L'information dans les environnements virtuels distribués

Yannick Willame

10 juillet 2003

Abstract

Les environnements virtuels connaissent un intérêt certain ces deux dernières décennies et trouvent des applications dans une multitude de domaines. Les projets de recherche abondent, chacun apportant de nouvelles approches et méthodes afin de développer des environnements virtuels partagés supportant un nombre important de participants et, par conséquent, d'interactions. A travers quelques projets innovants, ce rapport présente les objectifs et les contraintes actuels et fait le point sur les moyens mis en œuvre pour filtrer et diffuser l'information entre les participants.

Introduction

L'utilisation d'environnement virtuel (EV) s'est rapidement répandue, parallèlement à l'évolution des «*Personal Computers*». L'augmentation continue de la puissance de calcul des ordinateurs, les cartes graphiques toujours plus performantes ainsi que la standardisation de librairies graphiques telles que OpenGL, Direct3D et VRML permettent à tout un chacun de s'immerger dans ces représentations en trois dimensions (3D) de mondes virtuels, tantôt imaginaires, tantôt reproductions de lieux existants, que sont les EVs. Le grand public peut en effet trouver toutes sortes d'applications virtuelles, que ce soit ludique, éducatif, culturel, scientifique,...comme par exemple les jeux de rôles, les simulateurs, les visites de musée.
Très vite, l'envie d'interagir non seulement avec ces EVs mais aussi avec d'autres participants au sein d'un même EV s'est faite ressentir. C'est donc dans cette continuité que sont apparus les environnements virtuels partagés (EVPs) dans lesquels se regroupent des personnes, créant ainsi des communautés virtuelles. Pour ce faire, chaque machine hôte utilise une réplication

de tout ou d'une partie de l'EVP et des objets qui le composent, ce qui implique des nouveaux problèmes de cohérence entre objet source et réplications. Par ailleurs, afin de permettre à un grand nombre de participants de se réunir au sein d'un même EVP, de s'y mouvoir, d'y manipuler des objets de manière concurrente ou non, il est indispensable de faire des choix selon le ou les objectifs poursuivis et de mettre en place des mécanismes assurant la cohérence de l'EV entre tous les participants et leurs interactions.
Après l'identification des objectifs, des exigences et des contraintes liés aux EV, ce rapport présente ce qui a déjà été mis en œuvre concernant les EVPs distribués à grande échelle (de l'ordre de centaines voire de milliers de participants). Nous verrons succinctement pourquoi distribuer et non centraliser. Ensuite, nous nous concentrerons sur les moyens existants pour réduire les besoins en bande passante permettant leur déploiement sur Internet. Les mécanismes de synchronisation et d'ordonnancement des événements qui sont également un point essentiel dans le développement d'EVPs ne seront pas abordés dans ce rapport.

1 Objectifs des EVPs

En toute généralité, l'objectif actuel des recherches dans le domaine des EVPs est double, les avancées technologiques et leur accessibilité offrent d'une part un haut degré d'interactivité et, d'autre part, la perspective d'un nombre de plus en plus important de participants. Reprenons l'abréviation EVGEs utilisée par E. Léty [1] pour désigner ces environnements virtuels partagés à grande échelle. Il est donc nécessaire que l'architecture de communication soit *scalable* comme il le rappelle, c'est-à-dire «*qu'elle doit supporter le déploiement à grande échelle sur l'Internet sans pour autant impliquer d'importantes modifications dans sa structure et dans les mécanismes qu'elle met en œuvre*».
Le deuxième objectif est la tolérance à l'hétérogénéité aussi bien au niveau des machines hôtes, des réseaux et des applications. Comme le souligne A. Felon dans [2], il faut éviter le nivellement par le bas de l'ensemble du système. La présence d'une machine hôte peu performante et ou à faible connectivité ne doit pas se faire ressentir par les autres participants mais doit être autant que possible tolérée au sein de l'architecture tout en circonscrivant la dégradation des performances.

2 Exigences et contraintes

Les applications qui se veulent interactives exigent un temps de réponse minime car chaque participant doit percevoir les événements du monde virtuel en temps réel. L'interaction peut se faire tant entre les participants qu'au niveau d'objets inclus dans l'EVP, de manière concurrente ou non. Chaque participant doit pouvoir interagir dans l'EVP, et ce en fonction des interactions des autres participants. Cependant, la communication et le traitement des informations ne sont hélas pas instantanés et impliquent donc un délai entre le moment où un participant interagit et le moment où les autres participants perçoivent cette interaction.

Concernant la communication à travers le réseau, le délai doit être borné sans quoi les utilisateurs ressentiront une gêne et l'immersion dans le monde virtuel ne sera pas probante. La limite conseillée par le standard DIS[1] est un délai maximum de 150 ms mais ce n'est qu'une recommandation et cette borne doit bien sûr être pensée en fonction des besoins de l'application et de l'interactivité souhaitée. Par ailleurs, la bande passante est également un élément limitatif important au niveau des réseaux car chaque participant est à la fois émetteur et récepteur d'informations et le passage à un grand nombre d'utilisateurs ne fait qu'accroître les besoins en termes de volume de communication.

Le traitement de l'information au niveau des machines hôtes est également un élément contraignant. D'une part, il s'agit de donner l'illusion de la réalité et de la fluidité. Par exemple, l'affichage des textures 3D doit se faire à un taux de rafraîchissement de l'ordre de 60 Hz pour garantir une perception de continuité des mouvements. Aujourd'hui, les cartes graphiques assurent très bien ce rôle. Mais il faut également traiter les différents flux d'informations provenant des autres participants le plus rapidement possible étant donné le côté interactif. La puissance de calcul des machines hôtes doit donc bien être prise en compte, d'autant plus que le nombre de participants et donc le volume d'informations reçues est élevé.

En résumé, le défi est de développer des architectures de communication pour les EVPs permettant un haut degré d'interactivité entre un grand nombre d'utilisateurs, et ce en prenant en considération les contraintes liées à la transmission et au traitement de l'information.

[1] Distributed Interactive Simulation, projet lancé par le *Department of Defense* américain en 1989 et devenu standard IEEE en 1995

3 La distribution des données et leur diffusion

3.1 L'architecture et la couche réseau

Les EVPs commerciaux actuels sont en majorité conçus pour un petit nombre de participants (quelques dizaines tout au plus) et sont basés sur une architecture centralisée de type client serveur. Sans entrer dans les détails, nous pouvons facilement nous rendre compte des limites de ce type d'architecture. En effet, les utilisateurs désirant entrer dans l'EVP doivent se connecter au serveur et lorsqu'un participant agit dans l'EVP, l'information est envoyée au serveur qui centralise les données envoyées par tous les utilisateurs. Ensuite, le serveur redistribue les informations reçues à tous les autres participants connectés. La connexion est donc de type point-à-point (*unicast*) entre le serveur et chacun des participants. Le serveur renverra donc la même information autant de fois que le nombre de connectés. C'est de par ce caractère centralisé que le nombre de participants est fortement limité car l'expansion à un nombre arbitrairement grand d'utilisateurs entraînerait un goulot d'étranglement sur ce nœud du réseau qu'est le serveur de par la convergence des informations de chaque utilisateur. De plus, le traitement centralisé de ces informations et leur redistribution nécessitent des capacités de calcul importantes.
Il est donc primordial d'éviter l'utilisation d'une architecture centralisée au profit d'une architecture distribuée dans laquelle chaque participant se doit d'envoyer les informations à tous les autres. L'utilisation du point-à-point entre chacun des participants obligerait l'envoi d'autant de paquets d'information que d'utilisateurs présents dans l'EVGE et n'apporte pas d'amélioration notable par rapport à une architecture centralisée. Une alternative serait le *broadcast* où tout le monde reçoit les informations de tout le monde, mais ce n'est guère possible sur un réseau tel que l'Internet. C'est pourquoi les recherches s'orientent vers le multipoint (*multicast*) qui peut être vu comme un sous-ensemble du *broadcast* et qui permet une communication de type un-vers-plusieurs et plusieurs-vers-plusieurs, ce qui correspond mieux au caractère à la fois émetteur et récepteur de chacune des machines hôtes. Ce service utilise la notion de groupe pour désigner un ensemble d'utilisateurs désirant communiquer entre eux et une adresse IP multipoint est associée à chaque groupe et l'identifie. Afin d'être membre d'un groupe, une procédure d'abonnement est préalable et le départ d'un groupe suppose donc le désabonnement. La mise en place d'un protocole multipoint plus adapté aux problèmes de mise à l'échelle que le MBONE[2] est en cours, notamment le

[2]Multicast Backbone, réseau multipoint virtuel de l'Internet

Simple Multicast. Nous n'en parlerons pas dans ce rapport mais le lecteur intéressé peut se référer à [3].
La solution la plus prometteuse pour l'utilisation à grande échelle des EVPs est par conséquent une architecture distribuée avec une communication multipoint.

3.2 Les services de transport

Il est nécessaire de présenter brièvement les services offerts par les protocoles de transport avant d'aborder le problème du volume des communications. Mettre en œuvre un protocole multipoint fiable tel que TCP pour le point-à-point n'est pas possible pour des groupes comprenant plusieurs centaines de membres. Garantir la réception correcte d'un message nécessite des mécanismes d'acquittement et de retransmission qui introduisent des délais supplémentaires, délais qui ne sont déjà pas supportables pour les EVPs basés sur une architecture centralisée car ils pénalisent fortement les performances. Au sein d'un groupe important, il est impensable qu'un hôte doive maintenir des états de contrôle pour chacun des autres membres du groupe. La version multipoint du protocole UDP offre le même service que la version point-à-point à savoir une transmission non fiable et sans garantie de l'ordre de réception et elle n'apporte aucune vision des groupes. La pluralité de la nature des échanges d'informations induite par le caractère interactif (manipulations d'objets, structures, mouvements, flux audio et vidéo,...) amène à déterminer les besoins en terme de délai, de bande passante et de fiabilité de transmission propres à chaque type de données. Ainsi, certaines données peuvent se contenter des ressources disponibles à un moment donné alors que d'autres sont critiques et demandent des délais bornés et ou une transmission fiable. Communément, l'identification des besoins permet d'élaborer un protocole au-dessus d'UDP multipoint ou d'IP multipoint afin de définir des diffusions plus raffinées, avec un compromis entre fiabilité et latence. Il existe actuellement plusieurs propositions de protocoles multipoint offrant un certain degré de fiabilité [4]. Celles-ci sont cependant souvent destinées à un usage restreint ou menacées par des risques potentiels d'implosions des messages d'acquittement. D'autres solutions telle que XTP[3] [5] proposent une approche centralisée afin de faire face au problème de retransmission ou, à l'inverse, RBP[4] [6] utilise une répartition des tâches de retransmission. Plus particulièrement dans SRM[5] [7], un numéro de séquence est associé aux

[3]Xpress Transport Protocol
[4]Reliable Broadcast Protocol
[5]Scalable Reliable Multicast

informations transmises permettant de détecter une perte et dans ce cas, la retransmission sera effectuée par le plus proche voisin. Enfin, une dernière approche consiste à jouer sur la probabilité de perte en retransmettant plusieurs fois la même information.
Quoiqu'il en soit, le problème de la retransmission est toujours d'actualité et une alternative est de transmettre l'information de manière à accroître la tolérance du système aux pertes. Par exemple, la mise à jour de la position d'un participant se fera par la transmission de sa position relative à une origine définie au sein de l'EVP plutôt que par rapport à sa position précédente.

4 Du monde global au monde personnel

Outre l'aspect du transport des informations à travers le réseau proprement dit, le filtrage de celles-ci est requis dans le but de diminuer les besoins en volume de communication pour la mise à l'échelle. En effet, une réplication de la totalité de l'EVP chez chacun des hôtes n'est pas possible car la quantité d'informations utiles pour maintenir l'EVP cohérent serait trop importante. De plus, un participant n'interagit jamais avec l'ensemble de l'EVP et de ses membres mais seulement avec un sous-ensemble. Seule une mise à jour de ce sous-ensemble est donc réellement requise et l'espace virtuel peut être partitionner de manière statique ou dynamique.

4.1 Le partitionnement statique

L'EVP peut être découpé en cellules (*Areas of Interest*) définies une fois pour toute selon des critères d'ordre spatial, temporel ou encore fonctionnel. A chacune des cellules est associée une adresse multipoint identifiant le sous-ensemble des membres de cette cellule. Par conséquent, les cellules ne se chevauchent pas et une entité de l'EVP n'émet généralement des informations que dans une seule cellule à un instant donné. Par contre, elle reçoit les informations des cellules qui intersectent sa zone d'intérêt afin de connaître les autres entités avec lesquelles elle peut éventuellement interagir. Par conséquent, tout ce qui se trouve en dehors de cette zone de proximité ne peut pas interagir car l'entité n'en a même pas connaissance. Cette solution repose sur le caractère local des interactions et la métaphore du monde réel (pièces d'un bâtiment). Le pionnier de cette méthode est NPSNET[6] [8] où la découpe se fait à l'aide d'un motif hexagonal. La répartition des participants

[6] Groupe de recherche de la Naval Postgraduate School

à travers les cellules se fait par l'intermédiaire d'un Gestionnaire des Centres d'Intérêts et le passage d'une cellule à une autre nécessite des primitives d'abonnement et de désabonnement. La référence en EVGE, à savoir DIS[7], a profité de cet apport de NPSNET et est passé d'une réplication globale au découpage en cellules. Il a inclut la notion de *Group Leader*, rôle assuré par le doyen du groupe et chargé d'informer les nouveaux venus de l'état présent. Le concept d'*Areas of Interest* est également repris dans VREng[8] [2]. Ici, le participant est présent dans une unique cellule et le choix de celle-ci, c'est-à-dire du groupe multipoint, est complètement déterminé par la position spatiale dans l'EVP. L'entrée dans une cellule (i.e. l'inscription à un groupe de diffusion) est vu comme une collision entre l'avatar [9] et l'*Area of Interest*. A. Felon souligne que le caractère statique des cellules ne donne pas lieu à une véritable réplication de celles-ci et ne nécessite aucun échange d'information une fois les cellules définies dans la configuration du monde virtuel.
Un autre concept proche des cellules est celui de zone appelée *Locale* introduit dans HIVE [9], [10]. Selon le même principe, l'EVP est découpé de manière statique en *Locale* possédant chacune leur propre système de coordonnées où la communication entre les participants se fait en multipoint. La relation entre les *Locales* est définie localement par les frontières communes et le monde virtuel est structuré ainsi, par les liens entre les *Locales*. Un participant choisit la *Locale* qu'il veut rejoindre selon ses intérêts. Cette méthode est implémentée dans Diamond Park, basé sur une architecture multiserveurs appelée SPLINE [10], où l'utilisateur obtient les informations sur le système de coordonnées depuis ces serveurs. Dans HIVE, il existe plusieurs stratégies permettant d'offrir un déplacement continu entre les *Locales*. Par exemple, une machine hôte recevra les informations et répliquera toutes les *Locales* à moins de N pas de sa *Locale* courante (*N Step Replication*) ou bien les N *Locales* les plus proche (*N Nearest Replication*). La valeur de N peut être adaptée selon les capacités de calcul et la connectivité.
Le partitionnement statique permet d'améliorer la mise à l'échelle mais repose sur l'hypothèse d'une répartition uniforme des participants à travers les différentes cellules ou zones. De plus, cette répartition est généralement supposée être stable car le passage d'une cellule à une autre à un coût en

[7]Distributed Interactive Simulation, projet lancé par le *Department of Defense* américain en 1989 et devenu standard IEEE en 1995

[8]Virtual Reality Engine, projet de l'Ecole Nationale Supérieure des Télécommunications datant de 1997

[9]Représentation d'un participant au sein de l'EVP

[10]Scalable Platform for Large Interactive Networks Environments

terme de gestion des groupes multipoints. Le paramètre critique de cette approche est la taille des cellules car une taille trop petite accentue les coûts liés à l'abonnement et au désabonnement des groupes multipoints. A l'inverse, une taille trop importante introduit à nouveau les problèmes de surinformation des participants.

4.2 Le partitionnement dynamique

Pour s'affranchir des hypothèses d'uniformité et de stabilité de la répartition des participants, Benford et al. [11] proposent un modèle spatial d'interaction reposant sur le concept clé d'aura. A l'image de la perception du monde par nos sens, l'interaction avec l'EVP et ses participants est limitée par la distance spatiale qui nous en sépare. Nous pouvons ainsi définir, pour chaque média (comme par exemple l'ouïe et la vue), un sous-espace dans lequel nous avons une perception plus ou moins précise d'événements et de la présence d'entités. Les auteurs parlent d'un niveau de conscience mutuelle entre deux entités et propre à chaque média selon le *nimbus* et le *focus*

- Le *nimbus* représente le champ d'émission d'information. Plus A est dans le *nimbus* de B, plus A a conscience de B.
- Le *focus* modélise le champ de réception d'information. Plus A est dans le *focus* de B, plus B a conscience de A
- Le *niveau de conscience* qu'un objet A a d'un objet B sur un média M est une fonction du focus de A lié à M et du nimbus de B lié à M.

Par conséquent, les auras sont attachées à leur entité et seule l'intersection de celles-ci suite à une collision permet une interaction plus ou moins forte selon le degré de conscience mutuelle. Au-delà de l'aura, le participant n'a aucune conscience de ce qui existe. Quatre moyens de manipuler la portée des auras sont suggérés. Le premier est implicite de par son attachement à l'entité et la portée suit les mouvements et la direction du participant. L'aura peut également être concentrée sur un point précis comme nous le faisons lorsque nous lisons un panneau. Des moyens d'amplification ou d'atténuation (*adapteur*) peuvent être utilisés tels qu'un microphone ou des jumelles. Enfin, les frontières de l'environnement (murs,...) influencent aussi la perception. La métaphore au comportement social de l'individu peut être encore plus forte si on réduit dynamiquement la portée de l'aura lorsque l'EVP devient de plus en plus peuplé (Dans un endroit vide, notre attention porte plus loin que dans un endroit surpeuplé où nous communiquons uniquement avec les gens les plus proches). Cette réduction dynamique favorise le passage à l'échelle.

Benford et al. [11] présente deux implémentations du modèle : DIVE[11] qui est le premier a avoir utilisé le concept d'aura et MASSIVE[12]. DIVE utilise une arborescence d'objets selon la relation composant/composé, l'entité étant l'unité élémentaire de distribution. Le partitionnement et la réplication consiste à ne considérer qu'un sous arbre (i.e. une partie du monde virtuel) auquel est attaché une adresse multipoint, ce qui le différencie d'un partitionnement spatial classique en faveur d'un partitionnement lié à l'organisation des objets. Une caractéristique particulière est la manipulation locale d'objets et de la hiérarchie de ceux-ci avant la diffusion dans l'EVP.

Plutôt que de parler de MASSIVE en tant que tel, nous allons voir comment fonctionne la détection des collisions des auras à travers cet exemple ainsi que son successeur CVE[13]. Nous retombons dans la nécessité d'inclure des serveurs dont le rôle est de détecter les collisions. Étant donné qu'un aura est spécifique à un média particulier, il en est de même pour les détecteurs de collisions. Par ailleurs, cette duplication en fonction des médias permet un meilleur passage à l'échelle car le risque de goulot d'étranglement revient de par la nécessité de passer par ces détecteurs. Remarquons que ceux-ci peuvent être instanciés dynamiquement si aucun n'est présent pour le média utilisé. Dans MASSIVE (voir figure 1), les participants communiquent en fait par des connexions point-à-point et des groupes multipoints sont utilisés au niveau des détecteurs de collision. Lorsqu'un participant entre dans le système, il se connecte à un *master trader* afin de connaître les adresses multipoints des détecteurs dont il a besoin (i.e. selon les types de médias voulus). Le rôle du détecteur est donc de surveiller les auras des entités dont il a la connaissance et lorsqu'une collision a lieu, il communique à chacune des parties concernées l'adresse IP permettant d'établir une connexion point-à-point entre les deux entités. A partir de ce moment, l'interaction est possible via le média commun. Notons que le niveau de conscience mutuel est de la responsabilité des deux entités et non du détecteur de collisions. Cette approche répond bien au problème de l'hétérogénéité car, selon sa connectivité et sa capacité de calcul, un participant pourra choisir un ou plusieurs médias déterminé (terminal texte, flux audio et ou flux vidéo).
Dans Community Place[14], les détecteurs sont répliqués et organisés selon une hiérarchie primaire/secondaire afin de repousser les limites de passage à

[11] Distributed Interactive Virtual Environment de la Swedish Institute of Computer Sciences - 1990

[12] Model Architecture and System for Spatial Interaction in Virtual Environment

[13] CRG (Communication Research Group) Virtual Environment de l'University of Nottingham - 1994

[14] Virtual Society Project of Sony Computer Sciences Lab - 1995

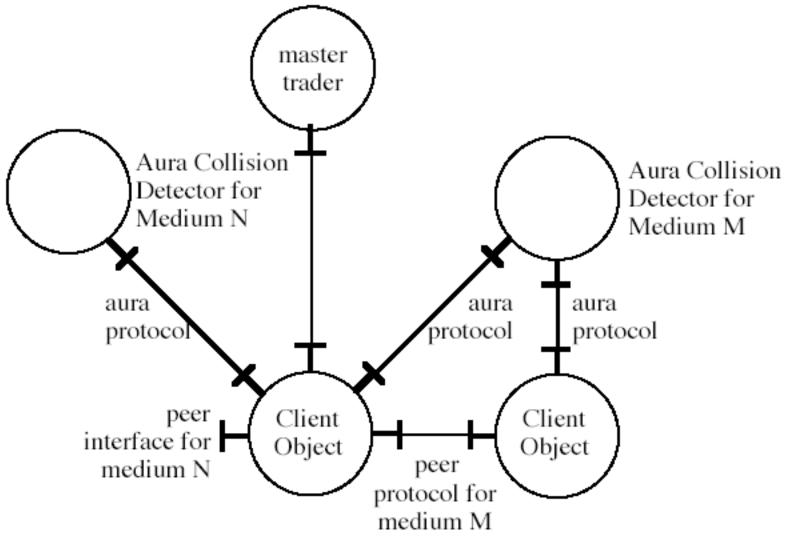

Figure 1: Architecture de la détection des collisions avec MASSIVE

l'échelle. Les détecteurs primaires sont responsables des collisions entre auras de différents détecteurs secondaires, ceux-ci communiquent donc périodiquement des informations aux primaires.
Le concept d'aura apporte un point de vue différent du filtrage des informations, celui-ci se fait donc en fonction de l'entité et non plus uniquement en fonction de la configuration du monde virtuel. Ce type de partitionnement dynamique permet également une meilleur hétérogénéité mais son point faible est la nécessité de détecteurs de collisions centralisés et des gestionnaires de ceux-ci.

4.3 Un partitionnement hybride

Après avoir identifier les problèmes du découpage statique en cellules, E. Léty décrit dans sa thèse [1] une architecture appelée SCORE qui intègre une approche à la fois statique et dynamique du partitionnement de l'EVP. Le monde virtuel est donc préalablement découpé (voir figure 2) en fonction de ses caractéristiques intrinsèques (étages, pièces, murs,...). Cette première découpe définit des zones de grande taille nommée *Start-Zones* et une adresse multipoint définitive est associée à chacune d'entre elles. Dans le but de facilité le partitionnement dynamique en *Zones*, les *Start-Zones* sont elles-mêmes redécoupées en *Zone-Units* qui, comme son nom l'indique, seront l'unité élémentaire des *Zones*. Une *Zone* (voir figure 3) est par conséquent une partie de l'EVP dans laquelle les cellules (représentant un groupe multipoint d'intérêt commun) auront toutes la même taille. Comme nous le verrons par la suite, cette découpe, à première vue complexe, permet un filtrage de l'information des données de par l'appartenance à un groupe multipoint au niveau des cellules, mais également un filtrage de l'information de signalisation grâce au concept plus général des *Start-Zones*.

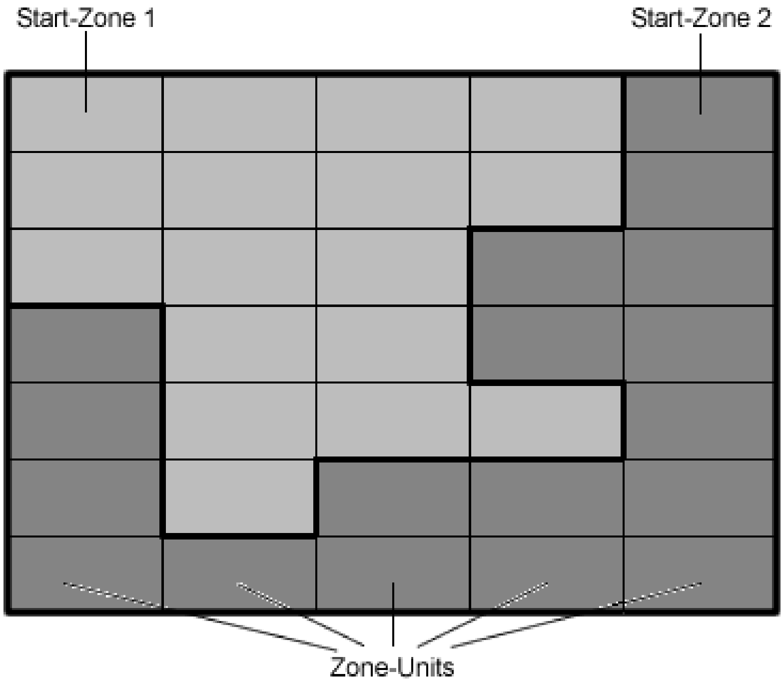

Figure 2: Découpage statique en *Start-Zones* et en *Zone-Units* selon l'architecture SCORE

A des endroits stratégiques du réseau (proximité des utilisateurs potentiels en terme de délai) sont placés des agents[15] dont le rôle est la définition dynamique des *Zones* et de la taille des cellules en fonction de la répartition actuelle des participants. Pour ce faire, les agents sont connectés en point-à-point avec un ensemble de participants via UDP et sont responsables d'éva-

[15] Le terme agent est utilisé plutôt que serveur car leur but n'est aucunement de redistribuer l'information ni de calculer l'état global de l'EVP.

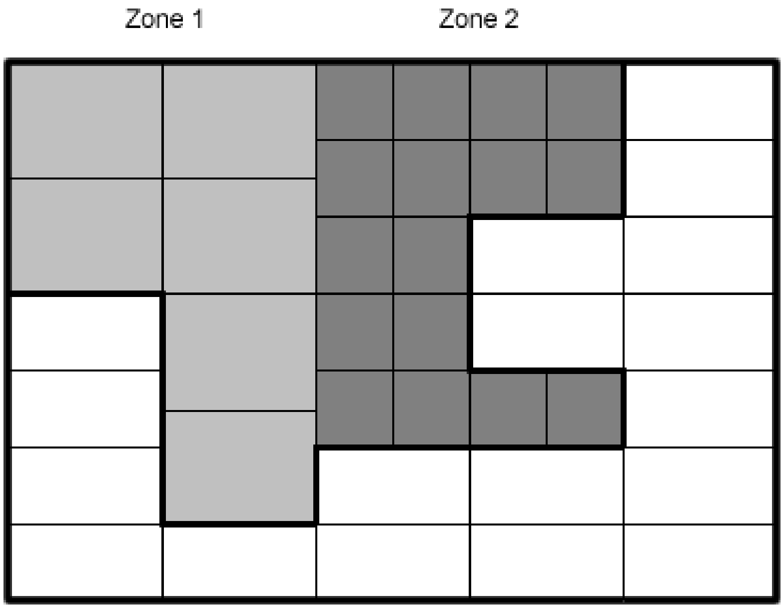

Figure 3: Exemple d'un découpage dynamique possible en *Zones* et en cellules dans l'architecture SCORE

luer périodiquement la satisfaction[16] des participants et d'en assurer l'équité max-min [17] en recalculant les *Zones* et en modifiant la taille des cellules si nécessaire. Le résultat de cette opération d'ajustement dynamique (informations de mise en correspondance) est communiqué aux participants via le groupe multipoint qu'est la *Start-Zone*. Notons que seuls les agents émettent dans ce groupe de signalisation, les participants étant de simple récepteur, et qu'un groupe (i.e. une adresse multipoint) réservé aux agents leur permet d'échanger de l'information afin de calculer la densité des participants dans les différentes *Zone-Units* par exemple.
Au niveau des participants, ceux-ci communiquent entre eux par l'adresse multipoint au niveau des cellules et dont ils ont la connaissance par les informations de mise en correspondance. Un participant ne transmet que dans un groupe (la cellule où il se trouve) mais reçoit les données des cellules limitrophes. Comme mentionné au niveau des agents, chaque participant utilise

[16]métrique en fonction de l'information utile reçue et traitée, du débit total de l'information utile et du débit maximal supporté par le participant

[17]De l'anglais *max-min fairness* et qui consiste à maximiser la satisfaction de l'utilisateur ayant la plus petite satisfaction

une connexion UDP avec un agent unique, à savoir le plus proche en délai, pour communiquer périodiquement (et aussi lorsqu'il change de *Zone-Unit*) diverses informations permettant à l'agent de calculer la satisfaction et la répartition des participants.
Certes, l'architecture de communication SCORE introduit un coût en terme de trafic de signalisation et de messages de contrôle mais celui-ci semble être bien circonscrit au stricte nécessaire. En outre, elle permet de mieux approximer les zones d'intérêt des participants lorsque leur densité est élevée, réduisant ainsi le trafic superflu et le problème de surinformation. Cette architecture tient également compte l'hétérogénéité de capacité des participants de par son partitionnement dynamique selon une stratégie d'équité max-min.

5 Conclusion

L'objectif actuel des environnements virtuels partagés est leur déploiement à grande échelle à travers un réseau tel que l'Internet. Le choix de ce réseau n'est pas anodin, il s'agit de développer des architectures qui pourront être utilisées pour des applications destinées au grand public. Cependant, il apporte ses contraintes en terme de délai et de volume de communication. En effet, les flux d'informations liés à l'environnement virtuel partagé doivent cohabiter avec ceux des applications que nous connaissons actuellement.
Afin d'y arriver, nous avons vu qu'une architecture distribuée et une utilisation des groupes multipoints s'imposent pour éviter l'apparition de goulots d'étranglement sur le réseau. La réduction du volume d'informations émises, reçues et traitées par chaque participant est également apparue comme inévitable pour le déploiement à grande échelle. Pour ce faire, les recherches se sont orientées vers des méthodes de partitionnement du monde virtuel. Ce découpage peut être établi une fois pour toute, selon les caractéristiques du monde virtuel ou alors de manière tout à fait arbitraire. Cependant, cette solution ne tient pas compte des phénomènes de regroupement des participants et n'était donc pas satisfaisante. Une autre idée était de modéliser le comportement des individus en associant un espace de conscience propre à chaque participant : l'aura. La complexité de cette approche rend malheureusement son implémentation difficile et nécessite à nouveau la présence de serveurs. La dernière architecture abordée, SCORE, propose un découpage statique préalable qui sert de guide au partitionnement dynamique effectué en fonction de la répartition des participants en cours de session.
Il y a donc encore beaucoup de possibilités dans le développement des envi-

ronnements virtuels partagés, ce qui en fait un domaine de recherche ouvert. La solution unique et idéale n'existe certainement pas mais la combinaison d'une architecture distribuée, de la communication point-à-point et multi-point ainsi que les techniques de filtrage par partitionnement semble être à la base des solutions les plus prometteuses.

References

[1] Léty, E. : Une architecture de communication pour environnements virtuels distribués à grande-échelle sur l'Internet. PhD thesis, Université de Nice-Sophia-Antipolis (2000)

[2] Felon, A. : Diffusion, cohérence et contraintes temporelles dans les mondes virtuels 3d système vreng. Master's thesis, Université de Paris VI / CNAM / ENST (1999)

[3] Perlman, C., Lee, C.Y., Ballardie, A., Crowcroft, J., Wang, Z., Maufer, T. : Simple multicast : A design for simple, low-overhead multicast. Internet-Draft (1998)

[4] Macedonia, M.R., Zyda, M.J. : A taxonomy for networked virtual environments. Technical report, Fraunhofer Center for Research in Computer Graphics and Computer Science Department Naval Postgraduate School (1996)

[5] Strayer, W., Dempsey, B., Weaver, A. : XTP : The Xpress Transfer Protocol. Addison-Wesley (1992)

[6] Chang, J.M., Maxemchuck, N. : Reliable broadcast protocols. ACM Transactions on Computer Systems (1984)

[7] Westerlund, A., Hagsand, O., Grönvall, B. : Wader - wide-area adaptative distribution environment - an open srm implementation. Technical report, Swedish Institute of Computer Sciences (1997)

[8] Macedonia, M.R., Zyda, M.J., Pratt, D.R., Barham, P.T., Zeswitz, S. : Npsnet : A network software architecture for large scale virtual environments. MIT Presence (1994)

[9] Greenhalgh, C. : Hivek : the hive project distributed vr runtime kernel. (1996)

[10] Purbrick, J. : Hive kernel locale facilities. (1997)

[11] Benford, S., Bowers, J., L.E., F., C., G. : Managing mutual awareness in collaborative virtual environments. (1994)