirés et reconstruits, et le flux TCP se termine sur l'un ou l'autre des côtés. Cela signifie que même la somme de contrôle TCP est retirée et recalculée au sein de la passerelle. Pour assurer la protection des commandes, des données, et de l'état, le protocole iSCSI DOIT inclure un CRC ou autre mécanisme de résumé qui est calculé sur le bloc de données SCSI lui-même, ainsi que sur chaque message de commande et d'état. Comme les passerelles peuvent ôter les en-têtes iSCSI et les reconstruire, un CRC d'en-tête séparé est exigé. Deux résumés d'en-tête, un pour les portions invariantes de l'en-tête (les adresses) et un pour la portion variable, vont assurer la protection contre les changements des portions de l'en-tête qui ne devraient jamais être changées par les boîtiers de médiation (par exemple, les adresses).

Le format d'en-tête iSCSI DEVRAIT être extensible pour inclure d'autres méthodes de calcul de résumé.

5.2 Récupération

Le protocole SCSI a été à l'origine conçu pour un transport de bus parallèle qui était très fiable. Les applications SCSI tendent à supposer que les erreurs de transport n'arrivent jamais, et quand il y en a, la récupération d'application SCSI tend à être coûteuse en termes de temps et de ressources de calcul.

La conception du protocole iSCSI, bien qu'elle mette l'accent sur la simplicité, DOIT conduire à la récupération à temps des défaillances de l'initiateur, de la cible, ou de l'infrastructure du réseau de connexion (câblage, équipements du chemin des données tels que les routeurs, etc.).

iSCSI DOIT spécifier des méthodes de récupération pour les demandes non idempotentes, telles que les opérations sur les pilotes de bandes.

Le mécanisme de récupération d'erreur du protocole iSCSI DEVRAIT prendre en compte les schémas de récupération pour les cibles reflétées ou les configurations de mémorisation à forte disponibilité qui fournissent des chemins vers les données cibles à travers plusieurs "serveurs de mémorisation". Cela doit fournir une base pour les technologies en couches comme la forte disponibilité et la mise en grappes.

Le protocole iSCSI DEVRAIT aussi fournir une méthode pour que les sessions soient terminées et redémarrées en douceur qui puisse être initiée par l'initiateur ou par la cible. Cela donne la capacité de récupérer en douceur à l'initiateur ou à la cible, ou de rétablir une cible après des tâches de maintenance comme la mise à niveau d'un logiciel.

6. Interopérabilité

Il doit aussi être possible aux initiateurs et cibles qui mettent en œuvre les portions requises de la spécification iSCSI d'interopérer. Bien que cette exigence soit si évidente qu'il semble à peine besoin de la mentionner, si la spécification du protocole contient des formulations ambiguës, des mises en œuvre différentes pourraient ne pas interopérer. Le document du protocole iSCSI DOIT être clair et sans ambiguïté.

6.1 Infrastructure Internet

Le protocole iSCSI DOIT :
-- être compatible avec IPv4 et IPv6.
-- utiliser avec prudence les connexions TCP, en se souvenant qu'il peut y avoir de nombreux autres utilisateurs de TCP sur une même machine.

Le protocole iSCSI NE DOIT PAS exiger de changements aux protocoles Internet existants et DEVRAIT minimiser les changements requis aux mises en œuvre TCP/IP existantes.

iSCSI DOIT être conçu de façon à permettre une future substitution de SCTP (à TCP) comme protocole de transport IP avec des changements minimes au fonctionnement du protocole iSCSI, aux structures et formats des unités de données de protocole (PDU). Bien que non encore largement mis en œuvre aujourd'hui, SCTP a de nombreuses caractéristiques de conception qui en font un choix désirable pour les futures améliorations de iSCSI.

6.2 SCSI

Afin d'être considérée comme un transport SCSI, la norme iSCSI doit se conformer aux exigences du modèle d'architecture SCSI [3] pour un transport SCSI. Toute caractéristique exigée par SAM2 dans une transposition valide de transport DOIT être spécifiée par iSCSI. Le document du protocole iSCSI DOIT spécifier pour chaque caractéristique si elle est de mise en œuvre et/ou utilisation FACULTATIVE, RECOMMANDÉE ou EXIGÉE.

Le modèle d'architecture SCSI [3] indique qu'il attend que le transport SCSI assure le rangement des commandes à la granularité de numéro d'unité logique (LUN, *Logical Unit Number*) initiateur cible. Il y a eu beaucoup de discussions sur la liste de diffusion IPS et dans les réunions du groupe de travail sur les moyens de s'assurer de ce rangement. Le consensus est en gros que iSCSI DOIT spécifier une livraison strictement ordonnée des commandes SCSI sur une session iSCSI entre une paire initiateur/cible, même en présence d'erreurs de transport. Ce mécanisme de rangement des commandes DEVRAIT chercher à minimiser la quantité de communications nécessaire à travers plusieurs adaptateurs qui font le transfert du transport. Si une mise en œuvre iSCSI n'exige pas le rangement, elle peut instancier plusieurs sessions par paire d'initiateur-cible.

iSCSI est destiné à être un nouveau "transport" SCSI [3]. Comme transposition de SCSI sur TCP, iSCSI exige l'interaction à la fois avec le T10 et l'IETF. Cependant, le protocole iSCSI NE DOIT PAS exiger de changements aux ensembles de commandes SCSI-3 et au code client SCSI sauf lorsque les spécifications SCSI pointent sur des champs et des comportements "dépendants du transport". Par exemple, des changements aux documents SCSI seront nécessaires pour refléter des noms de cibles iSCSI plus longs et des temporisations potentiellement plus longues. La collaboration avec le

T10 sera nécessaire pour satisfaire cette exigence.

Le protocole iSCSI DEVRAIT garder trace des changements à SCSI et au modèle d'architecture SCSI.

Le protocole iSCSI DOIT être capable de prendre en charge tous les ensembles de commandes SCSI-3 et types d'appareil. L'accent est mis principalement sur la prise en charge des 'plus grands' appareils : les ordinateurs d'hôte et les contrôleurs de mémorisation (armoires de disques, bibliothèques de bandes). Cependant, d'autres ensembles de commandes (imprimantes, scanneurs) doivent être pris en charge. Ces exigences NE DOIVENT PAS être interprétées comme signifiant que iSCSI doive être mis en œuvre dès le départ sur tous les appareils SCSI d'aujourd'hui, qui peuvent avoir des capacités de puissance de traitement ou de mémoire limitées.

L'obéissance autolimitée (ACA, *Auto Contingent Allegiance*) est un mécanisme SCSI facultatif qui arrête l'exécution d'une séquence de commandes SCSI lorsque l'une d'elles échoue. La situation environnante est complexe – le comité T10 spécifie ACA dans SAM2, et donc iSCSI doit la prendre en charge et entreprendre de s'assurer que ACA est suffisamment mise en œuvre (deux mises en œuvre indépendantes interopérables) pour éviter d'abandonner l'ACA dans la transition entre proposition de norme à projet de norme. Cela implique que iSCSI DEVRAIT prendre en charge la mise en œuvre d'ACA.

Le protocole iSCSI DOIT permettre la construction de passerelles avec d'autres transport SCSI, y compris SCSI parallèle [7] et FCP SCSI [8], [9]. Il DOIT être possible de construire des passerelles de "traduction" afin que les hôtes iSCSI puissent interopérer avec les appareils SCSI-X ; afin que les appareils SCSI-X puissent communiquer sur un réseau iSCSI ; et afin que les hôtes SCSI-X puissent utiliser des cibles iSCSI (où SCSI-X se réfère à SCSI parallèle, SCSI-FCP, ou SCSI sur tout autre transport). Cette exigence est impliquée par la prise en charge de SAM-2, mais mérite d'être explicitée. Il y a de vraies passerelles de protocole d'application, qui ne sont pas simplement des ponts/routeurs. Les différentes normes ont seulement la couche d'ensemble de commandes SCSI-3 en commun. Ces passerelles ne sont pas de simples transmetteurs de paquets.

Le protocole iSCSI DOIT transporter fiablement les commandes SCSI de l'initiateur à la cible. Selon [3] page 17, "la fonction du sous-système de livraison de service est de transporter une copie sans erreur de la demande ou réponse entre l'envoyeur et le receveur". Le protocole iSCSI DOIT traiter correctement l'abandon de paquet iSCSI, la duplication, la corruption, les paquets périmés, et le réarrangement.

7. Considérations pour la sécurité

Dans le passé, les systèmes de mémorisation directement rattachés ont mis en œuvre des vérifications de sécurité minimales parce que la connexion physique offrait peu de chances d'attaque. Le transport de mémorisation de blocs (SCSI) sur IP ouvre une nouvelle opportunité pour diverses attaques malveillantes. Les attaques peuvent prendre une forme active (usurpation d'identité, par interposition) ou passive (espionnage).

7.1 Sécurité extensible

Les services de sécurité nécessaire pour les communications dépendent des configurations et environnements individuels de réseau. Les organisations établissent des réseaux privés virtuels (VPN, *Virtual Private Network*), aussi appelés Intranets, qui vont exiger un ensemble de fonctions de sécurité pour les communications au sein du VPN et éventuellement de nombreuses fonctions de sécurité différentes pour les communications en dehors du VPN pour la prise en charge de composants géographiquement séparés. Le protocole iSCSI est applicable à une large gamme d'environnements de travail Internet qui peuvent employer des politiques de sécurité différentes. iSCSI DOIT fournir une authentification forte lorsque une sécurité accrue est exigée. Le protocole DEVRAIT exiger une configuration et une redondance minimales dans le fonctionnement non sécurisé, et permettre l'intégration de nouveaux mécanismes de sécurité sans casser le fonctionnement rétro compatible.

7.2 Authentification

Le protocole iSCSI PEUT prendre en charge divers niveaux de sécurité d'authentification, allant de pas d'authentification à une authentification sûre qui utilise des clés publiques ou privées.

Le protocole iSCSI DOIT prendre en charge la connexion authentifiée privée.

La connexion authentifiée aide la cible à bloquer l'utilisation non autorisée des ressources SCSI. La connexion authentifiée "privée" rend obligatoire l'échange d'identités protégées (au minimum, pas de mot de passe en clair). Comme la

confidentialité de la mémorisation de bloc est considérée comme critique dans les entreprises et que de nombreux réseaux IP peuvent avoir des trous d'accès, les organisations vont vouloir protéger leurs ressources iSCSI.

La connexion iSCSI authentifiée DOIT être résistante contre les attaques car de nombreux réseaux IP sont vulnérables à l'inspection de paquets.

De plus, le protocole iSCSI DOIT prendre en charge l'authentification d'origine des données de ses communications ; l'authentification d'origine des données PEUT être d'utilisation facultative. L'authentification d'origine des données est critique car les réseaux IP sont vulnérables à l'usurpation d'identité de source, où un tiers malveillant prétend envoyer des paquets depuis l'adresse IP de l'initiateur. Ces exigences devraient être satisfaites en utilisant des protocoles standard de l'Internet tels que IPsec ou TLS. Les points d'extrémité peuvent négocier la méthode d'authentification, et facultativement, aucune.

7.3 Intégrité des données

Le protocole iSCSI NE DEVRAIT PAS empêcher l'utilisation de protocoles supplémentaires de protection de l'intégrité des données (IPsec, TLS).

7.4 Confidentialité des données

La mémorisation de bloc est utilisée pour mémoriser des informations sensibles, lorsque la confidentialité des données est critique. Une application peut chiffrer les blocs de données avant de les écrire dans la mémorisation – cela donne la meilleure protection pour l'application. Même si la mémorisation ou les communications sont compromises, l'attaquant aura des difficulté à lire les données.

Dans certains environnements, le chiffrement peut être désiré pour fournir une assurance de confidentialité supplémentaire. Une mise en œuvre iSCSI DOIT assurer l'utilisation d'un protocole de chiffrement de données tel que TLS ou IPsec ESP pour fournir la confidentialité des données entre les points d'extrémité iSCSI.

8. Gestion

Les mises en œuvre iSCSI DEVRAIENT être gérables en utilisant les protocoles de gestion standard fondés sur IP. Cependant, le document du protocole iSCSI NE DOIT PAS définir l'architecture de gestion pour iSCSI au sein de l'infrastructure réseau. iSCSI sera déjà un autre service de ressource au sein d'un environnement complexe de ressources réseau (imprimantes, serveurs de fichiers, NAS, serveurs d'application, etc.). Il y aura certainement des efforts pour concevoir comment le "service de mémorisation de bloc" que fournissent les appareils iSCSI est intégré dans un environnement complet de gestion de réseau, au modèle réparti. Un "administrateur de réseau" (ou "administrateur de mémorisation") va souhaiter avoir des applications intégrées pour allouer les noms d'utilisateur, les noms de ressource, etc. et indiquer les droits d'accès. Les appareils iSCSI vont probablement vouloir interagir avec ces applications de gestion intégrées au réseau. Le document du protocole iSCSI ne va pas tenter de résoudre cet ensemble de problèmes, ou de spécifier les moyens par lesquels les appareils fournissent des agents de gestion. En fait, il ne devrait pas y être mentionné de MIB ou tout autre moyen de gestion des appareils iSCSI comme références explicites dans le document du protocole iSCSI, parce que les données et protocoles de gestion changent avec les besoins de l'environnement et les modèles commerciaux des applications de gestion.

8.1 Désignation

Chaque fois que possible, iSCSI DOIT prendre en charge l'architecture de désignation de SAM-2. Les déviations et incertitudes DOIVENT être rendues explicites, et les commentaires et résolutions élaborés entre le groupe de travail T10 de l'ANSI T10 et le groupe de travail IPS.

Le moyen par lequel est localisée une ressource iSCSI DOIT utiliser ou étendre des méthodes existantes de localisation de ressource des normes Internet. La RFC 2348 [12] spécifie la syntaxe et la sémantique des URL qui devrait être suffisamment extensible pour les ressources iSCSI.

Le protocole iSCSI DOIT fournir un moyen d'identifier un appareil de mémorisation iSCSI par un identifiant univoque qui est indépendant du chemin sur lequel il se trouve. Ce nom sera utilisé pour corréler les chemins de remplacement pour le même appareil. Le format des noms iSCSI DOIT utiliser les autorités de désignation existantes, pour éviter de créer de nouvelles tâches administratives centrales. Un nom iSCSI DEVRAIT être une chaîne lisible par l'homme dans un codage de jeu de caractères international.

Les services de recherche standard de l'Internet DEVRAIENT être utilisés pour résoudre les noms. Par exemple, le service des noms de domaine (DNS, *Domain Name Service*) PEUT être utilisé pour résoudre la portion <nom_d'hôte> d'un URL en une ou plusieurs adresses IP. Lorsque un nom d'hôte se résout en plusieurs adresses, ces adresses devraient être équivalentes pour les besoins fonctionnels (éventuellement pas de performances). Cela signifie que les adresses peuvent être utilisées de façon interchangeable tant que les performances ne sont pas en jeu. Par exemple, le même ensemble de cibles SCSI DOIT être accessible à partir de chacune de ces adresses.

Un schéma de dénomination des appareils iSCSI DOIT interagir correctement avec l'architecture de sécurité SCSI proposée [6]. Une attention particulière doit être portée à l'architecture de désignation des mandataires définie par le nouveau modèle de sécurité. Dans ce nouveau modèle, un hôte est identifié par un identifiant d'accès, et des numéros d'unité logique (LUN, *Logical Unit Number*) SCSI peuvent être transposés d'une manière qui donne à chaque identifiant d'accès une LU transposée unique. Donc, une certaine LU au sein d'une cible peut être adressée par des LUN différents.

L'architecture de dénomination iSCSI DOIT s'occuper de la prise en charge d'opérations SCSI de tiers telles que EXTENDED COPY. La question clé ici se rapporte à l'architecture de désignation des LU SCSI – iSCSI doit fournir un moyen de passer un nom ou un lien entre les parties. iSCSI doit spécifier un moyen de fournir un nom ou lien qui pourrait être utilisé dans la commande XCOPY et convenir dans l'espace disponible alloué par cette commande. Et il doit être possible, bien sûr, à la cible XCOPY (le tiers) de déréférencer le nom en la cible et LU correctes.

8.2 Découverte

iSCSI DOIT n'avoir aucun impact sur l'utilisation des techniques actuelles de découverte de réseau IP. Les plateformes de gestion de réseau découvrent les adresses IP et ont diverses méthodes de sondage des services disponibles à travers ces adresses IP. Un service iSCSI devrait évidemment utiliser des techniques similaires.

Les spécifications iSCSI DOIVENT fournir des moyens pour déterminer si un service iSCSI est disponible à travers une adresse IP. Il est prévu que iSCSI soit un point de service dans un hôte, tout comme SNMP, etc. sont des points de services, associés à un numéro d'accès bien connu.

Les techniques qui dépendent du protocole SCSI DEVRAIENT être utilisées pour d'autres découvertes au delà de la couche iSCSI. La découverte est un processus complexe, multi couches. Les spécifications du protocole SCSI fournissent des commandes spécifiques pour découvrir les LU et les commandes associées à ce processus vont aussi fonctionner sur iSCSI.

Le protocole iSCSI DOIT fournir une méthode de découverte, pour un point d'extrémité IP sur son accès bien connu, la liste des cibles SCSI disponibles pour le demandeur. L'utilisation de ce service de découverte DOIT être facultative.

Aller plus loin dans les lignes directrices sur la découverte sortirait du domaine d'application du présent document et pourra être traité dans des documents pour information distincts.

9. Accessibilité de l'Internet

9.1 Déni de service

Comme avec tous les services, le déni de service soit par des mises en œuvre incorrectes, soit par des agents malveillants est toujours un souci. Tous les aspects du protocole iSCSI DEVRAIENT être examinés avec attention à la recherche de problèmes potentiels de déni de service, et en être protégés autant que possible.

9.2 NAT, pare-feu et serveurs mandataires

Les traducteurs d'adresse réseau (NAT, *Network Address Translator*), pare-feu, et serveurs mandataires sont une réalité dans l'Internet d'aujourd'hui. Ces appareils présentent un certain nombre de défis aux méthodes d'accès aux appareils qui ont été développés pour iSCSI. Par exemple, spécifier une syntaxe d'URL pour la connexion de ressources iSCSI permet à un initiateur de s'adresser à un appareil cible iSCSI à la fois directement et à travers un serveur iSCSI mandataire ou un NAT. iSCSI DEVRAIT permettre le déploiement lorsque c'est fonctionnel et d'optimiser des boîtiers de médiation tels que des pare-feu, des serveurs mandataires et des NAT lorsqu'il y en a.

L'utilisation de l'adressage IP et des numéros d'accès TCP par le protocole iSCSI DOIT être compatible avec les pare-feu. Cela signifie que toutes les demandes de connexion devraient normalement être adressées à un accès TCP bien connu spécifique. De cette façon, les pare-feu peuvent filtrer sur la base des adresses IP de source et de destination, et le numéro

d'accès de destination (cible). Des connexions TCP supplémentaires exigeraient des numéros d'accès de source différents (pour l'unicité) mais pourraient être ouvertes après un dialogue de sécurité sur le canal de contrôle.

Il est important que iSCSI fonctionne à travers un pare-feu pour fournir des moyens possibles de se défendre contre des assauts de déni de service (DoS) provenant de domaines du réseau moins dignes de confiance. On suppose qu'un pare-feu aura des pouvoirs de traitement beaucoup plus grands pour éliminer des demandes de connexion boguées que les nœuds d'extrémité.

9.3 Contrôle d'encombrement et choix du transport

Le protocole iSCSI DOIT être un bon citoyen du réseau avec un contrôle d'encombrement prouvé (comme défini dans la RFC2914 [13]). De plus, les mises en œuvre iSCSI NE DOIVENT PAS utiliser plusieurs connexions comme moyen d'éviter le contrôle d'encombrement de couche transport.

10. Définitions

Certaines définitions sont données ici, avec des références au document d'origine lorsque applicable, afin de préciser la discussion des exigences. Les définitions sans référence sont données par les auteurs et réviseurs de ce document.

Unité logique (LU, *Logical Unit*) : entité résidant à la cible qui met en œuvre un modèle d'appareil et exécute les commandes SCSI envoyées par une application client (paragraphe 3.1.50, p. 7 de [3]).

Numéro d'unité logique (LUN, *Logical Unit Number*) : identifiant de 64 bits d'une unité logique cal unit (paragraphe 3.1.52 p. 7 de [32]).

Appareil SCSI : appareil connecté au sous-système de livraison d'un service et qui prend en charge un protocole d'application SCSI (paragraphe 3.1.78, page 9 de [3]).

Accès de livraison de service SDP, *Service Delivery Port*) : interface résidente d'un appareil utilisée par l'application client, le serveur d'appareil, ou le gestionnaire de tâche pour entrer et restituer les demandes et réponses provenant du sous-système de livraison de service. Synonyme d'accès (paragraphe 3.1.61 de [3]). (Paragraphe 3.1.89, page 9 [3]).

Cible : Appareil SCSI qui reçoit une commande SCSI et la dirige sur une ou plusieurs unités logiques pour exécution (paragraphe 3.1.97, page 10 de [3]).

Tâche : Objet au sein de l'unité logique qui représente le travail associé à une commande ou un groupe de commandes qui y sont liées (paragraphe 3.1.98, page 10 de [3]).

Transaction : Interaction coopérative entre deux objets, impliquant l'échange d'informations ou l'exécution de certains services par un objet au nom de l'autre (paragraphe 3.1.109, page 10 de [3]).

11. Références

[1] RFC2026 S. Bradner, "Le processus de normalisation de l'Internet -- Révision 3", (BCP0009) octobre 1996. (*MàJ par* RFC3667, RFC3668, RFC3932, RFC3979, RFC3978, RFC5378, RFC6410)

[2] RFC2119 S. Bradner, "Mots clés à utiliser dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997.

[3] SAM-2, ANSI NCITS. Weber, Ralph O., editor. "SCSI Architecture Model -2 (SAM-2)". T10 Project 1157-D. rev 23, 16 mars 2002.

[4] SPC-2, ANSI NCITS. Weber, Ralph O., editor. "SCSI Primary Commands 2 (SPC-2)". T10 Project 1236-D. rev 20, 18 juillet 2001.

[5] CAM-3, ANSI NCITS. Dallas, William D., editor. "Information Technology - Common Access Method - 3 (CAM-3))". X3T10 Project 990D. rev 3, 16 mars 1998.

[6] 99-245r9, Hafner, Jim. "A Detailed Proposal for Access Controls". T10/99-245 revision 9, 26 avril 2000.

[7] SPI-X, ANSI NCITS. "SCSI Parallel Interface - X".

[8] FCP, ANSI NCITS. "SCSI-3 Fibre Channel Protocol" [ANSI X3.269:1996].

[9] FCP-2, ANSI NCITS. "SCSI-3 Fibre Channel Protocol - 2" [T10/1144-D].

[10] Paxon, V. "End-to-end internet packet dynamics", IEEE Transactions on Networking 7,3 (juin 1999) p. 277-292.

[11] Stone J., Partridge, C. "When the CRC et TCP checksum disagree", ACM Sigcomm (Sept. 2000).

[12] RFC2348 G. Malkin, A. Harkin, "Option TFTP Taille de bloc", mai 1998. *(D.S.)*

[13] RFC2914 S. Floyd, "Principes du contrôle d'encombrement", BCP 41, septembre 2000.

12. Remerciements

Un grand merci à Julian Satran, IBM et David Black, EMC, pour leurs commentaires sur leur relecture.

13. Adresse des auteurs

Adresser les commentaires à :

Marjorie Krueger
Hewlett-Packard Corporation
8000 Foothills Blvd
Roseville, CA 95747-5668, USA
téléphone : +1 916 785-2656
mél : marjorie_krueger@hp.com

Randy Haagens
Hewlett-Packard Corporation
8000 Foothills Blvd
Roseville, CA 95747-5668, USA
téléphone : +1 916 785-4578
mél : Randy_Haagens@hp.com

Costa Sapuntzakis
Stanford University
353 Serra Mall Dr #407
Stanford, CA 94305
téléphone : 650-723-2458
mél : csapuntz@stanford.edu

Mark Bakke
Cisco Systems, Inc.
6450 Wedgwood Road
Maple Grove, MN 55311
téléphone : +1 763 398-1054
mél : mbakke@cisco.com

14. Déclaration complète de droits de reproduction

Remerciement

Le financement de la fonction d'édition des RFC est actuellement fourni par l'Internet Society.