

RESEAUX A HAUT DEBIT : EVOLUTION TECHNOLOGIQUE

Auteur : Thierry USO

Version : 1.4

Date : 18 Aout 2001

Table des matières

CHAPITRE 1	INTRODUCTION	1-1
CHAPITRE 2	EVOLUTION TECHNOLOGIQUE	2-1
2.1	CONNECTIQUE	2-1
2.2	PROCESSEUR ET MEMOIRE	2-4
CHAPITRE 3	IMPACT SUR LE RESEAU	3-1
3.1	LATENCE VS BANDE PASSANTE	3-1
3.2	COMMUTATION VS MEDIUM PARTAGE	3-2
ANNEXE A	LATENCE D'UN RESEAU	A-1
ANNEXE B	BIBLIOGRAPHIE	B-1
GLOSSAIRE		Gloss.-1
TABLEAUX		
2-1	Evolution du coût de la transmission sur fibre optique	2-2
2-2	Comparaison du coût de la fibre optique et de la paire torsadée dans un LAN	2-2
2-3	Blindage des câbles en paires torsadées	2-2
2-4	Classification des câbles en paires torsadées par l'EIA/TIA	2-3
2-5	Spécification de câbles en paires torsadées	2-3
2-6	Caractéristiques de quelques processeurs	2-5
2-7	Hiérarchie de la mémoire	2-6
3-1	Analyse de la latence	3-1

Chapitre 1

INTRODUCTION

Ce document est le premier d'une série consacrée aux réseaux à haut débit.

Le chapitre 2 présente l'évolution technologique des 15 dernières années (connectique, processeur et mémoire). Le chapitre 3 montre l'impact de cette évolution technologique sur le réseau (problème de la latence, introduction des commutateurs dans les LANs).

Un glossaire à la fin du document explicite les nombreux acronymes utilisés.

Chapitre 2

EVOLUTION TECHNOLOGIQUE

2.1 CONNECTIQUE

La performance d'un médium de transmission s'exprime non seulement par sa capacité à transmettre (exprimée en Mbit/s.km) mais aussi par son taux d'erreur bit. Dans ces 2 domaines, les progrès réalisés durant les 15 dernières années ont été considérables et ont permis l'effondrement du ratio prix/performance de média de transmission tels que la fibre optique et la paire torsadée.

La fibre optique s'est révélée être le médium de transmission idéal pour les opérateurs Telecoms en raison de son coût (tableau 2-1) et de ses propriétés :

- bande passante forte
Les technologies les plus récentes permettent des capacités de plusieurs Tbit/s.
- atténuation faible
L'atténuation due aux phénomènes d'absorption et de diffusion de la lumière dans la matière a une valeur de quelques dB/km pour les fibres à base de silicium.
- taux d'erreur faible
Le taux d'erreur bit d'une liaison numérique de type Transfix est de 10^{-8} ce qui est 1000 fois plus faible qu'une liaison analogique comparable.
- absence d'interférence électromagnétique et radio
- faible encombrement

Tableau 2-1: Evolution du coût de la transmission sur fibre optique

Année	Technologie	Gbit/s.km	Coût relatif par Mbit/s.km
1975	0.8 um, MMF	1	2500
1980	1.3 um, SMF	30	300
1985	1.55 um, SMF	400	40
1990	1.55 um, SMF	4000	4

En conséquence, tous les opérateurs Telecoms se sont lancés dans un processus de remplacement de leurs lignes analogiques à base de cuivre par des lignes numériques à base de fibre optique monomode. Par exemple, le taux de numérisation du réseau de transmission de France Telecom est proche de 100 %.

En opposition, les entreprises utilisent plutôt la paire torsadée pour leurs LANs et réservent la fibre optique multimode ou monomode à l'interconnexion de ces LANs. En effet, si la fibre optique est relativement bon marché, il n'en est pas de même de son installation (main d'oeuvre, connecteurs, câblage) et des adaptateurs opto-électroniques nécessaires (FOIRL Ethernet, PMD FDDI...). Le tableau 2-2 montre que le coût de la fibre optique dans un LAN était nettement supérieur en 1993 à celui de la paire torsadée (d'après SAUNDERS S., 1993).

Tableau 2-2: Comparaison du coût de la fibre optique et de la paire torsadée dans un LAN

	Fibre optique (\$)	Paire torsadée (\$)
Installation d'une connexion	1800	350
Adaptateur Ethernet	500	130
Adaptateur SAS FDDI	2600	1800

Il est habituel de distinguer les câbles en paires torsadées par leur blindage (voir tableau 2-3), leur impédance (100, 120 ou 150 ohms) ou leur nombre de paires torsadées (2, 4,...). Cependant, ces caractéristiques ne fournissent aucune indication sur la performance de ces câbles.

Tableau 2-3: Blindage des câbles en paires torsadées

Type	Blindage	Remarque
UTP	non blindé	-
FTP	écrané	feuille métallique
STP	blindé	feuille métallique + fil de masse

Les organismes tels que l'EIA/TIA ou l'ISO se servent de plusieurs paramètres pour mesurer la performance des câbles en paires torsadées. L'ACR exprimé en dB est le plus important d'entre eux. L'ACR est la différence entre l'atténuation et la diaphonie NEXT pour une distance et une fréquence de signal données. Un câble est d'autant plus performant que son ACR est élevé.

L'EIA/TIA classe les câbles en paires torsadées en différentes catégories (tableau 2-4) et l'ISO en classes.

Tableau 2-4: Classification des câbles en paires torsadées par l'EIA/TIA

Catégorie	Fréquence max (MHz)	Applications
3	16	Téléphonie, 10baseT, TR 4Mbit/s
4	20	idem, TR 16Mbit/s
5	100	idem, 100baseT, TP-PMD
5e	100	idem, 1000baseT

Actuellement, le choix majoritaire des entreprises est l'UTP 100 ohms 4 paires catégorie 5e qui permet d'évoluer sans risque vers des débits de 100 à 200 Mbit/s et qui supporte même le GigaEthernet.

Les catégories 6 et 7 plus performantes que la catégorie 5e sont en cours de définition au sein de l'EIA/TIA et de l'ISO (tableau 2-5). Toutefois, il n'est pas sûr que la paire torsadée de catégorie 6 ou 7 conserve un avantage significatif de coût par rapport à la fibre optique.

Tableau 2-5: Spécification de câbles en paires torsadées

Paramètres	Cat 5	Cat 5e	Cat 6	Cat 7
Fréquence (MHz)	0-100	0-100	0-250	0-600
Atténuation à 100 MHz (dB)	24	24	21,7	20,8
NEXT à 100 MHz (db)	27,1	30,1	39,9	62,1
Power-sum NEXT à 100 MHz (db)	-	27,1	37,1	59,1
ACR à 100 MHz (db)	3,1	6,1	18,2	41,3
Power-sum ACR à 100 MHz (db)	-	3,1	15,4	38,3
FEXT à 100 MHz (db)	17	17,4	23,2	?
Power-sum FEXT à 100 MHz (db)	14,4	14,4	20,2	?
Perte en retour à 100 MHz (db)	8	10	12	14,1
Délai de propagation à 100 MHz (nsec)	548	548	548	504
Délai inter-paire à 100 MHz (nsec)	50	50	50	20

2.2 PROCESSEUR ET MEMOIRE

Les 5 dernières années ont vu l'effondrement du ratio prix/performance des systèmes. Cet effondrement a touché micros, stations de travail, minis et grands systèmes bien qu'il soit plus marqué pour les micros et les stations de travail. Ce phénomène qui se poursuit s'explique en grande partie par les progrès réalisés dans le domaine des composants électroniques :

- **intégration**
Le nombre de transistors sur un processeur et sur une mémoire à semi- conducteurs DRAM augmente respectivement de 35% et 50% par an. La densité d'un disque augmente de 25% par an.
- **performance**
La vitesse d'horloge d'un processeur augmente de 50% par an. Le temps de cycle d'une mémoire DRAM diminue de 3% par an. Le temps d'accès d'un disque diminue de 3% par an.

La vitesse d'horloge d'un processeur n'est pas seulement liée à la technologie CMOS mise en oeuvre mais aussi à l'architecture du processeur (jeu et format des instructions, modes d'adressage, registres...). Les architectures actuelles sont de type CISC, RISC ou VLIW.

Les architectures CISC se caractérisent par un nombre élevé d'instructions de taille variable et un nombre réduit de registres dont certains sont spécialisés. Par exemple, l'architecture VAX a 304 instructions de taille variable et 15 registres de 32 bits dont 4 spécialisés. En opposition, les architectures RISC se caractérisent par un nombre réduit d'instructions de taille fixe et un nombre élevé de registres non spécialisés. Par exemple, l'architecture Alpha a 4 instructions de 32 bits et 64 registres de 64 bits.

Un processeur RISC exécute en plusieurs instructions ce qui est exécuté en une seule instruction complexe sur un processeur CISC. Par conséquent, un processeur RISC est plus performant qu'un processeur CISC si les instructions RISC équivalentes à une seule instruction CISC complexe s'exécutent plus vite que celle-ci. C'est souvent le cas. En effet, pour un nombre de transistors donné, une architecture RISC permet à la fois une vitesse d'horloge supérieure et un plus grand nombre d'instructions exécutées par cycle d'horloge (tableau 2-6).

Tableau 2-6: Caractéristiques de quelques processeurs

	Pentium	21064AA	21064A	21264
Architecture	CISC	RISC	RISC	RISC
Pipeline	-	2	2	4
Intégration [um]	0.8	0.75	0.5	0.35
Transistors [10 ⁶]	3.1	1.68	2.8	15.2
Horloge [MHz]	60-66	133-200	225-275	500-600
Cache [K]	8+8	8+8	16+16	64+64
Date	6/92	7/92	10/93	1/97

Les processeurs CISC et RISC utilisent le plus souvent la technique dite superscalaire pour exécuter plusieurs instructions en parallèle. En superscalaire de niveau N, le hardware par un mécanisme dit de prédiction de branchement essaie tant que faire ce peut de choisir N instructions à traiter dans N pipelines durant un cycle d'horloge. Par conséquent, l'efficacité de la technique superscalaire est liée à la qualité des algorithmes de prédiction de branchement. Or, ces algorithmes se sont révélés :

- coûteux à implémenter
La prédiction de branchement représente 30% des transistors d'un processeur Alpha 21264.
- d'une efficacité limitée
Quelque soit le nombre N de pipelines, il est difficile de dépasser en moyenne 2 instructions exécutées durant un cycle d'horloge.

Les architectures VLIW ont été conçues récemment pour permettre une meilleure parallélisation de l'exécution des instructions à un moindre coût. Les N instructions à exécuter en parallèle sont choisies par le compilateur et regroupées dans un bloc d'instructions. C'est ce bloc qui est traité par le processeur. Actuellement, les processeurs VLIW n'existent qu'à l'état de prototypes et leur avantage par rapport aux processeurs RISC restent à démontrer. L'architecture qui doit succéder à l'architecture x86 chez Intel est de type VLIW.

La mémoire est devenue un périphérique lent par rapport au processeur. Il en résulte que les accès mémoire sont un goulot d'étranglement s'ajoutant à celui traditionnel des accès disque (tableau 2-7). Cette situation devrait perdurer puisque toutes les analyses prévoient dans les années à venir que l'amélioration de la performance continuera à être beaucoup plus rapide pour les processeurs que pour la mémoire ou les disques.

Tableau 2–7: Hiérarchie de la mémoire

	Instruction	Registre	Cache	DRAM	Disque
Taille	-	<1Ko	<512Ko	<512Mo	>xGo
Délai [ns]	2.5	10	20	100	20000000

Puisque la mémoire est devenu un périphérique lent par rapport au processeur, les concepteurs de systèmes tentent de réduire le goulot d'étranglement qui en résulte en utilisant :

- des caches de plus en plus grands et sophistiqués
- des architectures superscalaires avec scoreboarding

Dans certains cas, la mémoire flash peut remplacer les disques ce qui supprime le goulot d'étranglement correspondant. Par exemple, ce type de mémoire est déjà utilisée par la plupart des équipements nécessitant des firmwares performants et de grande taille (routeurs multi-protocoles...). Son utilisation devrait se généraliser avec l'essor de l'informatique nomade.

Chapitre 3

IMPACT SUR LE RESEAU

3.1 LATENCE VS BANDE PASSANTE

Des média de transmission de capacité élevée et de taux d'erreur bit faible ainsi que des processeurs très rapides sont désormais disponibles à un coût raisonnable. Les technologies à haut débit qui en découlent nécessitent une approche totalement différente de la part des constructeurs informatiques.

KLEINROCK (1992) illustre ce basculement en analysant la latence du réseau (voir définitions dans l'annexe A).

Soit un WAN constitué de 2 ES dialoguant grâce à 3 commutateurs reliés entre eux par 2 LS de 300 km. Supposons que les commutateurs ont un temps moyen de traversée de 5 ms et que la taille moyenne des PDUs est de 825 octets. Considérons 2 cas :

- WAN1
La capacité des LS est de 9.6 Kbit/s et le taux d'erreur bit est de 10^{-5} .
- WAN2
La capacité des LS est de 100 Mbit/s et le taux d'erreur bit est de 10^{-8} .

Tableau 3-1: Analyse de la latence

Réseau	Dem, Dre (ms)	Dis (ms)	Dpr (ms)	Dtr (ms)	D (ms)
WAN1	2-20	5	2	688	709-745
WAN2	2-20	5	2	0.06	21-57

WAN1 a une latence dont la principale composante est le délai de transmission (92-97%). Par conséquent, le goulot d'étranglement est la bande passante.

Le concepteur de réseaux se doit de proposer des protocoles économisant la bande passante (taille des PDUs variable, entêtes de petite taille, reprise sur erreur au niveau des commutateurs...). Améliorer les implémentations de ces protocoles dans les ES et les commutateurs n'est pas prioritaire car l'impact est faible sur la latence du réseau. Cette stratégie est celle des années 70-80.

WAN2 a une latence dont les principales composantes sont les temps de traversée des ES et des commutateurs (90-96%). Par conséquent, le goulot d'étranglement est la vitesse de traitement des ES et des commutateurs. Un LAN à haut débit tel que FDDI conduit à une conclusion similaire.

Le concepteur de réseaux se doit de proposer non seulement des protocoles permettant un traitement rapide par les ES et les commutateurs (taille des PDUs et des entêtes fixe, suppression de la reprise sur erreur dans les commutateurs...) mais aussi des implémentations efficaces de ces protocoles. Cette stratégie est celle des années 90.

3.2 COMMUTATION VS MEDIUM PARTAGE

L'utilisation de commutateurs au sein des LANs offre un certain nombre d'avantages parmi lesquels :

- une gestion souple de la bande passante à travers l'ajout de ports et de commutateurs
- une latence réduite par l'absence de protocoles d'accès au médium
- une disponibilité accrue par l'introduction de chemins redondants entre ES
- une organisation par réseaux virtuels

Les commutateurs amènent également quelques inconvénients jusqu'ici inconnus du monde des LANs. Les applications distribuées des LANs se caractérisent par un trafic sporadique constitué de bursts émis à la capacité de la liaison de donnée (JAIN R. et ROUTHIER S.A., 1986). Les LANs à médium partagé sont conçus pour supporter ce type de trafic; le protocole d'accès au médium assure l'absence de congestion au niveau liaison de données. Ce n'est plus le cas dans les LANs utilisant des commutateurs.

La plupart des constructeurs informatiques considèrent que les commutateurs auront un poids de plus en plus important au sein des LANs. Cette idée est confortée par la constatation que tous les protocoles d'accès connus à ce jour sont peu performants à des capacités au delà du Gbit/s (LAMAIRE R.O., 1991). Les commutateurs devront offrir non seulement le service asynchrone des LANs actuels mais aussi les services synchrone et isochrone nécessaires aux nouvelles applications (multimédia...).

Annexe A

LATENCE D'UN RESEAU

Définissons D , la latence d'un réseau (temps de traversée), comme étant le temps moyen mis par une TPDU partant d'un ES émetteur pour atteindre l'ES destinataire. D se compose des délais suivants :

- Dem
 Dem est le délai nécessaire à l'ES émetteur pour construire la PDU à travers les couches du réseau. Ce délai dépend des protocoles utilisés et de leur implémentation.
- Dtr
 Dtr est le délai de transmission. Il se calcule en faisant le rapport *taille de la PDU/capacité du réseau*, la capacité étant le débit maximal possible sur le médium de transmission.
- Dpr
 Dpr est le délai de propagation. Il se calcule en faisant le rapport *longueur du médium de transmission/vitesse de propagation*. La vitesse de propagation du médium ne peut pas dépasser la vitesse de la lumière qui est d'environ 300000 km/s.
- Dis
 Dis est le délai nécessaire au commutateur pour traiter la PDU. Ce délai comprend le traitement proprement dit de la PDU (checksum, routage...) mais aussi l'attente dans la file de PDUs du commutateur. Ce délai dépend des protocoles utilisés et de leur implémentation.
- Dre
 Dre est le délai nécessaire à l'ES récepteur pour analyser la PDU à travers les couches du réseau. Ce délai dépend des protocoles utilisés et de leur implémentation.

Par conséquent, la latence $D = Dem + Dtr + Dpr + n*Dis + Dre$ est fonction :

- des protocoles utilisés
- de l'implémentation logicielle et matérielle de ces protocoles
- de la capacité du réseau
- de la taille du réseau
- de la vitesse de transmission du médium

Annexe B

BIBLIOGRAPHIE

HENNESSY J.L. et PATERSON D.A. (1990) Ed. Morgan Kaufmann
Computer architecture: A quantitative approach

JAIN R. et ROUTHIER S.A. (1986) JSAC 4(6) pp.1162-1165
Packet trains: measurements and a new model for computer network traffic

KLEINROCK L. (1992) IEEE Communications Magazine pp. 36-40
The latency/bandwidth tradeoff in gigabit networks

LAMAIRE R.O. (1991) Proceedings of ICC'91 pp. 1043-1048
FDDI performance at 1 Gbit/s

PARTRIDGE G. (1994) Ed. Addison-Wesley
Gigabit Networking

SAUNDERS S. (1993) Data Communications 21 sept. pp. 58-70
Choosing high-speed LANs

Glossaire

ACR Attenuation-to-Crosstalk Ratio

ANSI American National Standard Institute

CISC Complex Instructions Set Computer

DRAM Dynamic Random Access Memory

ES End System

EIA/TIA Electronic Industry Association/Telecommunication Industry Association

FDDI Fiber Distributed Data Interface

FEXT Far-End CrossTalk

FOIRL Fiber Optic Inter Repeater Link

FTP Foiled Twisted Pair

IEEE Institute of Electrical and Electronics Engineers

IS Intermediate System

ISO International Standard Organization

LAN Local Area Network

MMF MultiMode Fiber

NEXT Near-End CrossTalk

PDU Protocol Data Unit

PMD Physical layer Medium Dependent

RISC Reduced Instructions Set Computer

SAS Single Attachment Station

SMF Single Mode Fiber

STP Shielded Twisted Pair

TR Token Ring

UTP Unshielded Twisted Pair

VLIW Very Long Instruction Word

WAN Wide Area Network