

Expérimentation ATM du loria ; pourquoi ATM ?

G.Feltin, A.Valdenaire, T.Dautcourt

Les besoins actuels

La puissance toujours croissante des stations de travail, les applications spécifiques, tel le calcul parallèle, demandant de plus en plus de performance, le passage du texte à l'image, l'intégration de la voix et de la vidéo demandent des **débits importants** et des **délais de transfert** des équipements réseaux extrêmement **courts (latence)**.

Les besoins immédiats

Les besoins croissants de moyen de calcul performant font que les chercheurs se sont rapidement tournés vers des solutions exploitant les puissances cumulées des stations de travail.

Dans ce but, ils utilisent des bibliothèques de passage de message comme PVM (*Parallel Virtual Machine*) et MPI (*Message Passing Interface*).

Dans un premier temps, l'utilisation du réseau local Ethernet a permis de valider les outils. Mais l'utilisation d'Ethernet avec son temps de latence élevé et non garanti ne permettait pas d'obtenir des résultats constants. D'autre part, le déséquilibre entre la puissance cpu et le débit des entrées-sorties est trop important dans le cas d'Ethernet.

Une solution haut-débit était souhaitée.

Le choix d'ATM

Plusieurs technologies sont en concurrence : FDDI (100 Mb/s), Fast-Ethernet (100 Mb/s), 100VG-AnyLan (100 Mb/s), Ethernet Commuté, ATM (100 Mb/s, 155 Mb/s,...), Fibre Channel...

ATM (*Asynchronous Transfer Mode*) apparaît comme une solution aussi bien en ce qui concerne les réseaux locaux que les réseaux étendus, et offre des classes de service. Il résout les problèmes d'augmentation de bande passante avec des temps de latence courts, et aussi ceux du multimédia avec des applications synchrones incluant voix, données et images.

La commutation de cellules présente l'avantage de temps de latence très faible et constant, et la garantie de la bande passante.

Le choix du matériel

Le choix s'est porté sur du matériel de la société FORE pour essentiellement deux raisons :

- le switch FORE est celui qui, au premier trimestre 95, offrait le meilleur temps de latence et la meilleure capacité à gérer des flux importants (réf. TELECOM RESEAUX Mars 1995)

- FORE fournit une bibliothèque qui permet de développer des applications directement sur les couches AAL3/4 et AAL5.

ATM en 10 μ s

La commutation de cellules

ATM est basé sur la commutation de petits paquets de taille fixe, nommés cellules. Il est orienté connexion, et nécessite l'établissement de circuits virtuels avant la transmission des données.

Une cellule est composée de 53 octets, dont 5 octets pour l'en-tête.

Le modèle de référence

Higher Layers
ATM Adaptation Layer (AAL) Adaptation des données au format des cellules
ATM Layer Commutation et multiplexage des cellules
Physical Layer Adaptation des cellules au transport physique

Les AAL

Elles sont chargées du découpage des données en cellules et du réassemblage.

Il en existe actuellement 4 :

- AAL1 : classe A , voix (service en mode connecté - synchronisation entre la source et la destination - débit constant)
- AAL2 : classe B, vidéo (service en mode connecté - synchronisation entre la source et la destination - débit variable)
- AAL3/4 : classe C/D, emploi limité (service en mode connecté classe C ou non connecté classe D - asynchrone - débit variable - réservé au trafic qui ne tolère pas les pertes de cellules)
- AAL5 : classe D, données (service en mode non connecté - asynchrone - débit variable)

Chaque AAL est divisée en deux parties :

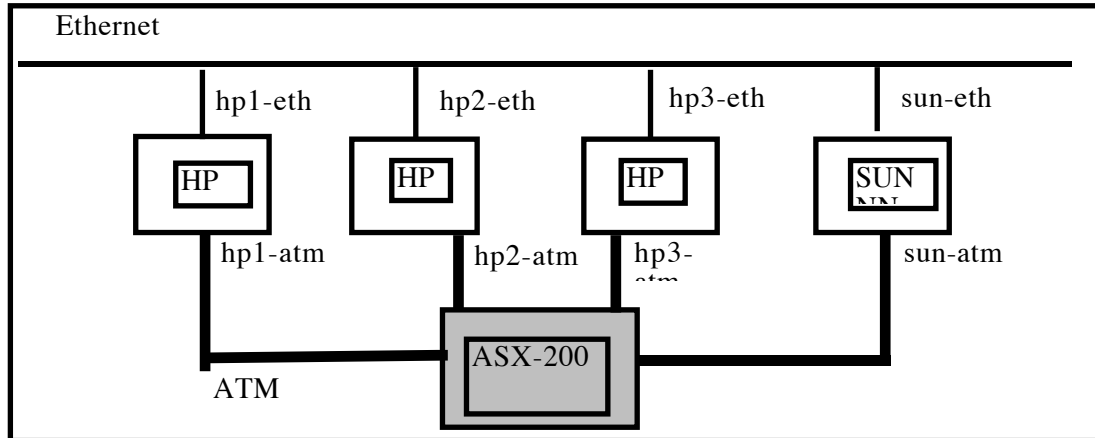
- une sous couche de segmentation et de réassemblage (SAR)
- une sous-couche de convergence (CS), chargée de gérer les échanges de données.

L'expérimentation ATM

Le Cadre

L'expérimentation ATM s'est réalisée au CRIN-CNRS/INRIA-Lorraine sur un réseau ATM à 100 Mb/s, composé d'un switch ATM FORE ASX-200, prêté par la société SILICOMP, de 3 stations HP (un 755 et deux 715) et d'une station SUN (Sparestation 10), connectées au switch à 100 Mb/s, en fibre optique par des cartes d'interface ATM FORE HPA-200 pour les HP et SBA-200 pour le SUN. Le réseau ATM ainsi constitué n'est pas isolé : chaque

station est restée connectée au réseau local Ethernet, pour bénéficier des services réseaux utilisant le mode broadcast (NIS) qui ne fonctionnait pas directement sur ATM. Les tests ont eu lieu au mois de mars durant une semaine.



La grappe ATM

La plate-forme ATM

L'intégration au réseau local existant

Le premier but est de valider la transparence des protocoles TCP/IP sur ATM. L'installation s'est faite rapidement avec reconfiguration des noyaux et installation des «drivers» correspondants.

FORE fournit le logiciel nécessaire pour véhiculer des datagrammes IP sur un réseau ATM, défini par le RFC 1483 (*Request For Comment*).

Pour valider les connexions physiques et logiques, l'utilitaire Unix ping est utilisé. L'utilitaire ftp, utilisé pour l'échange de fichiers, a été validé. Le driver supporte l'interface des sockets Berkeley, et permet donc d'utiliser NFS sans changer les applications.

Les tests se sont déroulés sans difficulté particulière.

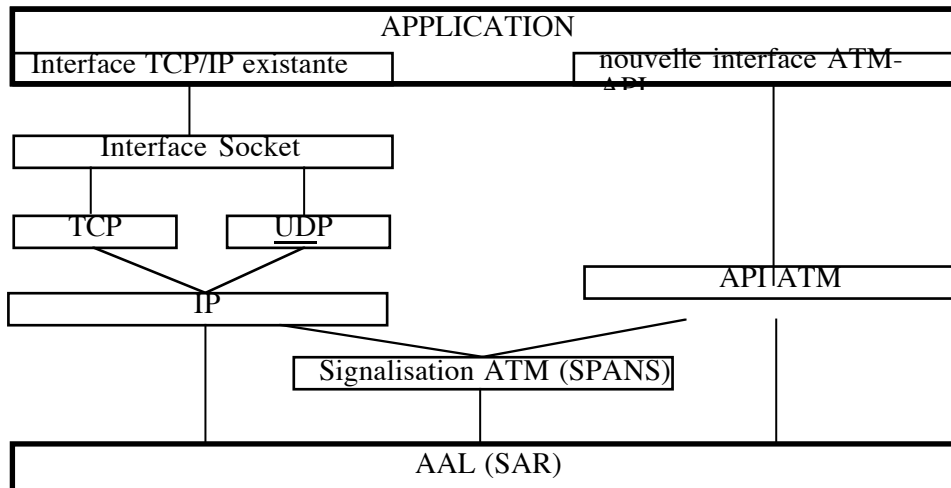
Les performances sur IP

Les tests sont réalisés au-dessus des protocoles TCP-UDP/IP.

Netperf 2.0, développé par Hewlett-Packard Company, est l'outil choisi pour tester les performances réseau (<http://www.cup.hp.com/netperf/NetperfPage.html>).

En faisant varier la taille des buffers d'émission et de réception des sockets et la taille des messages, les meilleures performances pour TCP sont de l'ordre de 65 Mb/s, entre un HP 755 et un HP 715, non isolés. Il est à noter que ces résultats dépendent de la puissance des machines, et que les machines de la plate-forme d'expérimentation ne sont pas les plus puissantes du réseau. Les débits sont un peu plus faibles entre deux HP 715.

Des tests ont été également réalisés sur PVM et MPI dans le cadre du calcul parallèle. Ils ont donné lieu à une publication (*Homogenous and Heterogenous Networks of Workstations : message Passing Overhead* - E.Dillon, C. Gamboa Dos Santos, J. Guyard, <http://www.cse.nd.edu/mpidc95/proceedings/papers/html/dillon>)



Architecture des couches

Les performances sur ATM

Les tests sont réalisés directement sur ATM, en utilisant l'API 2.3 (*Application Programming Interface*).

L'outil utilisé est également Netperf adapté en standard à l'API FORE.

Les performances enregistrées sont de l'ordre de 80 Mb/s.

Conclusion

Tant au niveau transparence que performance, cette expérimentation a démontré la possibilité immédiate de constituer un réseau haut débit ATM entre machines spécifiques, tout particulièrement dans le cadre du calcul parallèle. Un montage stable d'une plateforme ATM à 155 Mb/s, sur paires torsadées (câble ATT 1061 de catégorie 5), entre trois stations HP, dans le cadre du Centre Charles Hermite (*Centre Lorrain de Compétence en Modélisation et Calcul à Hautes Performances*, <http://www.loria.fr/CCH>), est opérationnelle. Bien que les prix des commutateurs et des cartes ATM restent élevés, cette solution est évolutive et permet d'utiliser la puissance de calcul local répartie pour un coût acceptable. Cette réalisation montre que l'utilisation en réseau local est viable, et nous a permis d'acquérir une compétence dans ce domaine.

Cependant actuellement, le tout ATM n'est pas encore envisageable, car techniquement tous les problèmes ne sont pas résolus. Citons les principales difficultés :

- l'inter-opérabilité entre constructeurs, due à la phase d'implémentation de la signalisation dynamique et la normalisation en cours du contrôle de flux
- la gestion du réseau
- la normalisation du multimédia.