

N° d'ordre : 1270

50376  
1985  
266

50376  
1985  
266

UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

# THÈSE

Présentée à

L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

pour obtenir le titre de

DOCTEUR DE 3ème CYCLE

(Informatique)

par

Larbi KZAZ



## ETUDE ET DEFINITION DE TOPOLOGIES NOUVELLES DE COMMUNICATION A HAUTES PERFORMANCES DANS UN SYSTEME MULTIPROCESSEUR

Thèse soutenue le 7 juin 1985, devant la Commission d'Examen

Membres du Jury : M. DAUCHET  
V. CORDONNIER  
F. BOERI  
C. CARREZ  
M.P. LECOUFFE

Président  
Rapporteur  
Examinateur  
Examinateur  
Examinateur

P R O F E S S E U R S   C L A S S E   E X C E P T I O N N E L L E

M. CONSTANT Eugène	I.E.E.A.
M. FOURET René	Physique
M. GABILLARD Robert	I.E.E.A.
M. MONTREUIL Jean	Biologie
M. PARREAU Michel	Mathématiques
M. TRIDOT Gabriel	Chimie
M. VIVIER Emile	Biologie
M. WERTHEIMER Raymond	Physique

P R O F E S S E U R S l è r e c l a s s e

---

M. BACCHUS Pierre	Mathématiques
M. BEAUFILS Jean-Pierre (dét.)	Chimie
M. BIAYS Pierre	G.A.S.
M. BILLARD Jean (dét.)	Physique
M. BOILLY Bénoni	Biologie
M. BOIS Pierre	Mathématiques
M. BONNELLE Jean-Pierre	Chimie
M. BOUGHON Pierre	Mathématiques
M. BOURIQUET Robert	Biologie
M. BREZINSKI Claude	I.E.E.A.
M. CELET Paul	Sciences de la Terre
M. CHAMLEY Hervé	Biologie
M. COEURE Gérard	Mathématiques
M. CORDONNIER Vincent	I.E.E.A.
M. DEBOURSE Jean-Pierre	S.E.S.
M. DYMENT Arthur	Mathématiques

PROFESSEURS 1ère classe (suite)

M. ESCAIG Bertrand	Physique
M. FAURE Robert	Mathématiques
M. FOCT Jacques	Chimie
M. GRANELLE Jean-Jacques	S.E.S.
M. GRUSON Laurent	Mathématiques
M. GUILLAUME Jean	Biologie
M. HECTOR Joseph	Mathématiques
M. LABLACHE COMBIER Alain	Chimie
M. LACOSTE Louis	Biologie
M. LAVEINE Jean Pierre	Sciences de la Terre
M. LEHMANN Daniel	Mathématiques
Mme LENOBLE Jacqueline	Physique
M. LHOMME Jean	Chimie
M. LOMBARD Jacques	S.E.S.
M. LOUCHEUX Claude	Chimie
M. LUCQUIN Michel	Chimie
M. MIGEON Michel Recteur à Grenoble	E.U.D.I.L.
M. MIGNOT Fulbert (dét.)	Mathématiques
M. PAQUET Jacques	Sciences de la Terre
M. PROUVOST Jean	Sciences de la Terre
M. ROUSSEAU Jean-Paul	Biologie
M. SALMER Georges	I.E.E.A.
M. SEGUIER Guy	I.E.E.A.
M. SIMON Michel	S.E.S.
M. STANKIEWICZ François	S.E.S.
M. TILLIEU Jacques	Physique
M. VIDAL Pierre	I.E.E.A.
M. ZEYTOUNIAN Radyadour	Mathématiques

P R O F E S S E U R S 2ème classe

M. ANTOINE Philippe	Mathématiques (Calais)
M. BART André	Biologie
Mme BATTIAU Yvonne	Géographie
M. BEGUIN Paul	Mathématiques
M. BELLET Jean	Physique
M. BERZIN Robert	Mathématiques
M. BKUCHE Rudolphe	Mathématiques
M. BODARD Marcel	Biologie
M. BOSQ Denis	Mathématiques
M. BRASSELET Jean-Paul	Mathématiques
M. BRUYELLE Pierre	Géographie
M. CAPURON Alfred	Biologie
M. CARREZ Christian	I.E.E.A.
M. CAYATTE Jean-Louis	S.E.S.
M. CHAPOTON Alain	C.U.E.E.P.
M. COQUERY Jean-Marie	Biologie
Mme CORSIN Paule	Sciences de la Terre
M. CORTOIS Jean	Physique
M. COUTURIER Daniel	Chimie
M. CROSNIER Yves	I.E.E.A.
M. CURGY Jean-Jacques	Biologie
Mle DACHARRY Monique	Géographie
M. DAUCHET Max	I.E.E.A.
M. DEBRABANT Pierre	E.U.D.I.L.
M. DEGAUQUE Pierre	I.E.E.A.
M. DELORME Pierre	Biologie
M. DELORME Robert	S.E.S.
M. DE MASSON D'AUTUME Antoine	S.E.S.
M. DEMUNTER Paul	C.U.E.E.P.

PROFESSEURS 2ème classe (Suite 1)

M. DENEL Jacques	I.E.E.A.
M. DE PARIS Jean-Claude	Mathématiques (Calais)
Mlle DESSAUX Odile	Chimie
M. DEVRAINNE Pierre	Chimie
M. DHAINAUT André	Biologie
Mme DHAINAUT Nicole	Biologie
M. DORMARD Serge	S.E.S.
M. DOUKHAN Jean-Claude	E.U.D.I.L.
M. DUBOIS Henri	Physique
M. DUBRULLE Alain	Physique (Calais)
M. DUBUS Jean-Paul	I.E.E.A.
M. FAKIR Sabah	Mathématiques
M. FONTAINE Hubert	Physique
M. FOUQUART Yves	Physique
M. FRONTIER Serge	Biologie
M. GAMBLIN André	G.A.S.
M. GLORIEUX Pierre	Physique
M. GOBLOT Rémi	Mathématiques
M. GOSSELIN Gabriel (dét.)	S.E.S.
M. GOUDMAND Pierre	Chimie
M. GREGORY Pierre	I.P.A.
M. GREMY Jean-Paul	S.E.S.
M. GREVET Patrice	S.E.S.
M. GUILBAULT Pierre	Biologie
M. HENRY Jean-Pierre	E.U.D.I.L.
M. HERMAN Maurice	Physique
M. JACOB Gérard	I.E.E.A.
M. JACOB Pierre	Mathématiques
M. JEAN Raymond	Biologie
M. JOFFRE Patrick	I.P.A.

PROFESSEURS 2ème classe (suite 2)

M. JOURNEL Gérard	E.U.D.I.L.
M. KREMBEL Jean	Biologie
M. LANGRAND Claude	Mathématiques
M. LATTEUX Michel	I.E.E.A.
Mme LECLERCQ Ginette	Chimie
M. LEFEVRE Christian	Sciences de la Terre
Mle LEGRAND Denise	Mathématiques
Mle LEGRAND Solange	Mathématiques (Calais)
Mme LEHMANN Josiane	Mathématiques
M. LEMAIRE Jean	Physique
M. LHENAFF René	Géographie
M. LOCQUENEUX Robert	Physique
M. LOSFELD Josph	C.U.E.E.P.
M. LOUAGE Francis(dét.)	E.U.D.I.L.
M. MACKE Bruno	Physique
M. MAIZIERES Christian	I.E.E.A.
M. MESSELYN Jean	Physique
M. MESSERLIN Patrick	S.E.S.
M. MONTEL Marc	Physique
Mme MOUNIER Yvonne	Biologie
M. PARSY Fernand	Mathématiques
Mle PAUPARDIN Colette	Biologie
M. PERROT Pierre	Chimie
M. PERTUZON Emile	Biologie
M. PONSOLLE Louis	Chimie
M. PORCHET Maurice	Biologie
M. POVY Lucien	E.U.D.I.L.
M. RACZY Ladislas	I.E.E.A.
M. RAOULT Jean François	Sciences de la Terre
M. RICHARD Alain	Biologie

PROFESSEURS 2ème Classe (suite 3)

M. RIETSCH François	E.U.D.I.L.
M. ROBINET Jean-Claude	E.U.D.I.L.
M. ROGALSKI Marc	Mathématiques
M. ROY Jean-Claude	Biologie
M. SCHAMPS Joël	Physique
Mme SCHWARZBACH Yvette	Mathématiques
M. SLIWA Henri	Chimie
M. SOMME Jean	G.A.S.
Mle SPIK Geneviève	Biologie
M. STAROSWIECKI Marcel	E.U.D.I.L.
M. STERBOUL François	E.U.D.I.L.
M. TAILLIEZ Roger	Institut Agricole
Mme TJOTTA Jacqueline (dét.)	Mathématiques
M. TOULOTTE Jean-Marc	I.E.E.A.
M. TURRELL Georges	Chimie
M. VANDORPE Bernard	E.U.D.I.L.
M. VAST Pierre	Chimie
M. VERBERT André	Biologie
M. VERNET Philippe	Biologie
M. WALLART Francis	Chimie
M. WARTEL Michel	Chimie
M. WATERLOT Michel	Sciences de la Terre
Mme ZINN JUSTIN Nicole	Mathématiques

CHARGES DE COURS

M. ADAM Michel S.E.S.

CHARGES DE CONFERENCES

M. BAF COP Joël I.P.A.

M. DUVEAU Jacques S.E.S.

M. HOF LACK Jean I.P.A.

M. LATOUCHE Serge S.E.S.

M. MALAUSSENA DE PERNO Jean-Louis S.E.S.

M. NAVARRE Christian I.P.A.

M. OPIGEZ Philippe S.E.S.



A Mes parents,  
A Mes frères.

إلى أُمِّي الحَسِينِ، اعترافاً مِنِّي لهُ عَلَى  
تَسْجِيْعَاتِهِ الْمُسْتَمِرَّةِ.

إلى وَالِدَيْيَ.

*Je tiens à remercier*

*Monsieur M. Dauchet, Professeur à l'Université de Lille 1, qui m'a fait l'honneur de présider le Jury de cette thèse.*

*Monsieur V. Cordonnier, Professeur à l'Université de Lille 1, qui a bien voulu m'accueillir dans son laboratoire, et qui est à l'origine de ce travail et l'a suivi constamment en m'aidant de ses précieux conseils.*

*Monsieur F. Boëri, Professeur à l'Université de Nice, d'avoir bien voulu se déplacer et examiner ce travail.*

*Monsieur C. Carrez, Professeur à l'Université de Lille 1, pour sa participation à ce Jury.*

*Madame M.P. Lecouffe, Maître-Assistante à l'Université de Lille 1, d'avoir bien voulu examiner ce travail.*

*Madame B. Vandroemme et Monsieur H. Glanc qui ont assuré avec compétence, la réalisation matérielle de ce document.*

*Il m'est très agréable de saluer tous mes camarades et amis du laboratoire, pour leur sympathie.*

*Enfin, je tiens à exprimer au corps enseignant en Informatique, ma profonde reconnaissance et mon admiration pour leurs enseignements que je suis fier d'avoir suivis.*

## INTRODUCTION GENERALE

La multiplication des ressources d'un ordinateur est un moyen privilégié pour accroître ses performances. Les ordinateurs MIMD occupent, dans ce domaine, une place de premier rang. Et cela en raison de leur capacité d'exploiter diverses formes de parallélisme existants dans des domaines d'application aussi variés que ceux des applications scientifiques, de l'intelligence artificielle ou encore celui de l'informatique graphique.

L'apport de la technologie VLSI, dans le domaine de la conception d'unités fonctionnelles spécialisées - Processeurs réseaux ou Processeurs systoliques -, de plus en plus performantes, contribue à l'utilisation massive de ces unités dans les ordinateurs MIMD.

Cette tendance à la multiplication des ressources fonctionnellement, voire physiquement hétérogènes, rend particulièrement cruciale la fonction de l'outil de communication.

C'est aux outils de communication dans des environnements à unités hétérogènes, et de type MIMD, qu'est consacré l'essentiel de ce travail. Nous nous intéressons exclusivement à la commutation de ligne comme stratégie de transfert et proposons dans ce contexte de nouvelles topologies d'interconnexion, qui répondent aux exigences de la technologie VLSI.

Cette étude se présente en quatre chapitres :

Le chapitre I, est consacré à l'étude des communications internes à un ordinateur en fonction de sa structure. Il présente les principaux outils de communication existants et fixe les objectifs de ce travail.

Le chapitre II, présente l'outil de communication qui nous servira de modèle de base. Sa topologie, ainsi que son protocole d'interconnexion y sont définis.

Le chapitre III, étudie les possibilités d'extension du modèle de base. Deux outils y sont développés : le réseau ARBRE et le MULTI-ANNEAU. Un troisième outil, le réseau CELLULAIRE, montre les limites des extensions de notre modèle de base.

Le chapitre IV, est consacré à l'évaluation des outils proposés dans les chapitres II et III. L'évaluation porte sur les deux aspects : complexité et performances. Des comparaisons avec des outils existants, lorsque cela est possible, y sont également effectuées.

## CHAPITRE I : INTRODUCTION AUX COMMUNICATIONS INTERNES AUX CALCULATEURS

### 1. COMMUNICATIONS INTERNES ET STRUCTURE DES CALCULATEURS

#### 1.a. Les calculateurs SISD

##### 1.a.1. Structure et fonctionnement

##### 1.a.2. Communications internes

##### 1.a.3. Conclusion

#### 1.b. Les calculateurs évolués

##### 1.b.1. Les calculateurs SIMD

###### 1.b.1.a. Structure et fonctionnement

###### 1.b.1.b. Communications internes

##### 1.b.2. Les calculateurs MIMD

###### 1.b.2.a. Structure et fonctionnement

###### 1.b.2.b. Communications internes

###### 1.b.2.c. Variantes

### 2. LES OUTILS DE COMMUNICATION

#### 2.a. Le bus partagé

##### 2.a.1. Protocole d'accès

##### 2.a.2. Protocole de transfert

##### 2.a.3. Performances dans un environnement multiprocesseurs

#### 2.b. Le cross-bar

#### 2.c. Les bus-multiples et les multi-étages

##### 2.c.1. Les bus-multiples

##### 2.c.2. Les multi-étages

###### 2.c.2.a. Le réseau CLOS sans blocage

###### 2.c.2.b. Un réseau CLOS réarrangeable : BENES

###### 2.c.2.b.1. Topologie

###### 2.c.2.b.2. Commande

###### 2.c.2.c. Les réseaux BANYAN

###### 2.c.2.c.1. Définition

###### 2.c.2.c.2. Le réseau BASELINE

###### 2.c.2.c.2.a. Topologie

###### 2.c.2.c.2.b. Commande

###### 2.c.2.c.3. Le réseau OMEGA

###### 2.c.2.c.3.a. Topologie

###### 2.c.2.c.3.b. Commande

**2.c.2.d. Comparaison des réseaux multi-étages**

**2.c.2.d.1. Comparaison des réseaux BANYAN**

**2.c.2.d.2. Comparaison du réseau BENES aux réseaux BANYAN**

**3. PERFORMANCES, COMPLEXITE ET ENVIRONNEMENT TECHNOLOGIQUE**

**4. OBJECTIFS DE L'ETUDE**

## CHAPITRE II : LE MODELE DE BASE : L'ANNEAU

### 1. DEFINITION DU MODELE DE BASE

- 1.a. La topologie d'interconnexion
- 1.b. La cellule de commutation
- 1.c. La commutation de ligne dans l'ANNEAU

### 2. LE PROTOCOLE DE COMMUNICATION

- 2.a. Une décomposition du protocole de communication
- 2.b. Le protocole d'interconnexion
  - 2.b.1. Le contrôle d'accès
  - 2.b.2. Le contrôle d'une connexion
  - 2.b.3. Le module de contrôle
    - 2.b.3.a. Environnement d'interaction du module de contrôle
    - 2.b.3.b. Etude des états du module de contrôle
  - 2.b.4. Configuration générale de l'ANNEAU
- 2.c. Le protocole de transfert
  - 2.c.1. Transfert séquentiel
    - 2.c.1.a. Transfert séquentiel synchrone
    - 2.c.1.b. Transfert séquentiel asynchrone
  - 2.c.2. Transfert pipeline
    - 2.c.2.a. Structure des modules de transfert
    - 2.c.2.b. Contrôle synchrone du transfert pipeline
    - 2.c.2.c. Contrôle asynchrone du transfert pipeline
  - 2.c.3. Comparaison

### 3. CONCLUSION

## CHAPITRE III : EXTENSIONS DU MODELE DE BASE

### 1<sup>ère</sup> PARTIE : EXTENSIONS PAR GRAPHE

#### 0. INTRODUCTION

#### 1. LE RESEAU ARBRE

- 1.a. La topologie d'interconnexion
- 1.b. La fonction de numérotation
- 1.c. Le protocole d'interconnexion
  - 1.c.1. La procédure d'accès
  - 1.c.2. Conflits et blocage
- 1.d. Conclusion

#### 2. LE RESEAU CELLULAIRE

- 2.a. La topologie d'interconnexion
- 2.b. La commutation de ligne dans le réseau CELLULAIRE
- 2.c. Puissance fonctionnelle
  - 2.c.1. Conflits
  - 2.c.2. Blocage
  - 2.c.3. Simultanéité
  - 2.c.4. Critique de la cellule de base
- 2.d. Conclusion

#### 3. CONCLUSION

### 2<sup>ème</sup> PARTIE : EXTENSION PAR ANALOGIE ENTRE LE MODELE DE BASE ET LE BUS

#### 0. INTRODUCTION

#### 1. LA TOPOLOGIE DU MULTI-ANNEAU

#### 2. LA COMMUTATION DE LIGNE DANS LE MULTI-ANNEAU

#### 3. LE PROTOCOLE D'INTERCONNEXION

- 3.a. La fonction de numérotation
- 3.b. La procédure de routage

#### 4. CONCLUSION



## CHAPITRE IV : EVALUATION DES OUTILS PROPOSES

### 0. INTRODUCTION

### 1. L'ANNEAU

#### 1.a. Performances

##### 1.a.1. Performances potentielles

###### 1.a.1.a. Délai d'accès

###### 1.a.1.b. Débit potentiel

##### 1.a.2. Performances effectives

###### 1.a.2.a. Le logiciel de simulation

###### 1.a.2.b. Les hypothèses de la simulation

###### 1.a.2.c. Résultats de la simulation

##### 1.a.3. Conclusion

#### 1.b. Compléxité

#### 1.c. Comparaison de l'ANNEAU avec le BUS

### 2. LES OUTILS EXTENSION

#### 2.a. Le réseau ARBRE

##### 2.a.1. Performances

##### 2.a.2. Compléxité

##### 2.a.3. Conclusion

#### 2.b. Le MULTI-ANNEAU

##### 2.b.1. Performances

##### 2.b.2. Compléxité

##### 2.b.3. Comparaison du MULTI-ANNEAU au cross-bar

### 3. CONCLUSION

\*\*\*\*\*  
\* CHAPITRE I \*  
\*\*\*\*\*

COMMUNICATIONS INTERNES AUX CALCULATEURS

Les performances d'un ordinateur dépendent de la vitesse de ses organes essentiels, ainsi que de la capacité d'assurer des communications rapides entre ces organes. Ces communications résultent à la fois de son mode de fonctionnement et des mouvements d'information qu'implique l'élaboration d'un traitement donné.

## 1. Communications internes et structure des ordinateurs

Les principaux flux de communication internes à un ordinateur, concernent les instructions et les données. Il est donc naturel, dans le but d'étudier ces communications, de partir d'un modèle qui caractérise chacun de ces flux, c'est la raison pour laquelle nous utiliserons la classification de Flynn [FLY 72], et nous limiterons aux trois structures de ordinateurs suivantes : SISD, SIMD et MIMD\*.

### 1.a.1. Les ordinateurs SISD

Les ordinateurs SISD, de conception classique, comportent les quatre unités de base suivantes :

- L'unité de contrôle et de commande - u.c.c. - ,
- l'unité arithmétique et logique - u.a.l. - ,
- la mémoire, et
- l'unité d'échange - u.e. - .

Leur fonctionnement se résume en l'exécution séquentielle d'un ensemble d'instructions. Le processus d'exécution d'une instruction se décompose en cinq étapes, et s'effectue sous le contrôle de l'u.c.c. :

- $E_1$  : Recherche de l'instruction.
- $E_2$  : Décodage de l'instruction.
- $E_3$  : Recherche des opérandes données.
- $E_4$  : Exécution proprement dite.
- $E_5$  : Rangement de l'opérande résultat.

\* Selon la terminologie de Flynn : SISD : Single Instruction Single Data,  
SIMD : Single Instructions Multiple Data,  
MIMD : Multiple Instructions Multiple Data.

### 1.a.2. Communications internes

Parmi les cinq étapes du processus d'exécution d'une instruction, seule la quatrième peut faire l'objet d'un traitement par l'u.a.l., à condition que l'instruction ne soit pas elle-même une instruction d'entrée-sortie. Les étapes 1, 3 et 5 sont, quant à elles, toutes des opérations de transfert entre la mémoire et les autres unités du calculateur. Il en découle que :

- d'une part, une partie considérable de l'activité de la machine est consacrée aux communications entre ses organes, et que
- d'autre part, la plupart de ces communications ont lieu entre la mémoire et une autre unité du calculateur.

Cette dernière constatation a des implications sur le format d'un transfert élémentaire, qui est celui du mot mémoire, et sur la vitesse de transfert qui est limitée par le débit potentiel de cette même mémoire.

Par ailleurs, la longueur d'un échange, mesurée en nombre de mots mémoire, est spécifique à l'unité avec laquelle la mémoire effectue cet échange. Cette longueur est, en effet, constante pour l'u.c.c., elle est égale au nombre de mots mémoire sur lesquels sont codées les instructions ; elle est par contre variable pour l'u.e. et l'u.a.l., et dépend, pour la première, de la longueur et du nombre d'opérandes de l'instruction à exécuter, et des valeurs des opérandes de l'instruction d'entrée-sortie, pour la seconde.

### 1.a.3. Conclusion

En raison du nombre réduit de leurs organes, et du caractère séquentiel des transferts qu'impose le mode de fonctionnement des calculateurs SISD ; l'outil de communication ne représente pas une ressource critique. Un outil de type Bus répond en général aux besoins de ces calculateurs.

### 1.b. Les calculateurs évolués

L'accroissement des performances dans les calculateurs évolués résulte des progrès enregistrés dans le domaine de la technologie des composants qui

ont permis la conception, à des coûts réduits, de composants de plus en plus performants [BLG 77]. Cette évolution s'est traduite au niveau de l'organisation des calculateurs, par la multiplication du nombre de leurs ressources, organes de stockage et de traitement, ainsi que par la définition de nouveaux modes de fonctionnement permettant un meilleur coefficient d'utilisation de ces ressources et ce en autorisant des traitements parallèles.

Deux classes de calculateurs ont été définies. La première, les calculateurs SIMD, exploite un micro-parallélisme existant dans des applications scientifiques, telles que des opérations sur des vecteurs ou des matrices par exemple. La seconde, les calculateurs MIMD exploite un macro-parallélisme existant entre tâches ou instructions d'un même programme.

### 1.b.1. Les calculateurs SIMD

Ce sont des calculateurs spécialisés dans le traitement vectoriel, leur puissance de calcul peut atteindre le GFLOPS : performances théoriques de MPP [BAT 80].

#### 1.b.1.a. Structure et fonctionnement

Un calculateur SIMD comporte les principaux organes suivants :

- Un ensemble de processeurs d'exécution élémentaires (PE), tous identiques. Chaque PE est composé d'une unité arithmétique et logique, d'une mémoire locale et d'une logique de communication.
- Une unité de contrôle et de commande, beaucoup plus sophistiquée que celles dont disposent généralement les calculateurs conventionnels.

L'ensemble des PE sont synchronisés sur la même horloge et exécutent, si simultanément, la même instruction sur des données différentes. Cette instruction est préalablement diffusée à l'ensemble des PE par l'unique u.c.c. du calculateur.

#### 1.b.1.b. Communications internes

L'exécution simultanée de la même instruction par tous les PE exige

la disponibilité des opérandes dans la mémoire locale de chacun de ces PE, ce qui implique :

- d'une part, une répartition adéquate des données dans les différentes mémoires de manière à ce qu'elles puissent être accédées sans conflit [TAH 79],
- et d'autre part, la possibilité d'effectuer des transferts de données entre les PE.

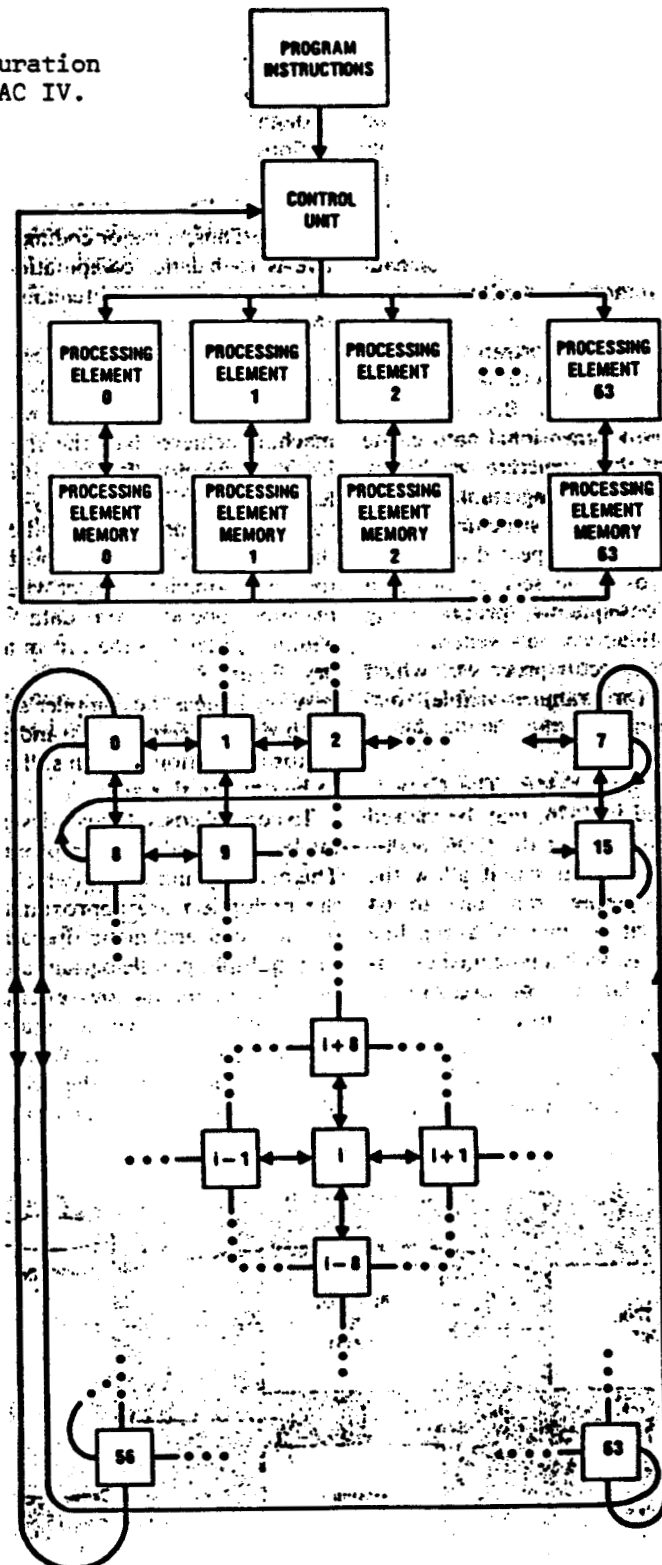
L'utilisation massive des PE dans ce genre d'architecture - le calculateur DAP par exemple, comporte 4096 PE, MPP en comporte 16384 -, ainsi que la fréquence très élevée des demandes de transfert entre ces PE, rendent particulièrement critique le rôle de l'outil de communication. L'élaboration de transferts rapides, donc parallèles, conditionne l'obtention d'un meilleur coefficient d'utilisation des PE et, par voie de conséquence une puissance de calcul élevée.

Notons, cependant, que certaines caractéristiques de la structure matérielle du calculateur et/ou de l'application envisagée peuvent être mises à profit dans la définition de l'outil de communication :

- i) Le caractère homogène des organes de la machine. Cette homogénéité porte aussi bien sur l'aspect fonctionnel que sur l'aspect physique : l'utilisation, par exemple, de microprocesseurs identiques comme PE.
- ii) Le format - nombre de bits par mot - des mots échangés entre les PE est fixe. Une proposition de classification des calculateurs selon la valeur du format et le nombre des mots traités en parallèles a été faite par FENG, et on peut considérer dans ce sens que cette valeur caractérise la structure du calculateur. Le calculateur ILLIAC IV [BAR 68], par exemple, travaille sur des mots de format 64 bits.
- iii) Les demandes de transfert sont formulées par les PE de manière synchrone.

iv) Lorsque le calculateur est destiné à un type d'applications bien défini, l'analyse du mouvement des données qu'il entraîne dans le calculateur permet d'adapter la configuration à son flux de communication. Dans ILLIAC IV, les communications d'un PE sont limitées à ses quatre PE voisins - figure 1-.

Figure 1 : Configuration d'ILLIAC IV.



Les opérations d'entrée-sortie sont, quant à elles, prises en charge par l'unité d'échange d'un ordinateur hôte, conventionnel en général, et auquel le ordinateur SIMD est attaché. Le ordinateur hôte supporte également l'environnement d'exploitation du ordinateur SIMD.

### 1.b.2. Les ordinateurs MIMD

Cette classe de ordinateurs permet de répondre aux besoins en puissance de calcul qu'exigent des domaines d'application aussi variés que ceux des applications scientifiques, de l'intelligence artificielle, du traitement d'images ou encore celui de la reconnaissance des formes. Le parallélisme qu'exploitent les ordinateurs MIMD est, contrairement aux ordinateurs SIMD, également varié :

- parallélisme dans l'évaluation des expressions arithmétiques,
- parallélisme entre tâches coopérantes,
- parallélisme ET et OU.

#### 1.b.2.a. Structure et fonctionnement

La structure des ordinateurs MIMD se présente, de manière générale, sous la forme d'un ensemble de processeurs et d'une mémoire commune, décomposée en un ensemble de modules indépendants. Chaque processeur dispose de ses propres unités de contrôle et de calcul et élabore, de manière asynchrone et indépendamment des autres processeurs, le cycle classique d'exécution d'une instruction. A un instant donné, donc, plusieurs processeurs exécutent en parallèle des instructions différentes sur des opérandes différents.

Deux configurations possibles d'un ordinateur MIMD sont présentées sur la figure 2 ; dans la première, la mémoire commune est centralisée - figure 2.a -, c'est le cas du ordinateur NYU [GKM 83] ; dans la deuxième, elle est répartie - figure 2.b -, c'est le cas des ordinateurs CHOPP [SUL 77] et CM\* [SFS 77] par exemple.



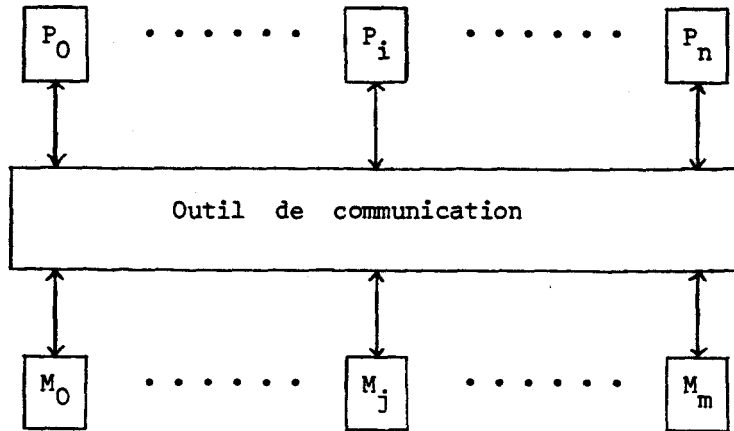
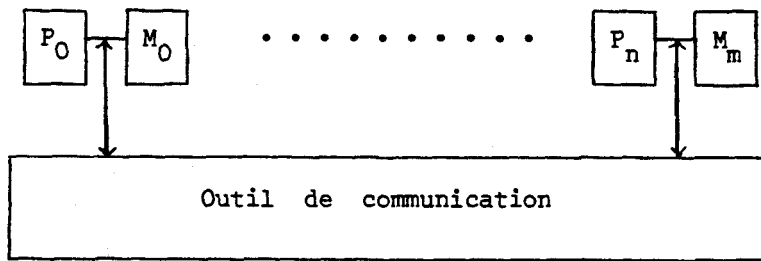
a - Mémoire commune centralisée ( $m \geq n$ ).b - Mémoire commune répartie ( $m = n$ ).

Figure 2 : Configurations de la mémoire commune dans les calculateurs MIMD.

### 1.b.2.b. Communications internes

Les communications internes aux calculateurs MIMD sont dominées par les opérations de transfert entre processeurs et modules de la mémoire ; en raison, en partie, des communications inter-processeurs qui se font via leur mémoire commune. Les demandes de transfert sont par ailleurs émises de manière asynchrone et la fréquence de leurs occurrences, très élevée par processeur, est accentuée par l'utilisation massive de ces processeurs ; le calculateur CHOPP [SUL 77], par exemple, est conçu pour comporter entre  $10^5$  et  $10^6$  processeurs. Afin de ne pas ralentir les performances du calculateur, l'outil de communication doit assurer un débit au moins égal à  $m$  fois celui de chacun des  $m$  modules de la mémoire commune. Le format des échanges est, quant à lui, fixé par celui du mot mémoire.

### 1.b.2.c. Variantes de la structure MIMD

La structure des calculateurs MIMD, telle que nous l'avons décrite dans un précédent paragraphe, peut être modifiée par :

- i) L'introduction d'unités fonctionnelles spécialisées ;
- ii) et/ou l'affectation de ressources privées à chacun des processeurs et à chaque groupe.

L'utilisation d'unités fonctionnelles spécialisées n'est pas, en soi, une approche propre aux calculateurs évolués. L'utilisation dans les calculateurs conventionnels d'unités d'échange chargées des opérations d'entrée-sortie, en est l'illustration. Dans les calculateurs MIMD, un certain nombre de fonctions de base telles que, l'inversion d'une matrice, le calcul de la transformée de fourrier, la résolution d'un système linéaire ... etc... sont élaborées par des processeurs spécialisés. Ces opérations pouvant être traitées beaucoup plus rapidement par des processeurs spécialisés, qu'elles ne peuvent l'être par les processeurs généraux, ce qui permet d'accroître la puissance de traitement du calculateur. En revanche, ceci se traduit au niveau de l'outil de communication par l'accroissement du nombre des ressources communes qu'il interconnecte et leur hétérogénéité. Il est alors nécessaire de mettre en place des protocoles de communication spécifiques entre les processeurs généraux et les processeurs spécialisés [WAH 84].

L'affectation de ressources privées aux processeurs permet d'accroître leur degré d'autonomie. Un exemple de telles ressources, est celui de la mémoire locale privée, utilisée par le calculateur FMP [BAR 81]. La structure de ce calculateur est présentée par la figure 3. Un autre exemple, est celui de l'affectation d'une unité d'échange à chaque processeur, c'est le cas du calculateur CM\* dont la structure est illustrée par la figure 4. L'introduction de ressources privées se traduit par la décomposition, en deux niveaux, des communications dans le système - figure 5 - :

- \* Un niveau de communication local, entre chaque processeur et ses ressources privées,

\*\* et un niveau de communication global, entre les processeurs et les ressources communes et, éventuellement entre les processeurs eux-mêmes, dans le cas où les communications directes entre ces processeurs sont prévues.

Après avoir présenté quelques aspects liés aux communications internes aux calculateurs, nous nous proposons, par la suite, d'étudier les principaux outils de communications utilisés dans les calculateurs.

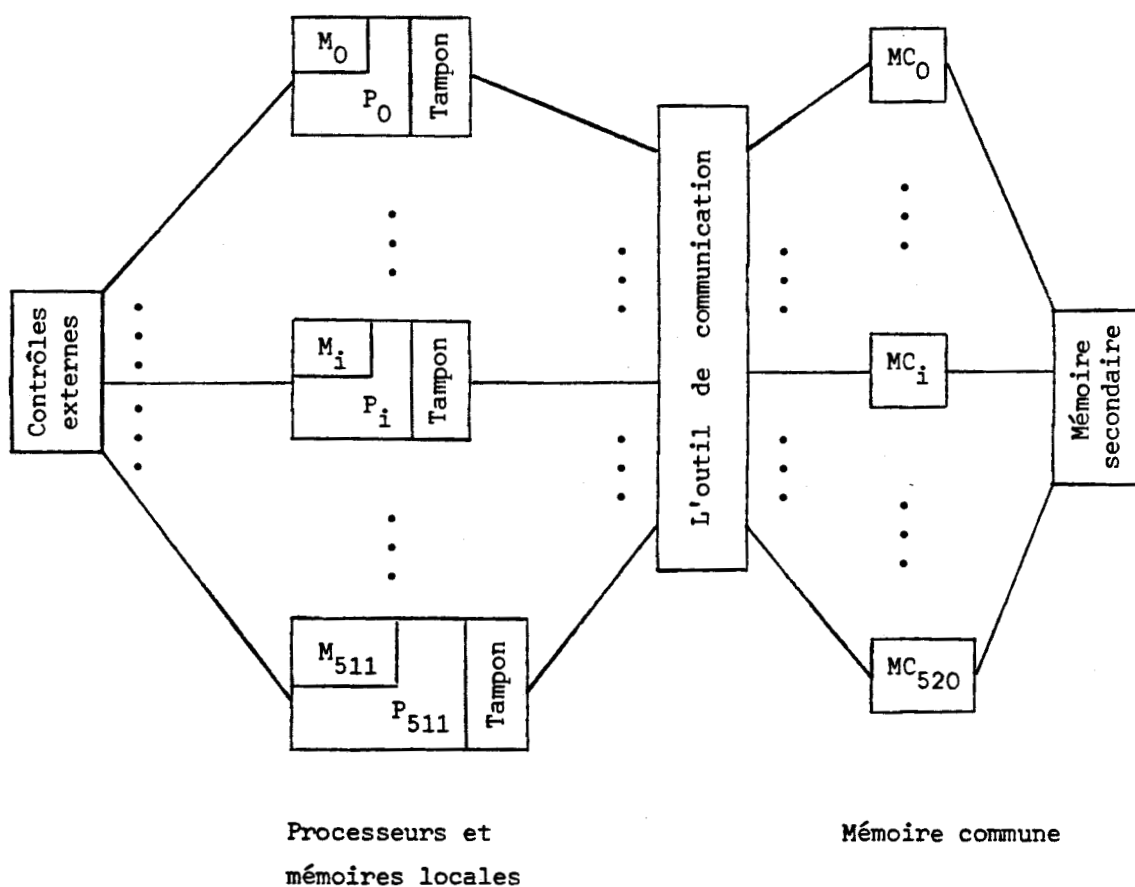
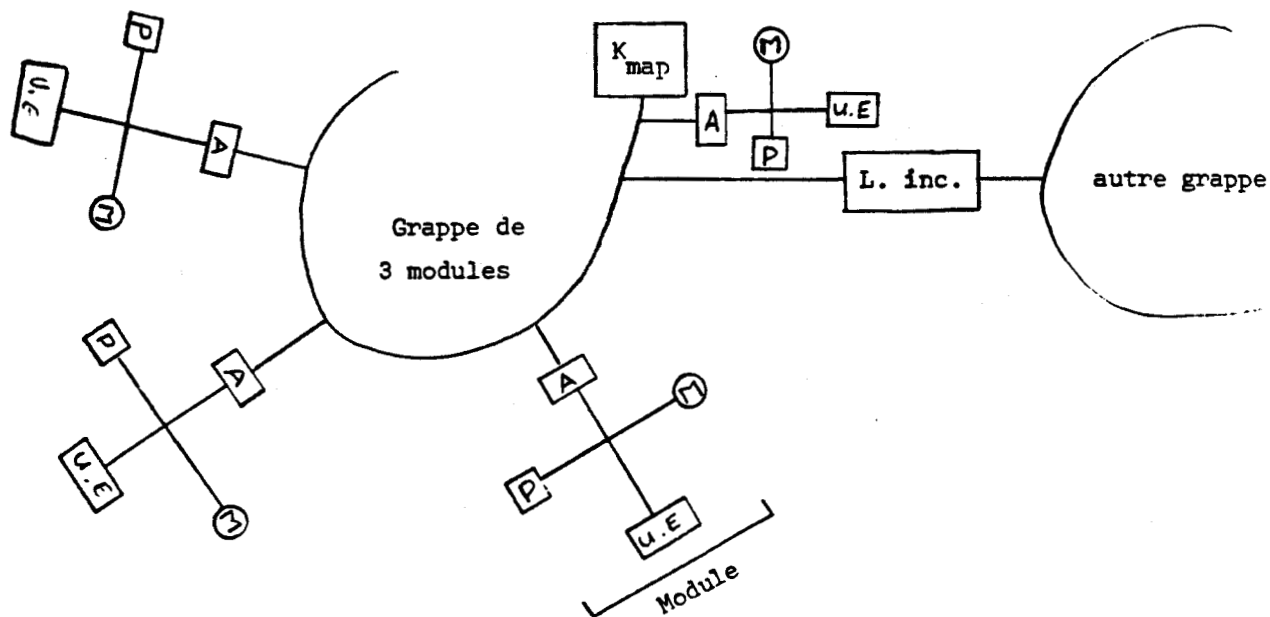


Figure 3 : Structure du calculateur Flow Model Processor.



A : adresseur.

Figure 4 : Structure du calculateur CM\*.

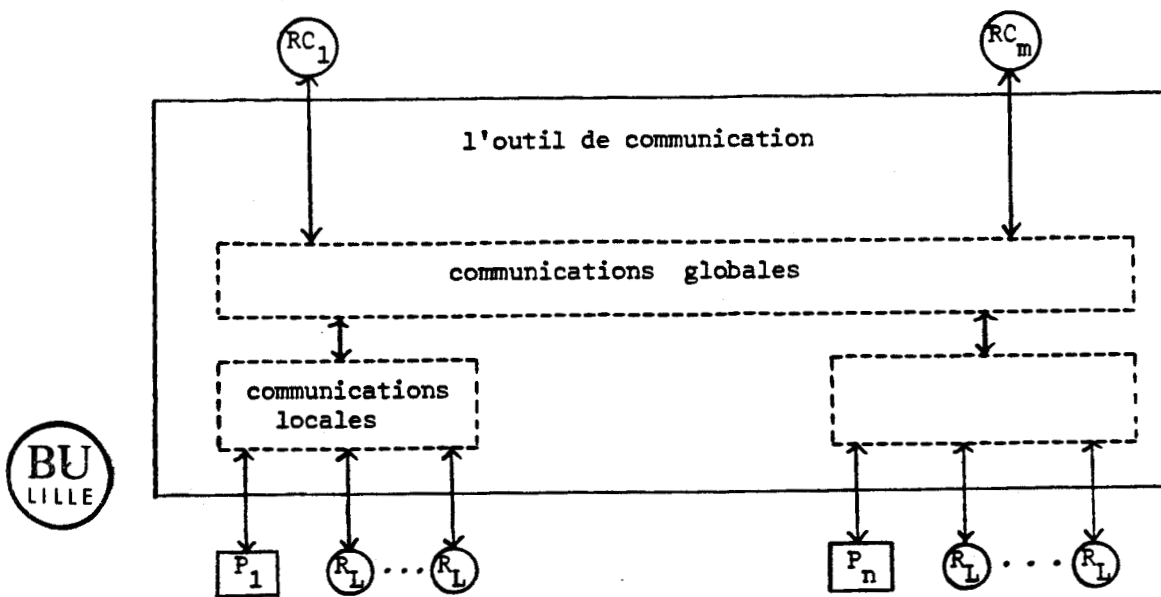


Figure 5 : Structure MIMD avec deux niveaux de communication.

## 2. Les outils de communication

### 2.a. Le bus partagé

Les connexions entre unités sont réalisées par un ensemble de fils conducteurs groupés en parallèle ; elles constituent ce support physique appelé BUS. Une seule communication au plus, peut être établie à un instant donné ; ce qui transforme le bus, lorsque le nombre des unités actives qui le partagent est supérieur à 1, en un outil de communication avec blocage.

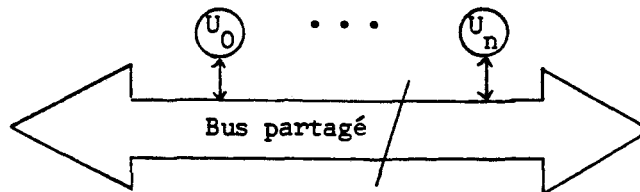


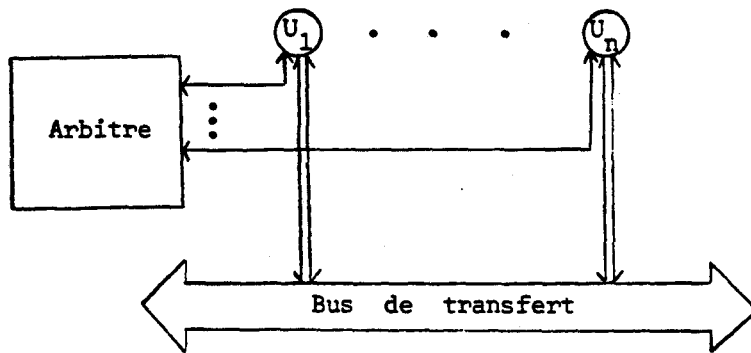
Figure 6 : Structure de bus partagé.

#### 2.a.1. Protocoles d'accès

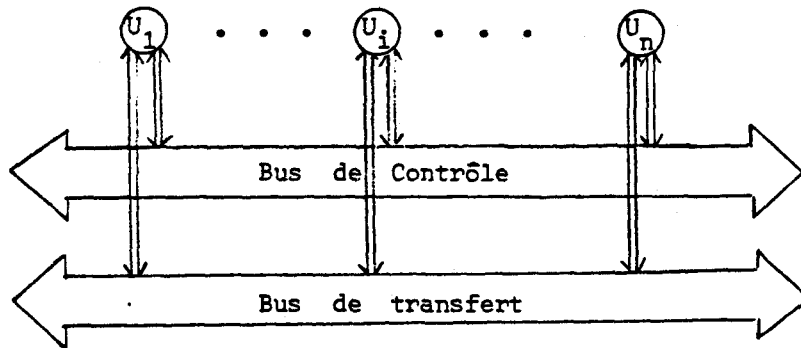
Le protocole d'accès, associé au bus, doit définir :

- i) La technique d'accès : la technique, dite d'accès par sélection [COR 81], est la plus généralement adoptée dans les systèmes fortement couplés.
- ii) La règle de gestion des demandes, selon une priorité logique ou temporelle.
- iii) La politique d'arbitrage, lorsque deux demandes d'accès au bus occurent simultanément. Il existe plusieurs politiques d'arbitrage ; les plus significatives d'entre-elles sont : l'arbitrage linéaire, circulaire, aléatoire et cyclique.

Le protocole d'accès peut être implanté de manière répartie ou centralisée - figure 7.a -, dans ce dernier cas le dispositif qui le définit est un circuit spécialisé appelé arbitre [SCG 83] - figure 7.b -. Dans les deux cas, des voies de communications réservées sont nécessaires, l'ensemble de ces voies est désigné dans la littérature par le nom de bus de contrôle.



a - Contrôle centralisé.



b - Contrôle réparti.

Figure 7 : Contrôle dans le bus partagé.

La solution centralisée a l'avantage de permettre la mise en oeuvre d'une politique d'arbitrage très évoluée, en revanche elle est plus sensible aux pannes et limite l'extensibilité du bus au nombre d'entrées qu'autorise l'arbitre. D'autre part, le fonctionnement en mode asynchrone de l'arbitre, s'il permet des délais de décision très courts, entre 20 ns et 50 ns [SCG 83], sa réalisation est plutôt difficile. Le délai d'accès au bus est déterminé par le protocole d'accès, afin de ne pas laisser le bus inoccupé durant ce temps, il est possible d'utiliser la technique d'anticipation, à condition que la règle de gestion des demandes le permette. Cette technique consiste à sélectionner une unité, parmi celles qui demandent l'accès au bus, avant que le transfert en court ne soit terminé [TSG 80].

### 2.a.2. Protocoles de transfert

Il définit les règles qui régissent le transfert dans le bus entre deux unités, une fois l'accès leur y a été autorisé. Il existe deux façons d'envisager le transfert dans un bus :

- i) Transfert synchrone : associée au bus, une horloge de référence unique cadence les transferts. Sa période, et par conséquent le débit potentiel du bus, dépend des caractéristiques physiques des fils qui composent le bus : conductivité et longueur. Ce mode de transfert est utilisé lorsque les débits des unités interconnectées sont compatibles. Il est par ailleurs aisé à réaliser, mais il a l'inconvénient de limiter l'extensibilité du bus.
  
- ii) Transfert asynchrone : lorsque les débits des unités interconnectées par le bus ne sont pas compatibles, on utilise la technique dite de l'accord confirmé [COR 81]. L'UNIBUS du PDP 11 [COR 81] utilise ce mode de transfert. Ce mode de transfert permet de mettre à profit l'extensibilité du bus, et est bien adapté à un environnement composé d'unités hétérogènes. Mais sa mise en oeuvre exige des liaisons supplémentaires pour l'échange des requêtes et des acquittements. Le débit de transfert dépend, pour une large part, de celui des unités interconnectées.

### 2.a.3. Performances dans un environnement multiprocesseurs

L'utilisation du bus partagé dans un environnement multiprocesseur le transforme vite en un goulot d'étranglement du système. En effet :

- i) La probabilité de conflit, pour accéder au bus, croît en fonction du nombre des processeurs et des fréquences de leurs requêtes - figure 8.a -. Au delà d'un certain nombre de processeurs, ils deviennent en conflit permanent.

- ii) Le rapport  $\frac{\text{délai d'accès}}{\text{temps d'activité total}}$  croît et tend asymptotiquement vers 1, en fonction du nombre de processeurs - courbe  $d_n$  de la figure 8.b - entraînant, en conséquence, une diminution des performances individuelles des processeurs - courbe  $P_n$  de la figure 8.b -, et des performances totales du système - courbe  $P_n^t$  de la figure 8.b -.

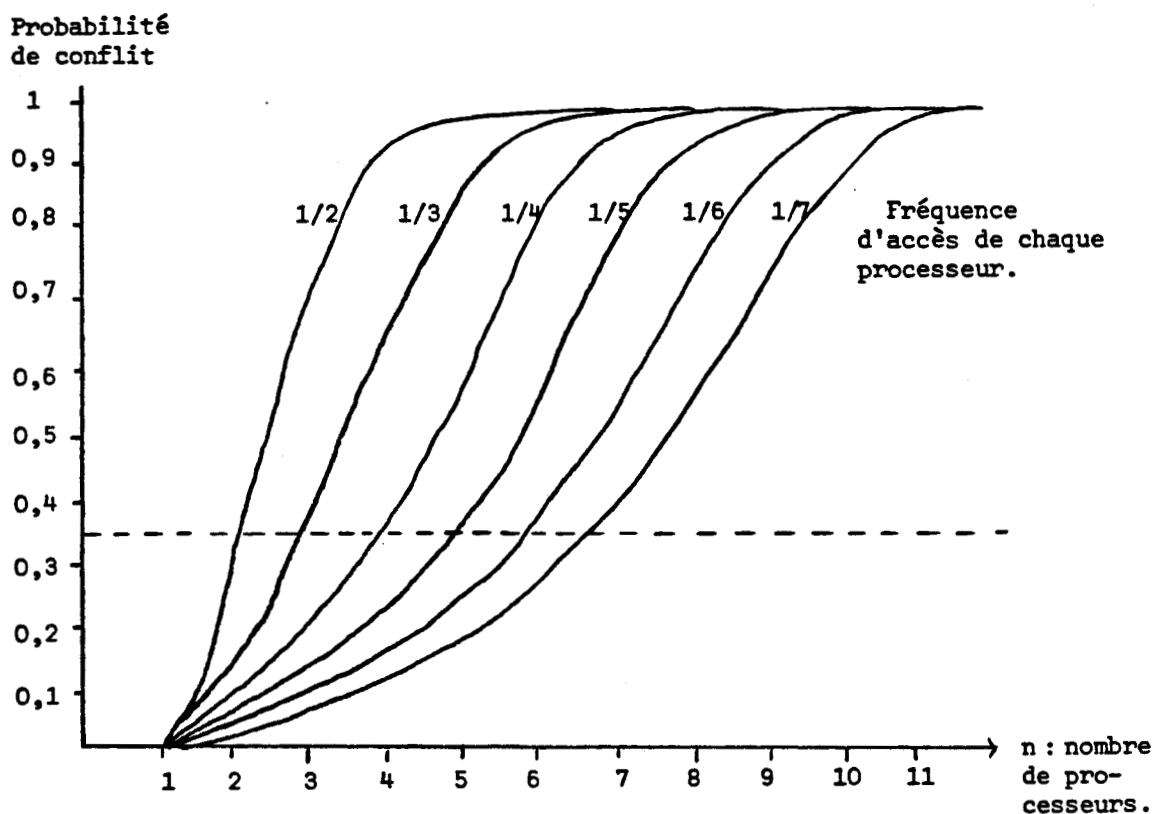


Figure 8.a. : Probabilité de conflit d'accès au bus en fonction du nombre de processeurs et de leurs fréquences d'accès au bus.





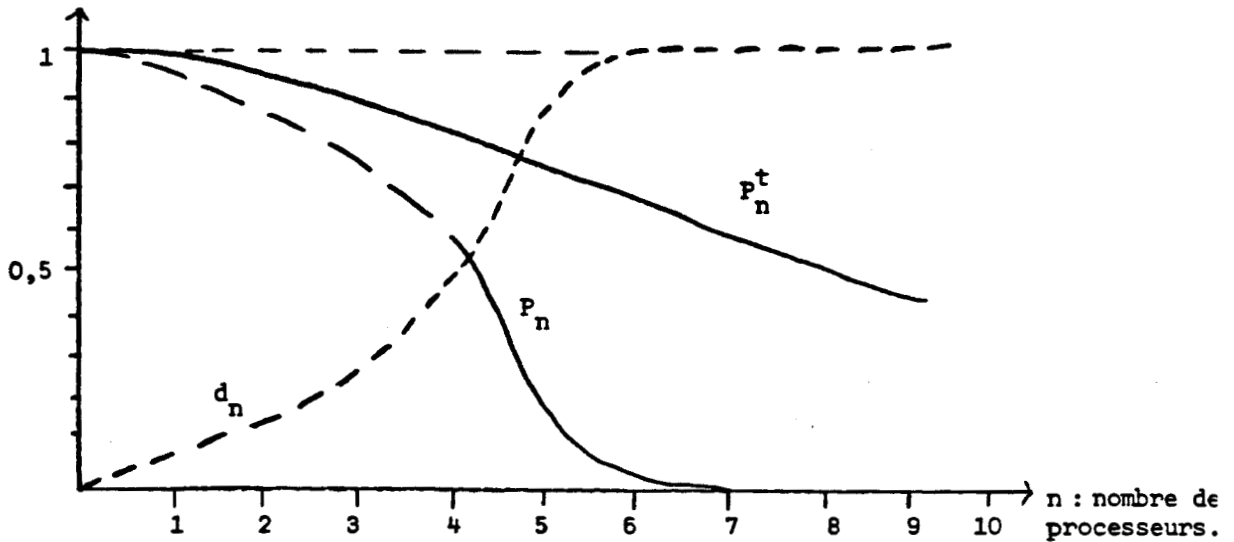


Figure 8.b. : Délai d'accès, performance et performance totale d'un système multiprocesseur utilisant un bus partagé et synchrone [HOR 77].

$d_n$  : Délai d'accès/temps total d'activité.

$P_n$  : Performance ou taux d'activité moyenne de chaque processeur.

$P_n^t$  : Performances totales des processeurs.

2.b. Le cross-bar

Topologiquement, le cross-bar se présente sous la forme d'une matrice  $N \times M$  de commutateurs - figure 9 -, reliés entre-eux par des liaisons bidirectionnelles.

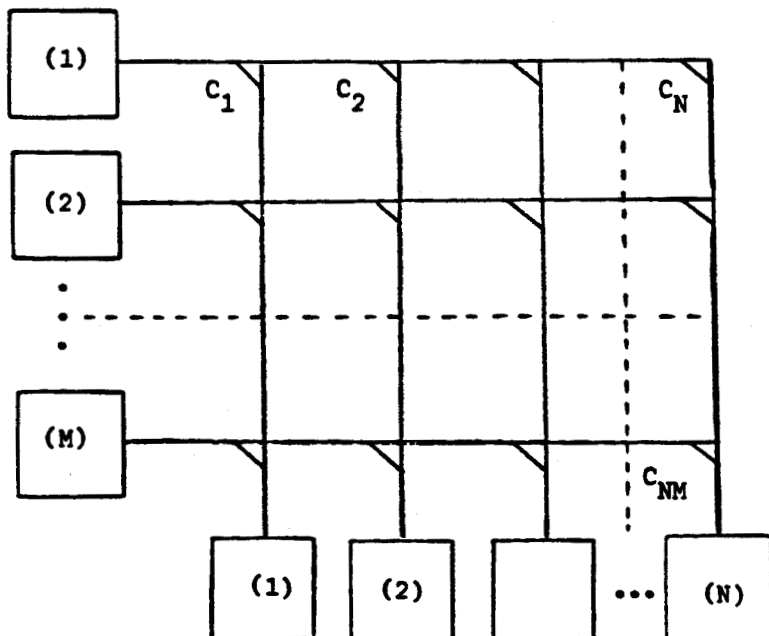


Figure 9 : Cross-bar  $N \times M$ .



Ordinairement placé entre un ensemble de processeurs et un ensemble de bancs de mémoires, le cross-bar permet d'établir une liaison entre un couple quelconque d'éléments inactifs ( $P_i, M_j$ ), sans modification des connexions existantes : c'est donc un outil sans blocage. Il dispose d'une capacité fonctionnelle maximale, car il permet l'établissement de toute correspondance biunivoque, parmi les  $N!$  pour  $N = M$ , entre l'ensemble des processeurs et l'ensemble des modules de la mémoire.

Si le cross-bar constitue l'outil de communication le plus puissant, il est également le plus coûteux. Sa complexité, évaluée en nombre de commutateurs, croît, en effet, en  $O(N^2)$ . Son utilisation est alors limitée aux systèmes de taille relativement réduite.

Notons enfin, que les deux modes de contrôle d'accès, centralisé [AHR 77] et réparti, sont possibles, mais c'est ce dernier mode qui est le plus souvent utilisé.

Le calculateur BSP "Burroughs Scientific Processor", composé de 16 unités arithmétiques et d'une mémoire parallèle de 17 modules, utilise le cross-bar comme outil de communication. Le calculateur Cmp utilise, lui aussi, le cross-bar pour interconnecter ses 16 processeurs PDP 11 et ses 16 modules mémoire.

## 2.c. Les bus multiples et les multi-étages

Deux approches principales ont inspiré les différentes études et réalisations en matière d'outils de communications. Chacune d'entre-elles opte pour l'une des deux solutions extrêmes précédentes, Bus ou Cross-bar, et se propose d'en réduire les inconvénients.

La première approche conserve le Bus partagé, l'accroissement des performances de cet outil de communication est obtenu par :

- Une multiplication du nombre de voies dans le système,
- et/ou, une organisation appropriée des unités du système autour de chacune de ces voies.

La seconde approche, quant à elle, se propose de réduire la complexité du Cross-bar, tout en tolérant une faible dégradation de sa puissance fonctionnelle.

L'une et l'autre des deux approches ont donné lieu à deux classes d'outils de communication : les bus-multiples et les multi-étages. Nous nous proposons d'étudier par la suite chacune de ces deux classes.

### 2.c.1. Les bus-multiples

Il existe essentiellement deux configurations de bus-multiples.

- i) La première, découle immédiatement de la structure à bus partagé. Elle consiste, tout simplement, à en multiplier le nombre - figure 10 -.

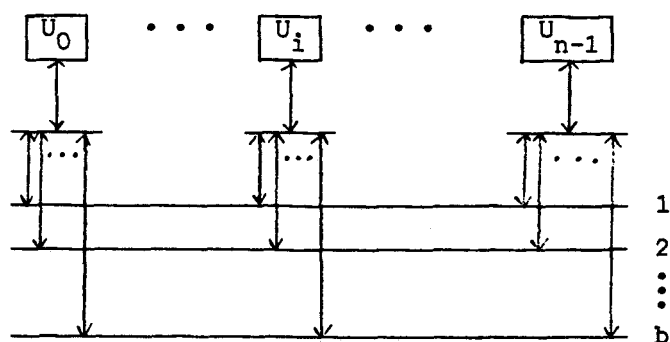


Figure 10 : Interconnexion par Bus-multiples.

$$1 \leq b \leq \min\left(\frac{n}{2}, a\right). \quad a : \text{nombre d'unités actives.}$$

Cette configuration permet à deux unités quelconques de communiquer entre-elles, en utilisant l'un quelconques des  $b$  Bus. Elle présente l'avantage d'autoriser jusque  $b$  communications simultanées. Le contrôle d'accès à ces bus doit être centralisé afin d'éviter les risques d'interblocage [LVA 82]. Elle a aussi l'avantage, contrairement au bus et au cross-bar, de permettre un fonctionnement en mode dégradé, la mise hors service de l'un des  $b$  bus laisse accessibles l'ensemble des unités du système.

Le coût de l'outil avoisine celui du cross-bar lorsque le nombre  $b$  des bus devient égal à  $\min(\frac{n}{2}, a)$ , il ne présente d'intérêt que lorsque ce nombre est faible devant  $\min(\frac{n}{2}, a)$ .

ii) La seconde configuration est obtenue en effectuant une partition sur l'ensemble des unités du système. On regroupe, dans un même élément de la partition, les unités qui ont entre-elles un degré d'interaction très élevé : un processeur et ses ressources privées par exemple. A chaque élément de la partition, appelé grappe, on associe un bus qui assure les communications entre ses unités. Les communications entre unités appartenant à deux partitions distinctes sont assurées par un Bus global : le bus inter-grappes - figure 11 -.

Eléments de  
la partition : 1

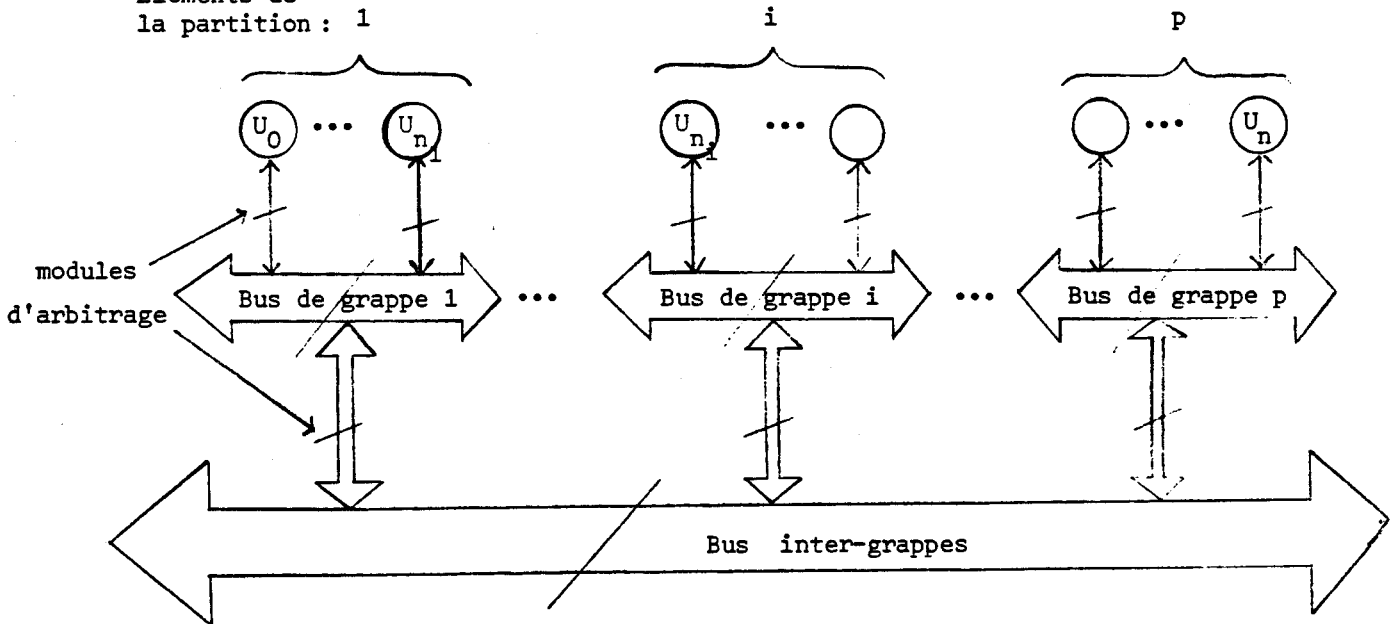


Figure 11

L'accroissement des performances dans cette configuration est obtenu :

\* d'une part, par la réduction de la charge de chaque bus de grappe, qui peut alors assurer un délai d'accès plus rapide toujours par rapport au bus unique ; et de la longueur physique, également plus réduite, permet d'accroître son débit potentiel,

\*\* d'autre part, par la diminution de la fréquence des communications globales.

Les communications globales, entre unités appartenant à deux grappes distincts, sont établies après trois accès successifs : d'abord au bus de grappe de l'unité source, au bus inter-grappes ensuite, et enfin au bus de grappe de l'unité destinataire. Les délais d'accès pour les communications globales sont donc assez élevés.

Le contrôle d'accès peut être défini en associant un module d'arbitrage à chaque bus de grappe, chargé de reconnaître et de distinguer entre-elles les requêtes locales et les requêtes externes. Un arbitre est également attaché au bus inter-grappes et vers lequel est émise toute requête externe, reconnue par l'un quelconque des arbitres des Bus de grappe. Le contrôle est donc localement centralisé mais globalement réparti.

Cette configuration est celle utilisée par le calculateur parallèle CM\*.

Remarque : D'autres configurations, obtenues à partir de combinaisons des deux précédentes, sont possibles. En effet, la technique de partition peut être appliquée de nouveau aux éléments de la première partition. Le schéma obtenu, est alors celui d'un outil de communication configuré en Arbre à plusieurs niveaux - figure 12 -

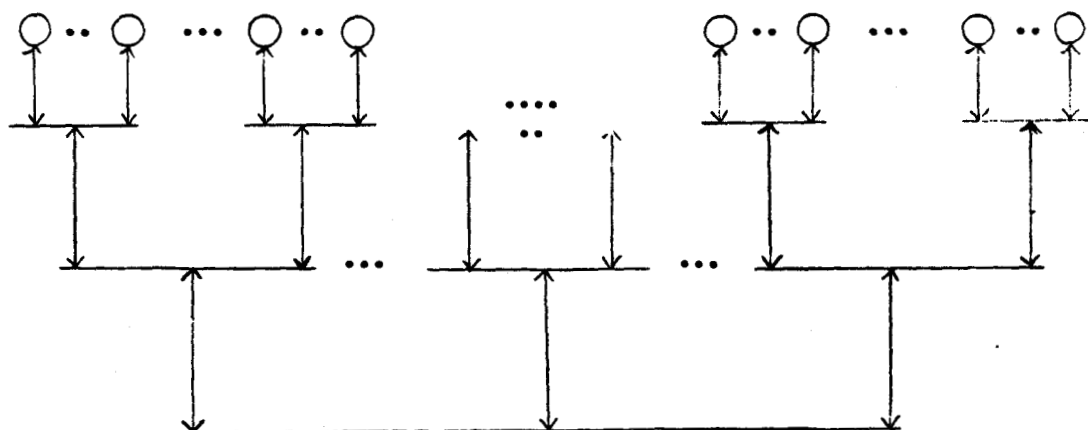


Figure 12 : Arbre à trois niveaux de communication.

Une autre solution est d'appliquer la technique de multiplication de bus au bus inter-grappes - figure 13 -.

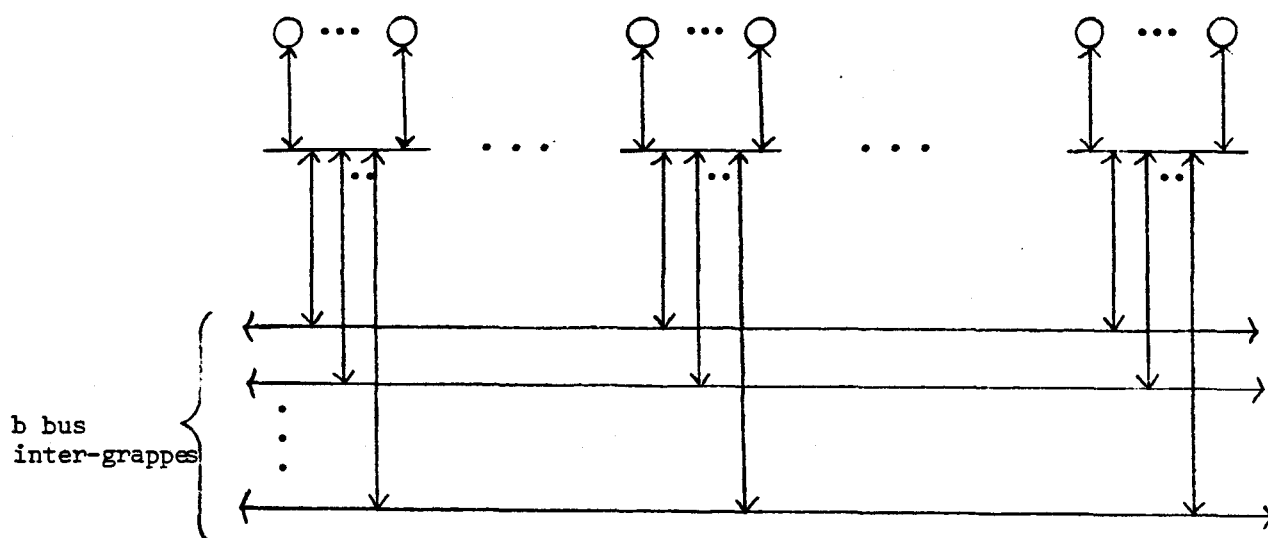


Figure 13

### 2.c.2. Les multi-étages

#### 2.c.2.a. Le réseau de CLOS sans blocage

La réduction de la complexité du cross-bar a été soulevée au départ pour répondre à des besoins en commutation téléphonique, c'est ainsi que CLOS proposa dès l'année 1953, un réseau  $N \times M$  composé de trois étages de commutateurs - figure 14 - :

- Un étage d'entrée comportant  $r_1$  commutateurs identiques  $n_1 \times m$ , chacune des entrées du réseau est reliée à un et un seul commutateur de cet étage,  $N = n_1 r_1$ .
- Un étage de sortie comportant  $r_2$  commutateurs identiques  $m \times n_2$ , les  $n_2 r_2$  sorties de cet étage sont également celles du réseau,  $M = n_2 r_2$ .
- Enfin, un étage de milieu, composé de  $m$  commutateurs également identiques  $r_1 \times r_2$ . Chacun des  $m$  commutateurs est relié, d'une part, à une seule

sortie de chacun des commutateurs de l'étage d'entrée et, d'autre part, à une seule entrée de chacun des commutateurs de l'étage de sortie.

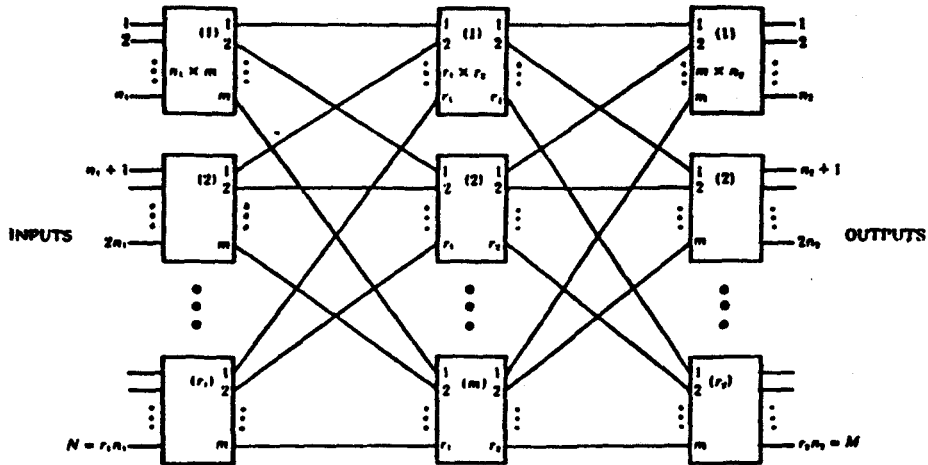


Figure 14 : Le réseau CLOS.

On montre [BEN 65] que, si la condition  $m \geq n_1 + n_2 - 1$  est vérifiée, le réseau CLOS est sans blocage, et que dans le cas particulier où  $N = M$  et  $m = n_1 = n_2 = n$ , sa complexité tend asymptotiquement vers  $(2n-1)(2n + \frac{N^2}{n^2})$ . Le réseau CLOS est donc aussi fonctionnel, mais sensiblement moins complexe, que le cross-bar [MAR 77].

Notons que le coût du réseau CLOS sans blocage malgré qu'il soit inférieur à  $N^2$ , reste cependant trop élevé. A cela s'ajoute le fait que sa commande est quelque peu difficile à mettre en oeuvre, en partie, en raison du nombre élevé des chemins initialement libres et que peut emprunter toute demande de connexion [BRH 83]. C'est pour ces raisons que les réseaux CLOS sans blocage n'ont pas fait l'objet de réalisations notables en tant qu'outils de communication dans les calculateurs.

### 2.c.2.b. Un réseau de CLOS réarrangeable : BENES

Lorsque la condition  $m \geq n_2$  est vérifiée [BEN 65], le réseau CLOS devient réarrangeable. Les réseaux CLOS réarrangeables ont été, contrairement aux réseaux CLOS sans blocage, intensivement étudiés [WAK 68], [OPT 71], [LEN 78], [NAS 80].

### 2.c.2.b.1. Topologie

La topologie du réseau BENES, se définit récursivement à partir du réseau CLOS avec  $N=M$  et  $n_1 = n_2 = m = 2$ , la condition de réarrangeabilité :  $m \geq n_2$  étant vérifiée, de la manière suivante :

Les deux commutateurs de l'étage du milieu du réseau CLOS  $N \times N$  - figure 15 -, considérés comme des réseaux CLOS  $\frac{N}{2} \times \frac{N}{2}$ , peuvent, à leur tour, être décomposés en trois étages. Au bout de  $\log_2 N - 1$  itérations, le réseau CLOS  $N \times N$  est entièrement composé de commutateurs  $2 \times 2$ , organisés en  $2 \log_2 N - 1$  étages - figure 16 - avec  $N = 8$ .

La complexité du réseau BENES est égale à  $N \log_2 N - \frac{N}{2}$ , elle croît donc en  $O(N \log_2 N)$ .

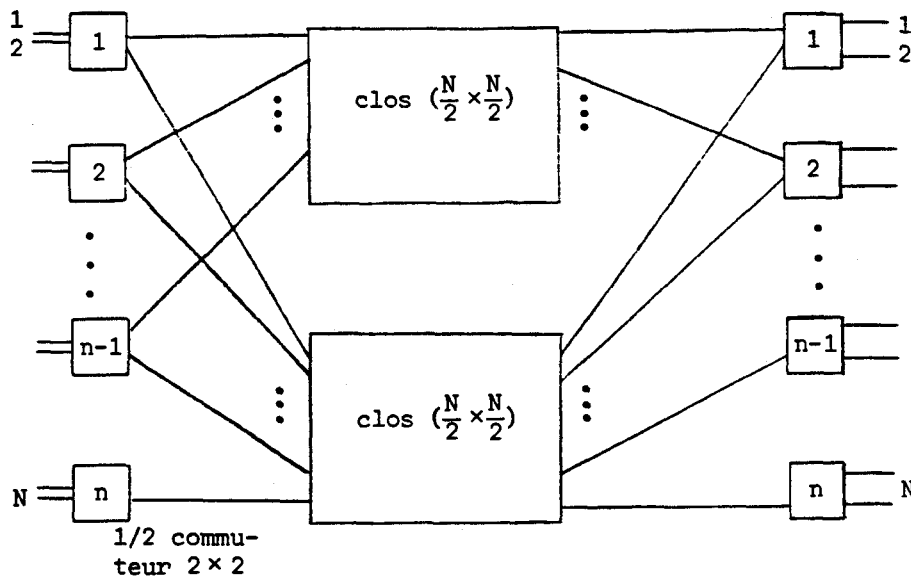


Figure 15 : CLOS ( $N \times N$ )

$$\left\{ \begin{array}{l} n_1 = n_2 = m = 2, N = 2^n \\ r_1 = r_2 = \frac{N}{2} \end{array} \right.$$



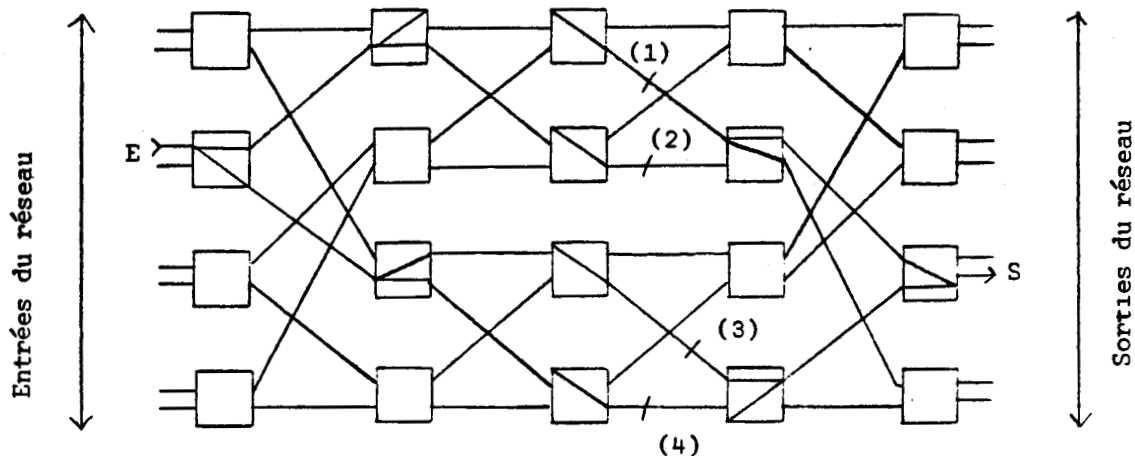


Figure 16 : Le réseau BENES ( $N = 8$ ) avec l'ensemble des chemins ( $4 = \frac{N}{2}$ ) reliant E et S.

#### 2.c.2.b.2. Commande

La commande du réseau BENES repose sur la notion de vecteur de routage. En effet, tout chemin dans le réseau est défini par les états des  $2n-1$  commutateurs qui le composent. Le vecteur de routage est par conséquent composé de  $2n-1$  éléments. Chaque élément de ce vecteur définit l'état que doit prendre le commutateur de l'étage correspondant.

Ainsi, le  $i^{\text{ème}}$  commutateur traversé par le vecteur de routage, relie l'entrée à la sortie haute, lorsque la valeur du  $i^{\text{ème}}$  composant est "0", et à la sortie basse, lorsqu'elle est "1" - figure 17 -

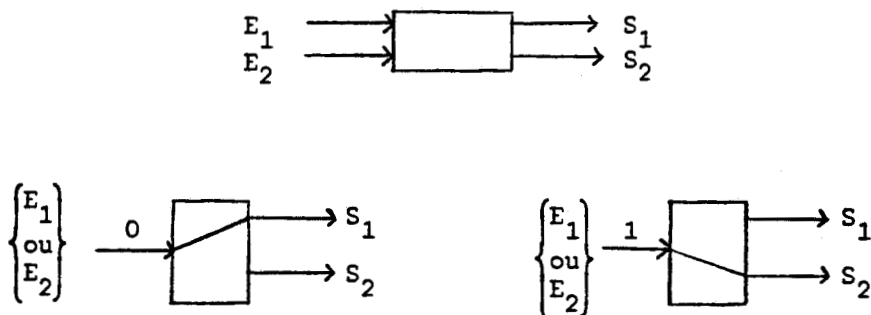


Figure 17 : Le commutateur du réseau BENES.

Il existe dans le réseau BENES  $\frac{N}{2}$  chemins possibles reliant une entrée quelconque à toute sortie du réseau - figure 16 -. L'algorithme de commande doit donc sélectionner un chemin libre : ce dernier existe toujours, lorsque les éléments à relier sont inactifs, puisque le réseau BENES est réarrangeable. La commande nécessite donc une connaissance instantanée, pour chaque demande, des états de l'ensemble des chemins possibles dans le réseau, condition qui exige que la commande soit centralisée. D'autre part, les demandes de connexion doivent être émises de manière synchrone. Il en est de même pour les opérations d'établissement et de libération des chemins, les échanges doivent donc être de durée fixe.

Les communications internes aux calculateurs SIMD répondent à de telles exigences :

- communications synchrones
- échanges de longueur fixe.

C'est pourquoi, l'essentiel des algorithmes de commande du réseau BENES [OPT 71], visent son utilisation dans les calculateurs SIMD. Le réseau BENES, muni de ces algorithmes, est un outil aussi fonctionnel que le cross-bar ; il permet l'établissement de l'une quelconque des  $N!$  permutations possibles. En revanche, son temps de commande est élevé ; la complexité de l'algorithme de commande croît, en effet, en  $N \log_2 N$ , à ce temps s'ajoute le délai d'établissement des chemins qui correspond à la traversée séquentielle des  $2n-1$  étages par les vecteurs de routage.

Afin de réduire le temps de commande du réseau BENES, LENFANT [LEN 78] a proposé une classification en cinq familles de permutations, jugées comme étant les plus fréquemment utilisées dans les algorithmes vectoriels. Pour chacune de ces cinq familles, un algorithme de commande du réseau BENES, moins complexe que l'algorithme général, a été développé.

D. NASSIMI et S. SAHNI [NAS 80] ont, pour leur part, proposé une implémentation parallèle de l'algorithme de commande général du réseau BENES. Cette version parallèle divise par  $N$  la complexité de l'algorithme de commande.

Enfin, une extension de l'utilisation du réseau BENES à un environnement asynchrone a été récemment proposé par AUGUIN et BOERI [AUB ]. Cependant, cette extension ne tient pas compte de la propriété de réarrangeabilité du réseau qui devient un réseau avec blocage. Sa mise en oeuvre nécessite alors un moyen pour la détection des conflits et une politique de gestion des demandes conflictuelles.

### 2.c.2.c. Les réseaux BANYAN

#### 2.c.2.c.1. Définition

Nous reprenons ici la définition informelle des réseaux BANYAN, donnée par PATEL [PAT 81] :

Un réseau d'interconnexion  $R$  - Figure 18 - est qualifié de BANYAN si et seulement si, il existe un et un seul chemin reliant entre-eux un processeur  $P_i$  et un module mémoire  $M_j$  quelconques.

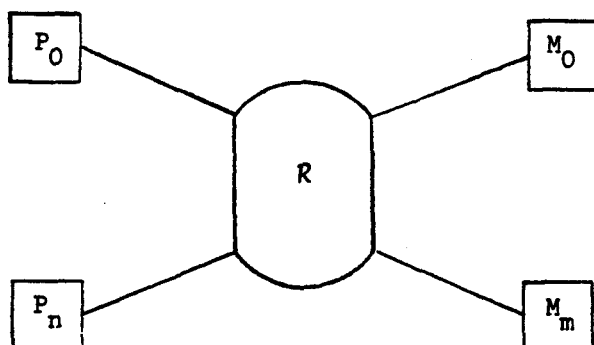


Figure 18

Dans le cadre de cette étude, nous nous intéressons à un sous-ensemble particulier des réseaux BANYAN, à savoir, les réseaux organisés en  $\log_2 N$  étages de commutateurs  $2 \times 2$ . Ces réseaux ont, en effet, fait l'objet de nombreuses études que ce soit au niveau de leurs topologies ou de leur commande. Nous citons, par exemple, les topologies BASELINE [WUF 80], OMEGA [LAW 75], CUBE [SIM 81]. Ces topologies sont toutes de complexités équivalentes, elles comportent en effet toutes  $\frac{N}{2} \log_2 N$  commutateurs, et diffèrent uniquement par les interconnexions de leurs étages et/ou des unités d'entrée (resp. sortie) avec le premier (resp. dernier) étage.

Nous nous limitons par la suite à la présentation de deux topologies  
BANYAN : BASELINE et OMEGA.

### 2.c.2.c.2. Le réseau BASELINE

#### 2.c.2.c.2.a. Topologie

BASELINE se définit récursivement [WUF 80] de la manière suivante : D'abord décomposé en deux étages de commutateurs, le premier est exclusivement composé de commutateurs  $2 \times 2$ , le second comporte deux commutateurs  $\frac{N}{2} \times \frac{N}{2}$ . Les deux sorties de chaque commutateur du premier étage sont reliées aux deux commutateurs du second étage - figure 19 -. Considérés comme des réseaux BASELINE  $\frac{N}{2} \times \frac{N}{2}$ , les deux commutateurs du second étage sont à leur tour décomposés en deux étages selon le même procédé que précédemment. Après  $\log_2 N$  itérations, on aboutit au réseau BASELINE composé entièrement de commutateurs  $2 \times 2$  - figure 20,  $N = 16$  -.

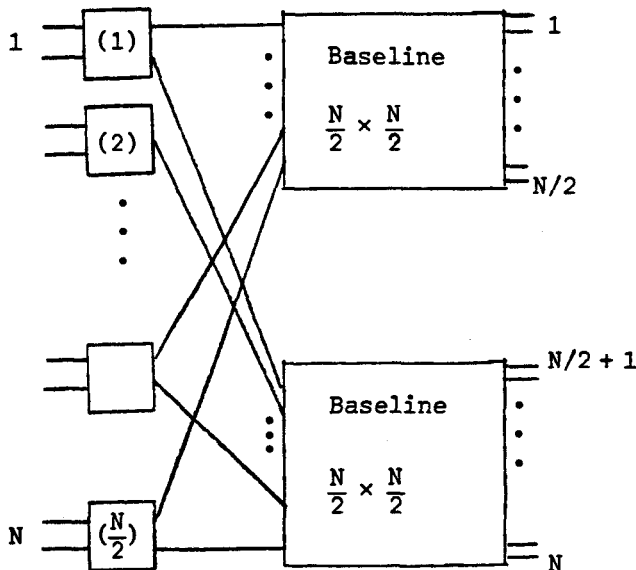


Figure 19 : Topologie Baseline  $N \times N$  : première décomposition

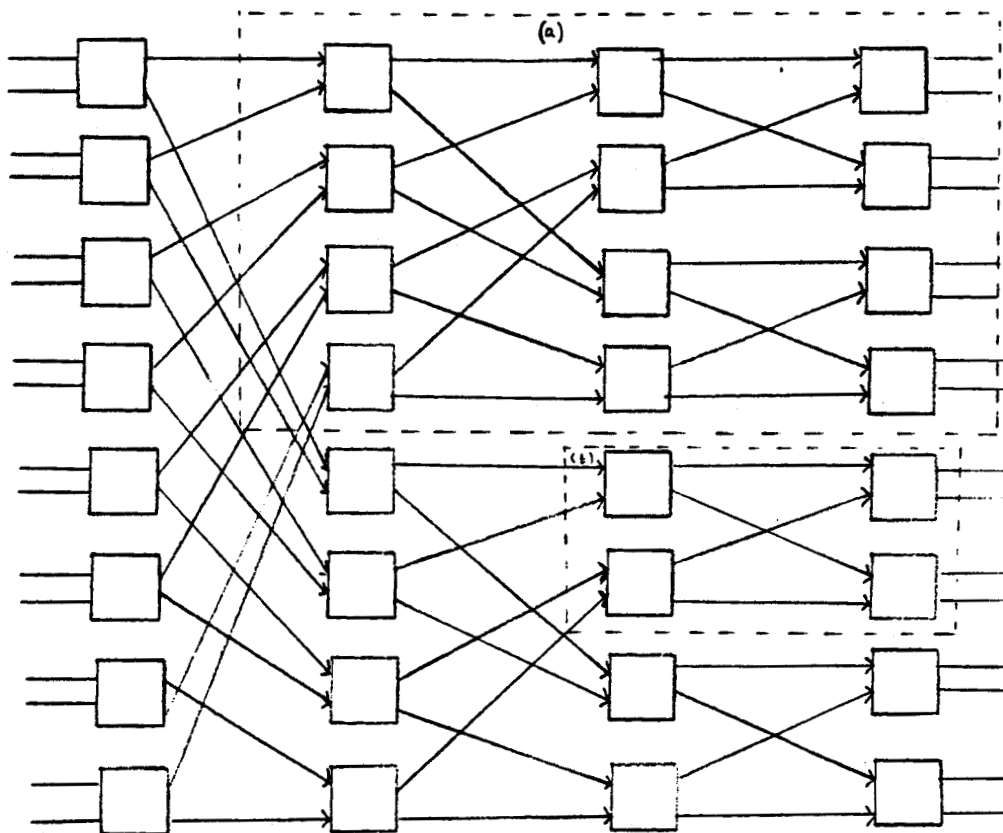


Figure 20 : Topologie Baseline  $16 \times 16$

(a) Baseline  $8 \times 8$ , (b) Baseline  $4 \times 4$ .

#### 2.c.2.c.2.b. Commande

La commande du réseau BASELINE, s'appuie sur une numérotation appropriée des  $N$  sorties du réseau. Un arbre binaire, dont la racine est une entrée quelconque du réseau et les feuilles sont ses  $N$  sorties, est construit de la manière suivante - figure 21 - :

- Les deux descendants d'un noeud père sont les deux entrées des deux commutateurs de l'étage suivant auxquelles ce noeud père est relié.
- Un lien de l'arbre est étiqueté "0", s'il traverse un commutateur, noeud père, par sa sortie haute, et "1" s'il le traverse par sa sortie basse.

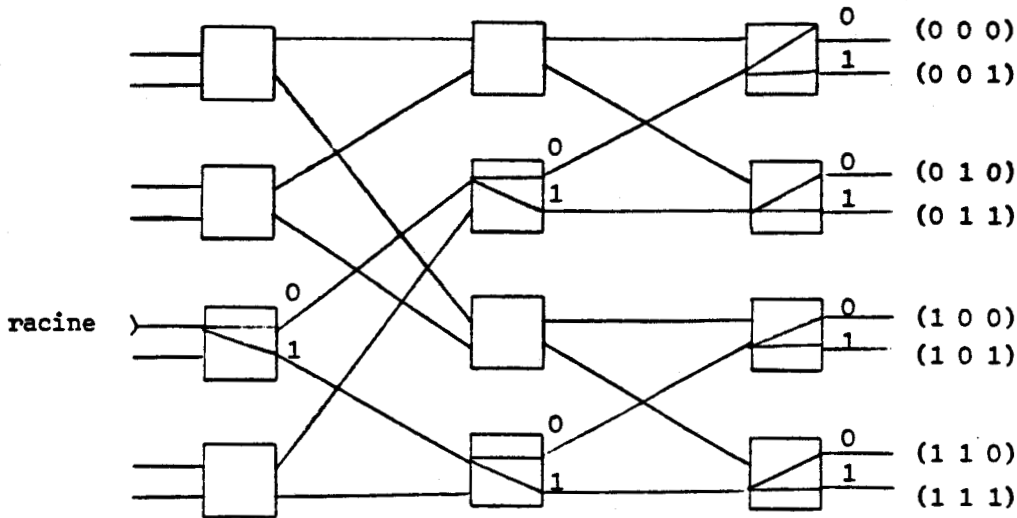


Figure 22 : Arbre de numérotation dans BASELINE (N = 8).

Le numéro d'une sortie est le nombre dont la représentation binaire, est obtenue en concaténant la suite des bits figurant sur les liens qui relient l'entrée racine à cette sortie feuille de l'arbre.

Notons que la numérotation obtenue, est indépendante de l'entrée choisie comme racine de l'arbre. Elle démontre par ailleurs, l'existence et l'unicité du chemin reliant une entrée à toute sortie du réseau.

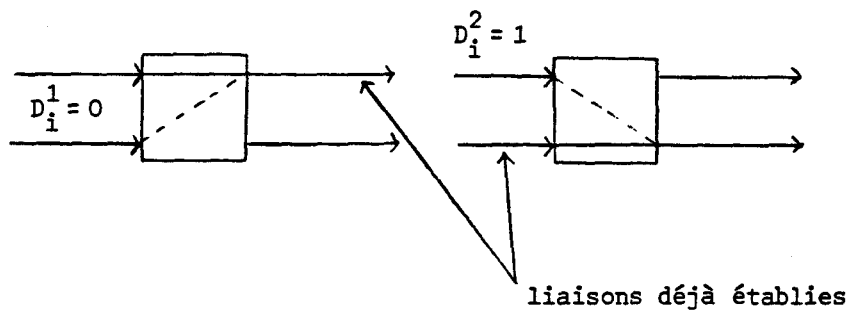
Cette numérotation permet, d'autre part, de déduire facilement la commande du réseau. En adoptant une définition des états du commutateur, identique à celle du commutateur du réseau BENES - figure 17 -, le vecteur de routage n'est autre que le numéro de la sortie destinataire. La commande du réseau BASELINE est donc entièrement répartie ; chaque entrée établit une connexion avec une des sorties du réseau en envoyant le numéro de cette dernière dans le réseau.

BASELINE, et d'une manière générale tout réseau à  $\log_2 N$  étages est un réseau avec blocage. En effet, chaque commutateur peut prendre deux états, un réseau à  $\log_2 N$  étages comporte  $\frac{N}{2} \log_2 N$  commutateurs. Le nombre d'états du réseau, et par conséquent le nombre de correspondances biunivoques qu'il peut établir, est donc égal à  $2^{\frac{N}{2} \log_2 N}$  ( $= \sqrt{N}^N$ ). Ce nombre est nettement plus faible que les  $N!$

correspondances possibles qui peuvent être effectuées entre les  $N$  entrées et les  $N$  sorties. Il est donc nécessaire de définir un mécanisme qui permet la détection des situations de conflit ou de blocage et leur résolution.

### a) Situation de blocage

La situation de blocage est liée au caractère asynchrone des demandes d'établissement des liaisons dans le réseau. Elle se produit lorsque l'état courant d'un commutateur, déjà traversé par une liaison, n'est pas compatible avec la demande qui lui est soumise.



Situation de blocage  $D^1$  et  $D^2$  sont bloquées.

3 solutions sont alors possibles :

- i) La nouvelle demande reste maintenue jusqu'à ce que la liaison blocante soit libérée, elle occupe inutilement une portion de chemin dans le réseau.
- ii) La seconde solution, consiste à renvoyer un acquittement négatif à l'entrée demandeur qui retire alors sa demande. Ceci à l'avantage de ne pas occuper inutilement une portion de chemin du réseau et de permettre à d'éventuelles liaisons de s'établir en utilisant la portion de chemin ainsi libérée. En revanche, cela nécessite que la demande retirée soit ultérieurement soumise de nouveau au réseau.

iii) La dernière solution consiste à ne retirer la demande bloquée que, si, ou lorsque, la liaison blocante est "verrouillée". Une liaison est verrouillée quand elle est entièrement établie. Cette technique présente le même avantage que la seconde solution en plus, elle élimine le risque de retirer une demande à cause d'une liaison qui n'est que partiellement établie et qui par conséquent peut être retirée à son tour.

#### b) Situation de conflit

Il y a conflit au niveau d'un commutateur lorsque deux demandes lui sont simultanément soumises et que ces demandes sont incompatibles entre elles - figure 23 -. Dans ce cas un arbitrage est à effectuer entre les deux demandes, sachant qu'une seule peut continuer sa progression dans le réseau.

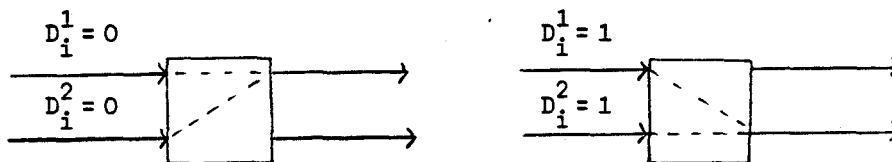


Figure 23 : Conflit entre  $D^1$  et  $D^2$ .

Dans un contexte où il n'existe pas de priorités logiques entre les entrées, on peut effectuer l'arbitrage entre les demandes conflictuelles selon l'une des techniques suivantes :

- i) Sélection aléatoire de l'une des deux demandes et blocage de l'autre.
- ii) Les deux entrées de chaque commutateur ont alternativement la priorité, l'une par rapport à l'autre, par exemple lors d'un cycle horloge réseau toutes les entrées hautes sont prioritaires ; lors du cycle horloge suivant ce sont les entrées basses des commutateurs qui ont la priorité.

Une fois le conflit résolu, seule une des deux demandes continue sa progression dans le réseau. La seconde est bloquée et l'une des trois solutions précédemment citées, peut alors lui être appliquée.



Notons enfin, que la mise en oeuvre de l'un quelconque de ces mécanismes accroît la complexité de la logique de contrôle associée à chaque commutateur.

### 2.c.2.c.3. Le réseau OMEGA

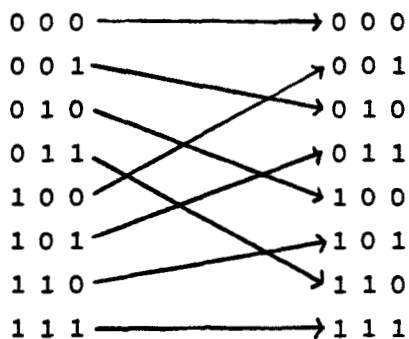
#### 2.c.2.c.3.a. Topologie

La description de la topologie OMEGA exige préalablement la numérotation de 0 à N-1, des N entrées et des N sorties du réseau ainsi que de chacun des  $\log_2 N = n$  étages.

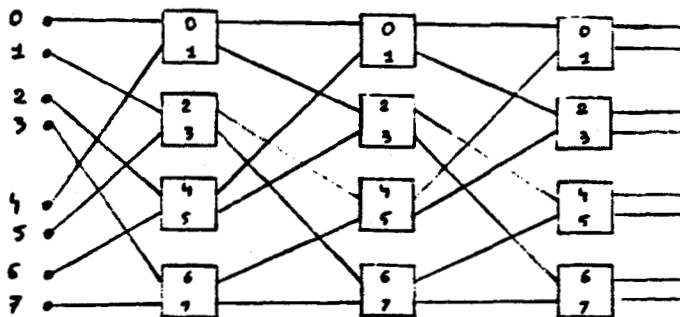
La topologie OMEGA est obtenue à partir d'une interconnexion dite "Perfect Shuffle", entre :

- Les entrées du réseau et les entrées du premier étage.
- Les sorties d'un étage et les entrées de l'étage suivant.

En considérant leurs représentations binaires, l'interconnexion "Perfect Shuffle" relie  $X = x_{n-1} x_{n-2} \dots x_0$  à  $Y = x_{n-2} \dots x_0 x_{n-1}$ . Pour  $N = 8$ , on obtient l'interconnexion en "Perfect Shuffle" suivante :



Il en résulte la topologie OMEGA suivante :



### 2.c.2.c.3.b. Commande

La commande répartie et asynchrone du réseau OMEGA est identique à celle décrite pour le réseau BASELINE et soulève les mêmes problèmes de blocage.

## 2.d. Comparaison des réseaux multi-étages

Les principaux éléments d'évaluation des réseaux sont :

- i) Le taux de parallélisme, qui traduit le nombre de liaisons pouvant être simultanément établies.
- ii) Le délai de propagation.
- iii) Le coût en composants.

### 2.d.1. Comparaison des réseaux BANYAN

Si on compare entre-eux les différentes topologies BANYAN, force est de constater qu'ils présentent tous les mêmes caractéristiques :

- i) Ce sont tous des réseaux avec blocage. Le parallélisme maximal ne peut être atteint que si la correspondance à établir entre l'ensemble des  $N$  entrées et celui des  $N$  sorties, fait partie des  $\sqrt{N}^N$  correspondances autorisées par le réseau. Notons, cependant, que si l'ensemble de ces correspondances a une cardinalité constante ( $\sqrt{N}^N$ ), quelle que soit la topologie, l'ensemble de ces correspondances diffère d'une topologie à l'autre.
- ii) Les délais d'établissement d'un chemin, ainsi que le temps de propagation sont proportionnels à  $\log_2 N$ . C'est la longueur, évaluée en nombre de commutateurs, de tout chemin dans le réseau.
- iii) Le coût total, évalué en nombre de commutateurs, est de  $\frac{N}{2} \log_2 N$ .

### 2.d.2. Comparaison du réseau BENES aux réseaux BANYAN

i) La commande asynchrone du réseau BENES, le transforme en un réseau avec blocage. Mais la probabilité de blocage est plus faible que dans les réseaux BANYAN. Une étude en simulation [AUB ] du réseau BENES et d'un réseau BANYAN : OMEGA, a permis de montrer que cette probabilité de blocage est réduite de 12 %.

ii) Le délai de propagation est deux fois plus élevé dans le réseau BENES que dans un réseau BANYAN.

iii) Il en est de même pour le coût. Le réseau BENES utilise pratiquement deux fois plus de commutateurs, que les réseaux BANYAN.

### 3. Performance, complexité et environnement technologique

Nous avons présenté dans §2.c.2.d, trois éléments pour l'évaluation des réseaux multi-étages :

- i) Le taux de parallélisme,
- ii) le délai de propagation, et
- iii) le coût.

Si le premier élément dépend plutôt de la topologie d'interconnexion et de sa commande, il importe, par contre, de placer les deux derniers dans leur contexte technologique.

L'unité qui a servi de base pour le calcul des délais de propagation et de la complexité a été le commutateur  $2 \times 2$ . Le choix de cette unité ne peut se justifier que si on se place dans un environnement d'intégration où seul un commutateur  $2 \times 2$  est implanté par module physique - boîtier -. Il en résulte que :

- d'une part, le temps de propagation d'un signal est prédominé par le délai de traversé d'un module physique et donc d'un commutateur. Les délais de propagation inter-modules sont négligés. Le délai de propagation le long d'un chemin est donc directement proportionnel au nombre de commutateurs qui le composent.

- et d'autre part, le coût du réseau est, lui aussi, dominé par le coût d'un module physique, et est donc directement proportionnel au nombre de commutateurs qui le composent.

Notons que les réseaux multi-étages ont été étudiés et développés en tenant compte de ces considérations technologiques. La réduction de coût revenait alors à réduire le nombre de commutateurs dans le réseau.

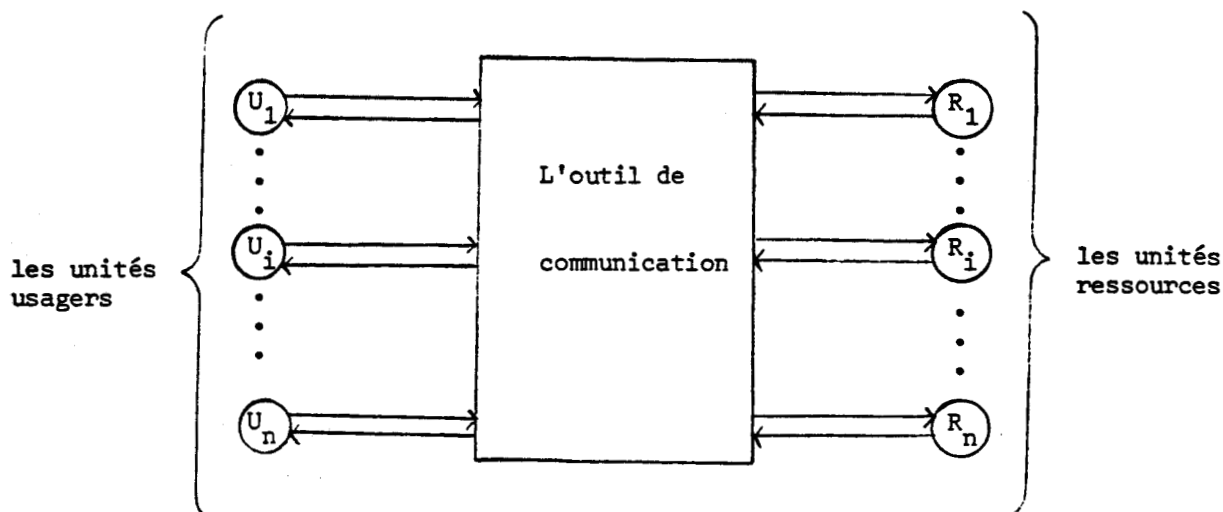
Dans un environnement d'intégration à très haute échelle - VLSI -, plusieurs commutateurs  $2 \times 2$  peuvent être implantés dans un même module [FRA 81, MAS 81, SIM 81]. Cet environnement introduit également de nouvelles contraintes et présente des limites dont il est nécessaire de tenir compte lors de la définition d'une topologie, et de l'évaluation de son coût et de ses performances.

Le nombre maximal de broches, évalué actuellement à une centaine par module est la principale limite à l'intégration VLSI d'un circuit [FWT 82]. Un réseau est d'autant mieux adapté à une intégration VLSI, que son rapport  $\frac{\text{nombre de broches}}{\text{complexité logique}}$  est faible, et que son coût de fabrication est d'autant plus réduit que sa topologie est régulière.

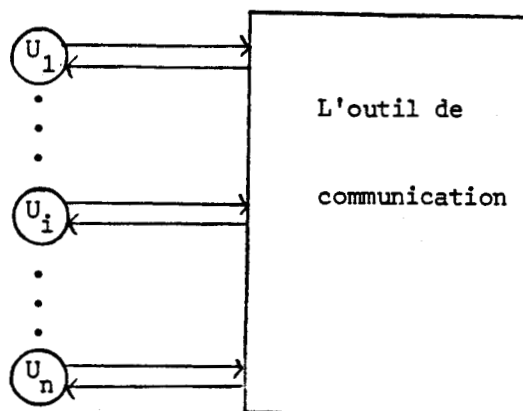
Une étude analytique [FRA 81], basée sur un modèle spacio-temporel, montre que le coût, évalué en surface d'intégration, des topologies BANYAN et cross-bar, croît en  $O(N^2)$  et que le délai de propagation intra-module croît en  $O(N^a(\log_2 N)^2)$ , avec  $0 < a < 1$ , pour BANYAN et en  $O(N)$  pour le cross-bar. Cette étude conclut que pour  $N < 64$ , le rapport coût-performance est sensiblement le même pour les deux topologies. La topologie BANYAN étant avec blocage, il est plus intéressant alors de concevoir des topologies dont chaque module élémentaire est un cross-bar  $4 \times 4$  par exemple, et une topologie BANYAN d'interconnexion entre ces modules [MAS 81, SIM 81, FWT 82].

#### 4. Objectifs de l'étude

Les outils présentés dans ce chapitre ont été classés en deux catégories, selon qu'ils dérivent du bus ou du cross-bar. Notons qu'un autre critère permet de les distinguer. Les réseaux dérivant du cross-bar sont des réseaux à double face - figure a -. Les réseaux dérivant du bus sont quant-à-eux, des réseaux à simple face - figure b -.



a - Réseau à double face.



b - Réseau à simple face.

Les réseaux à simple face présentent l'avantage d'être moins spécialisés que ceux à double face. Deux unités quelconques d'un réseau simple face peuvent en effet, établir une communication.

Notre travail se situe dans le cadre des réseaux à simple face et se propose de définir de nouvelles topologies de réseaux.

Nos hypothèses sont les suivantes :

- La commutation de ligne est la meilleure solution pour l'obtention de débits élevés et pour simplifier les problèmes de synchronisation.
- Des topologies d'interconnexion régulières conduisent à des densités d'intégration élevée.

Notre environnement cible est celui des systèmes MIMD, et plus généralement tout système dont les unités sont caractérisées par :

- Un fort degré de couplage.
- L'assynchronisme de leurs accès à l'outil de communication.
- Leur hétérogénéité fonctionnelle, voire physique.

Les outils proposés doivent permettre l'ajustement du parallélisme entre ses valeurs extrêmes.

\*\*\*\*\*  
\*  
\* CHAPITRE II \*  
\*  
\*\*\*\*\*

LE MODELE DE BASE : L'ANNEAU

Dans ce chapitre, nous décrivons l'outil de communication qui nous servira de modèle de base le long de cette étude.

Après définition de sa topologie d'interconnexion, nous optons, conformément au cahier des charges établi dans le précédent chapitre, pour la commutation de ligne comme stratégie de transfert. Nous proposons ensuite une décomposition en trois niveaux fonctionnels du protocole de communication. Seules les fonctions des protocoles d'interconnexion et de transfert seront développées ici. Une mise en oeuvre asynchrone et répartie du protocole d'interconnexion est d'abord proposée, différentes techniques de transfert sont ensuite décrites et leurs implications sur la structure physique du modèle de base analysées.

## 1. Définition du modèle de base

### 1.a. La topologie d'interconnexion

Le modèle de base s'appuie sur une configuration en anneau de  $n$  cellules de commutation  $\{C_i\}_{i=0}^{n-1}$  - figure 1 -.

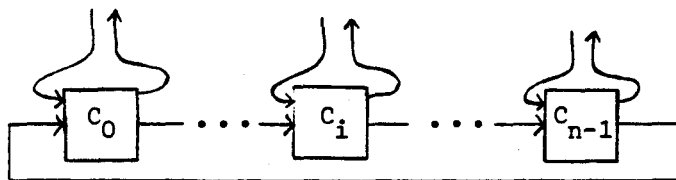


Figure 1 : Le modèle de base : l'ANNEAU.

### 1.b. La cellule de commutation

Chaque cellule  $C_i$ , est un commutateur à deux entrées  $\{e_i^1, e_i^2\}$  et deux sorties  $\{s_i^1, s_i^2\}$  unidirectionnelles - figure 2.a. -.

Une commande, notée  $!C_i$ , unique et externe, permet d'établir une correspondance biunivoque entre les ensembles  $\{e_i^1, e_i^2\}$  et  $\{s_i^1, s_i^2\}$ . Le nombre de correspondances réalisables par une cellule  $C_i$  est réduit aux deux suivantes :



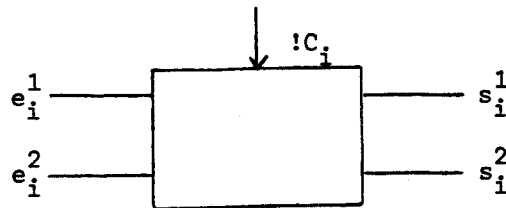
i) La première laisse transparente la cellule - figure 2.b. -

$$e_i^1 \longrightarrow s_i^1$$

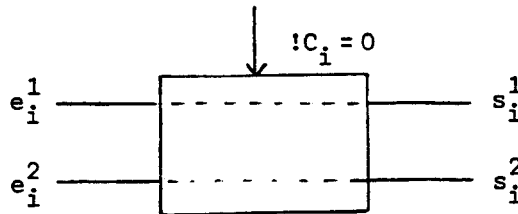
$$e_i^2 \longrightarrow s_i^2$$

ii) La seconde effectue un échange des entrées et des sorties de la cellule - figure 2.c. -

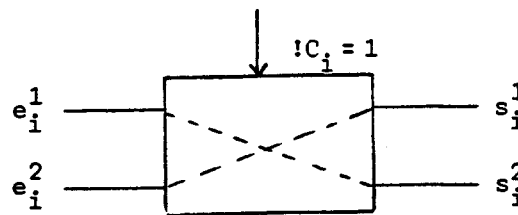
$$\begin{array}{l} e_i^1 \longrightarrow s_i^2 \\ e_i^2 \longrightarrow s_i^1 \end{array}$$



a - La cellule de commutation.



b - Mode transparent.



c - Mode échange.

Figure 2 : La cellule de commutation : ses états et sa commande.

Nous définissons l'état d'une cellule par la correspondance qu'elle réalise. Soit  $\eta$  la fonction qui à une cellule de l'ANNEAU, fait correspondre son état dans l'ensemble  $\{T (\underline{T}ransparent),  $E (\underline{E}$ change) $\}$ , et soit  $C = \{C_i\}_{i=0}^{n-1}$ , l'ensemble des cellules de l'ANNEAU.$

$$\eta : C \rightarrow \{T, E\}$$

L'ensemble des états d'une cellule étant réduit à deux, une valeur binaire de la commande  $!C_i$  suffit pour déterminer l'état que doit prendre la cellule  $C_i$ .

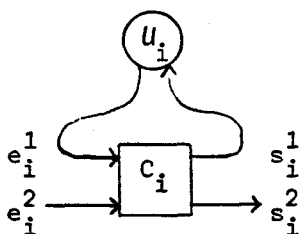
On pose par convention :

$$!C_i = 0 \Leftrightarrow \eta(C_i) = T.$$

$$!C_i = 1 \Leftrightarrow \eta(C_i) = E.$$

### 1.c. La commutation de ligne dans l'ANNEAU

Soit  $\{U_i\}_{i=0}^{n-1}$ , l'ensemble des unités interconnectées par l'ANNEAU, chaque unité  $U_i$  est reliée à la cellule  $C_i$  selon le schéma de la figure suivante :



Une connexion unidirectionnelle est établie depuis une unité  $U_i$  vers une unité  $U_j$ , et est notée  $U_i \rightarrow U_j$ , si et seulement si :

$$i) \eta(C_i) = \eta(C_j) = E, \text{ et}$$

$$ii) \text{ si } i < j \Rightarrow \forall k : i < k < j \quad \eta(C_k) = T$$

$$\text{si } i > j \Rightarrow \forall k : j < k \leq n-1 \quad \eta(C_k) = T$$

$$0 \leq k < i$$

Il résulte de cette définition que les valeurs prises par les composantes du vecteur  $(!C_0, \dots, !C_{n-1})$ , définissent l'ensemble des connexions établies entre les unités à travers l'ANNEAU.

$2^n$  valeurs différentes peuvent être prises par le vecteur  $(!C_0, \dots, !C_{n-1})$ .

Soit  $p$  le nombre des composantes non nulles du vecteur  $(!C_0, \dots, !C_{n-1})$ , alors  $p$  est le nombre de connexions unidirectionnelles établies ;  $0 \leq p \leq n$ .

Nous nous intéressons par la suite à une valeur particulière de  $p$ ,  $p = 2$ . Il existe  $\frac{n(n-1)}{2}$  configuration différentes telle que  $p = 2$ .

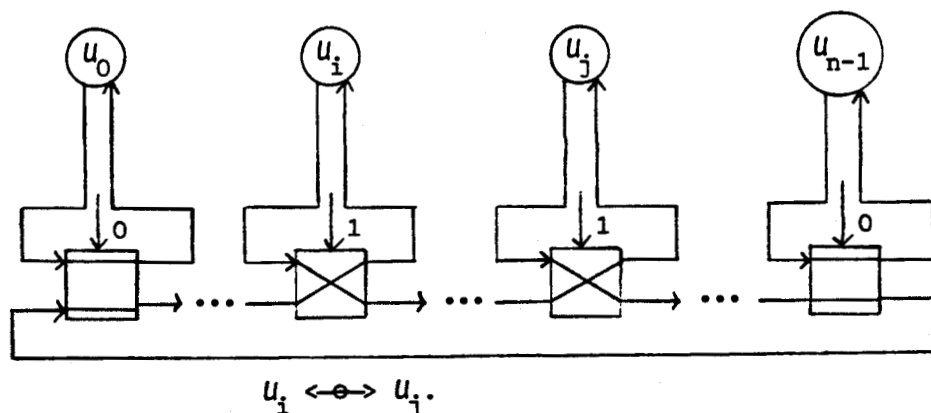
$$p = 2 \Rightarrow \exists i, j \quad i \neq j \text{ tels que } \begin{cases} !C_i = 1 \\ !C_j = 1 \end{cases} \Leftrightarrow \exists i, j \quad i \neq j$$

tels que  $\begin{cases} u_i \rightarrow u_j \\ \text{et} \\ u_j \rightarrow u_i \end{cases} \Leftrightarrow$  une liaison bidirectionnelle simultanée est établie entre les unités  $u_i$  et  $u_j$ , nous désignons par  $u_i \leftrightarrow u_j$ , une telle liaison.

Il résulte de ce qui précède

$$u_i \leftrightarrow u_j \Leftrightarrow \begin{cases} n(C_i) = n(C_j) = E, \text{ et} \\ \forall k \neq i, j \quad 0 \leq k \leq n-1 \quad n(C_k) = T. \end{cases}$$

La figure ci-dessous illustre une liaison entre  $u_i, u_j$



## 2. Le protocole de communication

Jusqu'ici, aucune hypothèse n'a été faite sur les unités interconnectées par l'ANNEAU. Nous envisageons, en effet, de présenter une définition générale de notre modèle de base. Cependant, la définition complète du protocole de communication nécessite une spécification précise et de l'activité interne et de l'activité de communication de chacune des unités. C'est la raison pour laquelle nous avons opté pour une décomposition du protocole de communication. Cette décomposition doit permettre :

- i) d'isoler les fonctions du protocole qui dépendent des unités,
- ii) de décrire les autres fonctions du protocole, sans faire d'hypothèse sur les caractéristiques physiques ou fonctionnelles de ces unités,
- iii) de proposer, en les analysant, les différentes options qui peuvent être prises pour la mise en oeuvre de chacune des fonctions du protocole.

Les unités interconnectées par l'ANNEAU sont supposées, à priori, asynchrones et fonctionnellement, voire physiquement, hétérogènes.

### 2.a. Une décomposition du protocole de communication

Nous proposons une décomposition du protocole selon les trois niveaux fonctionnels suivants :

- i) Le niveau "Echange" ; il définit les règles qui régissent les interactions directes entre les unités. Ses fonctions sont donc spécifiques aux unités et font abstraction de l'outil de communication proprement dit.
- ii) Le niveau "Interconnexion" ; il définit les mécanismes relatifs au contrôle de l'accès à l'outil de communication, d'adressage et de contrôle des connexions dans l'ANNEAU. Les fonctions de ce niveau sont

étroitement liées à la topologie de l'outil de communication et au mode de commutation adopté : la commutation de ligne dans notre cas.

- iii) Le niveau "Transfert" ; il assure la transmission d'une information élémentaire le long d'une connexion préétablie. Les fonctions à mettre en oeuvre, pour assurer une telle opération, dépendent à la fois du débit potentiel de l'ANNEAU, et des débits des unités impliquées dans une même opération de transfert.

Nous nous limitons à décrire les protocoles relatifs aux niveaux "Interconnexion" et "Transfert". Quant au niveau "Echange", il doit être possible d'implémenter différentes versions de ce niveau selon l'application envisagée et sur la base des choix des niveaux "Interconnexion" et "Transfert".

## 2.b. Le protocole d'interconnexion

### 2.b.1. Le contrôle d'accès

L'ANNEAU muni de la commutation de ligne bidirectionnelle et simultanée, est un outil purement séquentiel. A un instant donné, au plus une liaison  $u_i \leftrightarrow u_j$  peut être établie dans l'ANNEAU. Il est donc nécessaire de mettre en oeuvre un mécanisme qui résoud les conflits d'accès à l'ANNEAU.

Une méthode, généralement utilisée dans les réseaux locaux organisés en anneau, logique ou physique [COR 81, ACD 82, EST 83], consiste à faire circuler un message particulier, appelé jeton. La capture du jeton par une unité signifie que l'anneau est libre et que l'accès y est autorisé.

Dans le cas d'une implémentation de notre modèle de base, le jeton peut être un simple front, qui circule le long d'une liaison réservée.

### 2.b.2. Le contrôle d'une connexion

Le contrôle d'une connexion revient à définir les deux primitives qui permettent son établissement et sa libération.

$$\begin{aligned}
 U_i \leftrightarrow U_j &\Leftrightarrow \left\{ \begin{array}{l} \eta(C_i) = \eta(C_j) = E, \text{ et} \\ \forall k \neq i, j \quad 0 \leq k \leq n-1, \eta(C_k) = T \end{array} \right. \\
 &\Leftrightarrow \left\{ \begin{array}{l} !C_i = !C_j = 1 \\ \text{et } \forall k \neq i, j \quad 0 \leq k \leq n-1, !C_k = 0 \end{array} \right.
 \end{aligned}$$

Il s'agit donc de définir pour le contrôle d'une connexion, les règles de commande de chacune des cellules de commutation de l'ANNEAU. Nous proposons, pour cela, une solution répartie. A chaque cellule  $C_i$  est associé un module de contrôle, noté  $MC_i$ , qui définit, en plus de la commande de la cellule, le mécanisme de reconnaissance d'adresse et le protocole d'allocation et de libération de l'ANNEAU.

Pour la numérotation des cellules, nous confondrons le numéro d'une cellule avec l'adresse de l'unité à laquelle elle est attachée.

Notons  $v$  : la fonction de numérotation des cellules

$$v(C_i) = \text{adresse de } U_i. \quad \forall C_i \in C.$$

### 2.b.3. Le module de contrôle

Nous allons par la suite, définir les entités qui composent l'environnement d'interaction d'un module de contrôle  $MC_i$ , établir la liste des messages qu'il échange avec ces entités et enfin, étudier les états du module et les évènements qui provoquent leurs changements.

#### 2.b.3.a. Environnement d'interaction du module de contrôle

Nous définissons cet environnement comme étant composé de l'ensemble des entités avec lesquelles, le module de contrôle effectue un échange de messages - figure 3 -, et qui sont :

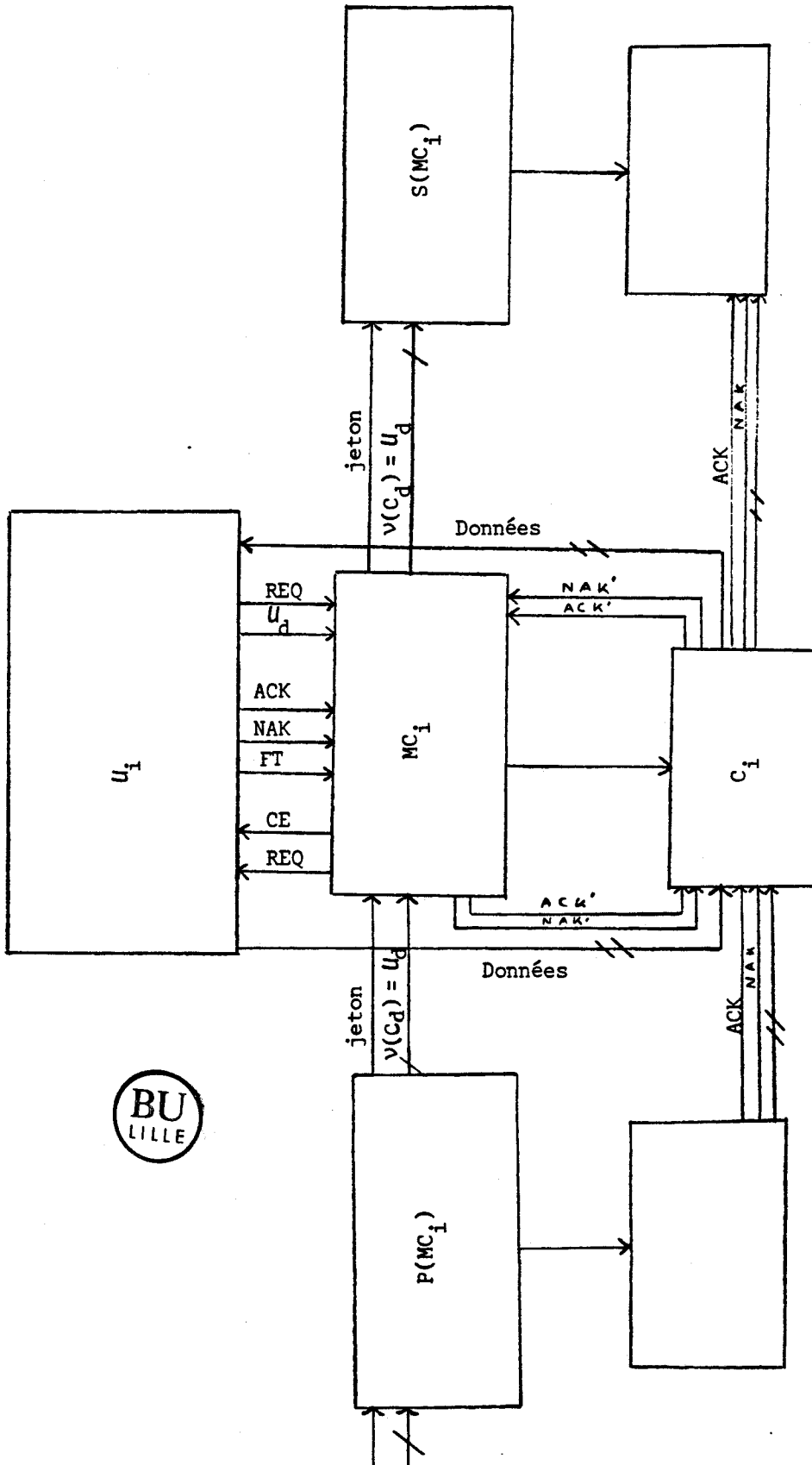


Figure 3 : Les messages échangés entre un module de contrôle et les composantes de son environnement.

- Son unité locale  $U_i$ ,
- sa cellule de commutation  $C_i$ ,
- les deux modules de contrôle prédécesseur et successeur, que nous désignons par  $P(MC_i)$  et  $S(MC_i)$ ,
- et tout autre module  $MC_j$  de l'ANNEAU, avec lequel il est impliqué dans une même opération d'interconnexion.

Nous distinguons deux types de messages, selon qu'ils sont échangés entre  $MC_i$  et son unité locale, nous parlerons dans ce cas d'évènements locaux, ou entre  $MC_i$  et un autre module  $MC_j$ , et nous parlerons alors d'évènements externes.

i) Evènements locaux : Ils sont au nombre de 6. Nous les répartissons en deux ensembles selon que leur source est le module de contrôle  $MC_i$  lui-même, ou bien son unité locale  $U_i$ .

a - {REQ, CE} : ensemble des messages émis par  $MC_i$  vers  $U_i$ .

- REQ, signale à l'unité l'occurrence d'une requête qui lui est destinée. Ce message est émis, après que le module  $MC_i$  ait reconnu l'adresse de son unité locale.

- CE, signale à l'unité que la connexion, dont la demande d'établissement lui a été auparavant soumise par cette même unité, est effectivement établie.

b - {REQ - ( $U_d$ ), ACK, NAK, FT} : ensemble des messages émis par une unité  $U_i$  vers son module de contrôle  $MC_i$ .

- REQ - ( $U_d$ ), demande au module de contrôle d'établir une connexion avec l'unité distante d'adresse  $U_d$ .

- ACK (resp. NAK), réponse positive (resp. négative) à une requête auparavant émise par le module  $MC_i$  à son unité locale.

- FT, signale au module de contrôle la fin de transfert.



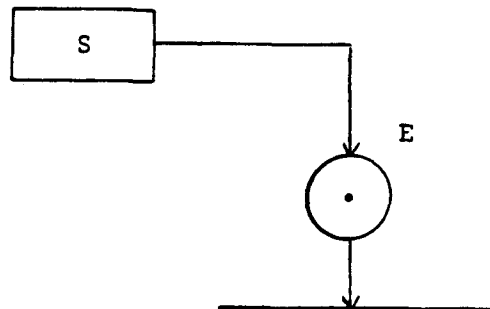
ii) Evènements externes : Ce sont les messages échangés entre deux différents modules de contrôle. Ils sont au nombre de quatre :

- Jeton, ce message est transmis du module prédécesseur de  $MC_i$  :  $P(MC_i)$  vers  $MC_i$  qui le retransmet, à son tour, au module successeur  $S(MC_i)$ .
- $V(C_i)$ , le numéro de la cellule  $C_i$ , reconnu par le module  $MC_i$ , et envoyé par un autre module  $MC_j$ .
- $ACK'$  (resp.  $NAK'$ ), réponse positive (resp. négative) à une demande de connexion, auparavant émise par un module de contrôle distant.

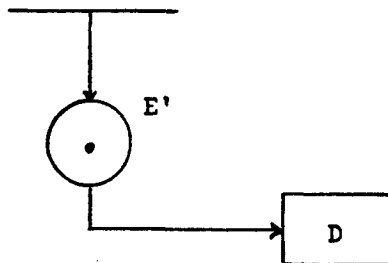
### 2.b.3.b. Etude des états du module de contrôle

La description des protocoles sera faite au moyen des réseaux de Pétri.

Notations : Un évènement  $E$  dont la source est une entité  $S$  est représenté de la façon suivante :



De la même façon, un évènement  $E'$  destiné à une entité  $D$  sera représenté :



Etats et évènements sont distingués en soulignant dans les réseaux de Pétri les noms des places qui correspondent à des états.

Lorsqu'une transition modifie la valeur de la commande de la cellule.

La nouvelle valeur de la commande figure avec cette transition.

Repos : C'est l'état initial du module de contrôle. Afin de donner une description simple du protocole d'interconnexion et de remédier, en particulier, au problème de l'occurrence simultanée de deux requêtes l'une locale et l'autre externe, et aux changements d'états que chacune d'elles provoque ; nous nous proposons de décomposer l'état Repos en deux états R-m et R-e. L'état du module de contrôle sera alors défini par le couple (E-m, E-e) où E-m désigne l'état atteint à partir de l'état initial R-m, et E-e celui atteint à partir de l'état initial R-e.

- i) R-m : Le module se maintient dans cet état initial tant qu'il n'a pas reçu une requête  $REQ - (U_d)$ , de son unité locale  $U_i$ . La réception d'une telle requête entraîne son passage à l'état attente de jeton Ajet - figure 4.a. -.

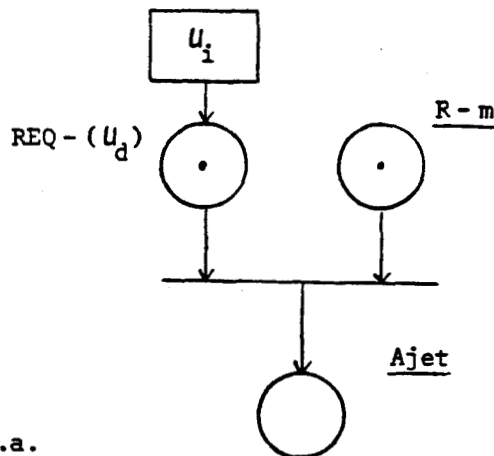


Figure 4.a.

- ii) R-e : Le module se maintient dans cet état jusqu'à la réception d'une requête externe, c'est-à-dire jusqu'à ce qu'il reconnaisse le numéro  $V(C_i)$  de sa cellule  $C_i$ . Il commute alors sa propre cellule ( $!C_i = 1$ ), transmet la requête reçue à son unité locale  $U_i - (REQ)$ , et se met en attente de l'acquittement de cette même unité AA-e - figure 5.a. -.

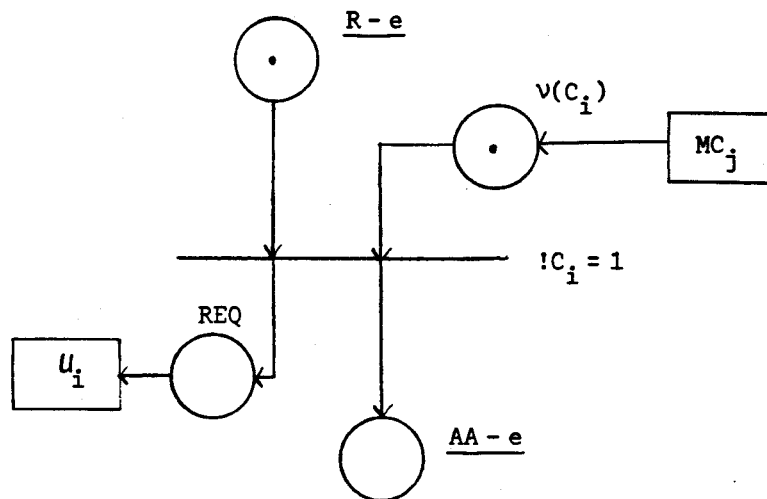


Figure 5.a.

Attente du jeton : Lorsqu'il est dans cet état - figure 4.a. -, le module de contrôle s'y maintient jusque la réception du jeton ; après quoi il commute sa cellule ( $!C_i = 1$ ), envoie le numéro  $v(C_j)$  de la cellule associée à l'unité destinataire  $U_d$ , et avec laquelle une connexion est à établir. Il se met ensuite en attente de l'acquittement AA-m du module de contrôle distant  $MC_j$  - figure 4.b. -.

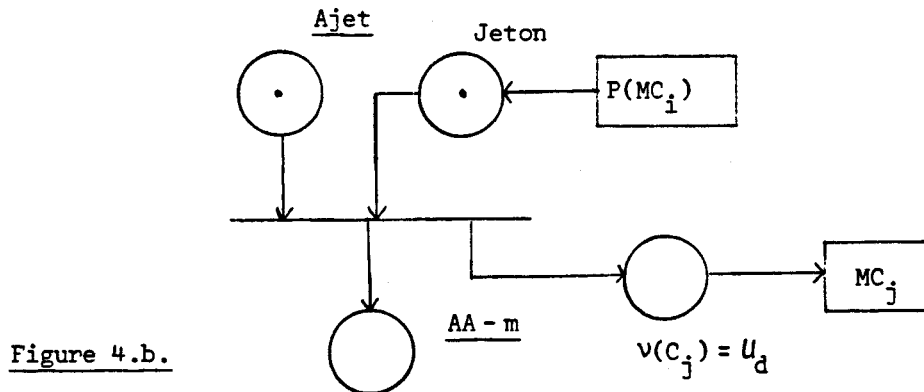


Figure 4.b.

La transparence du module de contrôle à la traversée du jeton, lorsqu'il ne se trouve pas dans l'état Ajet est décrite par la figure 4.c.

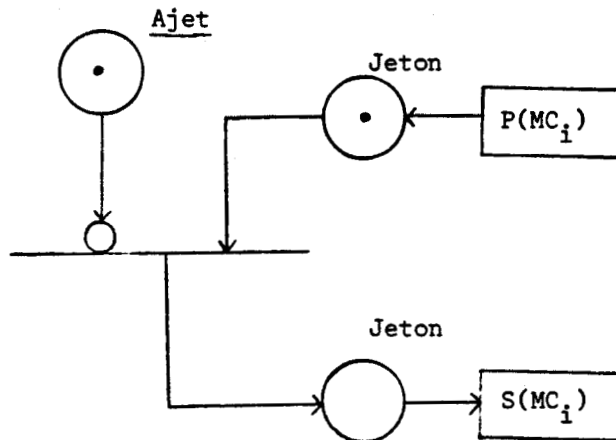


Figure 4.c.

Attente d'un acquittement :

- i) AA-m : attente d'un acquittement du module de contrôle distant  $MC_j$ , associé à la cellule  $C_j$  précédemment adressée - figure 4.b. -. La réception d'un acquittement positif entraîne le passage du module dans l'état transfert Tr-m, après avoir signalé à son unité locale  $U_i$  que la connexion qu'elle lui a auparavant soumise - figure 4.a. - est établie (CE) - figure 4.d. -.

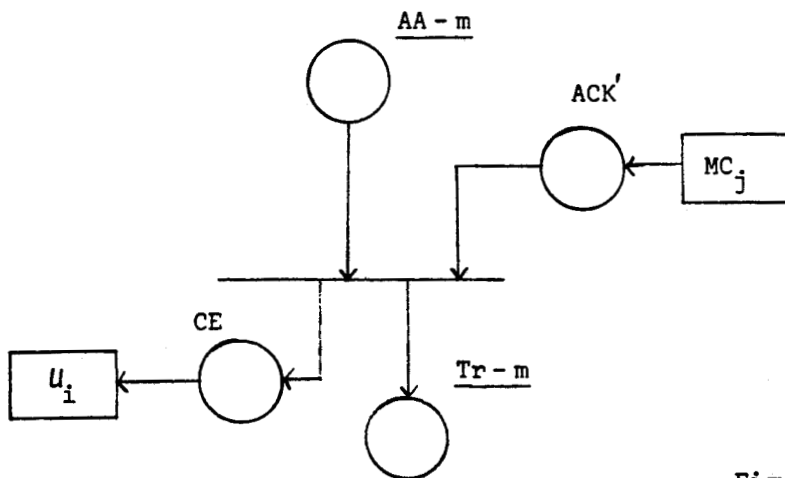


Figure 4.d.



Dans le cas où l'acquittement reçu est négatif, deux solutions sont envisageables :

La première, consiste à signaler à l'unité locale  $U_i$ , que sa demande de connexion est refusée - message CR -, le module de contrôle retrouve alors son état initial R-m, après avoir libéré le jeton et rendu transparente sa cellule de commutation - figure 4.e. -. L'unité  $U_i$  doit alors, pour établir la connexion avec l'unité  $U_d$ , réitérer sa demande auprès de son module de contrôle  $MC_i$ .

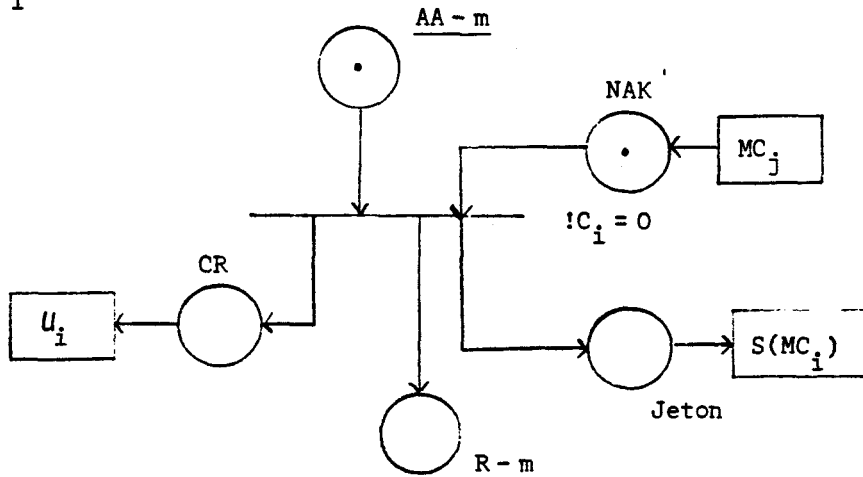


Figure 4.e.

La seconde solution, et c'est elle que nous avons retenue, consiste à libérer le jeton, rendre transparente la cellule de commutation et retourner à l'état attente de jeton : Ajet - figure 4.e'. -. Cette seconde solution, suppose que le module  $MC_i$  continue de conserver le numéro  $U_d$  de l'unité avec laquelle une connexion est à établir. L'unité  $U_i$  restera alors transparente à toutes les communications et aux actions qui les accompagnent et qui ont lieu dans l'ANNEAU et ce, jusqu'à ce que la connexion qu'elle a demandée soit définitivement établie.

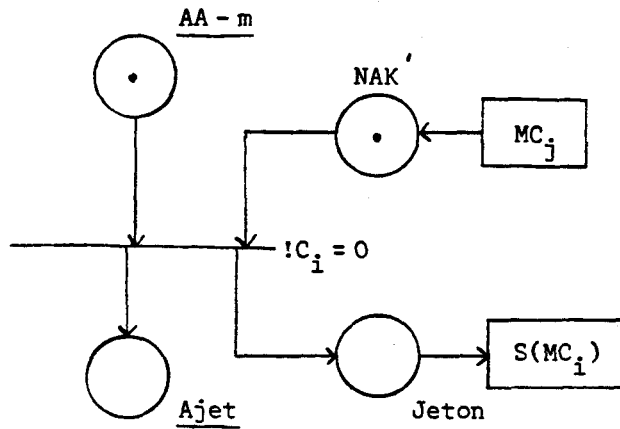


Figure 4.e'.

ii) AA-e, attente d'un acquittement de l'unité locale. Après avoir reconnu le numéro ( $V(C_i)$ ) de sa cellule de commutation, commuté sa cellule, et émis une requête à son unité locale  $U_i$ , le module  $MC_i$  se met en attente d'un acquittement en provenance de son unité  $U_i$  - figure 5.a.-.

La réception d'un acquittement entraîne sa transmission vers le module de contrôle distant  $MC_j$ .

Si l'acquittement est négatif, il rend sa cellule transparente et retrouve son état initial R-e - figure 5.b.-.

Si l'acquittement est positif, le module passe dans l'état de transfert Tr-e - figure 5.c.-.

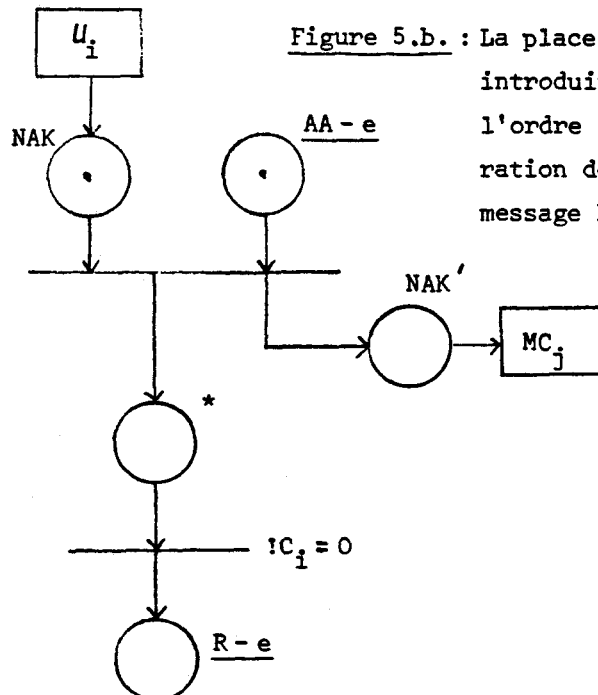


Figure 5.b. : La place étiquetée \* a été introduite pour expliciter l'ordre strict entre l'opération de transmission du message  $NAK'$  et l'action  $!C_i = 0$ .

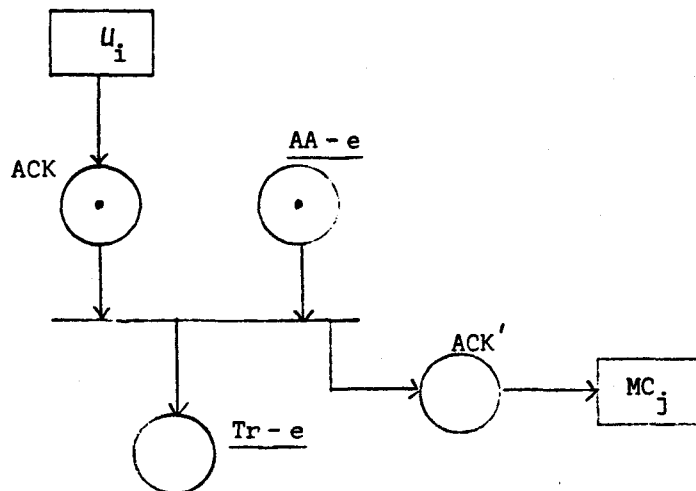


Figure 5.c.

Transfert : Lorsqu'il se trouve dans cet état Tr-m - figure 4.d. - ou Tr-e - figure 5.c. -, le module de contrôle s'y maintient jusqu'à la réception du message fin de transfert FT, émis par son unité locale  $U_i$ .

i) Tr-m : Lorsqu'il se trouve dans cet état, le module  $MC_i$  est l'allocateur de l'ANNEAU. La fin de transfert se traduit alors par : la mise en l'état transparent de sa cellule, la libération du jeton, et le retour à l'état initial R-m - figure 4.f. -.

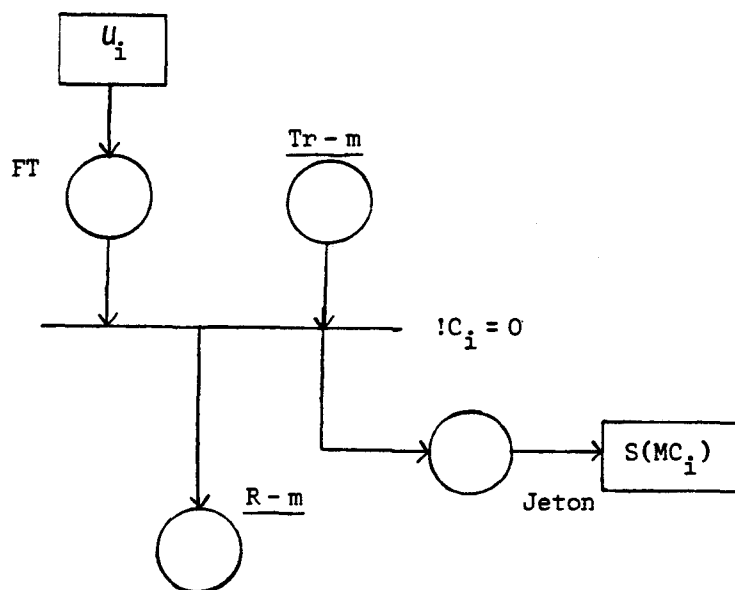
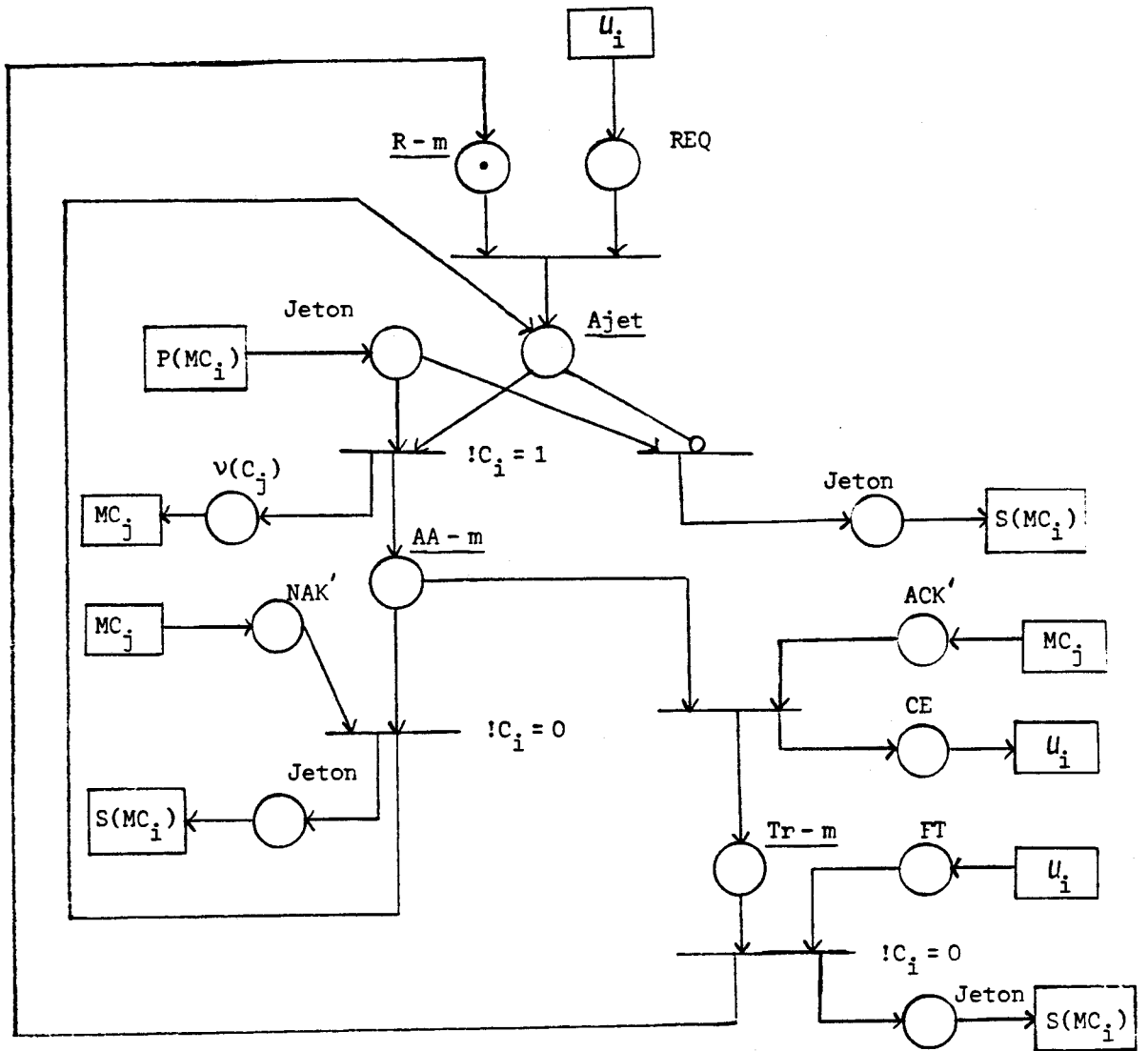


Figure 4.f.

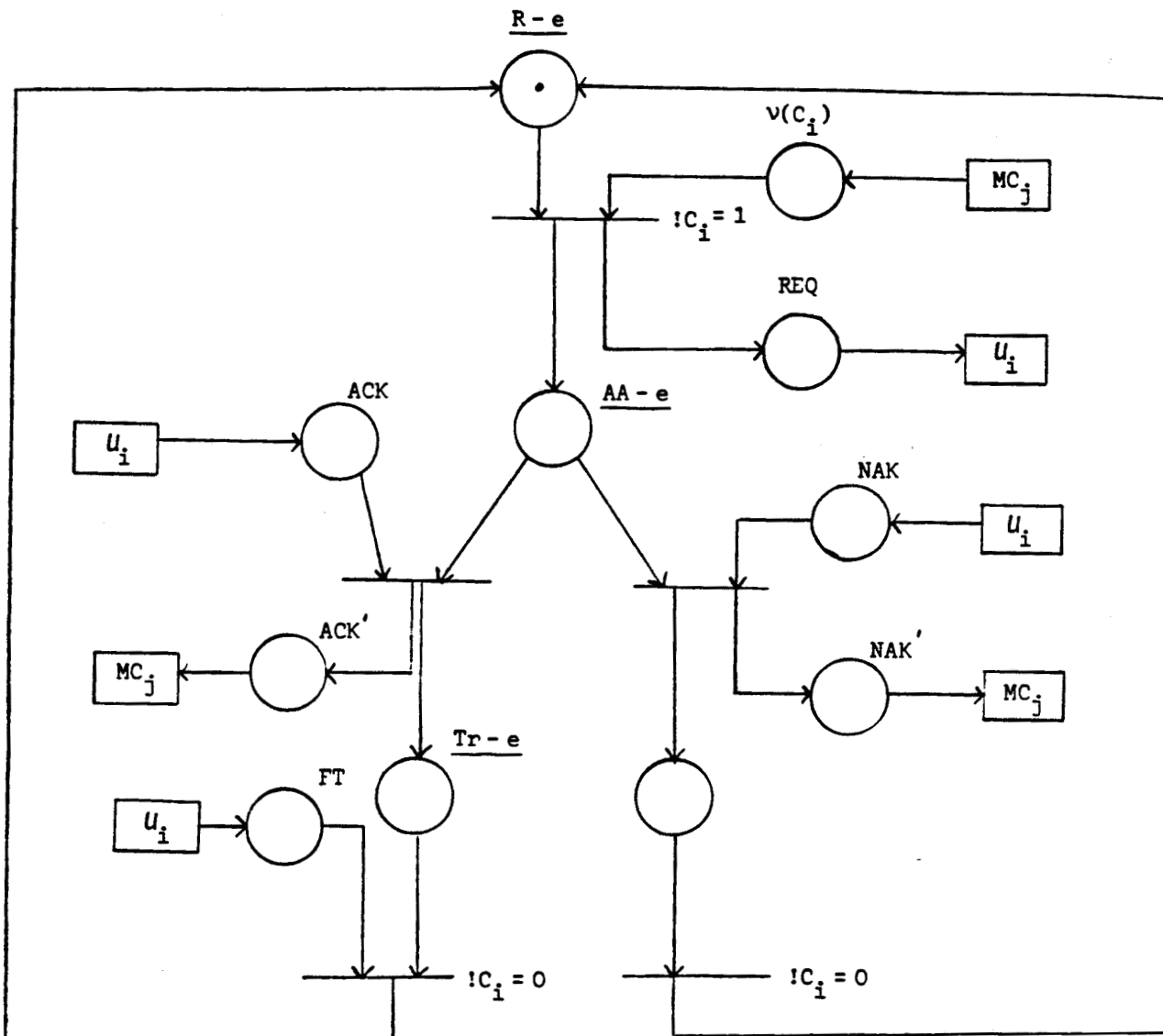






a - Cas où  $MC_i$  est l'allocateur de l'ANNEAU.





b - Cas où  $MC_i$  n'est pas l'allocateur de l'ANNEAU.



### 2.c. Le protocole de transfert

Le protocole d'interconnexion, défini dans le paragraphe précédent, permet d'établir une liaison  $U_i \leftrightarrow U_j$ , bidirectionnelle et simultanée entre deux unités  $U_i$  et  $U_j$ . Il s'agit maintenant de définir le protocole de transfert, dont la fonction est d'assurer la transmission d'une information élémentaire le long de chacune des deux connexions qui constituent cette liaison.

Il existe deux façons de concevoir le transfert sur une liaison

$U_i \leftrightarrow U_j$  :

- i) Transfert "séquentiel" - ou directe -. A un instant donné, une seule information élémentaire est transmise le long de chacune des deux connexions  $U_i \rightarrow U_j$  et  $U_j \rightarrow U_i$ .
- ii) Transfert "pipeline" - ou indirecte -. Le nombre d'informations élémentaires qui peuvent être transmises, à un instant donné, le long de chacune des deux connexions est égal à la longueur, évaluée en nombre de cellules de commutation, de cette connexion.

Quelle que soit la méthode de transfert adoptée, son contrôle peut être assuré selon un mode synchrone ou asynchrone. Il en résulte alors quatre options possibles pour établir le protocole de transfert. Nous nous proposerons par la suite, de décrire chacune de ces quatre options et d'en étudier les implications sur la structure physique de l'ANNEAU.

### 2.c.1. Transfert séquentiel

#### 2.c.1.a. Transfert séquentiel synchrone

Le transfert le long de chacune des deux connexions d'une liaison bidirectionnelle, est cadencé par une horloge de référence commune, ou par l'horloge de l'une des deux unités communicantes.

Evaluation de la période de l'horloge :

La période T doit être au moins égale au délai de propagation d'une information élémentaire le long d'une connexion. Ce délai est fonction des deux délais élémentaires suivants :

- le 1<sup>er</sup> est introduit par la traversée d'une cellule de commutation, nous le notons  $\delta_c$ ,
- le second est introduit par le parcours de la distance qui sépare deux cellules successives, nous supposons sa longueur constante - figure 8 - et nous la noterons  $\delta_l$ .

Notons  $\Delta_{i,j}$  le délai de propagation le long de la connexion  $U_i \rightarrow U_j$ , et  $N_{i,j}$  le nombre de cellules qui composent cette connexion.

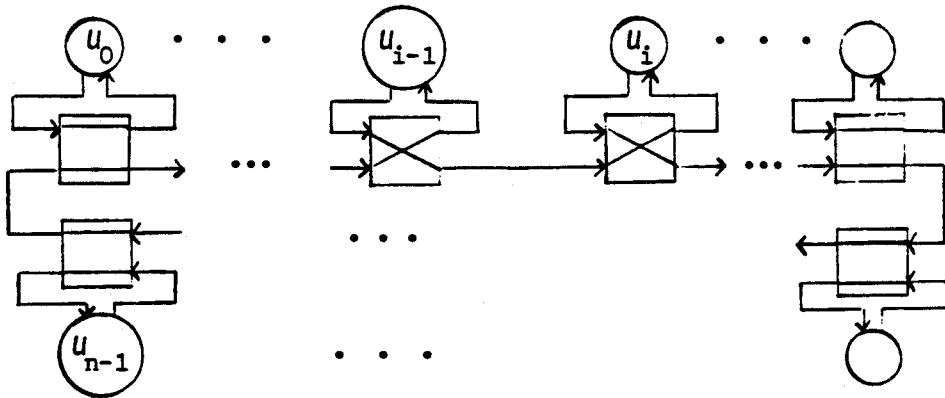
$$N_{i,j} = (j-i)+1 \quad 0 \leq i \leq j \leq n-1$$

$$N_{i,j} = N_{i,n-1} + N_{0,j} = n-i+j+1 \quad 0 \leq j \leq i \leq n-1$$

$$\Delta_{i,j} = N_{i,j} \cdot \delta_c + (N_{i,j} - 1) \cdot \delta_\ell.$$

Il est indispensable de prendre en compte, pour déterminer la période de l'horloge, le délai maximal introduit par la connexion la plus longue de l'ANNEAU.

La connexion la plus longue est celle qui relie une unité  $U_i$  à l'unité prédécesseur :  $U_i \rightarrow U_{i-1}$  ( $i \neq 0$ ) - figure 8 -.



**Figure 8** : Configuration en anneau des cellules de commutation avec une distance constante entre-elles, ( $n$  pair).  $U_i \rightarrow U_{i-1}$  est la connexion la plus longue.

$$\begin{aligned} \Delta_{i,i-1} &= N_{i,i-1} \cdot \delta_c + (N_{i,i-1} - 1) \cdot \delta_\ell \\ &= n \cdot \delta_c + (n-1) \cdot \delta_\ell \end{aligned}$$

$$\text{d'où } T > n \cdot \delta_c + (n-1) \cdot \delta_\ell$$

Cette inégalité doit être satisfaite par la période de l'horloge qui cadence les transferts dans l'ANNEAU.

### 2.c.1.b. Transfert séquentiel asynchrone

Un moyen de mise en oeuvre du transfert séquentiel asynchrone, est celui de l'accord confirmé ; il a été adoptée pour l'unibus du PDP-11 [DLV 79] et le microprocesseur 68000 [MIN 84]. Cette méthode fait intervenir deux messages :

- i) Un message de validation de transfert (MV), émis par l'unité source, après qu'elle ait transmis une information élémentaire, est interprété par l'unité destinataire comme une autorisation d'échantillonner cette information.
- ii) Un message d'acquiescement (MA), émis par l'unité destinataire. Il signale à l'unité source que l'information à été reçue et/ou que l'unité destinataire est en état de recevoir un autre message.

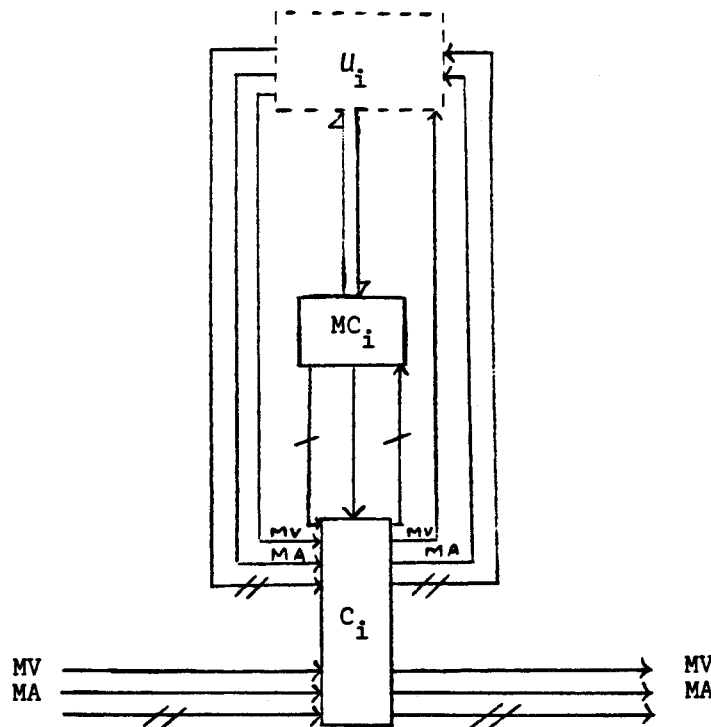


Figure 9 : Transfert séquentiel asynchrone. Liaisons entre  $U_i$ ,  $MC_i$  et  $C_i$ .

### 2.c.2. Transfert pipeline

L'intérêt du transfert pipeline est de rendre le débit potentiel dans l'ANNEAU indépendant de sa longueur.

Dans ce mode de transfert, un élément appelé "Module de Transfert", noté  $MT_i$ , est interposé entre chaque couple de cellules voisines,  $C_i$  et  $C_{i+1 \bmod n}$ , de l'ANNEAU - figure 10 -.

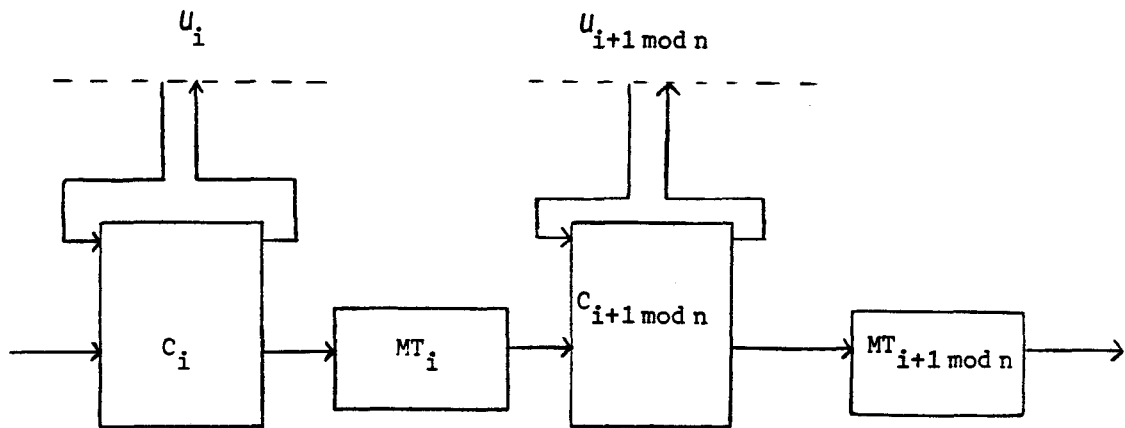


Figure 10 : Configuration de l'ANNEAU dans le cas d'un transfert pipeline.

Le contrôle s'applique au transfert d'une information élémentaire entre deux modules successifs  $MT_i$  et  $MT_{i+1 \bmod n}$ . En nous plaçant au niveau de ces modules, le transfert pipeline peut être considéré comme étant séquentiel entre deux modules voisins. D'une façon analogue au transfert séquentiel, les deux modes de contrôle, synchrone et asynchrone, peuvent être retenus pour le contrôle du transfert entre deux modules voisins. Nous nous proposons d'abord d'étudier la structure des modules de transfert, avant de leur appliquer chacun des deux modes de contrôle précédents.

#### 2.c.2.a. Structure des modules de transfert

Chaque module de transfert comporte deux éléments de mémoire ou registres, dont la capacité, est égale au nombre de bits d'une information élémentaire. Notons ces deux registres  $RE_i$  et  $RS_i$ . Le transfert le long d'une connexion se ramène à la séquence de transfert de registre à registre suivante :

- i) Transfert interne à chaque module  $MT_i$ , du contenu de son registre d'entrée  $RE_i$  dans le registre de sortie  $RS_i$ .
- ii) Transfert du contenu du registre de sortie  $RS_i$  de  $MT_i$  dans le registre d'entrée  $RE_{i+1 \bmod n}$  du module suivant  $MT_{i+1 \bmod n}$ .

### 2.c.2.b. Contrôle synchrone du transfert pipeline

Une horloge biphasée [MEC 80] contrôle la séquence de transferts de registre à registre. Sa période  $T$ , ainsi que la durée de chacune de ses deux phases  $\phi_1$  et  $\phi_2$ , sont déterminées par les intervalles de temps suivants :

$\Delta_1$  : L'intervalle de temps pendant lequel une information élémentaire, disponible à l'entrée du registre  $RE_i$ , est échantillonnée et mémorisée dans ce registre.

$\Delta_2$  : L'intervalle de temps nécessaire au transfert du contenu du registre d'entrée  $RE_i$  dans le registre de sortie  $RS_i$ .

$\Delta_3$  : L'intervalle de temps nécessaire à la traversée de la logique de la cellule de commutation séparant deux modules voisins  $MT_i$  et  $MT_{i+1 \bmod n}$ .

$\tau$  : Le retard maximum du signal d'horloge sur sa ligne de distribution.

La période  $T$  de l'horloge de synchronisation doit satisfaire l'inégalité suivante : - figure 11 -

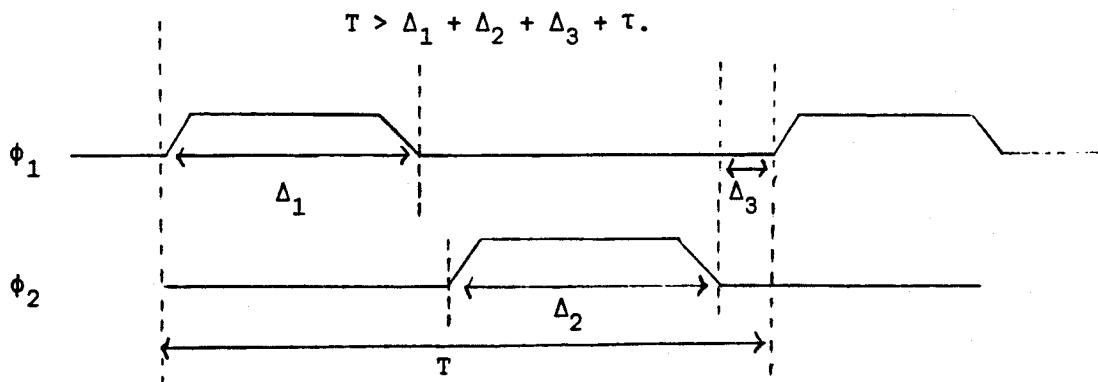


Figure 11 : Chronogramme des deux signaux de l'horloge biphasée.

Les deux phases  $\phi_1$  et  $\phi_2$  de l'horloge sont respectivement appliquées aux registres  $RE_i$  et  $RS_i$  des modules de transfert - figure 12 -. Durant  $\phi_1$  (resp.  $\phi_2$ ), les entrées des registres  $RE_i$  (resp.  $RS_i$ ) sont ouvertes alors que leurs sorties sont fermées. Inversement, pendant  $\overline{\phi_1}$  (resp.  $\overline{\phi_2}$ ), les entrées des registres  $RE_i$  (resp.  $RS_i$ ) sont fermées alors que leurs sorties sont ouvertes.

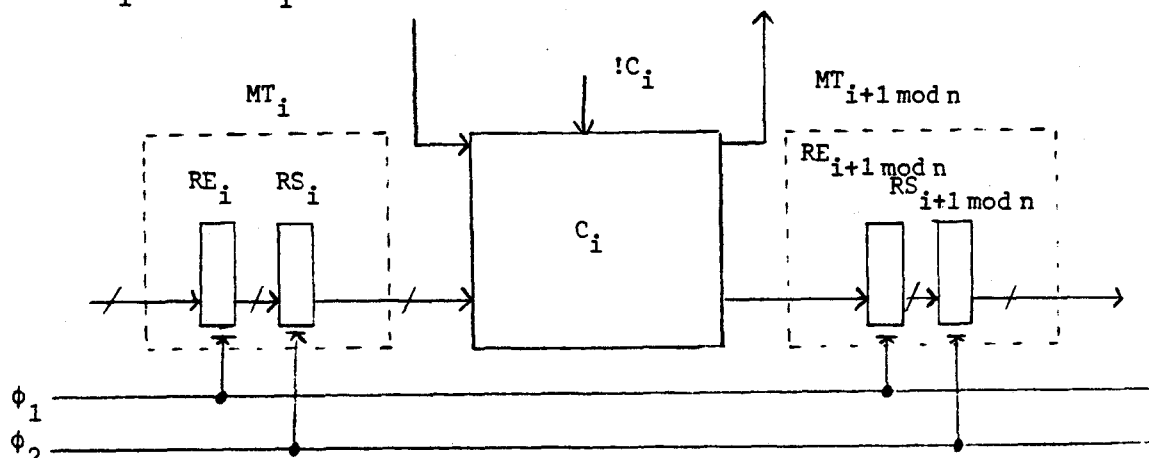


Figure 12 : Contrôle synchrone du transfert pipeline dans l'ANNEAU.

Dans ce mode de contrôle, le contenu d'un registre subit un transfert même lorsqu'il est vide. Il est donc nécessaire de disposer au niveau de l'unité réceptrice d'un moyen qui permette la distinction entre le contenu d'un registre lorsqu'il est vide et son contenu lorsqu'il est plein. Un moyen simple consiste à réserver à cet effet un bit de validation, qui est à 1 par exemple et qui accompagne une information dès qu'elle est introduite dans l'ANNEAU.

D'autre part, lorsque la vitesse d'émission des messages par une unité est supérieure à celle de l'absorption de ces messages par l'unité destinataire, il est indispensable de mettre en place une procédure de régulation du flux entre les deux unités. Il existe plusieurs méthodes de régulation de flux, la plupart d'entre-elles ont été développées dans le cadre des systèmes faiblement couplés [COR 81]. Dans le cadre de cette étude les deux suivantes nous semblent être les plus appropriées :

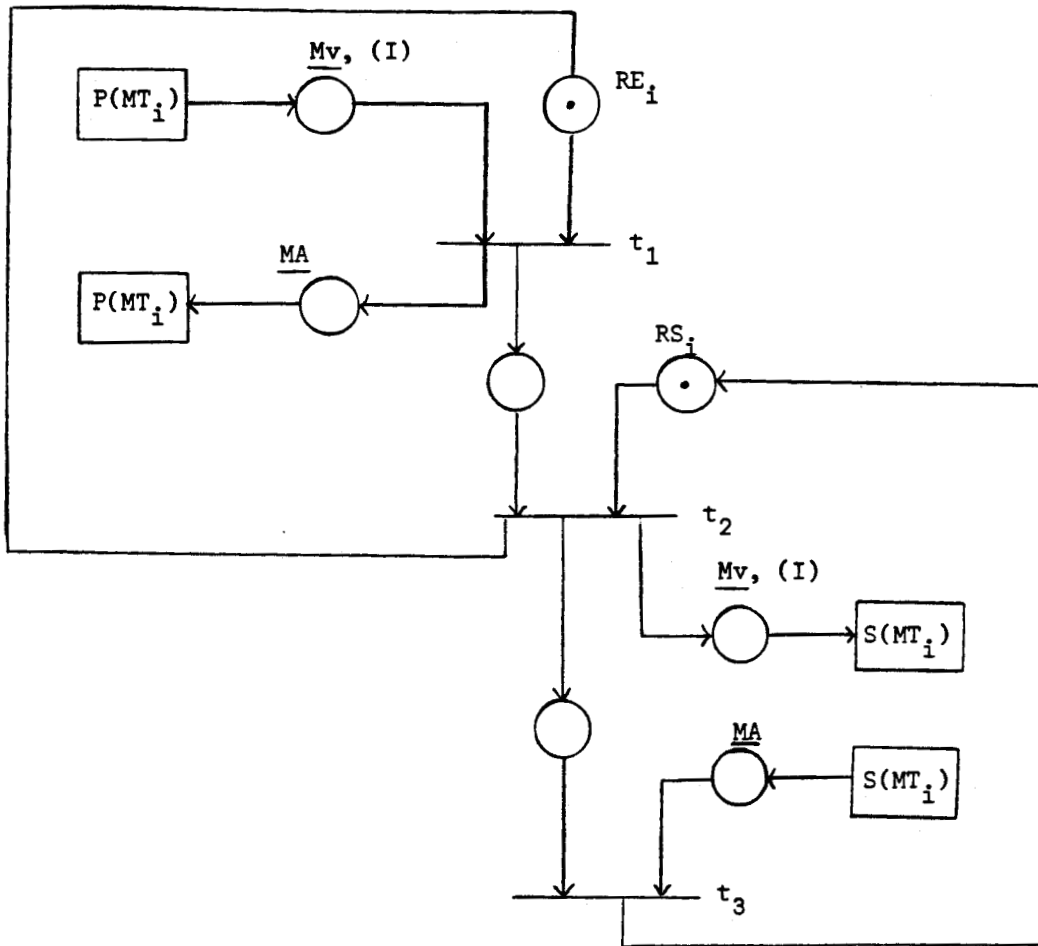


- i) Chaque unité dispose d'un tampon dont la taille est déterminée en fonction des caractéristiques de débit des unités et de l'ANNEAU. Cette méthode nécessite, pour remédier au problème de dépassement de la capacité du tampon, la mise en oeuvre d'une procédure d'interruption et de reprise de transfert.
- ii) La seconde technique est celle de réservation de registre. Cette méthode a été utilisée dans l'anneau synchrone CSB "Circular Synchronous Bus" [TBB 81]. L'opération de réservation d'un registre est effectuée par l'unité réceptrice et consiste à envoyer une information élémentaire vide accompagnée d'une en-tête, qui indique à l'unité émettrice qu'elle est autorisée à envoyer un nouveau message. La cadence de réservation des registres, par l'unité réceptrice, doit correspondre à sa cadence d'absorption des messages.

#### 2.c.2.c. Contrôle asynchrone du transfert pipeline

A chaque module de transfert  $MT_i$ , est associée une logique de contrôle qui définit la procédure de communication avec ses deux modules voisins : le module successeur  $S(MT_i)$  et le module prédécesseur  $P(MT_i)$ . Cette procédure de communication est décrite au moyen du réseau de Pétri de la figure ci-après, et fait intervenir les deux messages suivants :

- i) Un message de validation MV, qui accompagne toute information élémentaire (I), émise par  $MT_i$  vers son successeur  $S(MT_i)$ .
- ii) Un message d'acquiescement MA envoyé au module prédécesseur  $P(MT_i)$ , après réception d'une information élémentaire par  $MT_i$ .



Procédure de communication du module  $MT_i$ .

Les deux places étiquetées  $RE_i$  et  $RS_i$  contiennent une marque chacune initialement. Elles définissent les états des contenus des deux registres  $RE_i$  et  $RS_i$  du module  $MT_i$ . La présence d'une marque dans une de ces places, signifie que le registre correspondant est vide et qu'il peut, par conséquent, être chargé avec une information élémentaire, lorsque celle-ci est disponible à son entrée.



La transition  $t_1$  est franchie après réception du message de validation  $MV$  qui accompagne une information élémentaire (I), et lorsque le registre  $RE_i$  est vide. Cette transition provoque le chargement du registre  $RE_i$  par l'information (I) et l'émission, à destination du module prédécesseur  $P(MT_i)$ , du message d'acquiescement  $MA$ .

La transition  $t_2$  est, quant-à-elle, franchie lorsque le registre  $RS_i$  est vide et  $RE_i$  contient une information élémentaire. Elle provoque le transfert du contenu de  $RE_i$  vers  $RS_i$ ,  $RE_i$  devient à nouveau vide, et l'émission de l'information disponible dans  $RS_i$  vers le module successeur  $S(MT_i)$ , accompagnée du message de validation  $\underline{MV}$ .

La réception du message d'acquittement  $\underline{MA}$ , entraîne la transition  $t_3$ , qui rend à nouveau disponible le registre  $RS_i$ .

Le successeur  $S(MT_i)$  du module de transfert  $MT_i$ , dépend de l'état de la cellule de commutation  $C_i$ ,

$\eta(C_i) = T \Rightarrow$  toute information reçue par  $MT_i$  doit être transmise à destination de  $MT_{i+1 \bmod n}$  et l'acquittement sera, en conséquence, émis par  $MT_{i+1 \bmod n}$  vers  $MT_i$ .

$\eta(C_i) = E \Rightarrow$   $MT_i$  est connecté à l'unité  $U_i$ ; il constitue le dernier étage du pipeline. Toute information reçue par  $MT_i$  est transmise à l'unité  $U_i$ , qui à son tour envoie l'acquittement vers  $MT_i$ .

On supposera donc que chaque unité dispose d'une procédure de communication, identique à celle élaborée par chacun des modules de transfert de l'ANNEAU. Désignons par  $MT - U_i$  le module de l'unité  $U_i$  qui réalise une telle procédure.

$$S(MT_i) = \begin{cases} MT_{i+1 \bmod n} & \text{si } \eta(C_i) = T \\ MT - U_i & \text{si } \eta(C_i) = E \end{cases}$$

Il en résulte le schéma d'interconnexion, décrit par la figure 13, entre les modules de transfert de l'ANNEAU et des unités.

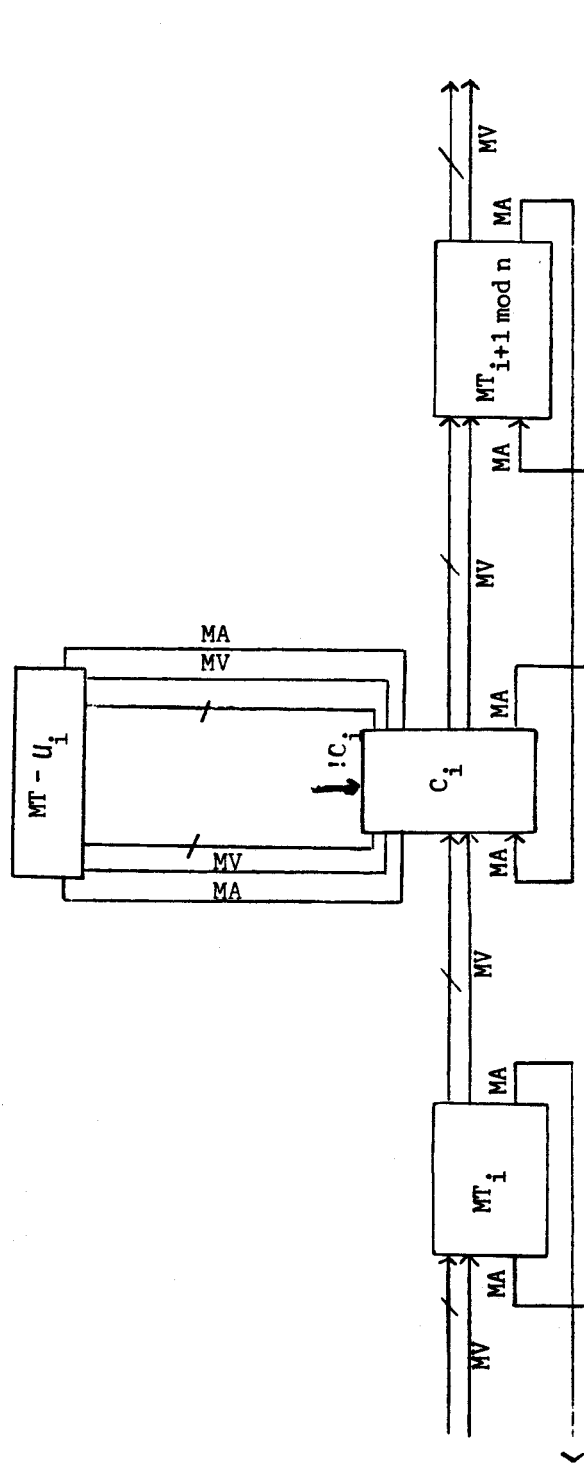


Figure 13 : Structure de l'ANNEAU, dans le cas du contrôle asynchrone du transfert pipeline.

Notons que, contrairement au transfert synchrone, le transfert asynchrone ne pose pas des problèmes de régulation de flux. En effet, une information élémentaire qui se présente à l'entrée du module de transfert d'une unité et qui ne peut être immédiatement retirée de l'ANNEAU, y reste en attente jusqu'à ce que l'unité réceptrice devienne en mesure de la retirer. Il est évident que cela se fait au détriment du débit effectif dans l'ANNEAU, qui est alors inférieur à son débit potentiel.

### 2.c.3. Comparaison

Le tableau ci-après compare les quatre modes de transfert. Deux éléments y sont présentés, le débit potentiel et le nombre de liaisons qu'exige chacun de ces quatre modes de transfert.

Contrôle Transfert	Synchrone		Asynchrone	
	Débit potentiel	nb de Liaisons	Débit potentiel	nb de Liaisons
Séquentiel	$\frac{2}{n \delta_c + (n-1) \delta_l}$	$\left\{ \begin{array}{l} 1/\text{unité} \\ H \\ n \end{array} \right\}$	$\frac{1}{(n+2) \delta_c + n \delta_l}$	$\left\{ \begin{array}{l} 4/\text{unité} \\ MV/MA \\ 4n \end{array} \right\}$
Pipeline	$\frac{1}{2\delta_r + \delta_c + \delta_l + \tau}$	$\left\{ \begin{array}{l} 2/\text{unité} \\ H(\phi_1, \phi_2) \\ 2n \end{array} \right\}$	$\frac{1}{\max(2\delta_r, \delta_r + 2(\delta_c + \delta_l))}$	$\left\{ \begin{array}{l} 4/\text{unité} \\ MV/MA \\ 4n \end{array} \right\}$

$\delta_c$  : délai de traversée d'une cellule de commutation,

$\delta_l$  : délai de propagation le long d'une liaison inter-cellules,

$\tau$  : retard du signal horloge le long de sa ligne de distribution,

$\delta_r$  : temps de chargement d'un registre.

### 3. Conclusion

La topologie du modèle de base répond aux conditions d'une intégration à très haute échelle. Les éléments qui la composent sont, en effet, tous identiques et les connexions entre-eux régulières.

Après avoir opté, conformément au cahier des charges, pour la commutation de ligne comme stratégie de transfert, une solution asynchrone et répartie du protocole d'interconnexion a été proposée. Le critère réseau à simple face a été également garanti, en autorisant un couple d'unités quelconques d'établir une liaison entre-elles. La méthode d'allocation de l'ANNEAU qui a été adoptée repose sur la technique du jeton circulant ; elle a l'avantage d'être simple quant à sa mise en oeuvre, cependant, elle ne permet pas d'établir une procédure évoluée de l'allocation de l'ANNEAU.

L'étude du protocole de transfert a permis, ensuite, de mettre en évidence la capacité de l'ANNEAU à supporter deux modes de transfert : séquentiel et pipeline et de définir, pour chacun d'entre-eux, les deux modes de contrôle synchrone et asynchrone.

Les protocoles d'interconnexion et de transfert ont été par ailleurs définis, sans que des hypothèses sur les caractéristiques physiques ou fonctionnelles n'aient été prises. Cela permet de considérer l'ANNEAU comme un outil de communication universel. Des protocoles d'échange spécifiques aux unités peuvent, en effet, être supportés par l'ANNEAU. Son utilisation pourra alors être envisagée pour connecter différentes unités d'un système telles que par exemples :

- Les blocs fonctionnels d'un même module physique. Dans le cas d'un microprocesseur, par exemple, ces blocs sont l'u.a.l, l'u.c.c, la mémoire interne et la logique de communication avec l'extérieur.
- Les unités fonctionnelles d'un ordinateur conventionnel.
- Ou encore, un ensemble de modules de communication, l'ANNEAU remplit, dans ce cas, la fonction de "passerelle" entre ces modules.

Notons par ailleurs que la liaison établie lors du protocole d'interconnexion est bidirectionnelle et simultanée. Ceci peut être mis à profit pour :

- accroître le débit dans l'ANNEAU, en autorisant des échanges simultanés le long des deux connexions qui composent une liaison,
- ou encore, garantir la sûreté de fonctionnement de l'ANNEAU, par retransmission systématique vers l'unité émettrice de toute donnée propagée le long de l'une des deux connexions.

Notons également qu'une liaison au plus, de ce type, peut être établie à un moment donné dans l'ANNEAU. C'est donc un outil de communication très séquentiel. L'objet du chapitre suivant est de proposer, à partir de l'ANNEAU, pris comme modèle de base, des topologies d'interconnexion qui autorisent la simultanéité et conservent des caractéristiques identiques à celles précédemment établies dans l'ANNEAU ; à savoir :

- La commutation de ligne bidirectionnelle et simultanée est la stratégie de transfert adoptée,
- l'asynchronisme des accès des unités à l'outil de communication, et
- le caractère simple face du réseau.

\*\*\*\*\*  
\*  
\* CHAPITRE III \*  
\*  
\*\*\*\*\*

EXTENSIONS DU MODELE DE BASE



Ce chapitre sera consacré à l'étude des possibilités d'extension qu'offre le modèle de base. L'objectif étant de définir des outils capables de prendre en charge des communications simultanées dans un environnement où les unités communicantes sont caractérisées par :

- leur hétérogénéité fonctionnelle, voire physique,
- l'assynchronisme de leurs accès à l'outil de communication.

Par ailleurs, les choix que nous avons adoptés lors de la définition du modèle de base, à savoir :

- la régularité de la topologie,
- la commutation de ligne comme stratégie de transfert et,
- la mise en oeuvre répartie du protocole d'interconnexion,

seront conservés lors de la définition des outils extension.

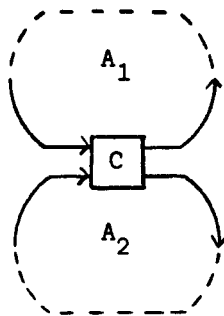
Deux directions se présentent à nous pour définir les extensions du modèle de base. Dans la première, nous considérons qu'un réseau extension est un graphe dont chaque noeud est un ANNEAU, et toute arrête est une cellule de commutation. Dans la seconde, nous partons d'une analogie entre l'ANNEAU et le bus et déduisons une extension de l'ANNEAU par analogie avec une des extensions du bus. Le chapitre est décomposé en deux parties ; chacune correspond à un type d'extension.

\*\*\*\*\*  
\*  
\* PREMIERE PARTIE \*  
\*  
\*\*\*\*\*

EXTENSIONS PAR GRAPHE

## 0. Introduction

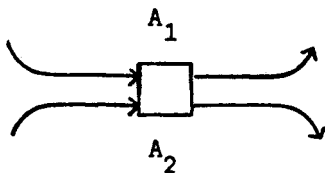
Topologiquement un réseau extension est constitué d'un ensemble d'ANNEAUX interconnectés. Deux ANNEAUX sont interconnectés si et seulement si, ils ont en commun l'une de leurs cellules de commutation.



Interconnexion de deux ANNEAUX.

La topologie d'un réseau extension, peut être décrite à l'aide de son graphe abstrait, défini comme suit :

- A tout ANNEAU du réseau correspond un et un seul noeud dans le graphe.
- A toute cellule de commutation, commune à deux ANNEAUX du réseau, correspond une arrête dans le graphe. Cette arrête relie les deux noeuds du graphe qui correspondent aux deux ANNEAUX ayant cette cellule en commun.



- a -



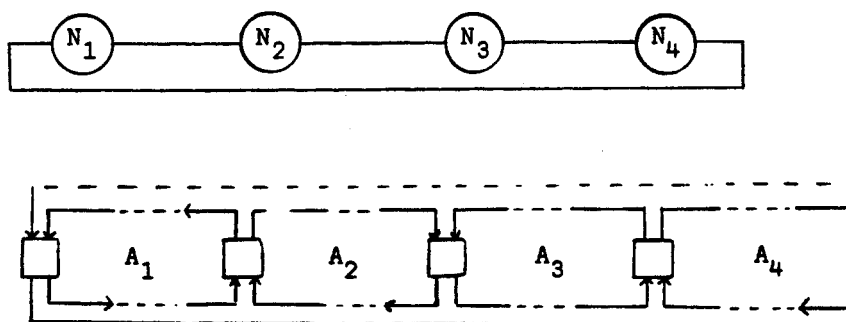
- b -

Interconnexion d'ANNEAUX et leur graphe abstrait.

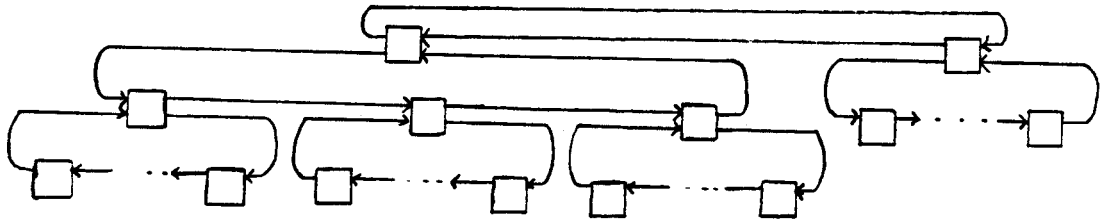
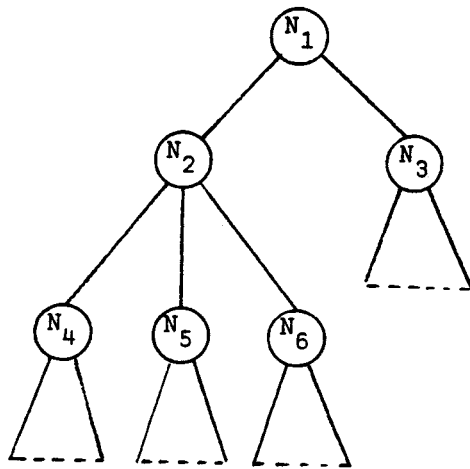
A partir d'un graphe, à priori quelconque, il est donc possible de définir un réseau extension du modèle de base. Cependant, parmi ces réseaux extensions, seuls peuvent être retenus ceux qui satisfont les contraintes suivantes, et qui sont de trois types :

- i - Des contraintes topologiques : la régularité de la topologie du réseau est une condition nécessaire pour permettre son intégration VLSI et réduire son coût de fabrication.
- ii - Des contraintes fonctionnelles : le réseau doit d'abord permettre l'établissement d'une liaison bidirectionnelle entre deux unités quelconques, il doit donc être un réseau à simple face, et ensuite permettre l'établissement d'un nombre élevé de liaisons en parallèle. Cette contrainte est une condition nécessaire pour envisager l'utilisation de l'outil dans un environnement à fort degré de couplage.
- iii - Des contraintes de contrôle : Le contrôle du réseau doit être simple. Il s'agit là, en effet, d'un critère fondamental, dont dépendra la mise en oeuvre programmée ou câblée de ce contrôle, ce qui, par conséquent, sera déterminant quant au temps de commande du réseau.

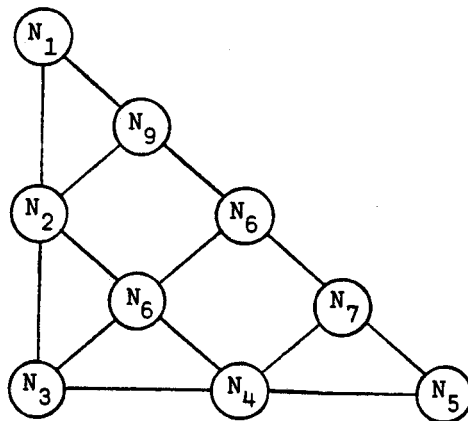
Dans les figures ci-après nous présentons quelques topologies de réseaux extension du modèle de base.



a - Interconnexion circulaire d'ANNEAUX.

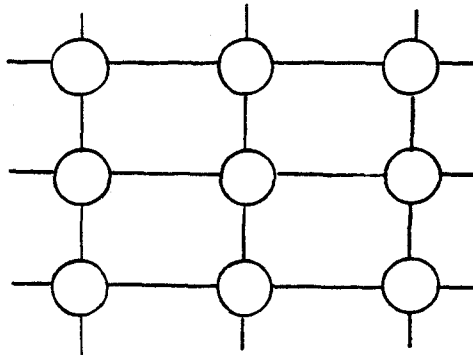


b - Interconnexion en ARBRE d'ANNEAUX.

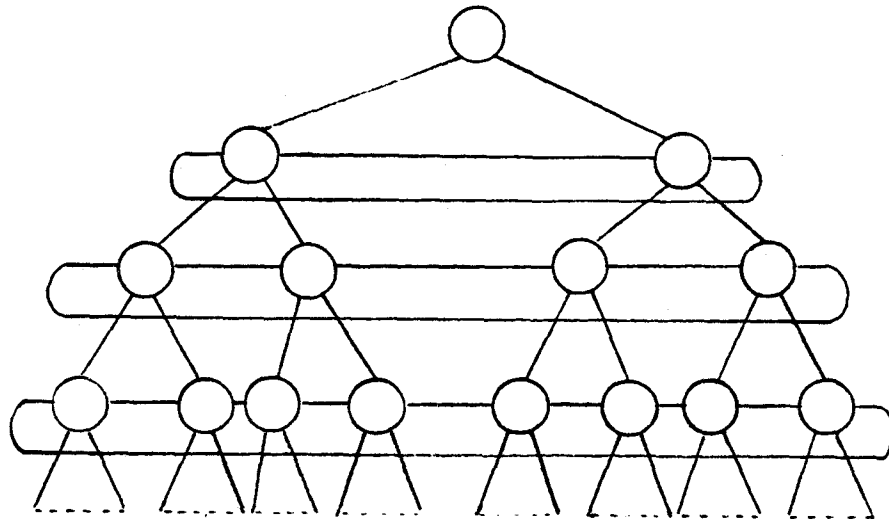


c - Interconnexion triangulaire d'ANNEAUX.





d - Interconnexion CELLULAIRE.

e - Interconnexion circulaire des ANNEAUX de même niveau  
du réseau ARBRE.

Quelques topologies extensions de l'ANNEAU - figures a à e -.

Remarque : L'interconnexion de systèmes géographiquement distants, au moyen de réseaux exclusivement composés d'anneaux, a eu un intérêt particulier au début des années 70 [PIE 72]. C'est ainsi qu'une technique d'adressage et un algorithme de routage, très généraux et indépendants de la topologie du réseau, ont été proposés [GRP 71], [BGK 72]. La reprise des résultats de ces études pour les extensions du modèle de base, nous semble être inefficace. Nous pensons qu'un contrôle simple passe nécessairement par une définition, qui soit spécifique à une topologie donnée, et qui tire profit de certaines caractéristiques de la topologie telles que par exemple, la constance des longueurs de liaisons, leur unicité ou encore la symétrie de la topologie.

## 1. Le réseau ARBRE

### 1.a. La topologie d'interconnexion

La structure de base du réseau est un arbre a-aire ( $a \geq 2$ ) complet et de profondeur  $n+1$  - figure 1-.

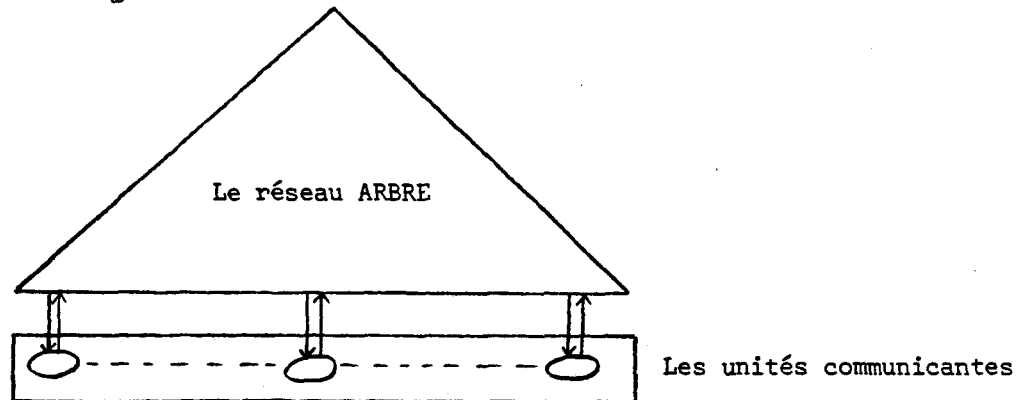


Figure 1 : Structure de base du réseau - ARBRE.

Les niveaux du réseau ARBRE sont numérotés de 0 à  $n$ , en les parcourant depuis la racine (niveau 0) jusqu'aux feuilles (niveau  $n$ ).

Chaque noeud du réseau ARBRE est un ANNEAU, que nous notons  $A_i^j$  où  $i$  désigne son niveau dans l'arbre et  $j$  son rang dans ce niveau. Les rangs des noeuds d'un niveau  $i$ , sont numérotés de 0 à  $a^i - 1$  en parcourant ce niveau depuis la gauche vers la droite.

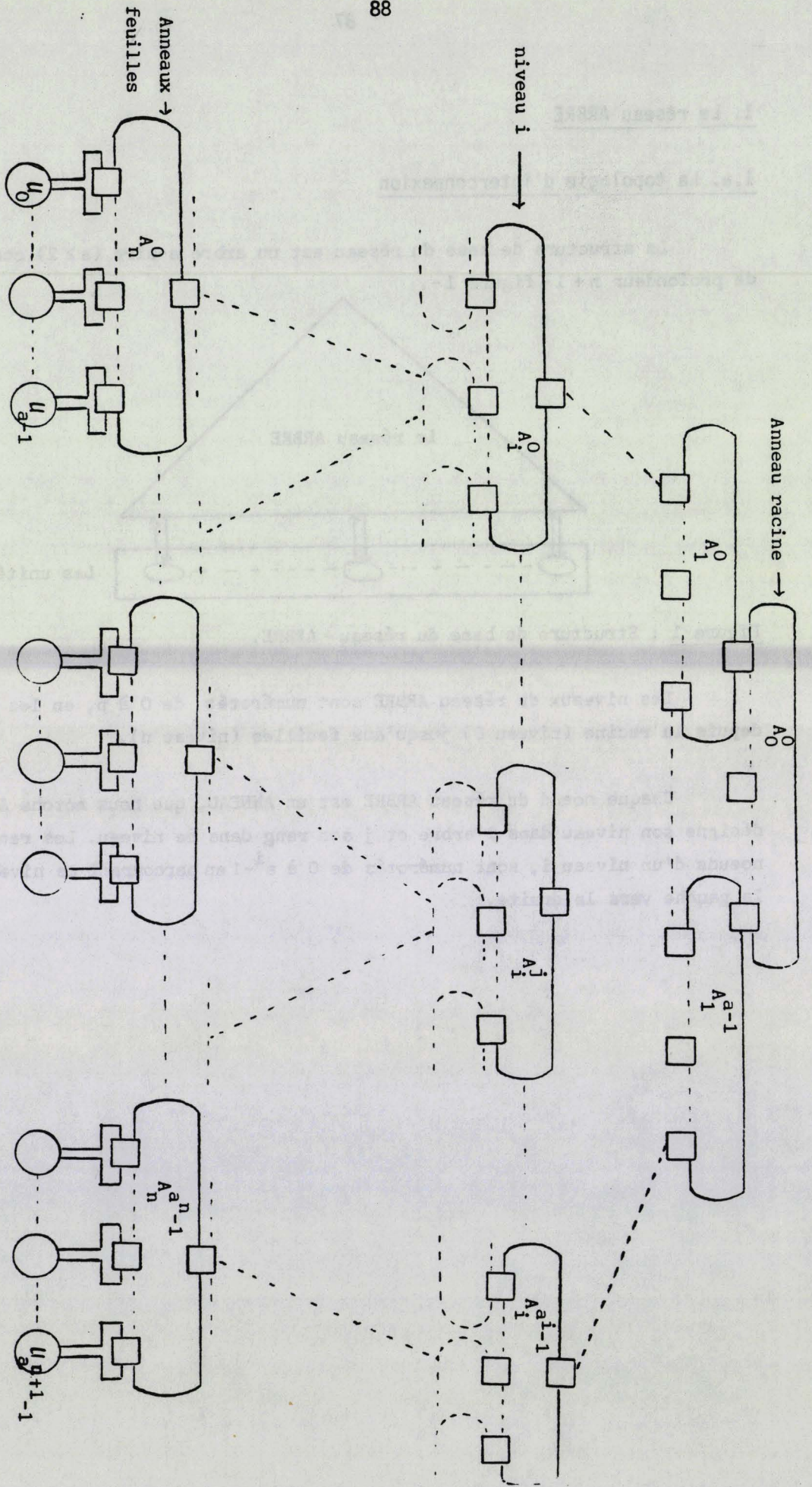


Figure 2 : Numérotation des ANNEAUX du réseau ARBRE.





Un ANNEAU  $A_i^j$  communique :

- (1) - à travers une suite de a-cellules  $(C_{i,j}^k)_{k=0}^{a-1}$  avec :
  - i) les a ANNEAUX  $(A_{i+1}^{j+k})_{k=0}^{a-1}$  du niveau i+1 si  $A_i^j$  n'est pas un ANNEAU feuille ( $i \neq n$ ), - figure 3-,
  - ii) avec un sous-ensemble de a unités, que nous désignons par la suite  $(U_{j+k}^a)_{k=0}^{a-1}$  si  $A_i^j$  est un ANNEAU du niveau feuille ( $i = n$ ) - figure 4-,
- (2) - à travers la cellule  $C_{i,j}^a$  avec l'ANNEAU  $A_{i-1}^{j \text{ div } a}$  du niveau i-1, si  $A_i^j$  n'est pas l'ANNEAU racine ( $i \neq 0$ ) - figure 3-.

Une cellule  $C_{i,j}^k$  est par conséquent commune aux deux ANNEAUX.

- (1)  $A_i^j$  et  $A_{i+1}^{j+k}$ ,  $k \neq a$  et  $C_{i,j}^k = C_{i+1,j+k}^a$
- (2)  $A_i^j$  et  $A_{i-1}^{j \text{ div } a}$ ,  $k = a$  et  $C_{i,j}^a = C_{i-1,j \text{ div } a}^{j \text{ mod } a}$

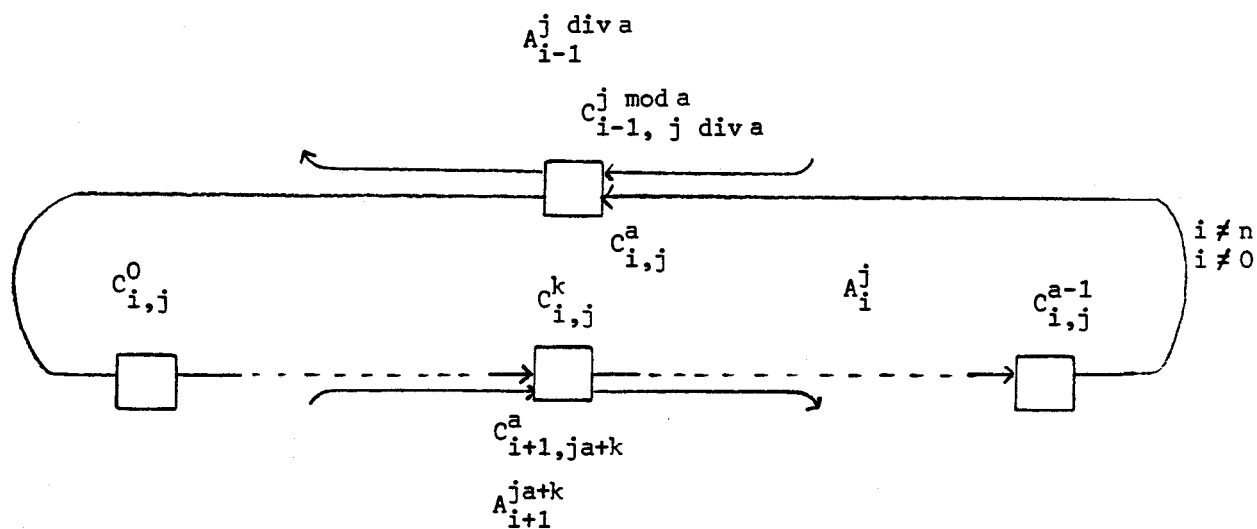


Figure 3

$$\begin{cases} C_{i,j}^a = C_{i-1,j \text{ div } a}^{j \text{ mod } a} \\ C_{i,j}^k = C_{i+1,j+k}^a \end{cases}$$

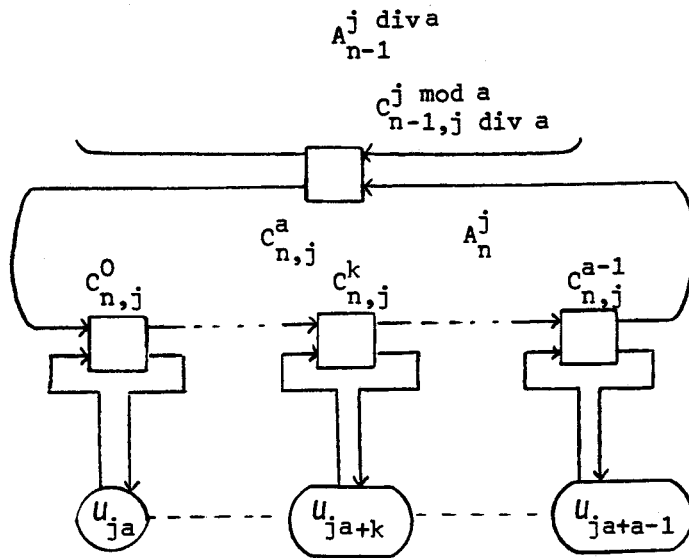


Figure 4

### 1.b. La fonction de numérotation

Cette fonction joue un rôle important dans la définition d'un réseau. Car, en plus de son rôle d'adressage des unités attachées au réseau et des entités, les ANNEAUX et/ou les cellules, qui composent ce réseau, elle intervient dans la phase de calcul de l'information de contrôle qui accompagne toute demande de connexion émise dans le réseau. La complexité de la procédure de routage dépend donc, en partie, de cette fonction.

La fonction de numérotation, désignée par  $v$ , associe à chaque cellule du réseau - ARBRE un numéro composé d'une suite de chiffres représentés dans la base  $a$  :

$$v : C \rightarrow \{0, \dots, a-1\}^{n+1}$$

où  $\{0, \dots, a-1\}^{n+1}$  est l'ensemble des suites de chiffres dans la base  $a$  et de longueur inférieure ou égale à  $n+1$ .

$C$  : ensemble des cellules du réseau - ARBRE

$$C = \{C_{i,j}^k \mid i=0, \dots, n, j=0, \dots, a^i-1 \text{ et } k=0, \dots, a-1\}.$$

a - Numérotation des cellules de l'ANNEAU racine

$$v(C_{0,0}^k) = k \text{ pour tout } k = 0, \dots, a-1.$$

b - Numérotation des cellules d'un ANNEAU de niveau  $i$  ( $i \neq 0$ ), en fonction des numéros de cellules de niveau  $i-1$ .

$$v(C_{i,j}^k) = v(C_{i-1,j}^{j \bmod a} \text{ div } a) \cdot k \quad k = 0, \dots, a-1$$

$\nwarrow$   
 concaténation

c - Extension de la numérotation aux cellules  $C_{i,j}^a$ .

Une cellule  $C_{i,j}^a$  est commune à deux ANNEAUX, le premier de niveau  $i$  et le second de niveau  $i-1$ , - figure 5 -.

$$C_{i,j}^a = C_{i-1,j}^{j \bmod a} \text{ div } a \Rightarrow v(C_{i,j}^a) = v(C_{i-1,j}^{j \bmod a} \text{ div } a)$$

La figure ci-après illustre le principe de numérotation des cellules du réseau - ARBRE.

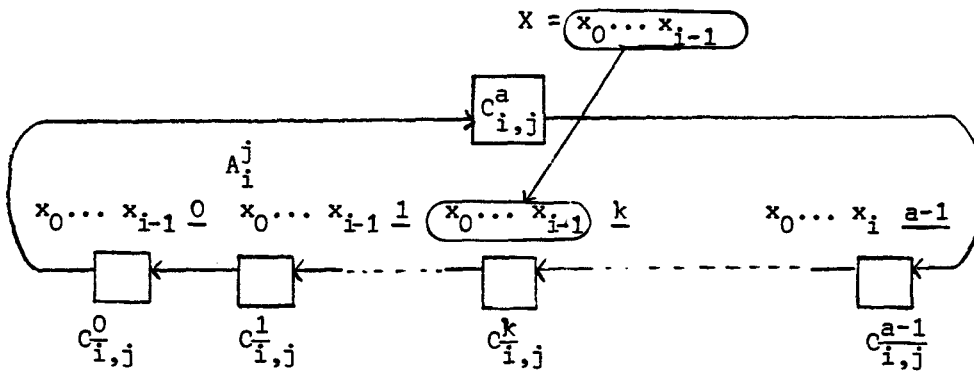


Figure 5 : Numérotation des cellules du réseau - ARBRE.

d - Numérotation des unités attachées au réseau.

Nous avons vu que les unités sont attachées aux cellules des ANNEAUX du niveau feuille ( $n$ ) Le numéro d'une unité sera en conséquence identique au numéro de la cellule de commutation à laquelle cette unité est attachée.

1.c. Le protocole d'interconnexion

Le protocole d'interconnexion définit les deux procédures suivantes :

- La procédure de routage qui, à partir de l'information de contrôle accompagnant une demande de connexion, définit l'état pris par chaque cellule de commutation traversée. Lorsque la demande a atteint sa destination, une liaison bidirectionnelle et simultanée est alors établie.

- La procédure d'accès aux ANNEAUX que doit emprunter une demande donnée. Elle doit en particulier détecter les situations de conflit, de blocage et définir les mécanismes permettant leur résolution.

### 1.c.1. La procédure de routage

#### - L'information de contrôle

Elle est constituée du seul numéro de l'unité destinataire  $U_D$ .

Posons :  $D = v(U_D)$

$= v(C_{n,j}^k)$  où  $C_{n,j}^k$  est la cellule auquel l'unité  $U_D$  est attachée.

$= d_0 \dots d_n$

La procédure de routage est entièrement répartie sur les cellules du réseau.

La notation  $D \xrightarrow{p} A_i^j$  désignera le fait que la demande  $D$  se propage dans l'ANNEAU  $A_i^j$ .

#### - Énoncé de la procédure de routage

Soit  $C_{i,j}^k$  une cellule du réseau-ARBRE traversée par la demande  $\underline{D} = d_0 \dots d_n$  et  $D \xrightarrow{p} A_i^j$ .

.  $k = a$

Si  $d_0 \dots d_{i-1} = v(C_{i,j}^a)$  alors

$$(R-1) : \begin{cases} n(C_{i,j}^a) = T \\ \text{et} \\ D \xrightarrow{p} A_i^j \end{cases}$$

sinon

$$(R-2) : \begin{cases} n(C_{i,j}^a) = E \\ \text{et} \\ D \xrightarrow{p} A_{i-1}^j \text{ div } a \end{cases}$$

fsi

$k \neq a \quad (0 \leq k \leq a-1)$

Si  $d_0 \dots d_i = v(C_{i,j}^k)$  alors

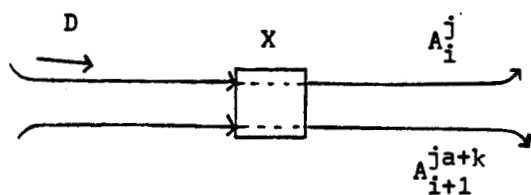
$$(R-3) : \begin{cases} n(C_{i,j}^k) = E \\ \text{et} \\ D \xrightarrow{p} A_{i+1}^{ja+k} \end{cases}$$

sinon

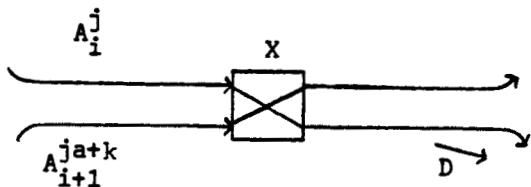
$$(R-4) : \begin{cases} n(C_{i,j}^k) = T \\ \text{et} \\ D \xrightarrow{p} A_i^j \end{cases}$$

fsi

Les figures 6 et 7 illustrent le principe de la procédure de routage.

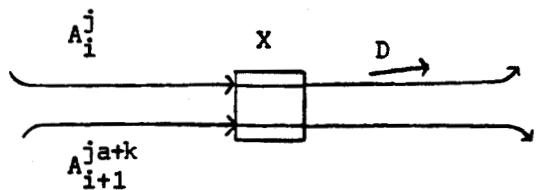


a - Avant traversée de la cellule X par la demande D.



b - Après traversée de la cellule X par D

$$d_0 \dots d_i = x_0 \dots x_i \quad (R-3)$$

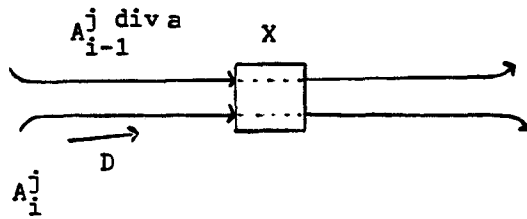


c - Après traversée de la cellule X par D

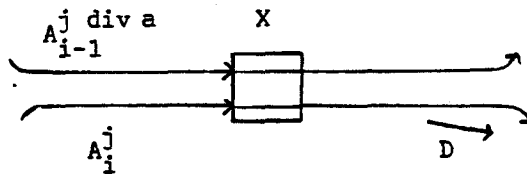
$$d_0 \dots d_i \neq x_0 \dots x_i \quad (R-4)$$

Figure 6 :  $X = x_0 \dots x_i = v(C_{i,j}^k)$  avec  $k \neq a$ .

. La traversée d'une cellule  $C_{i,j}^k$   $k \neq a$  par une demande D.

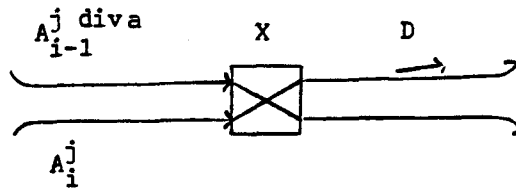


a - Avant la traversée de la cellule X par la demande D.



b - Après traversée de la cellule X par D.

$$d_0 \dots d_{i-1} = x_0 \dots x_{i-1} \quad (R-1)$$



c - Après traversée de la cellule X par D

$$d_0 \dots d_{i-1} \neq x_0 \dots x_{i-1} \quad (R-2)$$

Figure 7 : .  $X = x_0 \dots x_{i-1} = v(C_{i,j}^a)$ .

. Traversée de la cellule  $C_{i,j}^a$  par D.

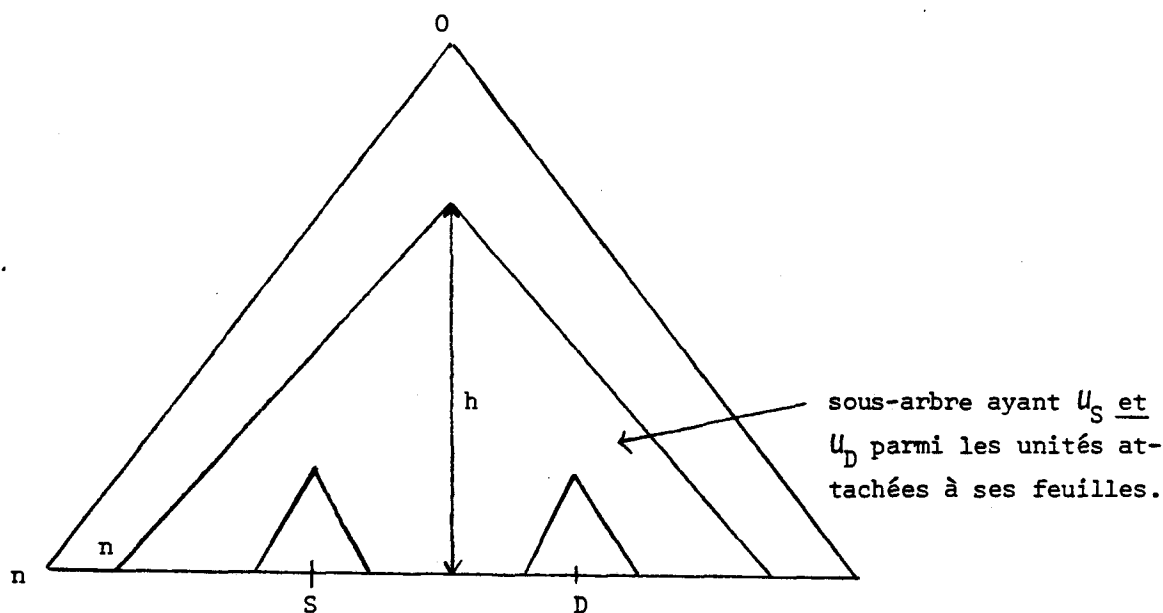
- Preuve de la procédure de routage.

Soient deux unités  $U_S$  et  $U_D$ , attachées respectivement aux deux cellules du niveau feuille, dont les numéros sont S et D

$$S = s_0 \dots s_n$$

$$D = d_0 \dots d_n$$

$U_S$  est l'unité source, et  $U_D$  l'unité destinataire.



La détermination de la hauteur du sous-arbre ayant  $U_S$  et  $U_D$  dans son niveau feuille, est immédiate, d'après la fonction de numérotation.

$$\text{Posons } m = \min_{0 \leq k \leq n} (s_k \neq d_k) \Leftrightarrow \begin{cases} s_0 \dots s_{m-1} = d_0 \dots d_{m-1} \\ \text{et } s_m \neq d_m \end{cases}$$

La hauteur  $h$ , du sous-arbre est :  $h = n - m + 1$

- Extension de la relation "est connecté à"

Cette relation a été définie dans le chapitre II, entre deux cellules de l'ANNEAU. Elle est étendue à deux cellules quelconques du réseau-ARBRE, de la manière suivante : soient X, Y deux cellules quelconques du réseau-ARBRE,  $X \leftrightarrow Y \Leftrightarrow \exists$  une suite de p cellules :  $(Z_i)_{i=1}^{i=p}$  telles que  $Z_1 = X$ ,  $Z_p = Y$  et  $\forall i = 1, p-1$   $Z_i, Z_{i+1}$  appartiennent au même ANNEAU du réseau-ARBRE et  $Z_i \leftrightarrow Z_{i+1}$ .

Le lancement d'une demande D dans le réseau, se traduit au départ par deux opérations :

1- l'unité source commute d'abord sa cellule

$$\eta(S) = E$$

2-  $D \xrightarrow{p} A_n^S \text{ div } a$ , la demande D est envoyée dans l'ANNEAU feuille  $A_n^S \text{ div } a$  et auquel l'unité  $U_S$  est attachée.

$$\text{donc au départ } \begin{cases} S \leftrightarrow S \\ D \xrightarrow{p} A_n^S \text{ div } a \end{cases} \quad (1)$$

La preuve de l'algorithme comporte deux parties, chacune d'elles correspond à une étape de progression de la demande D dans le réseau-ARBRE. La 1<sup>ère</sup> est l'étape de montée de la demande vers la racine du sous-arbre contenant  $U_S$  et  $U_D$ ; la seconde est l'étape de descente de la demande D depuis cette racine vers l'unité destinataire.

Etape 1

$\forall p : 0, \dots, n-m.$

$$E_1(p) : \begin{cases} S \leftrightarrow s_0 \dots s_{n-p} \\ D \xrightarrow{p} A_{n-S}^S \text{ div } a^{p+1} \end{cases}$$

Etape 2

$\forall p : 0, \dots, n-m.$

$$E_2(p) : \begin{cases} S \leftrightarrow d_0 \dots d_m \dots d_{m+p} \\ \text{et} \\ D \xrightarrow{p} A_{m+p+1}^D \text{ div } a^{n-(m+p)} \end{cases}$$



La preuve de chacune de ces deux étapes se fait par récurrence finie sur  $p$ , voir Annexe.

### 1.c.2. La procédure d'accès

Chaque cellule  $C_{i,j}^k$ , sauf celles auxquelles les unités sont attachées, est commune aux deux ANNEAUX  $A_i^j$  et  $A_{i+1}^{j+k}$ . Deux modules de contrôle, notés  $MC_{i,j}^k$  et  $MC_{i+1,j+k}^a$  sont alors associés à chacune de ces cellules. Ils définissent la procédure d'accès à leurs deux ANNEAUX respectifs et élaborent la commande  $!C_{i,j}^k$  de la cellule commune à ces deux ANNEAUX.

#### - Commande de la cellule $C_{i,j}^k$

Les demandes sont envoyées dans le réseau de manière asynchrone. Il doit être de même pour leur traitement par les modules de contrôle. L'état d'une cellule  $C_{i,j}^k$  est fonction des commandes  $!C_{i,j}^k$  et  $!C_{i+1,j+k}^a$ , générées par deux modules distincts et asynchrones - figure 8 -, ce qui soulève le problème de la compatibilité de ces deux commandes.

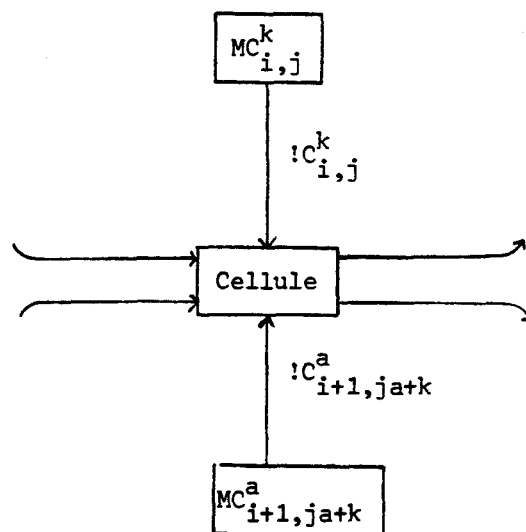


Figure 8 : Commande d'une cellule par deux modules distincts et asynchrones.

Le tableau ci-après détermine l'état d'une cellule en fonction des valeurs des deux commandes qui lui sont appliquées.

$!C_{i,j}^k$	$!C_{i+1,j+k}^a$	$n(C_{i,j}^k)$
0	0	T
1	0	E
0	1	E
1	1	Incompatible

Les conditions nécessaires pour qu'un module puisse faire passer la valeur de sa commande de 0 à 1 ; et mettre ainsi la cellule dans l'état Echange sont :

i) L'état courant de la cellule est Transparent

$$n(C_{i,j}^k) = T$$

ii) Les deux ANNEAUX  $A_i^j$  et  $A_{i+1}^{j+k}$  ayant  $C_{i,j}^k$  comme cellule commune, sont alloués.

iii) L'allocation des deux ANNEAUX a été effectuée pour satisfaire la même demande.

iv) Seul un des deux modules partageant la cellule  $C_{i,j}^k$  est allocateur de son propre ANNEAU.

L'ensemble des conditions précédentes i) à iv), nous a amené à doter les deux modules de moyens de coordination. Nous avons alors introduit une double liaison (REQ) entre ces deux modules - figure 9 -.

L'environnement d'interaction d'un module sera donc composé, en plus de sa cellule de commutation  $C_{i,j}^k$ , des trois modules de contrôle suivants :

1 -  $P(MC_{i,j}^k)$  : Prédécesseur dans  $A_i^j$  de  $MC_{i,j}^k$ .

2 -  $S(MC_{i,j}^k)$  : Successeur dans  $A_i^j$  de  $MC_{i,j}^k$ .

3 -  $v(MC_{i,j}^k)$  : Le second module avec lequel  $MC_{i,j}^k$  partage la cellule  $C_{i,j}^k$ .

$$v(MC_{i,j}^k) = MC_{i+1,j+k}^a \quad k = 0, -, a-1$$

$$v(MC_{i,j}^a) = MC_{i-1,j}^{j \bmod a}$$

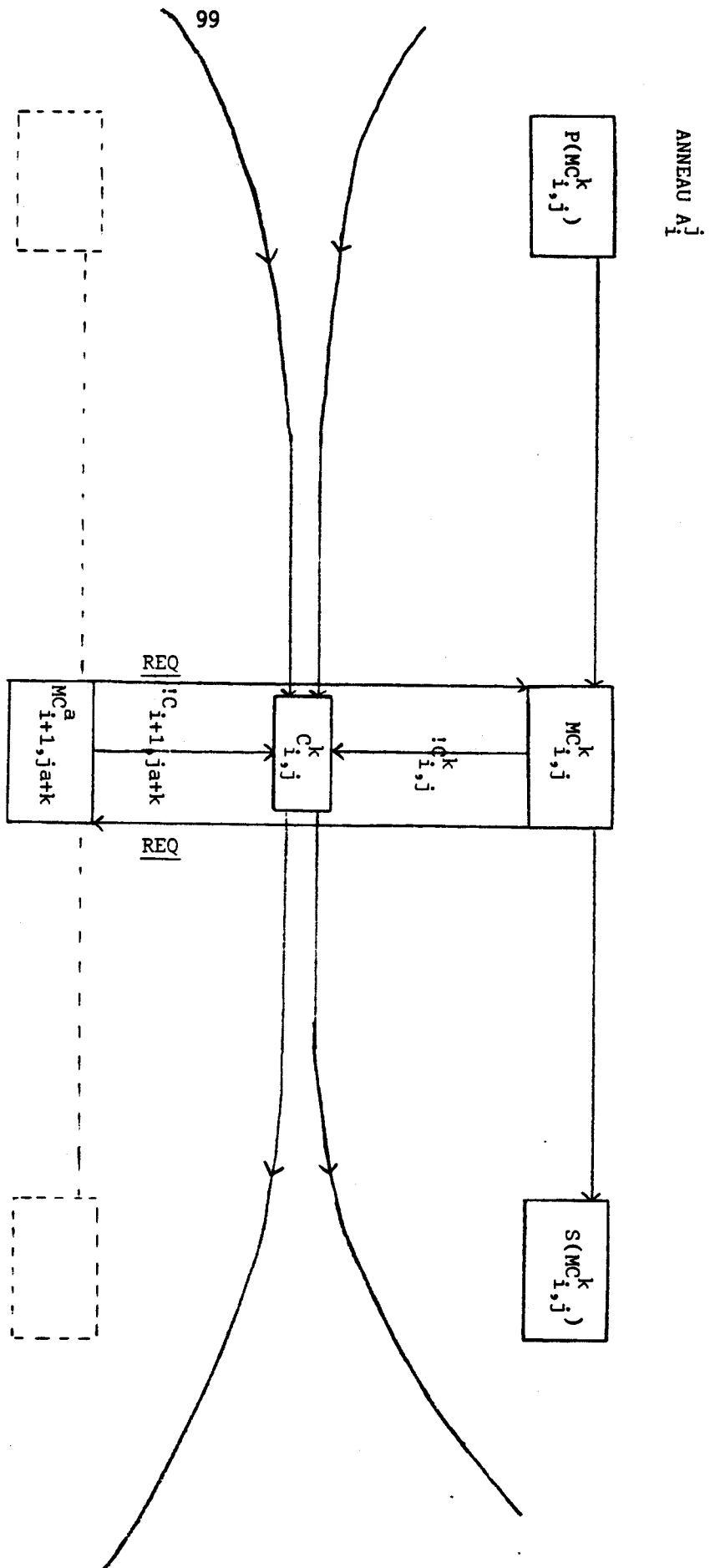
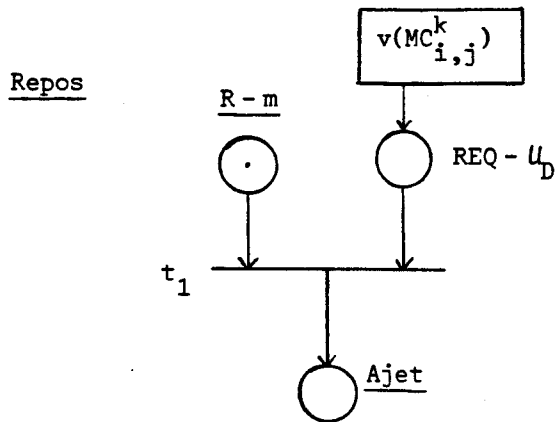


Figure 9 : Environnement d'interaction du module  $MC_{i,j}^k$

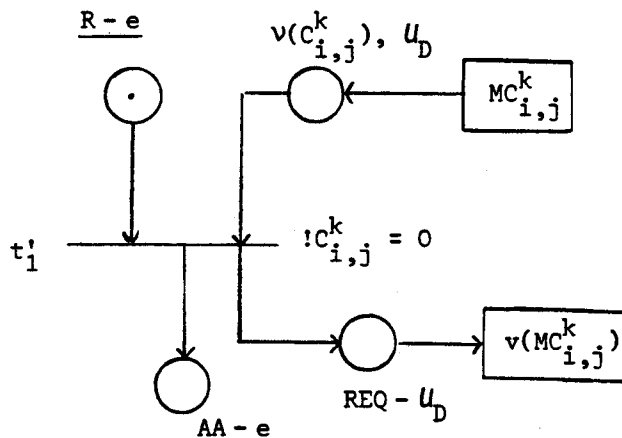
( $k \neq a$ )

Les opérations d'allocation et de libération d'un ANNEAU du réseau ARBRE sont élaborées, de la même manière que dans le modèle de base, au moyen d'un jeton circulant dans chaque ANNEAU. Cependant, la procédure d'accès a été légèrement modifiée afin d'assurer que les quatre conditions de commutation d'une cellule, soient vérifiées.

Nous reprenons ici, une nomenclature des états d'un module de contrôle et des évènements qui provoquent leurs changements, identique à celle utilisée lors de la définition du module de contrôle du modèle de base - cf. §II.2.b.3.b. -.



$$t_1 : !c_{i,j}^k = 0$$



$$t_1' : !c_{i,j}^k = 0$$

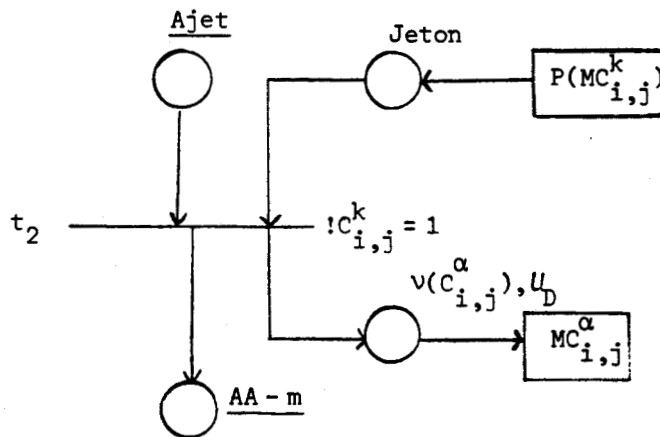
Attente du jeton

Le module  $MC_{i,j}^k$  passe dans cet état après réception d'une requête de son module voisin. Cela implique que l'ANNEAU  $A_{i+1}^{ja+k}$ , auquel appartient le module  $v(MC_{i,j}^k)$ , est alloué.

(1)  $!C_{i,j}^k = 0, \exists \beta \ 0 \leq \beta \leq a-1$  tel que

$$\text{Alloc}(A_{i+1}^{ja+k}) = MC_{i+1,ja+k}^\beta$$

On note  $\text{Alloc}(A_i^j)$  le module allocateur de l'ANNEAU  $A_i^j$ .



(2)  $t_2 : !C_{i,j}^k = 1$  et  $\text{Alloc}(A_i^j) = MC_{i,j}^k$ .

Vérifions que les quatre conditions de mise à l'état échange de la cellule  $C_{i,j}^k$  sont satisfaites lors de la transition  $t_2$ .

i)  $\eta(C_{i,j}^k) = T$

Condition vérifiée lors de la transition  $t_1$  qui précède  $t_2$ .

ii) Les deux ANNEAUX  $A_i^j$  et  $A_{i+1}^{ja+k}$  sont alloués,  $A_{i+1}^{ja+k}$  est alloué d'après (1), et l'arrivée du jeton provoquant la transition  $t_2$  entraîne l'allocation de  $A_i^j$  (2).

iii)  $A_i^j$  et  $A_{i+1}^{ja+k}$  sont alloués en vue de satisfaire une même demande.

L'allocation de  $A_i^j$  suit celle de  $A_{i+1}^{j+k}$  et elle a été effectuée suite à une requête émanante du module  $v(MC_{i,j}^k)$ , de l'ANNEAU  $A_{i+1}^{j+k}$ . Par conséquent, l'allocation des deux ANNEAUX a pour objet l'établissement d'une même connexion.

iv) Seul un des deux modules partageant la cellule  $C_{i,j}^k$  est allocateur de son propre ANNEAU. Cette condition est satisfaite, car :

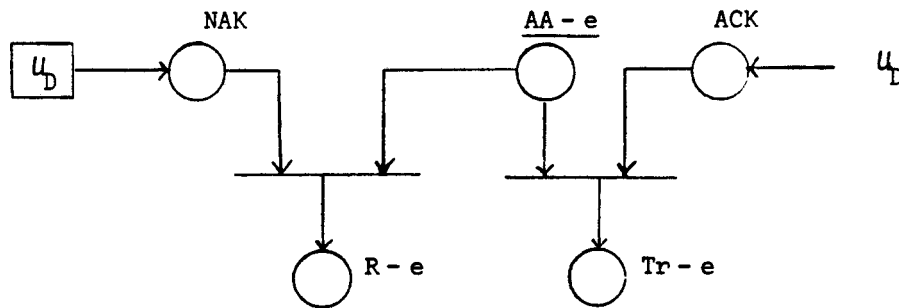
$$\text{Alloc}(A_i^j) = MC_{i,j}^k \quad \text{d'après (2)}$$

et

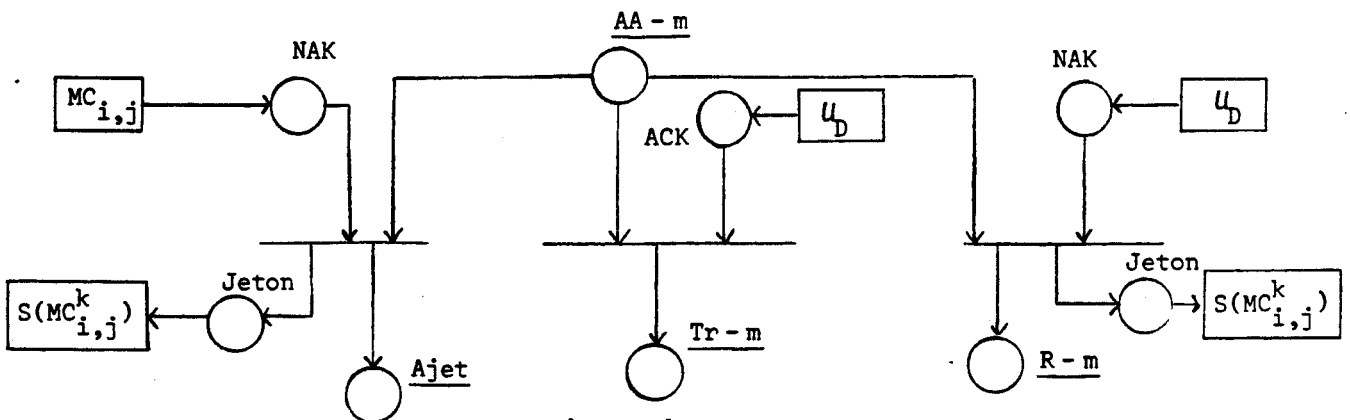
$$\begin{aligned} \text{Alloc}(A_{i+1}^{j+k}) &= MC_{i+1,j+k}^\beta & 0 \leq \beta \leq a-1 \\ &\text{d'après (1)} \\ &\neq v(MC_{i,j}^k) = MC_{i+1,j+k}^a \end{aligned}$$

Les quatre conditions sont donc vérifiées et la transition  $t_2$ , qui provoque la commutation de la cellule, peut par conséquent avoir lieu.

Attente acquittement

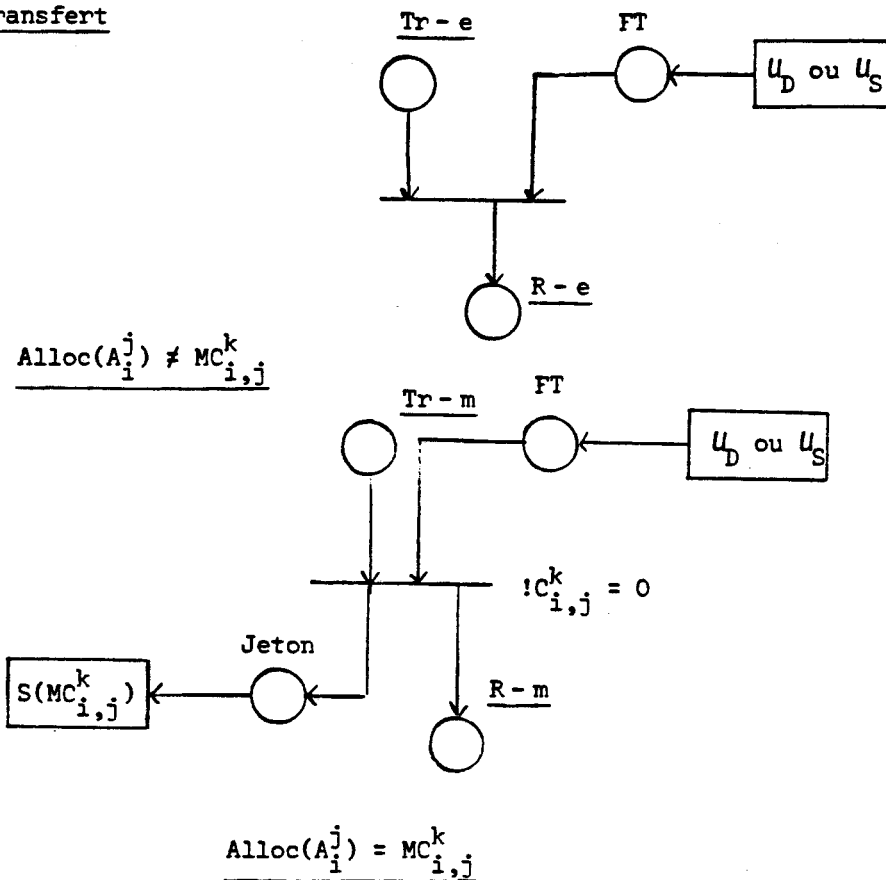


$$\underline{\text{Alloc}(A_i^j) \neq MC_{i,j}^k}$$



$$\underline{\text{Alloc}(A_i^j) = MC_{i,j}^k}$$

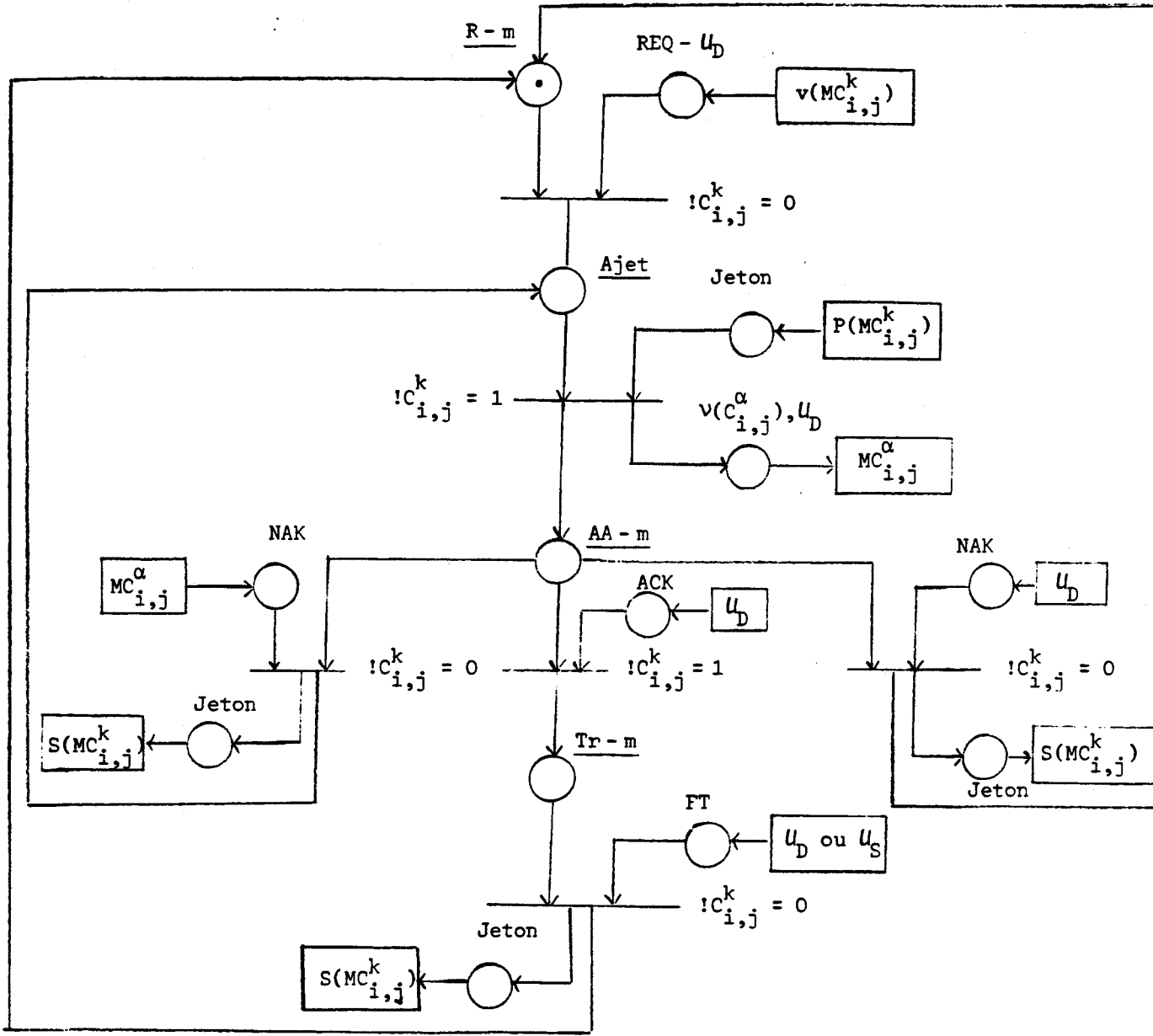
Transfert



Deux cas ont été distingués pour les états Attente acquittement et Transfert, selon que le module  $MC_{i,j}^k$  est ou non l'allocateur de son ANNEAU  $A_i^j$ .

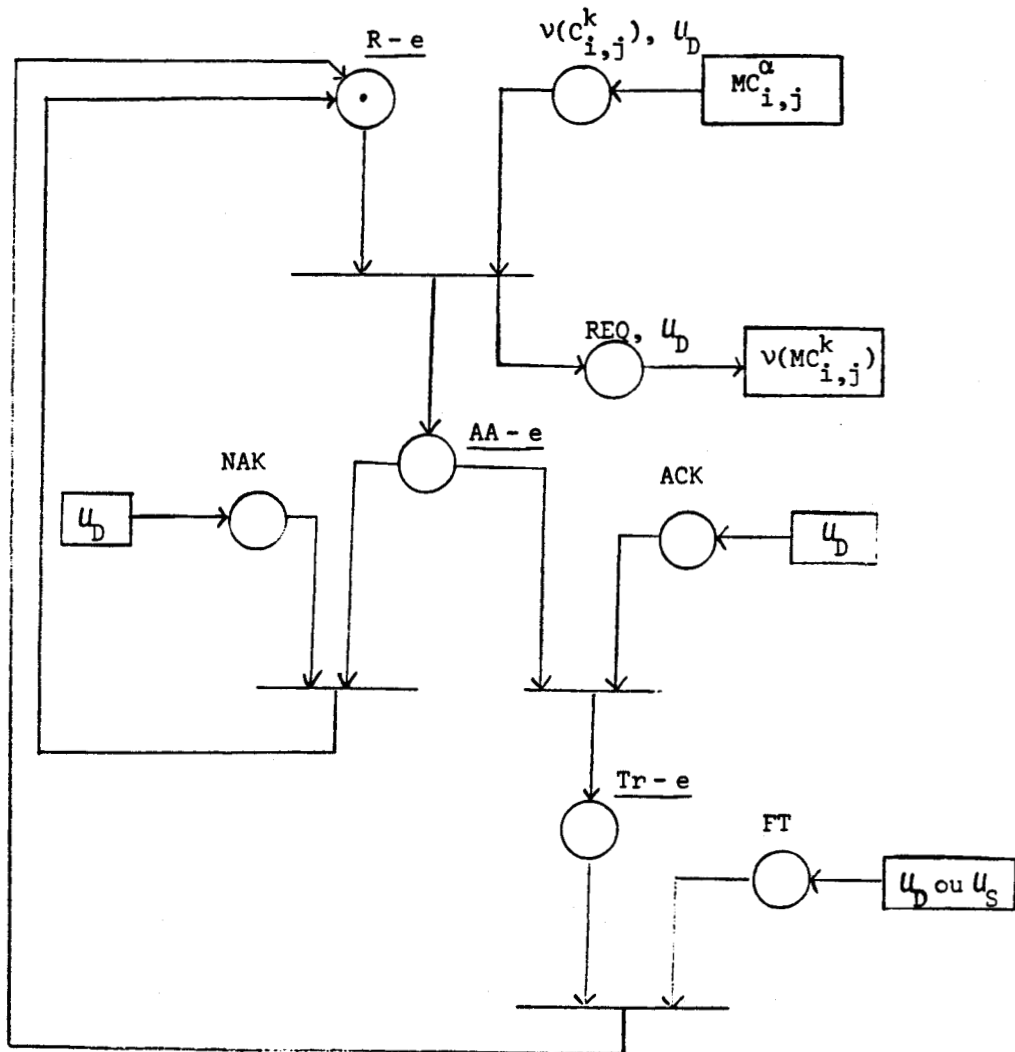
Le protocole d'accès complet d'un module de contrôle est représenté par les deux figures ci-après :





Protocole d'accès du module  $MC_{i,j}^k$  lorsqu'il est allocateur de l'ANNEAU  $A_i^j$ .





Protocole d'accès du module  $MC_{i,j}^k$  lorsqu'il n'est pas allocateur de son ANNEAU  $A_i^j$ .



### 1.c.3. Situations de conflit et de blocage

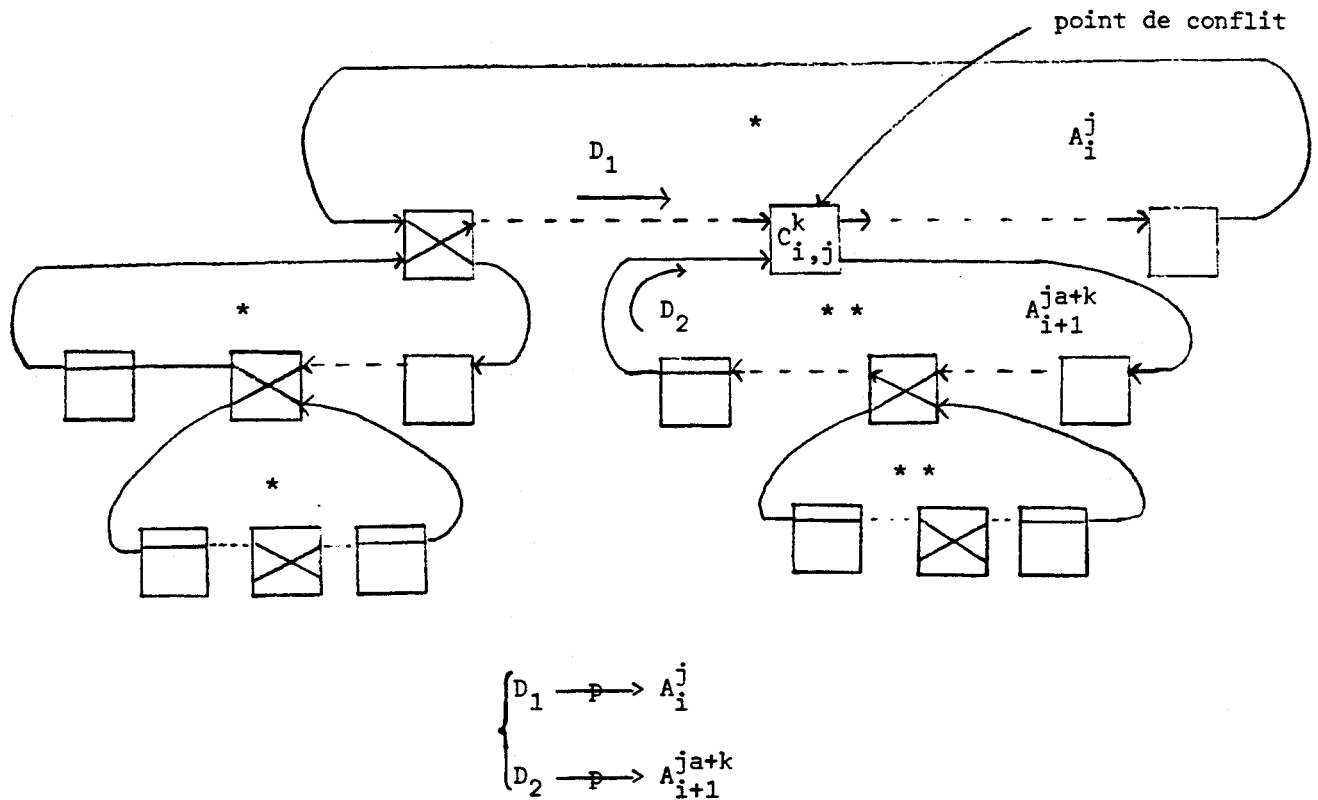
#### 1 - Situation de conflit

Deux demandes  $D_1, D_2$  se propageant simultanément dans les ANNEAUX  $A_i^j$  et  $A_{i+1}^{ja+k}$  du réseau-ARBRE ( $D_1 \xrightarrow{p} A_i^j, D_2 \xrightarrow{p} A_{i+1}^{ja+k}$ ) sont dites *conflituelles* si et seulement si l'établissement de chacune des deux demandes nécessite qu'à l'étape suivante on ait  $D_1 \xrightarrow{p} A_{i+1}^{ja+k}$  et  $D_2 \xrightarrow{p} A_i^j$ .

Autrement dit, les deux connexions s'intersectent aux anneaux  $A_i^j$  et  $A_{i+1}^{j+k}$ .

Le conflit se situe donc, au niveau de la cellule  $C_{i,j}^k$  commune aux deux ANNEAUX  $A_i^j$  et  $A_{i+1}^{j+k}$ . Les opérations de *détection* et de *résolution* du conflit doivent, par conséquent, être effectuées par les modules  $MC_{i,j}^k$ ,  $MC_{i+1,j+k}^a$ .

La figure ci-après illustre une telle situation.



Les ANNEAUX marqués d'une \* (resp. \*\*) sont ceux déjà alloués pour établir la demande  $D_1$  (resp.  $D_2$ ).

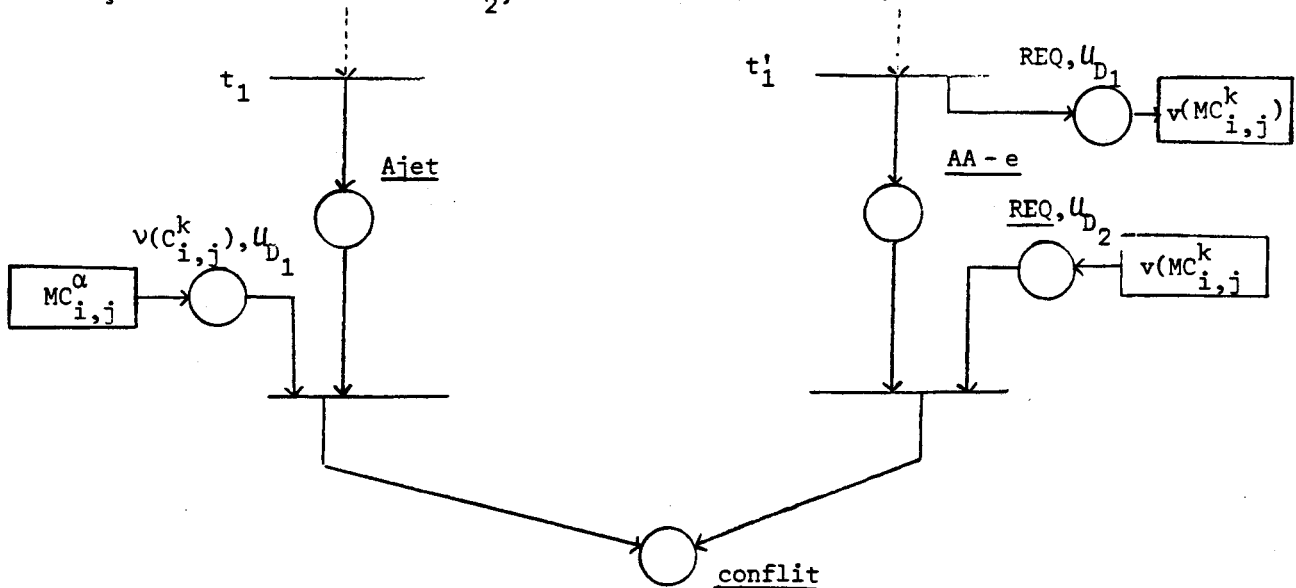
Les mécanismes de détection et de résolution de conflit sont également à la charge des modules de contrôle. Un nouvel état *conflit* est donc ajouté à l'ensemble des états d'un module, cet ensemble a été présenté lors de la définition du protocole d'accès.

### Détection de conflit

Un module  $MC_{i,j}^k$ , passe dans l'état *conflit* dans les deux cas suivants :

cas 1 : S'il se trouve dans l'état Attente de jeton, à la suite d'une première demande  $D_1$ , en provenance de son module voisin, lorsqu'il reçoit une seconde demande  $D_2$  en provenance d'un des modules  $MC_{i,j}^\alpha$  de son ANNEAU.

cas 2 : S'il se trouve dans l'état Attente d'acquiescement, après réception d'une première demande  $D_1$  et envoi d'une requête à son module voisin, lorsqu'il reçoit une seconde demande  $D_2$ , de ce même module voisin.



Protocole de détection de conflit par un module  $MC_{i,j}^k$ .

### Résolution de conflit

Une politique de résolution de conflit doit déterminer :

- a - La demande prioritaire et qui, par conséquent, est autorisée à continuer sa propagation dans le réseau.
- b - L'action à entreprendre vis-à-vis de la demande non prioritaire.

a - Détermination de la demande prioritaire :

La priorité d'une demande par rapport à une autre peut être établie selon deux critères.

Le premier critère est indépendant des caractéristiques du réseau d'interconnexion. Il est basé sur les priorités logiques prédéfinies pour chacune des unités. Dans ce cas, chaque demande est accompagnée de la valeur de la priorité logique de l'unité qui l'a émise, la procédure de détermination de la demande prioritaire consiste alors à comparer les priorités des deux demandes conflictuelles. Nous ne retiendrons pas ce critère.

Le second critère tient compte des caractéristiques du réseau d'interconnexion. Il vise en particulier la réduction du délai de séjour d'une demande dans le réseau. Les deux méthodes que nous proposons par la suite visent cet objectif. La première alterne la priorité entre les demandes conflictuelles des deux modules voisins  $MC_{i,j}^k$  et  $MC_{i+1,j+k}^a$ . La seconde attribue la priorité à la demande la plus "avancée" dans le réseau.

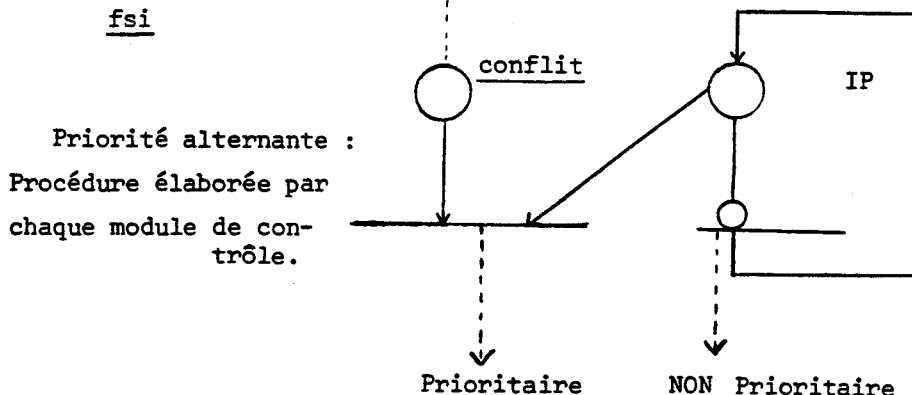
a-1 Priorité alternante

Chacun des deux modules  $MC_{i,j}^k$  et  $MC_{i+1,j+k}^a$  dispose d'un indicateur de priorité (IP). Quand un conflit se produit, un module reconnaît s'il est ou non prioritaire et modifie son propre indicateur.

si IP = vrai alors  $\left\{ \begin{array}{l} \text{IP} \leftarrow \text{Faux} \\ \text{Prioritaire} \end{array} \right.$

sinon  $\left\{ \begin{array}{l} \text{IP} \leftarrow \text{Vrai} \\ \text{NON-Prioritaire} \end{array} \right.$

fsi



Initialement, les priorités associées à deux modules voisins  $MC_{i,j}^k$  et  $MC_{i+1,ja+k}^a$  doivent être opposées :

$$IP(MC_{i,j}^k) = \overline{IP(MC_{i+1,ja+k}^a)}.$$

a-2 Priorité à la demande la plus "avancée"

Une demande est dite plus "avancée" par rapport à une autre, si le nombre d'ANNEAUX qui lui restent à acquérir avant d'atteindre sa destination, est plus faible.

Remarque : - Une situation dans laquelle deux demandes conflictuelles se trouvent dans des états d'avancement identiques, ne peut pas se produire dans le réseau-ARBRE.

En effet, deux demandes conflictuelles se trouvent nécessairement dans deux niveaux adjacents  $i$  et  $i+1$ . Une des demandes, soit  $D_1$ , se trouve au niveau  $i$  et doit passer au niveau  $i+1$ , et l'autre,  $D_2$ , se trouve au niveau  $i+1$  et doit passer au niveau  $i$ ; autrement dit,  $D_1$  a déjà atteint sa racine, et  $D_2$  ne l'a pas encore atteint - figure 10 -.

Notons  $\ell_i$  le nombre d'ANNEAUX acquis par  $D_i$ .

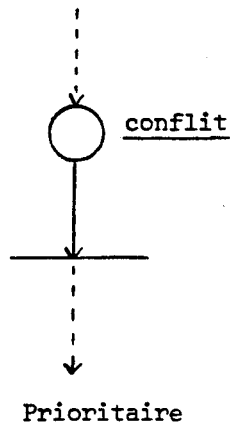
$\overline{\ell}_i$  le nombre d'ANNEAUX non encore acquis par  $D_i$ .

On a  $\overline{\ell}_1 = \ell_2$  et  $\overline{\ell}_2 \geq \ell_2 + 1$

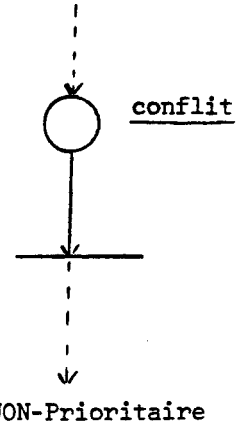
d'où  $\overline{\ell}_1 + 1 \leq \overline{\ell}_2 \Rightarrow \overline{\ell}_1 < \overline{\ell}_2$ .

$\overline{\ell}_1$  est donc différent de  $\overline{\ell}_2$ , de plus, c'est toujours la demande  $D_1$  (qui passe du niveau  $i$  au niveau  $i+1$ ) qui est la plus avancée ( $\overline{\ell}_1 < \overline{\ell}_2$ ).

La priorité est donc affectée à la descente, par rapport à la montée. Le module  $MC_{i,j}^k$ , du niveau  $i$ , est donc en permanence prioritaire par rapport à son module voisin  $MC_{i+1,ja+k}^a$ , du niveau  $i+1$ .



a - Module  $MC_{i,j}^k$  ( $k \neq a$ )



b - Module  $MC_{i+1,j+a}^a$

Priorité à la "descente".

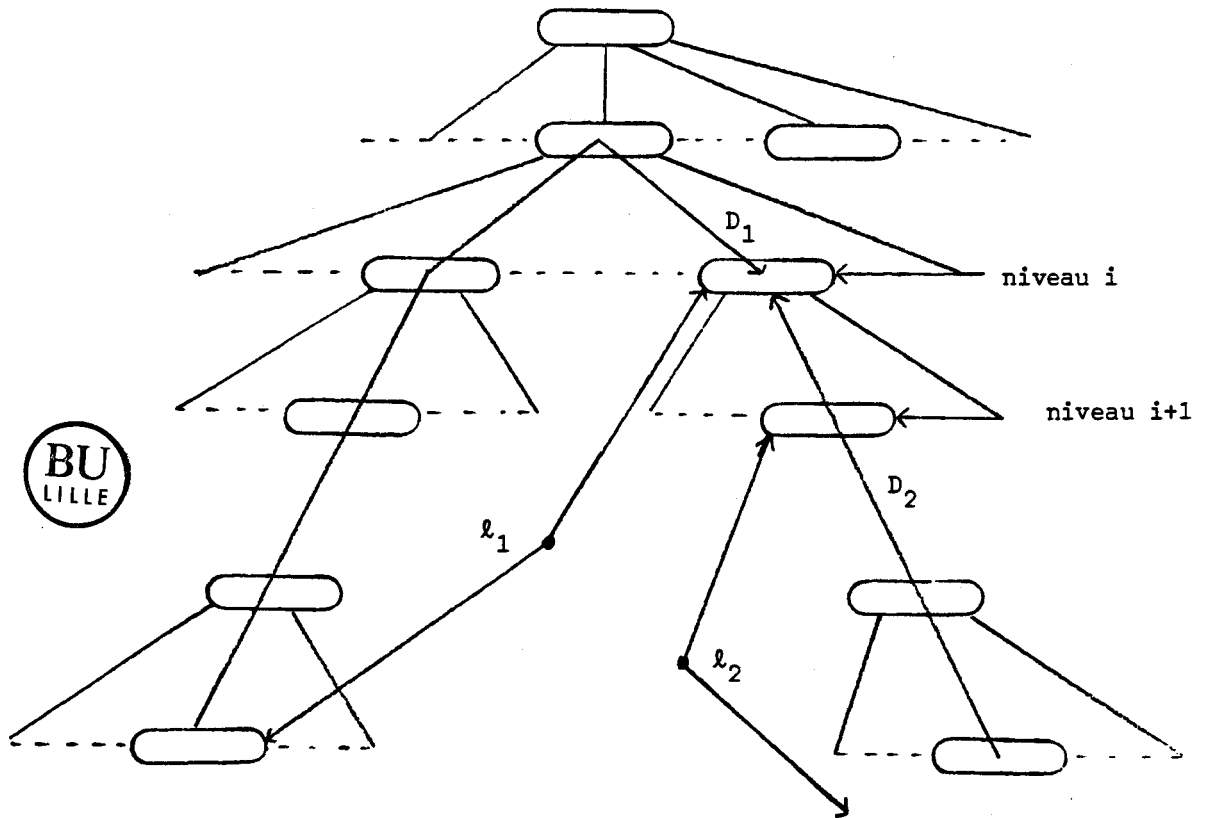


Figure 10 : Conflit entre  $D_1$  et  $D_2$ .  $D_1$  a atteint sa racine alors que  $D_2$  ne l'a pas encore atteint.

Une demande favorisée par cette méthode, est non seulement dans un état d'avancement supérieur, mais est en plus celle qui dispose, au moment de la détection de conflit, de plus de ressources de communication.

Si on note  $h_i$ , la hauteur du sous-arbre associé à la demande  $D_i$  :

$$l_i + \overline{l_i} = 2h_i - 1$$

$$l_1 \geq h_1 > l_2 \quad \Rightarrow \quad l_1 > l_2$$

b - Action vis-à-vis de la demande non prioritaire :

Afin que la demande prioritaire puisse continuer sa progression dans le réseau, il est indispensable que la demande non prioritaire soit d'abord retirée de l'ANNEAU dans lequel elle se trouvait au moment de la détection de conflit, qui est aussi le dernier ANNEAU qu'elle a acquis. Ensuite, il est possible d'adopter l'une des deux politiques suivantes, vis-à-vis de cette même demande.

b-1 Retrait du réseau - figure b-1 -

La demande non prioritaire, est retirée du réseau. Elle libère l'ensemble des ANNEAUX qu'elle a auparavant acquis. Elle pourra être soumise de nouveau au réseau dès que le module de contrôle attachée à l'unité émettrice de cette demande réobtient l'allocation de son propre ANNEAU.

b-2 Maintient dans le réseau - figure b-2 -

La demande non prioritaire, perd l'allocation de l'ANNEAU du niveau  $i+1$  et se met immédiatement en état d'attente d'allocation de cet ANNEAU en se plaçant dans le niveau  $i+2$ .

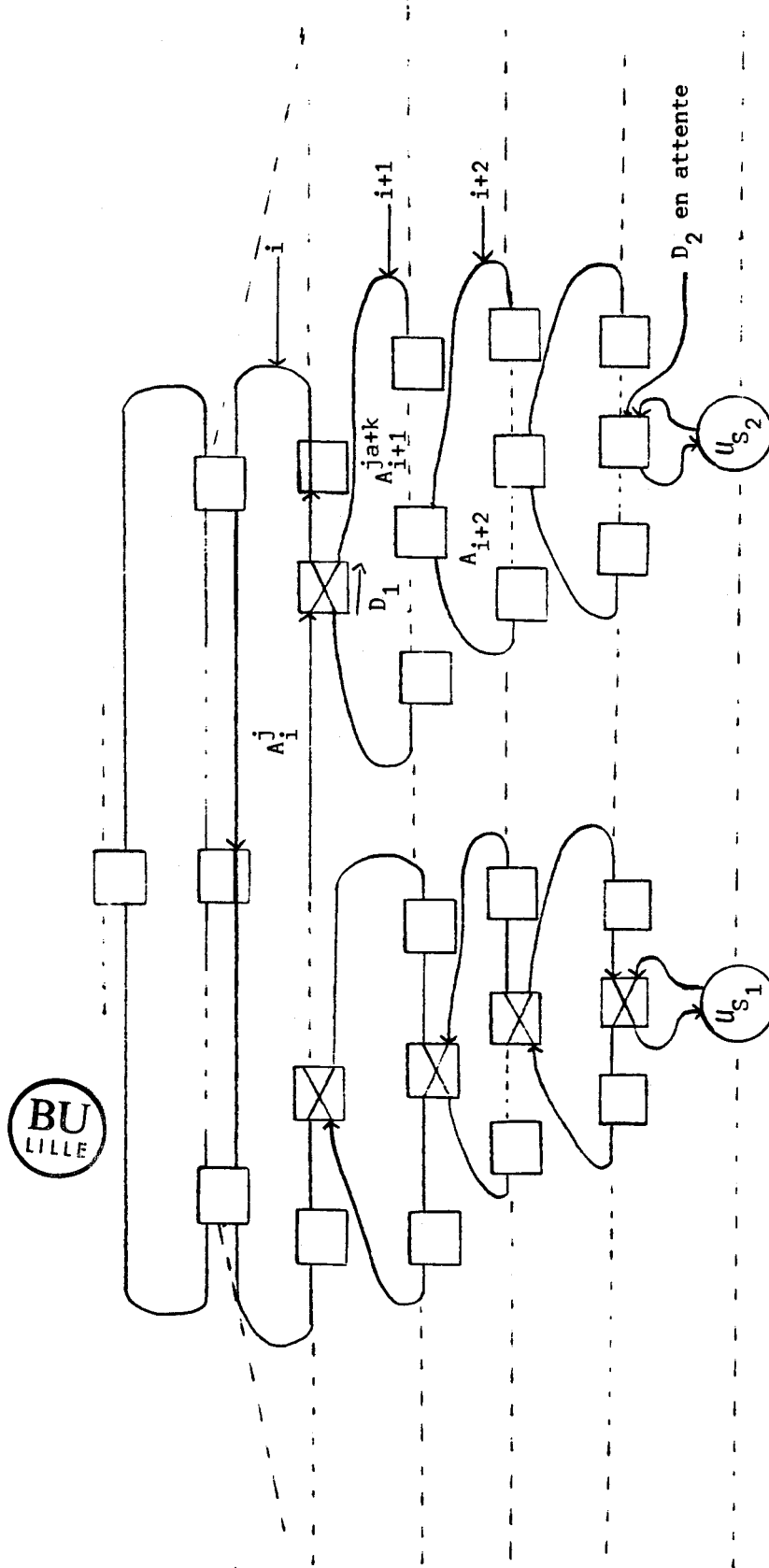


Figure b-1 : Retrait total de la demande non prioritaire, ici  $D_2$ .



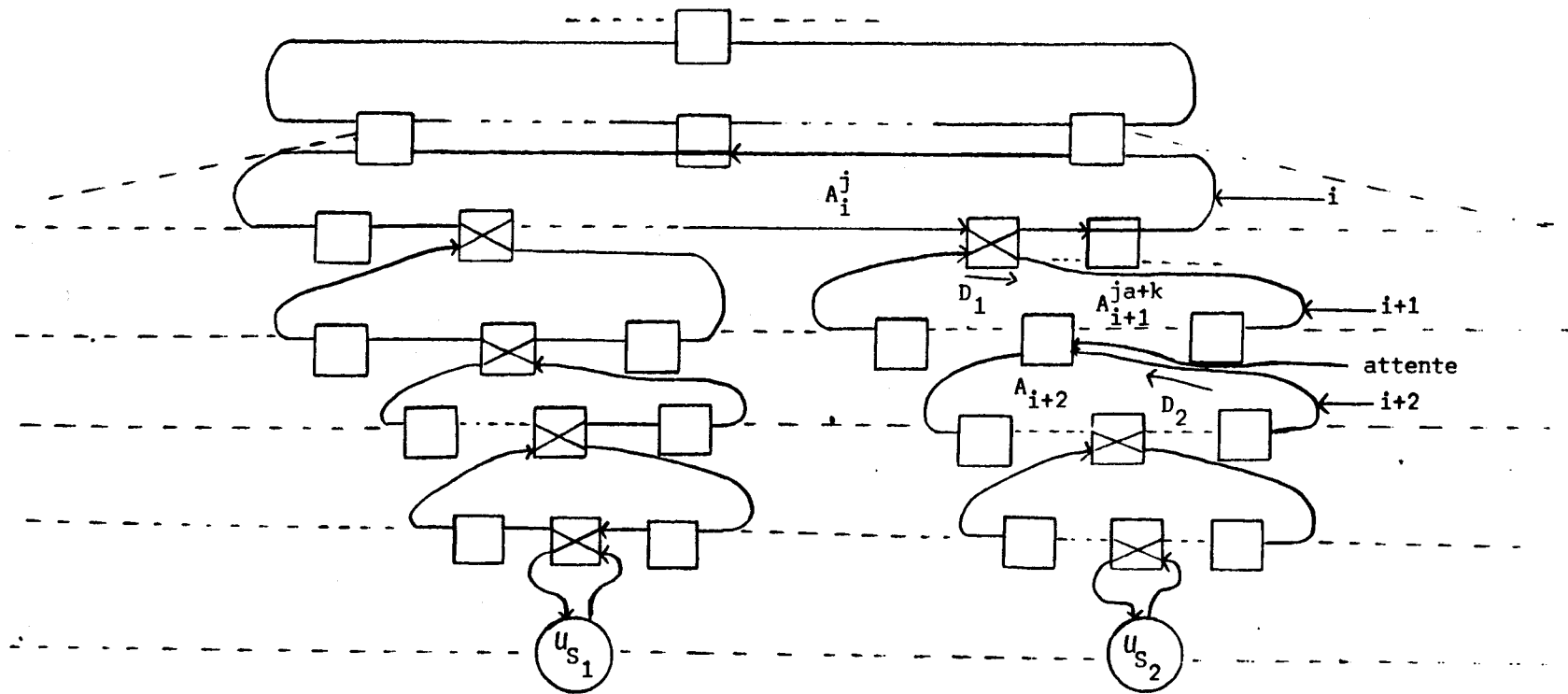


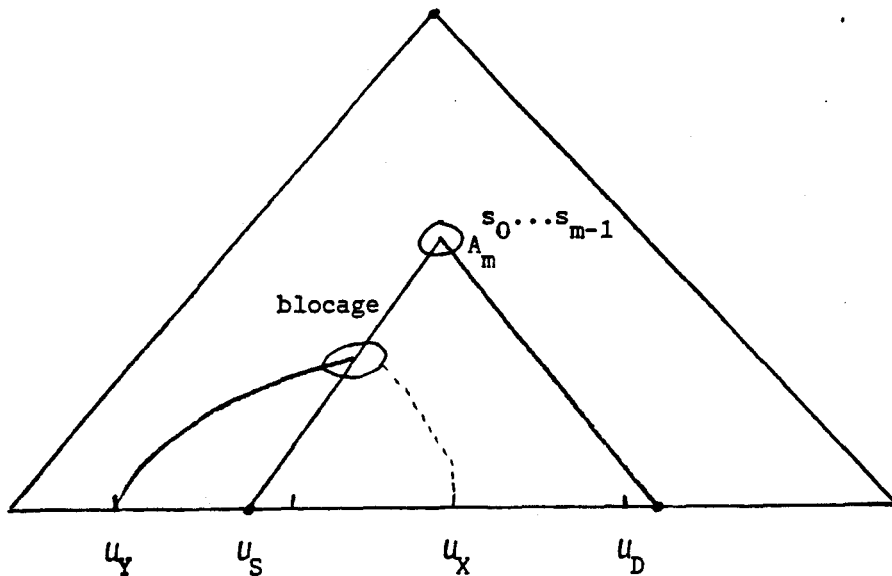
Figure b-2 : Maintient de la demande  $D_2$ , non prioritaire dans le réseau.



## 2) Situation de blocage

Le réseau-ARBRE est un outil de communication avec blocage.

En effet, l'établissement d'une liaison entre deux unités  $u_S$ ,  $u_D$ , empêche toute communication entre d'une part les unités du sous-arbre ayant  $u_S$  et  $u_D$  en feuilles et les autres unités du réseau.



Situation de blocage.

L'ANNEAU racine du sous-arbre contenant  $u_S$  et  $u_D$  est :

$$A_m^{d_0 \dots d_{m-1}} = A_m^{s_0 \dots s_{m-1}} \text{ avec } m = \min_{a \leq b \leq n} (s_k \neq d_k).$$

Les unités  $u_X$  attachées au sous-arbre de racine  $A_m^{s_0 \dots s_{m-1}}$  sont telles que  $x_0 \dots x_{m-1} = s_0 \dots s_{m-1}$  ( $X = x_0 \dots x_n$ ).

Les unités  $u_Y$  non attachées au sous-arbre de racine  $A_m^{s_0 \dots s_{m-1}}$  sont telles que  $y_0 \dots y_{m-1} \neq s_0 \dots s_{m-1}$ .

Il est impossible d'établir une liaison entre  $u_X$  et  $u_Y$ .

Une situation de blocage se produit donc lorsqu'une demande, soit  $D_2$ , pour pouvoir continuer sa progression dans le réseau, nécessite l'allocation d'un ANNEAU définitivement acquis par une autre demande, soit  $D_1$ . - Figure 11 -.

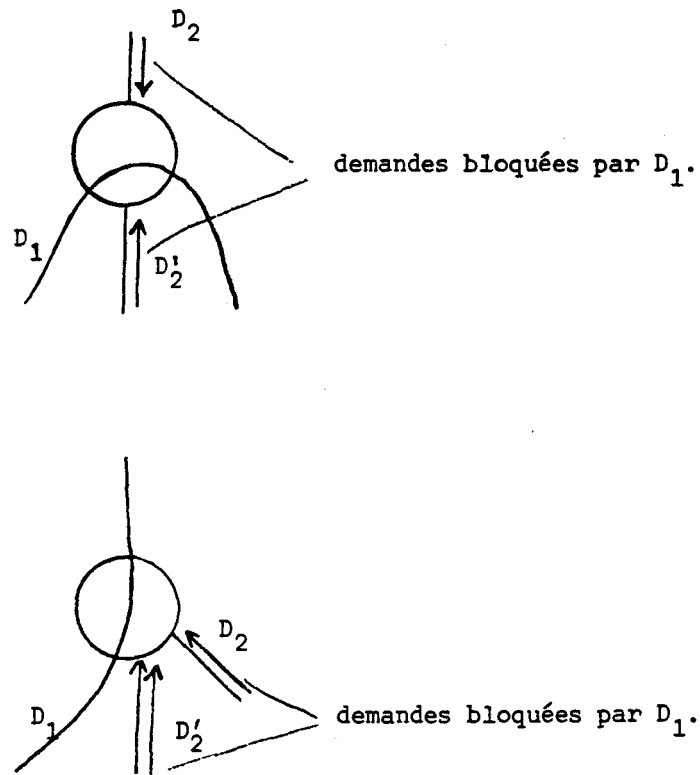
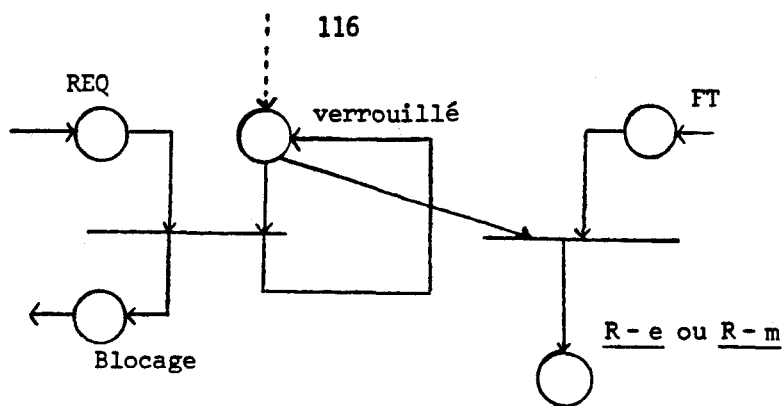


Figure 11 : Situation de blocage.

Le protocole élaboré par les modules de contrôle du réseau-ARBRE, ne nous permet pas de détecter une situation de blocage. Il est toutefois possible de remédier à ce problème, en verrouillant les modules de contrôle faisant partie d'une liaison définitivement établie. Par exemple, lors de la transmission d'un acquittement positif depuis une unité destinataire vers une unité source, le passage de cet acquittement par un ANNEAU, entraîne le verrouillage de ses modules de contrôle. Ce verrouillage est levé dès que la liaison est libérée.

Lorsqu'une demande d'allocation d'un ANNEAU est soumise à un module de contrôle verrouillé, ce module envoie le message "blocage" au module source de la demande.



#### Détection de la situation de blocage.

Les deux solutions qui consistent à maintenir ou retirer une demande et développées dans le cas d'une situation de conflit, peuvent être retenues lors de la détection d'un blocage. Cependant, l'opération de retrait ou de maintien n'est entreprise que vis-à-vis de la demande bloquée, celle qui ne peut pas continuer sa progression dans le réseau, la demande bloquante n'étant jamais suspendue.

Le retrait d'une demande du réseau, à la suite d'un conflit ou d'un blocage, a l'avantage de libérer les ressources acquises par une demande qui ne peut être immédiatement satisfaite. La demande retirée ne pourra en effet être satisfaite, qu'après un délai au moins égal à celui nécessaire pour que la demande maintenue puisse compléter sa connexion et terminer le transfert. Les ANNEAUX ainsi libérés, peuvent alors être utilisés pour satisfaire d'autres demandes. Cependant, ceci a l'inconvénient d'accroître le délai d'établissement de la demande retirée. Celle-ci doit, en effet, réobtenir l'allocation de toutes les ressources qu'elle a auparavant libérées, suite à la détection d'un conflit ou d'un blocage.

Le maintien dans le réseau d'une demande, bloquée ou non prioritaire, réduit en revanche son délai d'établissement mais a l'inconvénient de maintenir occupés, inutilement, des ANNEAUX, alors qu'ils peuvent être alloués pour satisfaire d'autres éventuelles demandes soumises au réseau.

Les solutions que nous avons présentées pour détecter et résoudre les situations de conflit ou de blocage sont suffisantes pour rendre impossible des situations d'interblocage.

En effet, un interblocage ne peut se produire qu'entre deux demandes n'ayant pas établie la totalité de leurs connexions, et qui, soit sont en conflit, soit l'une d'entre-elles bloque l'autre.

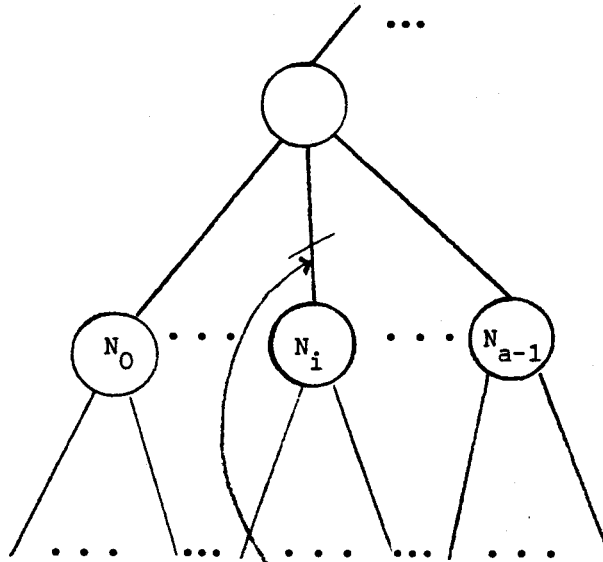
- Conflit : un conflit ne peut se produire qu'entre une demande en "montée" et une demande en "descente". Les solutions proposées permettent de détecter et lever cette situation.
- Blocage : la connexion établie par la demande bloquante est libérée tôt ou tard. Le délai fini devant être garanti par le niveau "Echange" du protocole de communication.

### 1.d. Conclusion

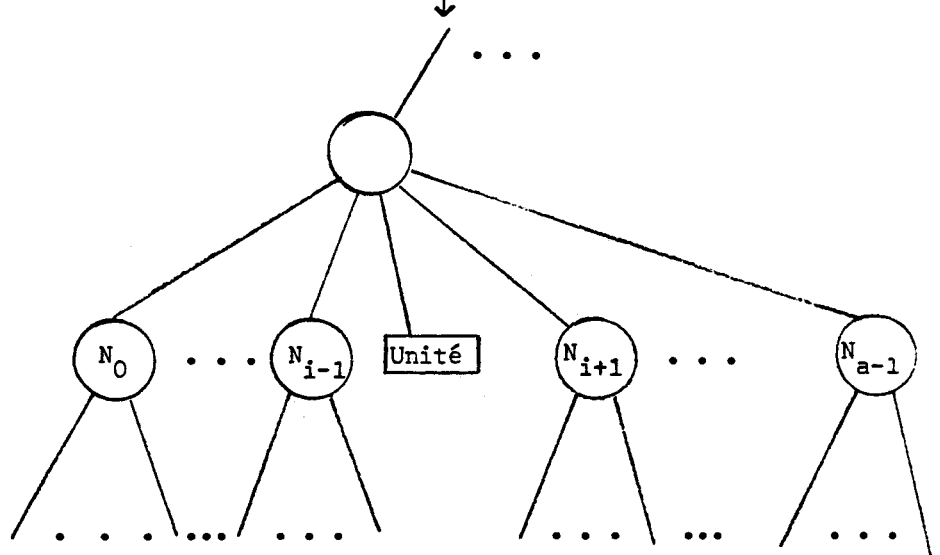
Le réseau-ARBRE, extension de l'ANNEAU, est un outil de communication simple face. Une fonction de numérotation de ses cellules de commutation ainsi qu'une mise en oeuvre répartie du protocole d'interconnexion ont été proposées.

Notons que ce réseau est un outil de communication avec blocage et que son utilisation ne peut se justifier que si les unités peuvent être réparties en  $n+1$  partitions, une partition par niveau de l'arbre, selon leurs degrés d'interaction - Principe de localité -.

Nous avons défini le réseau-ARBRE dans le cas d'une configuration a-aire équilibrée. Cette configuration peut être modifiée sans que cela ne remette en cause, ni la fonction de numérotation, ni le protocole d'interconnexion. En particulier, il est possible d'envisager d'attacher des unités à des noeuds du réseau appartenant à des niveaux autres que le niveau feuille. Cela privilégie les accès entre une unité attachée à un niveau non feuille et les autres unités, puisque la longueur des connexions les reliant se trouve réduite - figure ci-après -.



Suppression



Suppression du sous-arbre de racine  $N_i$  et son remplacement par une unité.

## 2. Le réseau CELLULAIRE

### 2.a. La topologie d'interconnexion

Une seconde extension du modèle de base, est le réseau CELLULAIRE. Il se présente sous la forme d'une matrice  $n \times m$  d'ANNEAUX - figure 1 -.

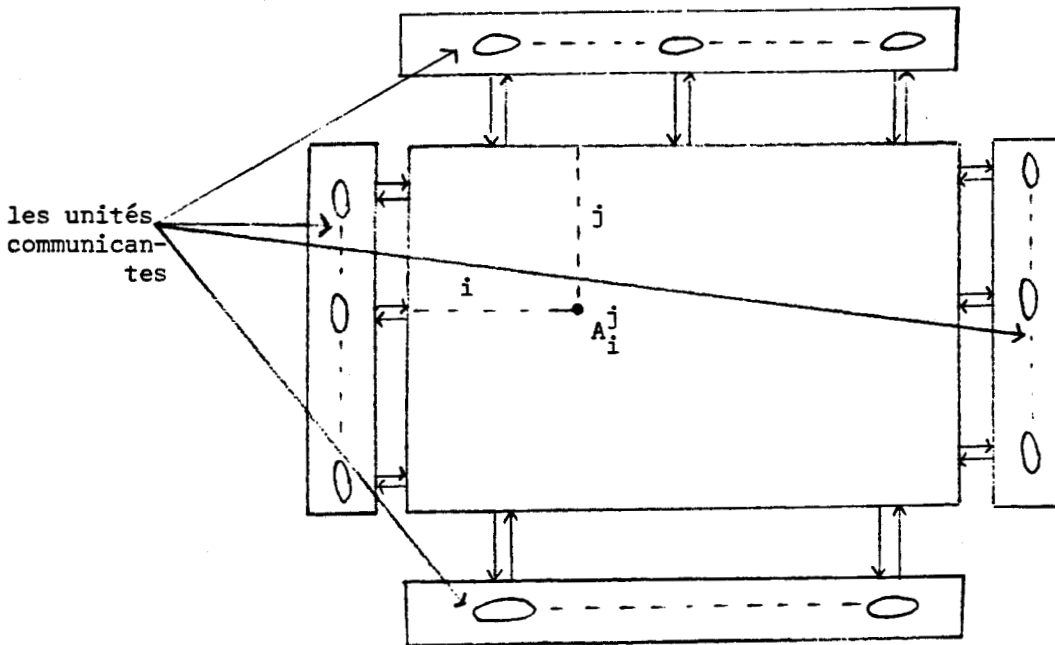


Figure 1 : Structure de base du réseau CELLULAIRE.



Les unités communicantes sont attachées aux ANNEAUX du réseau qui figurent sur chacun de ses quatre bords. Ainsi  $2(n+m)$  unités peuvent être interconnectées par le réseau CELLULAIRE.

Un ANNEAU  $A_i^j$   $1 \leq i \leq n-1$ ,  $1 \leq j \leq m-1$  dispose de quatre cellules de commutation, qu'il partage avec ses quatre ANNEAUX voisins :  $\{A_i^{j-1}, A_i^{j+1}, A_{i-1}^j, A_{i+1}^j\}$  - figure 2 -.

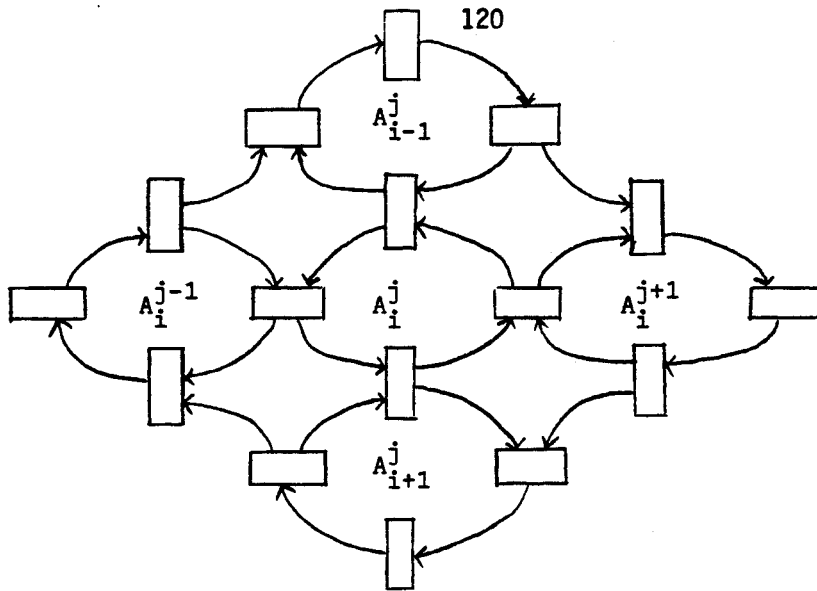


Figure 2 : Voisinage d'un ANNEAU  $A_i^j$  ( $1 \leq i \leq n-1$ ,  $1 \leq j \leq m-1$ ).

## 2.b. La commutation de ligne dans le réseau CELLULAIRE

Soient  $A_i^j$ ,  $A_i^{j'}$ , deux ANNEAUX du réseau CELLULAIRE :

$A_i^j \leftrightarrow A_i^{j'} \leftrightarrow \dots \leftrightarrow A_{i_\ell}^{j_\ell}$  une suite d'ANNEAUX  $(A_{i_\ell}^{j_\ell})_{\ell=1}^k$  distincts deux à deux

et tels que :

$$- A_i^j = A_{i_1}^{j_1}, A_i^{j'} = A_{i_k}^{j_k}$$

- si  $C_\ell$  est la cellule commune aux deux ANNEAUX

$$A_{i_\ell}^{j_\ell} \text{ et } A_{i_{\ell+1}}^{j_{\ell+1}} \quad 1 \leq \ell \leq k-1 \text{ alors } \eta(C_\ell) = E$$

- toute cellule  $C$  de  $A_{i_\ell}^{j_\ell}$  telle que  $C \neq C_\ell \quad 1 \leq \ell \leq k$   
 et  $C \neq C_{\ell-1} \quad 1 < \ell \leq k \quad \eta(C) = T$ .

Les deux figures 2-a, 2-b, présentent deux connexions différentes entre deux ANNEAUX  $A_0^0$  et  $A_1^3$  d'un réseau CELLULAIRE  $(4 \times 4)$ .



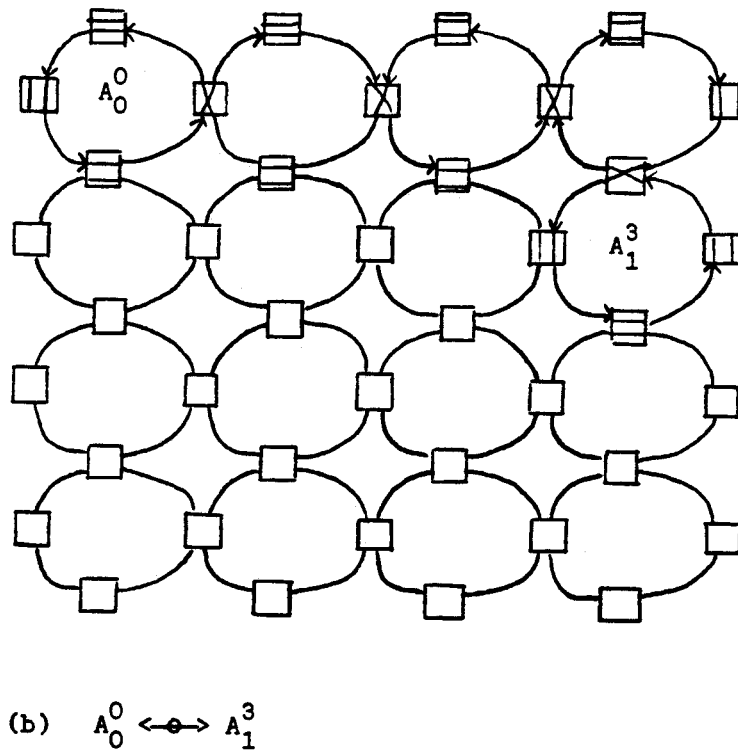
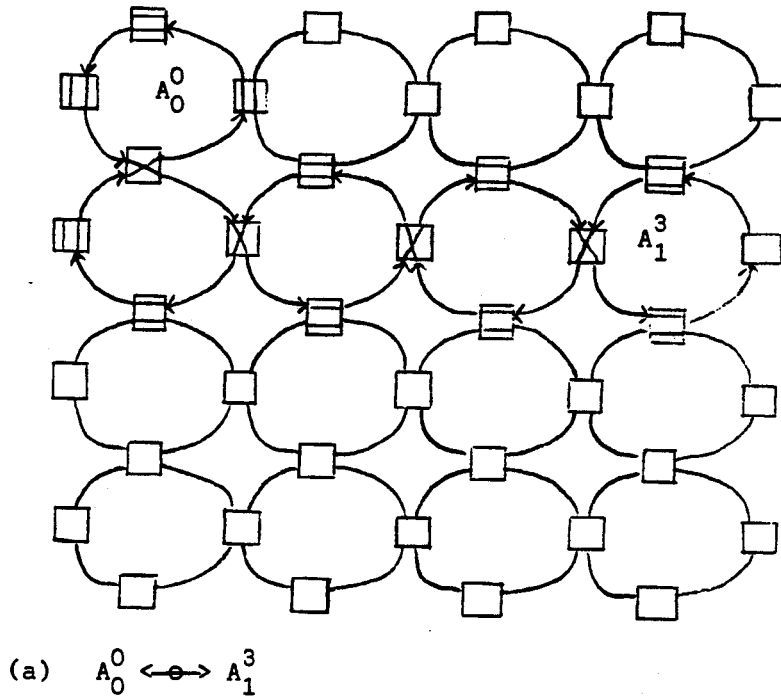
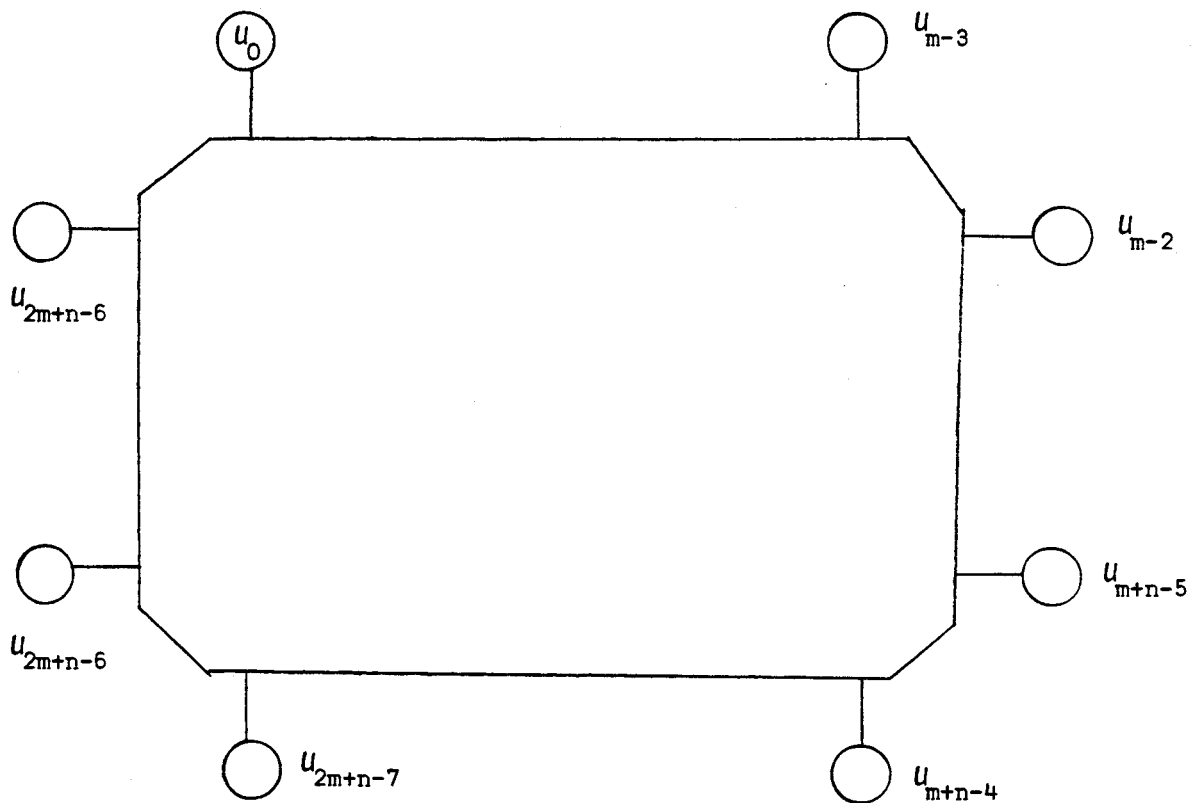


Figure 2 : Deux connexions différentes entre les ANNEAUX  $A_0^0$  et  $A_1^3$ .

### 2.c. Puissance fonctionnelle

A chaque ANNEAU de bord, sauf les quatre ANNEAUX  $\{A_0^0, A_0^{m-1}, A_{n-1}^0, A_{n-1}^{m-1}\}$ , une unité et une seule est supposée attachée.

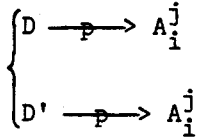
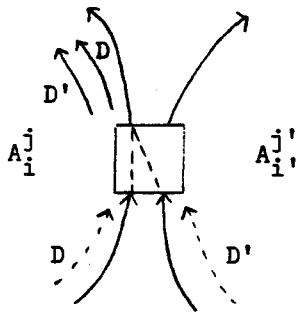
Le réseau CELLULAIRE interconnecte donc un ensemble de  $2(n+m) - 8$  unités  $(u_i)_{i=0}^{i=2(n+m)-9}$  - figure ci-dessous -.



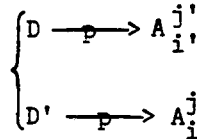
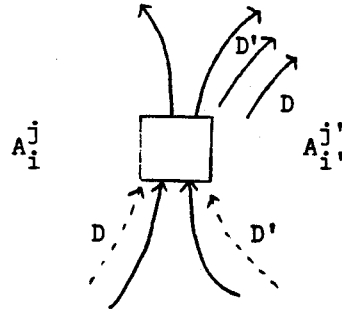
#### 2.c.1. Conflits

Soient deux ANNEAUX voisins  $A_i^j, A_i^{j'}$ , deux demandes  $D$  et  $D'$  telles que  $D \xrightarrow{p} A_i^j$  et  $D' \xrightarrow{p} A_i^{j'}$ , sont conflictuelles, lorsque le routage de l'une au moins, d'entre-elles nécessite la mise à l'état Echange de la cellule de commutation commune à  $A_i^j$  et  $A_i^{j'}$ .

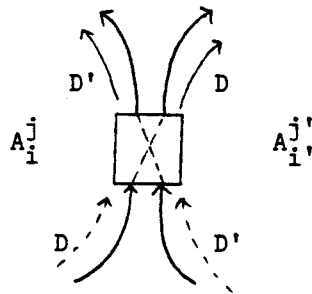
Par conséquent, parmi les quatre situations possibles, trois conduisent à une situation de conflit : - figure 3 -.



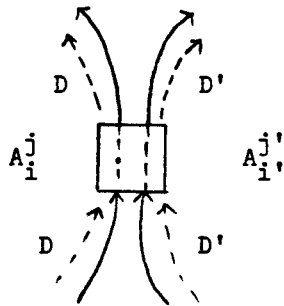
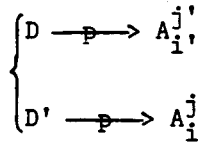
-a-



-b-



-c-



-d-

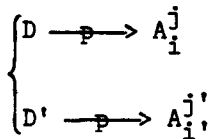


Figure 3 : Seule la situation (d) ne conduit pas à un conflit entre D et D'.

### 2.c.2. Blocage

L'établissement d'une liaison  $u_i \leftrightarrow u_j$  entraîne le partitionnement de l'ensemble des unités en deux sous-ensembles  $E_1$  et  $E_2$  - figure 4 -. Toute demande issue d'une unité de  $E_1$  vers une unité de  $E_2$ , et réciproquement, est bloquée par la liaison  $u_i \leftrightarrow u_j$ .

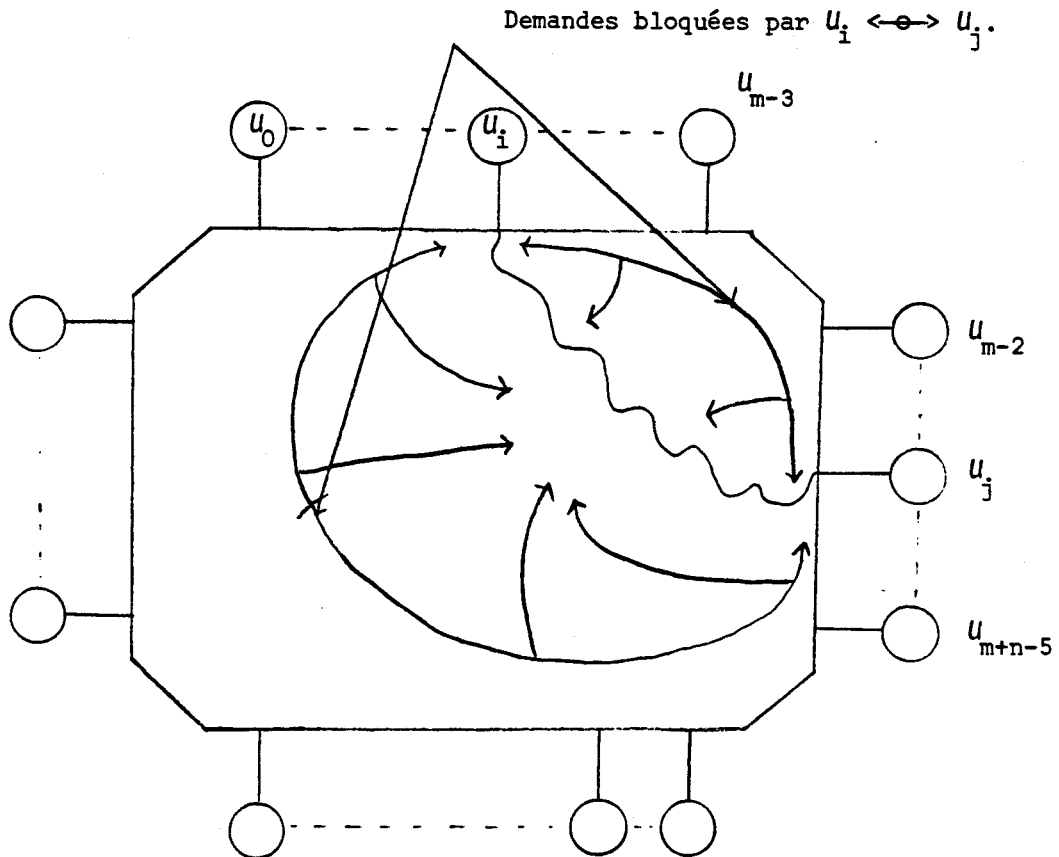


Figure 4 : Partitionnement des unités en deux sous-ensembles d'unités  $E_1$  et  $E_2$  par une seule connexion  $u_i \leftrightarrow u_j$ .

Le nombre de liaisons se trouvant ainsi interdites est égal au cardinal du produit cartésien de  $E_1$  et  $E_2$ .

$$C_B(i,j) = \text{nombre de connexions bloquées par une liaison } u_i \leftrightarrow u_j. \\ = \text{Card}(E_1 \times E_2).$$

Supposons  $i < j$

$$E_1 = \{u_\ell \mid i < \ell < j\}$$

$$E_2 = E \setminus E_1 \setminus \{u_i, u_j\}$$

$$E = \{(u_\ell)_{\ell=0}^{\ell=2(n+m)-9}\}$$

$$\text{Card}(E_1) = j-i+1$$

$$\text{Card}(E_2) = 2(n+m) - 8 - (j-i-1) - 2$$

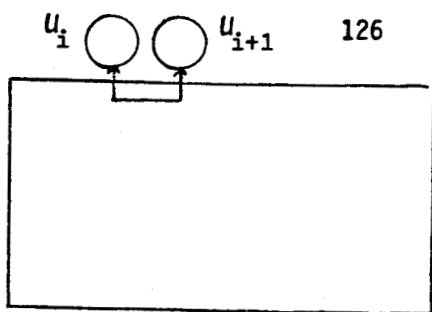
Posons  $\delta = j-i$  et  $\text{Card}(E) = 2(n+m) - 8 = N$

$$\text{Card}(E_1) = \delta - 1$$

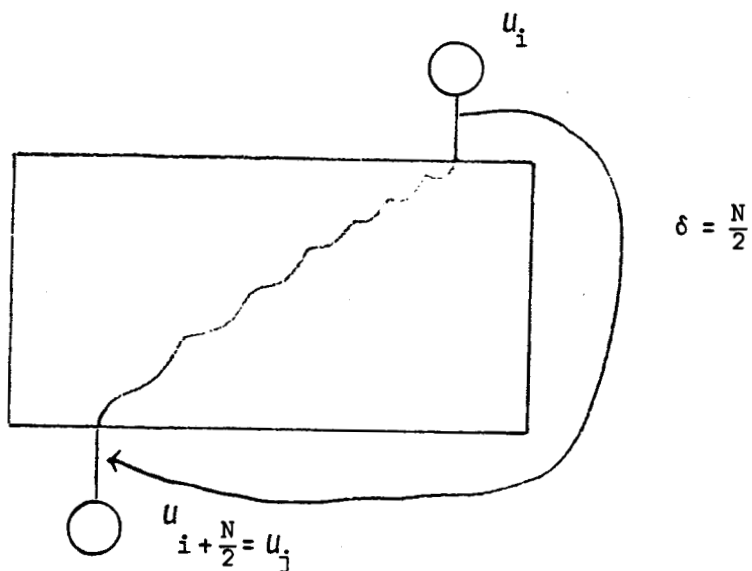
$$\text{Card}(E_2) = N - \delta - 1$$

$$\begin{aligned} C_B(i, j) &= \text{Card}(E_1) \times \text{Card}(E_2) \\ &= -\delta^2 + N \delta - N + 1. \end{aligned}$$

Le nombre de connexions bloquées par  $u_i \leftrightarrow u_j$ , dépend des positions relatives des unités  $u_i$  et  $u_j$ , l'une par rapport à l'autre. Ce nombre est minimal, il est réduit à zéro, lorsque la liaison est établie entre deux unités voisines,  $j = i+1$  ou  $j = i-1$  - figure 5 -. Le nombre de liaisons bloquées est maximal, et vaut  $(\frac{N}{2} - 1)^2$ , lorsque la liaison  $u_i \leftrightarrow u_j$  partage le réseau en deux parties symétriques,  $j = i + \frac{N}{2}$  - figure 5-b -.



(a) blocage minimal :  $C_B(i, i+1) = 0$ ,  $E_1 = \emptyset$ ,  $E_2 = E - \{u_i, u_{i+1}\}$



(b) Blocage maximal :  $C_B(i, i + \frac{N}{2}) = (\frac{N}{2} - 1)^2$

Figure 5

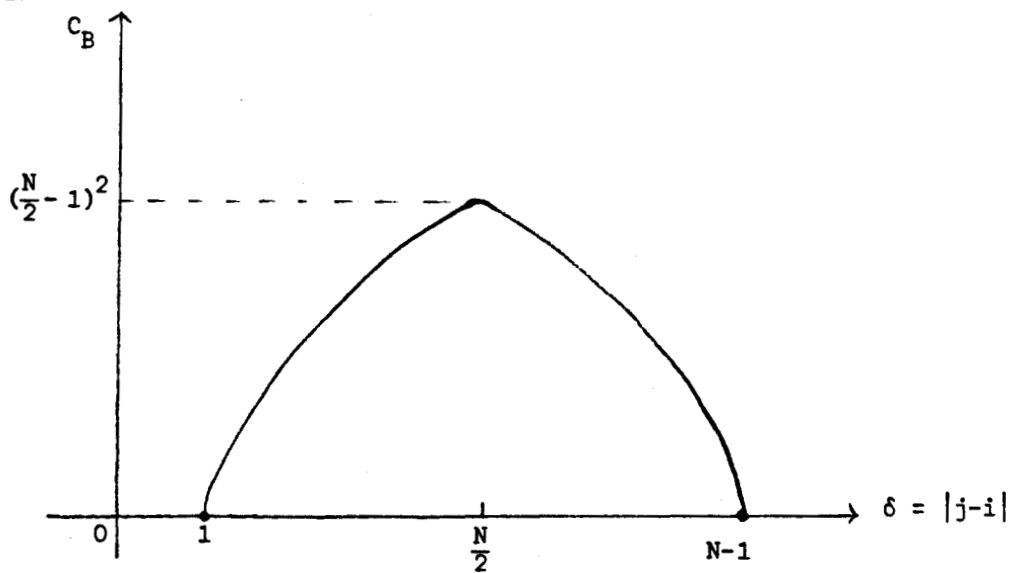


Figure 6 : Nombre de connexions bloquées par une connexion  $u_i \leftrightarrow u_j$ .

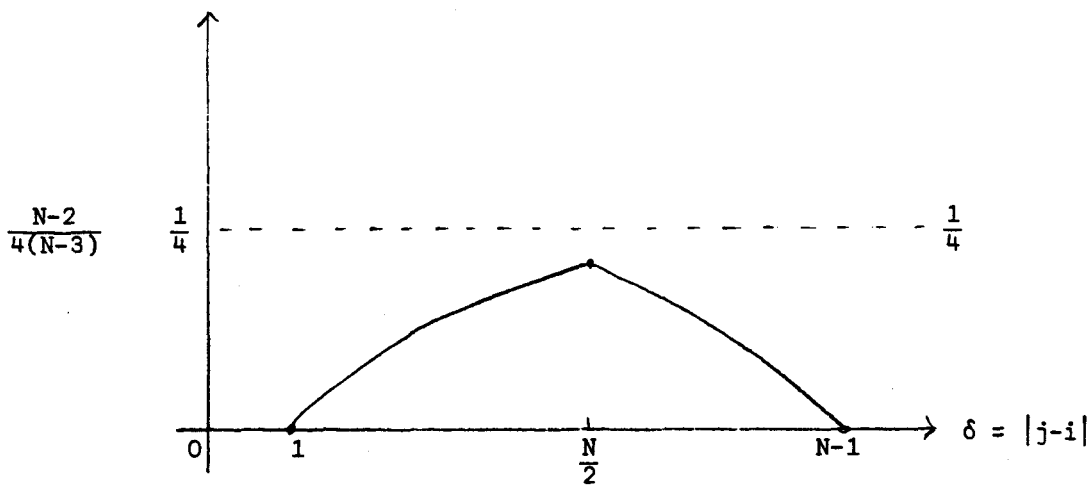
Le rapport entre le nombre des demandes bloquées par une connexion  $u_i \leftrightarrow u_j$ , et le nombre total des demandes de connexions pouvant être soumises au réseau exprime le taux de blocage dans ce réseau.

$$\begin{aligned} \tau &= \frac{\text{Nombre de connexions bloquées}}{\text{Nombre total de connexions}} = \\ &= \frac{C_B}{(\text{Card}(E) - 2)(\text{Card}(E) - 2 - 1)} \\ &= \frac{-\delta^2 + N\delta + -N + 1}{N^2 - 5N + 6} \end{aligned}$$

La valeur maximale de  $\tau$  est obtenue avec  $\delta = \frac{N}{2}$

$$\tau_{\max} = \frac{(N-2)}{4(N-3)} \sim \frac{1}{4}$$

Autrement dit, jusque 25 % des demandes sont bloquées du fait de l'établissement d'une seule connexion  $u_i \leftrightarrow u_j$ . ( $|j-i| = \frac{N}{2}$ ).



### 2.c.3. Simultanéité

On se propose ici d'évaluer le nombre moyen de liaisons pouvant être établies simultanément dans le réseau. Soit  $L_s$  le nombre de ces liaisons :

$$\bar{L}_s = \frac{n \cdot m}{\bar{l}}$$

$n \cdot m$  : est le nombre d'ANNEAUX du réseau CELLULAIRE.

$\bar{l}$  : la longueur moyenne, évaluée en nombre d'ANNEAUX, d'une liaison.

Nous supposons le réseau carré ( $n = m$ ) et désignons par  $N = 4n - 8$  le nombre des unités interconnectées

$$\bar{l} = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} l_{ij}}{N^2}$$

où  $l_{ij}$  est la longueur de la connexion  $u_i \leftrightarrow u_j$

$$\bar{l} \approx \frac{N}{4} + 2$$

$$\bar{L}_s \approx \frac{\left(\frac{N}{4} + 2\right)^2}{\left(\frac{N}{4} + 2\right)} \approx \frac{N}{4} + 2.$$

Le nombre maximal de liaisons simultanément établies dans un réseau à  $N$  unités étant  $L_s = \frac{N}{2}$

$$\bar{L}_s \approx \frac{L_s}{2}$$

seules 50 % des demandes soumises au réseau peuvent donc être satisfaites.

Notons que l'évaluation de  $\bar{L}_s$  est faite en considérant des hypothèses très optimistes :

- les liaisons sont établies à travers des chemins optimaux.
- Les demandes soumises ne sont pas conflictuelles et ne conduisent pas à des situations de blocage.

Il nous paraît donc évident qu'un taux des demandes soumises au réseau, nettement plus faible à 50 %, pourrait seulement être satisfait par le réseau CELLULAIRE.

#### 2.c.4. Critique de la cellule de base

Dans le réseau CELLULAIRE, la cellule de base est l'ANNEAU. Cette cellule présente l'inconvénient majeur d'avoir une capacité de commutation très réduite. En effet, elle ne peut être traversée, à un instant donné, que par une seule liaison au plus. La figure 6 décrit les états que peut prendre l'ANNEAU.



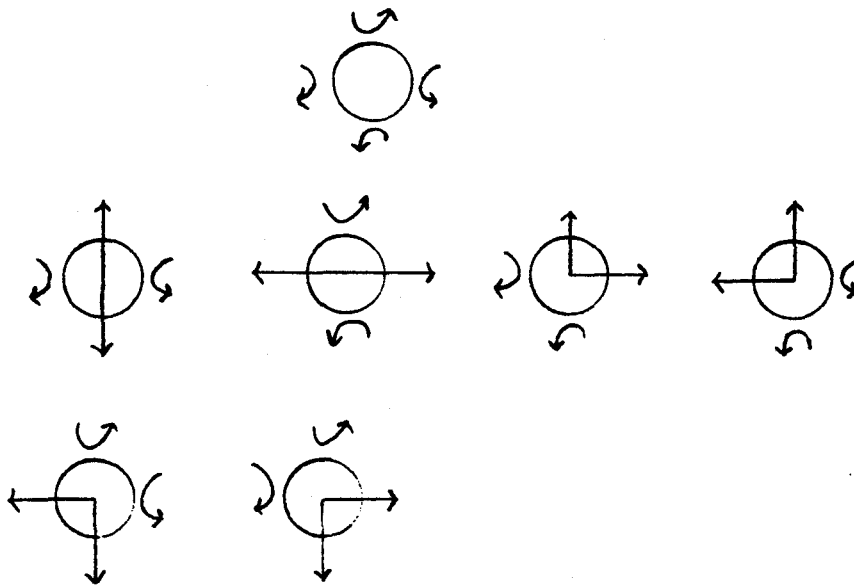


Figure 6 : L'ANNEAU ne peut être traversé que par une seule liaison au plus.

## 2.d. Conclusion

Le réseau CELLULAIRE présente les caractéristiques requises pour son intégration VLSI, grâce à la régularité de sa topologie, et laisse penser qu'une mise en oeuvre simple des procédures de routage et d'accès peut être aisément obtenue. Malheureusement, l'étude de la capacité fonctionnelle de ce réseau a montré que cette capacité est très réduite pour que son utilisation dans un environnement à fort degré de couplage puisse être envisagée. Nous n'avons donc pas jugé utile de poursuivre par l'étude du protocole d'interconnexion.

Notons par ailleurs, qu'une utilisation du réseau CELLULAIRE en commutation de message ou en commutation mixte, message et ligne, peut être envisagée. Nous pensons que cela permettra de réduire les conflits et d'éliminer les situations de blocage. Nous n'avons pas exploré entièrement cette voie, car cela déborderait du cadre de cette étude.

### 3. Conclusion

L'approche extension par graphe, permet d'engendrer de nombreuses topologies de réseau. Nous nous sommes limités à l'examen de deux d'entre-elles : ARBRE et CELLULAIRE. Les conclusions qui ont été tirées de l'étude de ce dernier, nous laissent penser qu'une redéfinition de la cellule de base, est nécessaire. En effet, l'accroissement de la simultanéité dans un réseau passe nécessairement par l'accroissement de la capacité de commutation de chacune des cellules qui le composent.

Une cellule idéale devra, si elle interconnecte  $p$  autres cellules voisines, prendre  $(p-1)(p-3)\dots \cdot 1 = \frac{(p-1)!}{\binom{p-2}{2} 2^2}$  états différents. Chacun de ces

états établissant  $\frac{p}{2}$  liaisons entre ses  $p$  cellules voisines.

\*\*\*\*\*  
\*  
\* DEUXIEME PARTIE \*  
\*  
\*\*\*\*\*

EXTENSION PAR ANALOGIE ENTRE L'ANNEAU ET LE BUS

## 0. Introduction

L'ANNEAU est un outil de communication avec blocage. Il ne permet d'établir qu'une seule liaison à la fois. On peut donc considérer qu'il possède des propriétés fonctionnelles analogues à celles d'un bus bidirectionnel.

La technique des bus multiples - cf. §I.2.c.1. - permet d'accroître la simultanéité, en multipliant le nombre de bus disponibles dans un système. Elle améliore en conséquence la capacité, très limitée, d'un bus unique. Par analogie, il est possible, moyennant une configuration de  $\frac{n}{2}$  ANNEAUX, de définir une extension du modèle de base, qui soit sans blocage et qui relie entre-elles deux unités quelconques d'un ensemble de  $n$  unités. Nous nous proposons, dans ce qui suit, de définir sa topologie et son protocole d'interconnexion.

### 1. La topologie du MULTI-ANNEAU

$\frac{n}{2}$  ANNEAUX :  $(A_i)_{i=1}^{\frac{n}{2}}$ , sont disposés en parallèle - figure 1 -. Chaque ANNEAU comporte  $n$  cellules  $(C_i)_{i=1}^n$ . L'accès d'une unité  $U_i$  au réseau est assuré en reliant respectivement, la voie d'entrée  $E_i$  de cette unité à la sortie de la cellule  $C_i$  de l'ANNEAU  $A_{\frac{n}{2}}$ , et la voie de sortie  $S_i$  de cette même unité à l'entrée de la cellule  $C_i$  de l'ANNEAU  $A_1$ .

De la même manière, chacune des cellules d'un ANNEAU  $A_k$ ,  $1 < k < \frac{n}{2}$  est reliée aux cellules des ANNEAUX adjacents  $A_{k-1}$  et  $A_{k+1}$ .

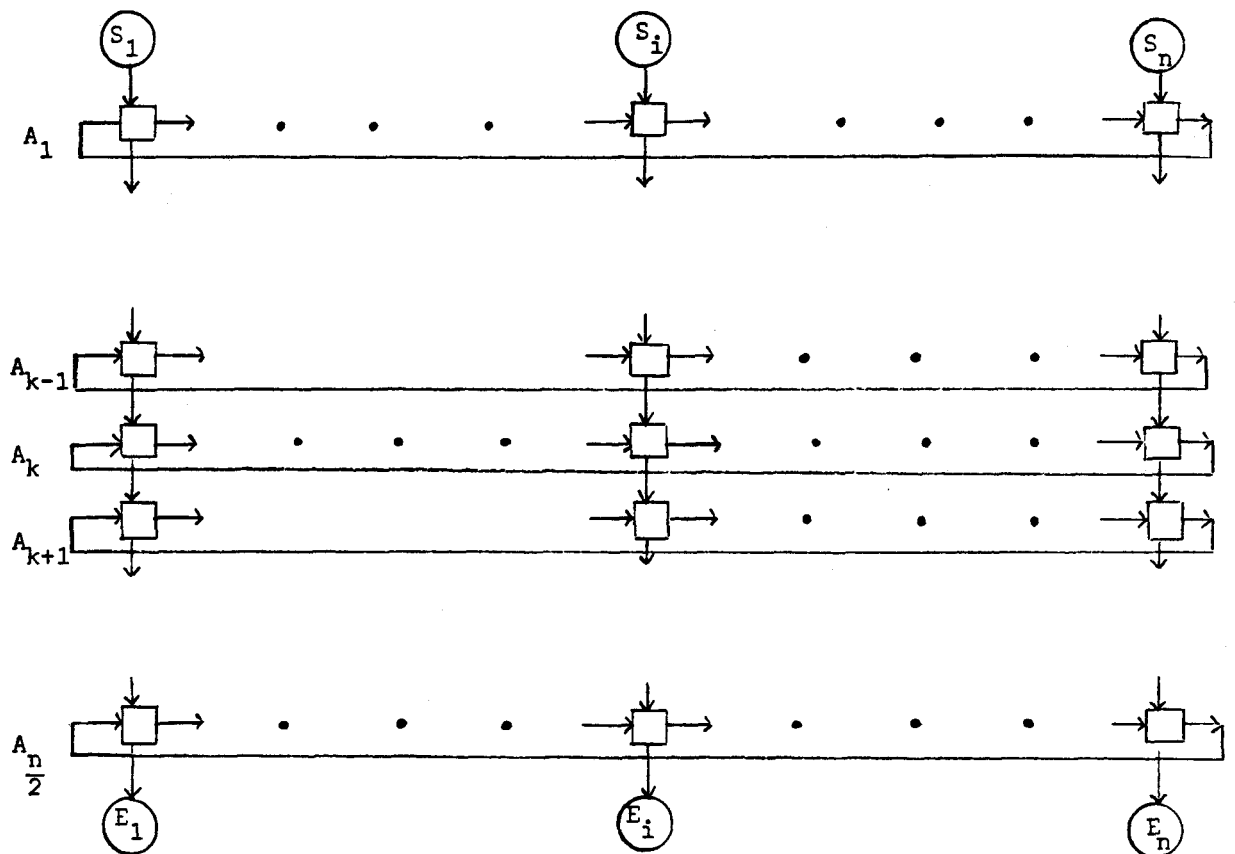


Figure 1 : La topologie du MULTI-ANNEAU.

## 2. La commutation de ligne dans le MULTI-ANNEAU

Initialement, la voie de sortie de chaque unité est reliée à sa voie d'entrée :  $u_i \leftrightarrow u_i$ . Cela suppose que les  $\frac{n}{2}$  cellules  $(c_k^i)_{k=1}^{n/2}$ , que nous appelons "cellule d'accès de l'unité  $u_i$ ", sont toutes à l'état "Transparent" - figure 2.

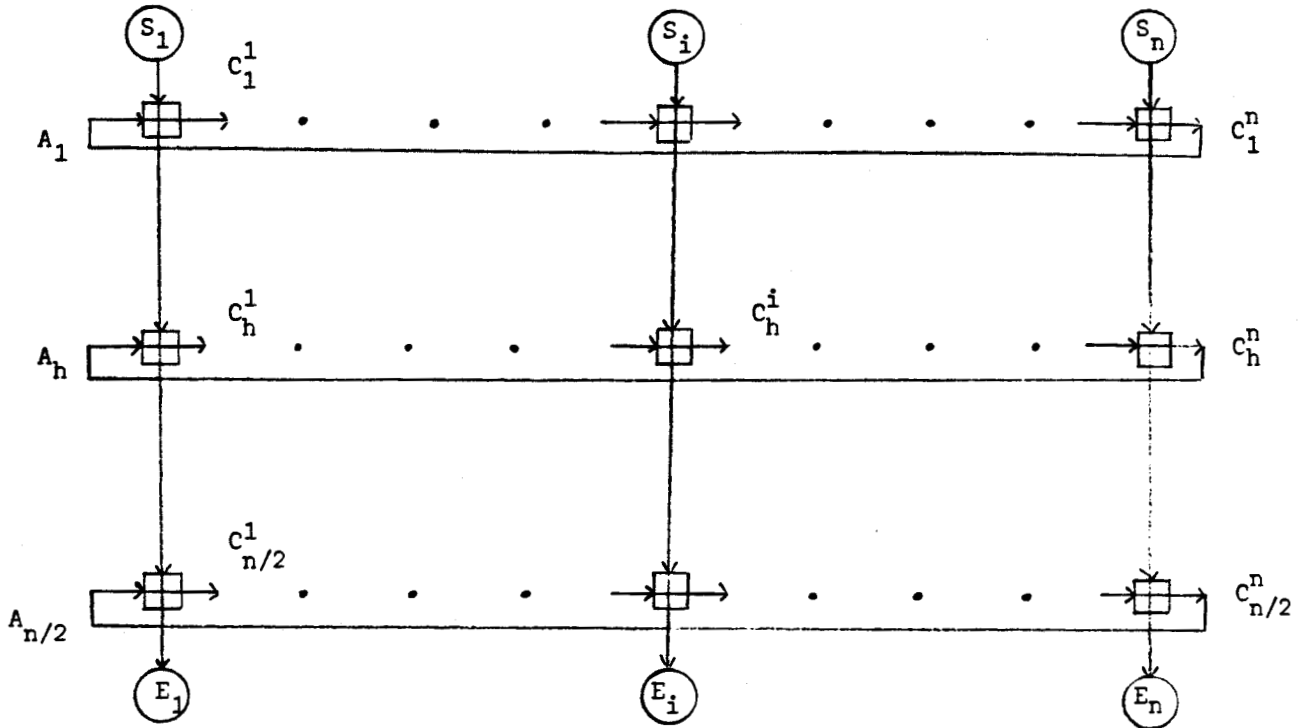


Figure 2 : Etat initial  $U_i \leftrightarrow U_j \quad i = 1, \dots, n.$

L'établissement d'une liaison bidirectionnelle et simultanée, entre deux unités  $U_i$  et  $U_j$  est alors obtenu en imposant l'état "Echange" aux deux cellules  $C_h^i$  et  $C_h^j$  d'un même ANNEAU  $A_h$ . Les cellules  $C_h^i$  et  $C_h^j$  font partie des cellules d'accès de chacune des unités à interconnecter.

Autrement dit,

$$U_i \leftrightarrow U_j \iff \exists h \quad 1 \leq h \leq \frac{n}{2} : \begin{cases} \cdot \eta(C_h^i) = \eta(C_h^j) = E \\ \cdot \forall k \neq i, j \quad \eta(C_h^k) = T \\ \cdot \forall h' \neq h \quad \eta(C_{h'}^i) = \eta(C_{h'}^j) = T \end{cases}$$



La figure 3 illustre une liaison bidirectionnelle, à travers le MULTI-ANNEAU, entre les unités  $U_i$  et  $U_j$ .

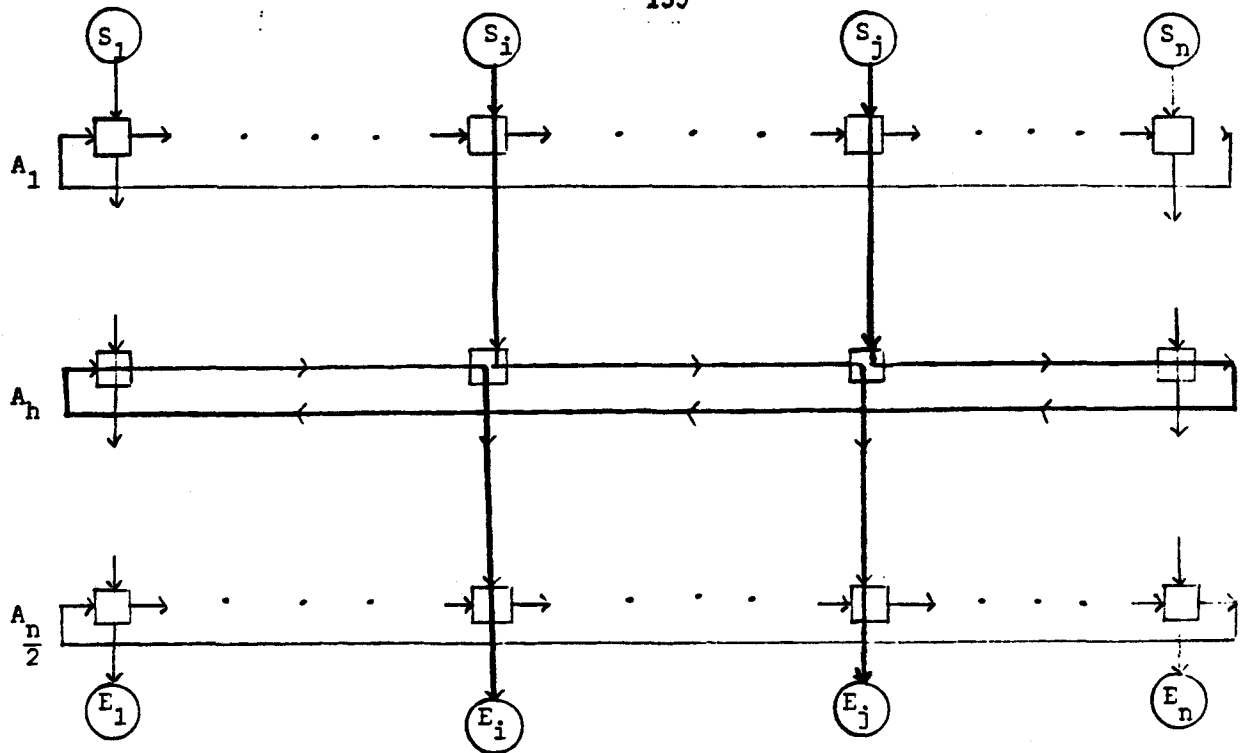


Figure 3 : En trait gras la liaison bidirectionnelle entre les unités  $U_i$  et  $U_j$  à travers l'ANNEAU  $A_h$ .

### 3. Le protocole d'interconnexion

#### 3.a. La fonction de numérotation

Trois types d'entités sont impliquées par la mise en place d'une communication à travers le MULTI-ANNEAU :

- i - les unités communicantes,
- ii - les ANNEAUX qui composent l'outil de communication, et
- iii - les cellules de commutation.

La mise en oeuvre des deux procédures de routage et d'accès, ultérieurement développées, exige l'identification des deux premières entités uniquement.

La fonction de numérotation est définie comme suit

$$v : \{U_i \mid 1 \leq i \leq n\} \cup \{A_h \mid 1 \leq h \leq \frac{n}{2}\} \rightarrow \{1, \dots, n\}$$

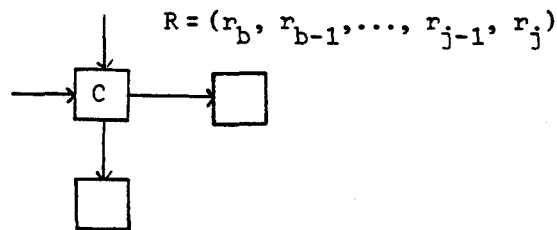
$$v(U_i) = U_i \quad 1 \leq i \leq n$$

$$v(A_h) = h \quad 1 \leq h \leq \frac{n}{2}.$$

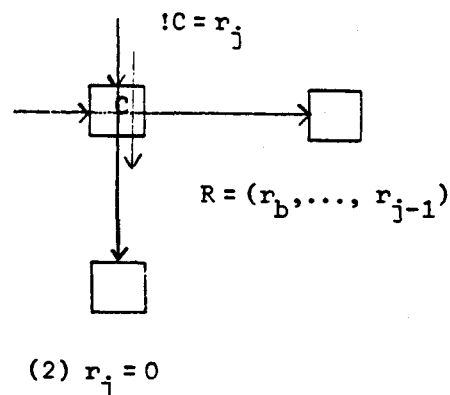
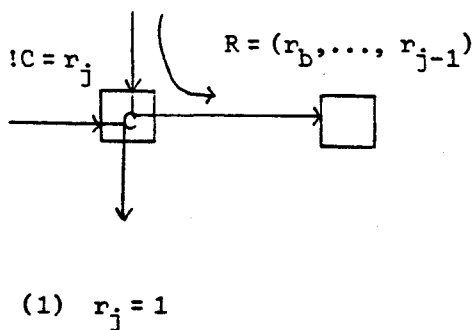
Une cellule de commutation sera désignée, sans qu'une numérotation ne lui soit attribuée, par  $C_{v(A)}^{v(U)}$  :  $C_{v(A)}^{v(U)}$  est la cellule intersection de l'ANNEAU  $v(A)$  et de l'axe vertical issu de l'unité  $v(U)$ .

### 3.b. La procédure de routage

Cette procédure repose sur la notion de "vecteur de routage". Ce vecteur se compose d'une suite de bits, dont la longueur est égale au nombre de cellules de la connexion à établir. Le bit en-tête de ce vecteur, définit, sous certaines conditions d'accès abordées dans le paragraphe suivant, la commande de la cellule de commutation qu'il traverse, après quoi, ce bit en-tête est supprimé, et le vecteur de routage progresse dans le réseau - figure 4 -.



-a- Avant traversée de la cellule C par R.



-b- Après traversée de la cellule C par R, le bit en-tête, ici  $r_j$ , définit la commande de la cellule C et est ensuite supprimé.

Figure 4 : Routage dans le MULTI-ANNEAU.



Remarque : L'établissement d'une connexion unidirectionnelle  $U_s \rightarrow U_d$ , à travers le MULTI-ANNEAU provoque, de la même manière que dans l'ANNEAU, l'établissement de la connexion réciproque  $U_d \rightarrow U_s$ . Il suffit, donc, que le vecteur de routage définisse la commande des cellules composants la seule connexion  $U_s \rightarrow U_d$ . La longueur  $||U_s \rightarrow U_d||$  de cette connexion, détermine donc le nombre de composantes du vecteur de routage, avant son envoi dans le réseau.

$$||U_s \rightarrow U_d|| = \frac{n}{2} + (n + (d - s)) \bmod n$$

### Calcul du vecteur de routage

Le vecteur de routage est déterminé à partir des données suivantes :

- Les numéros des unités source et destinataire :

$$v(U_s) = s, v(U_d) = d.$$

- Le numéro h de l'ANNEAU alloué

posons  $v = n + (d - s) \bmod n$

$$b = ||U_s \rightarrow U_d||$$

$$R = (r_j)_{j=1}^{j=b} : \begin{cases} r_j = 0 \quad \forall j \neq h, j \neq h+v \\ r_h = r_{h+v} = 1 \end{cases}$$

Le vecteur de routage ne peut donc être calculé qu'après allocation d'un ANNEAU - Aspect réparti du calcul du vecteur de routage -.

### 2.d.3. La procédure d'accès

L'accès au MULTI-ANNEAU est basé sur une simple extension de la technique du jeton circulant, et adoptée pour le modèle de base. A chaque ANNEAU, est associé un jeton. A un instant donné jusque  $\frac{n}{2}$  jetons se propagent dans le MULTI-ANNEAU. L'accès au réseau est conditionné par la capture de l'un quelconque de ces  $\frac{n}{2}$  jetons. Il doit y avoir, à un instant donné, au moins un d'entre-eux qui est libre. Sinon, cela signifierait que chacune des autres unités, et en particulier celle avec laquelle une communication est à établir, est occupée.

Remarque : Le calcul du vecteur de routage fait intervenir le numéro de l'ANNEAU alloué. L'opération de capture du jeton doit permettre l'identification de l'ANNEAU correspondant.

Etude des problèmes relatifs à l'accès au MULTI-ANNEAU

Dans sa configuration optimale, comportant  $\frac{n}{2}$  ANNEAUX, le MULTI-ANNEAU est un outil de communication sans blocage. En effet, chaque ANNEAU assure une liaison entre deux unités, dans un système à  $n$  unités,  $n$  supposé pair,  $\frac{n}{2}$  ANNEAUX sont suffisants pour satisfaire toute demande de connexion. Toutefois, le caractère asynchrone et indépendant des demandes d'accès au réseau, conduit, dans certains cas que nous explicitons par la suite, à des situations de conflit, de blocage, voire d'interblocage.

Une demande de connexion, émise dans le réseau, sera définie par le triplet  $(s, h, d)$ ; où  $s$  et  $d$  sont les numéros respectifs des unités source et destinataire, et où  $h$  est le numéro de l'ANNEAU alloué pour satisfaire cette demande.

i) Situation de conflit :

Il y a conflit dans le réseau, lorsque deux unités sources  $U_{s_1}, U_{s_2}$  adressent simultanément la même unité destinataire  $U_{d_1} = U_{d_2}$ . Cela se traduit dans le réseau par la soumission des deux demandes  $D_1 = (s_1, h_1, d_1)$  et  $D_2 = (s_2, h_2, d_2)$  aux deux entrées d'une même cellule  $C_{\max(h_1, h_2)}^{d_1=d_2}$  - figure 5 -.

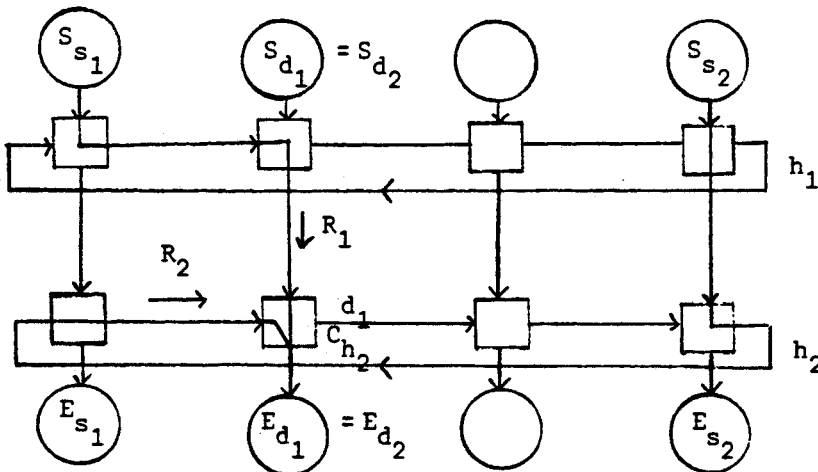
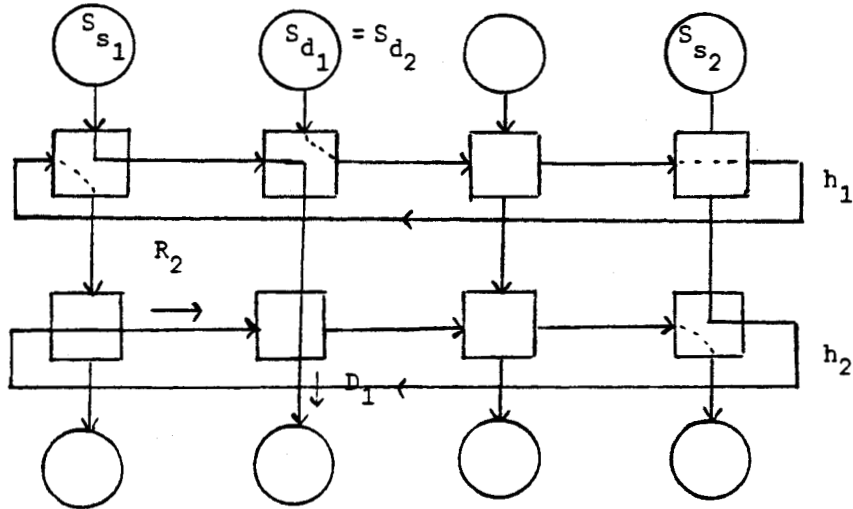


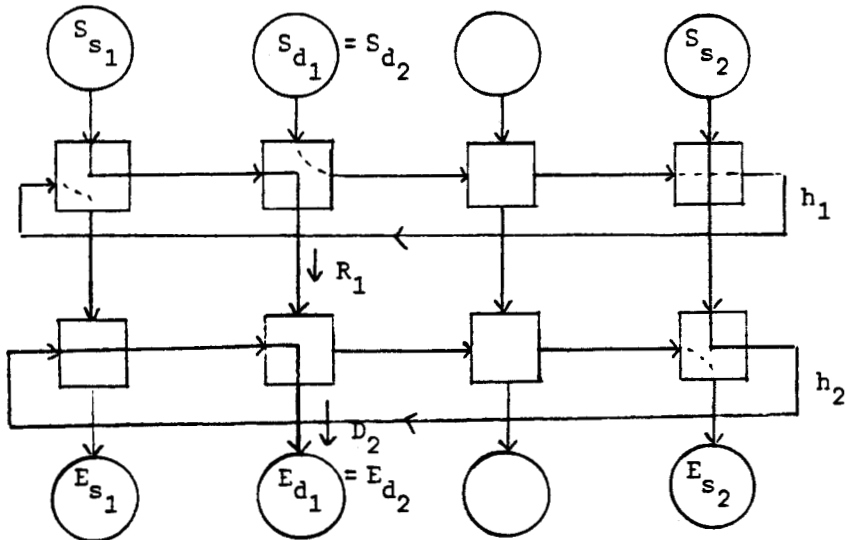
Figure 5 : Conflit : deux unités sources adressent la même unité destinataire.

ii) Situation de blocage :

A la différence de la situation de conflit, où les demandes de connexion se trouvent dans des états d'avancement identiques, la situation de blocage se produit lorsqu'une des demandes  $D_1$  ou  $D_2$  est en avance par rapport à l'autre. Les deux demandes adressant, bien entendu, la même unité - figure 6 -.



-a- La demande  $D_1$  est en avance par rapport à  $D_2$ .



-b-  $D_2$  est en avance par rapport à  $D_1$ .

Figure 6 : Blocage résultant de l'émission de deux demandes de connexion vers la même unité destinataire. L'une des deux demandes se trouvant en avance par rapport à l'autre.

## iii) Situation d'interblocage :

Une situation d'interblocage se produit entre deux unités dans ce cas il y a interblocage direct, ou entre plus de deux unités, l'interblocage est alors indirect.

- cas - 1 : Interblocage direct.

Il y a interblocage direct lorsque deux unités s'envoient mutuellement une demande de connexion :  $U_{s_1} = U_{d_2}$  et  $U_{s_2} = U_{d_1}$ .

Cela se traduit par l'existence dans le réseau de deux demandes  $D_1 = (s_1, h_1, d_1)$  et  $D_2 = (s_2, h_2, d_2)$  telle que  $s_1 = d_2$ ,  $s_2 = d_1$  et  $h_1 \neq h_2$  - figure 7-.

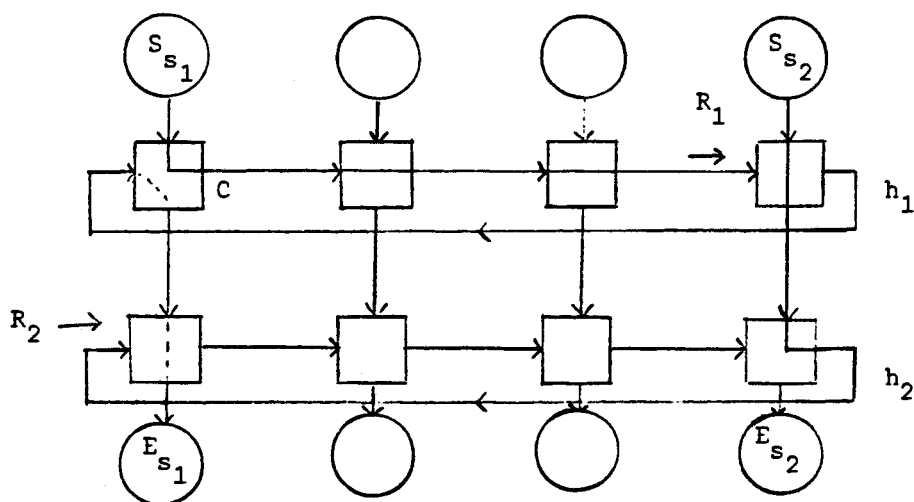


Figure 7 : Interblocage direct.

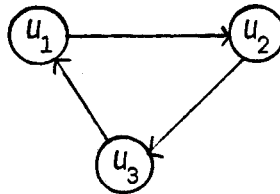
$R_1$  et  $R_2$  sont présentes sur les deux cellules d'accès

$C_{h_1}^{s_1}, C_{h_2}^{s_1}$  de l'unité  $U_{s_1}$ , ainsi que sur les cellules

$C_{h_1}^{s_2}, C_{h_2}^{s_2}$  de l'unité  $U_{s_2}$

- cas - 2 : Interblocage indirect.

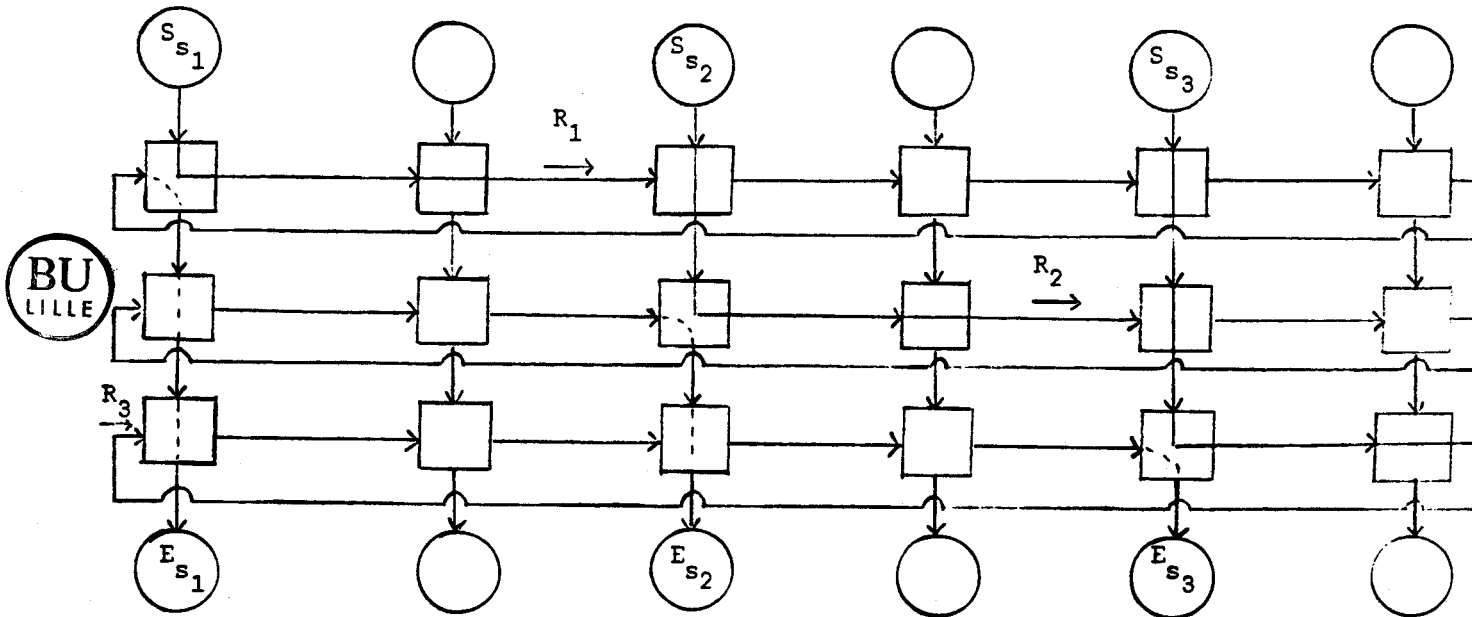
Dans ce cas au moins trois unités,  $u_1, u_2, u_3$  participent dans l'interblocage :



Trois requêtes  $R_1 = (s_1, h_1, d_1)$   $R_2 = (s_2, h_2, d_2)$   $R_3 = (s_3, h_3, d_3)$  sont alors soumises dans le réseau :

- $u_1 = u_{s_1} = u_{d_3}$
- $u_2 = u_{s_3} = u_{d_1}$
- $u_3 = u_{s_2} = u_{d_2}$

La figure ci-après illustre cette situation.



Interblocage indirect entre trois unités  $u_1, u_2$  et  $u_3$

Solutions aux problèmes d'accès au MULTI-ANNEAU

Les problèmes d'accès que nous venons de soulever résultent directement des demandes des unités et nullement de l'outil de communication qui, lui, est sans blocage. Les situations de conflit et de blocage se produisent, en effet, chaque fois que deux unités adressent "simultanément" une même unité destinataire, et il y a interblocage lorsque deux unités au moins adressent mutuellement l'une l'autre.

Une analyse centralisée des demandes, préalablement à leur envoi dans le réseau, permet de prévenir la production de ces situations dans le réseau. Un dispositif auquel sera assigné cette tâche est donc nécessaire, et des voies de communication ainsi qu'un protocole de communication, spécifiques, doivent être établis entre chaque unité et ce dispositif. Cette solution présente les mêmes avantages et inconvénients que les solutions du type arbitre, pour le contrôle d'accès à un bus partagé par exemple - cf. §I.2.a.

La solution que nous avons retenue pour remédier aux problèmes que soulève l'accès au réseau est entièrement répartie. Prévenir la production de chacune des situations de conflit, de blocage et d'interblocage est assuré, en partie, en garantissant la présence, à chaque instant, d'une seule demande au plus dans l'ensemble des cellules d'accès d'une même unité. Car, ces situations résultent toutes, de la présence simultanée d'au moins deux demandes distinctes dans les cellules d'accès d'une même unité.

Remarque : L'ensemble  $\{U_i, \text{cellules d'accès de } U_i\}$  forme un anneau - figure 8 -, nous notons  $A-U_i$  cet anneau.

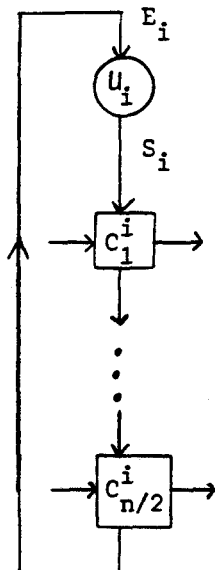


Figure 8 :  $A-U_i$ , l'anneau composé de l'unité  $U_i$  et de ses cellules d'accès  $(C_h^i)_{h=1}^{n/2}$ .

D'où l'idée d'assurer le contrôle d'accès aux cellules d'un anneau  $A-U_i$  au moyen d'un jeton circulant.

La commutation d'une cellule  $C_h^i$ , intersection des anneaux  $A_h$  et  $A-U_i$ , est précédée des opérations d'allocation de chacun de ces deux anneaux, et ce au profit d'une même demande, issue de, ou adressée à l'unité  $U_i$  - figure 9 -.

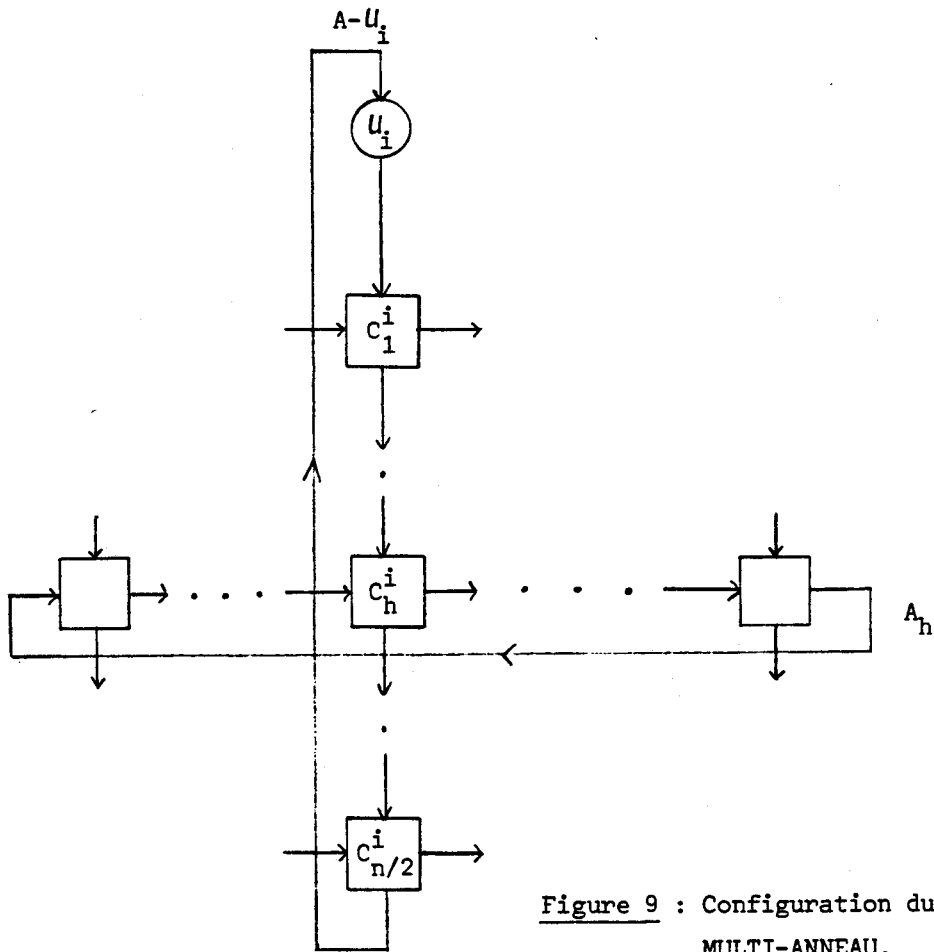


Figure 9 : Configuration du MULTI-ANNEAU.

Une requête peut se présenter à une cellule de commutation  $C_h^i$ , par l'une de ses deux entrées - figure 10 -, selon qu'elle se propage dans l'anneau  $A_h$  ou  $A-U_i$ .

- cas - 1 :  $R \xrightarrow{p} A-U_i$ , figure 10.a.

Dans ce cas, l'anneau  $A-U_i$  a été nécessairement alloué, soit par le module de contrôle d'une cellule  $C_h^i$ , si l'unité  $U_i$  est la destination de la requête, soit par  $U_i$ , si elle en est la source. Dans ce dernier cas, l'ANNEAU  $A_h$  est également alloué par  $U_i$ , et le bit en-tête du vecteur de routage, lorsqu'il a atteint  $C_h^i$  est à 1.

- cas - 2 :  $R \xrightarrow{p} A_h$ , figure 10.b

Dans ce cas,  $U_i$  ne peut pas être la source de la requête, elle peut par contre en être la destination. Le bit en-tête du vecteur de routage sera alors à 1 et l'anneau  $A-U_i$  doit par conséquent être alloué au profit de cette requête avant de commuter la cellule  $C_h^i$ ; l'ANNEAU  $A_h$  étant, lui, alloué auparavant par l'unité source de cette requête.

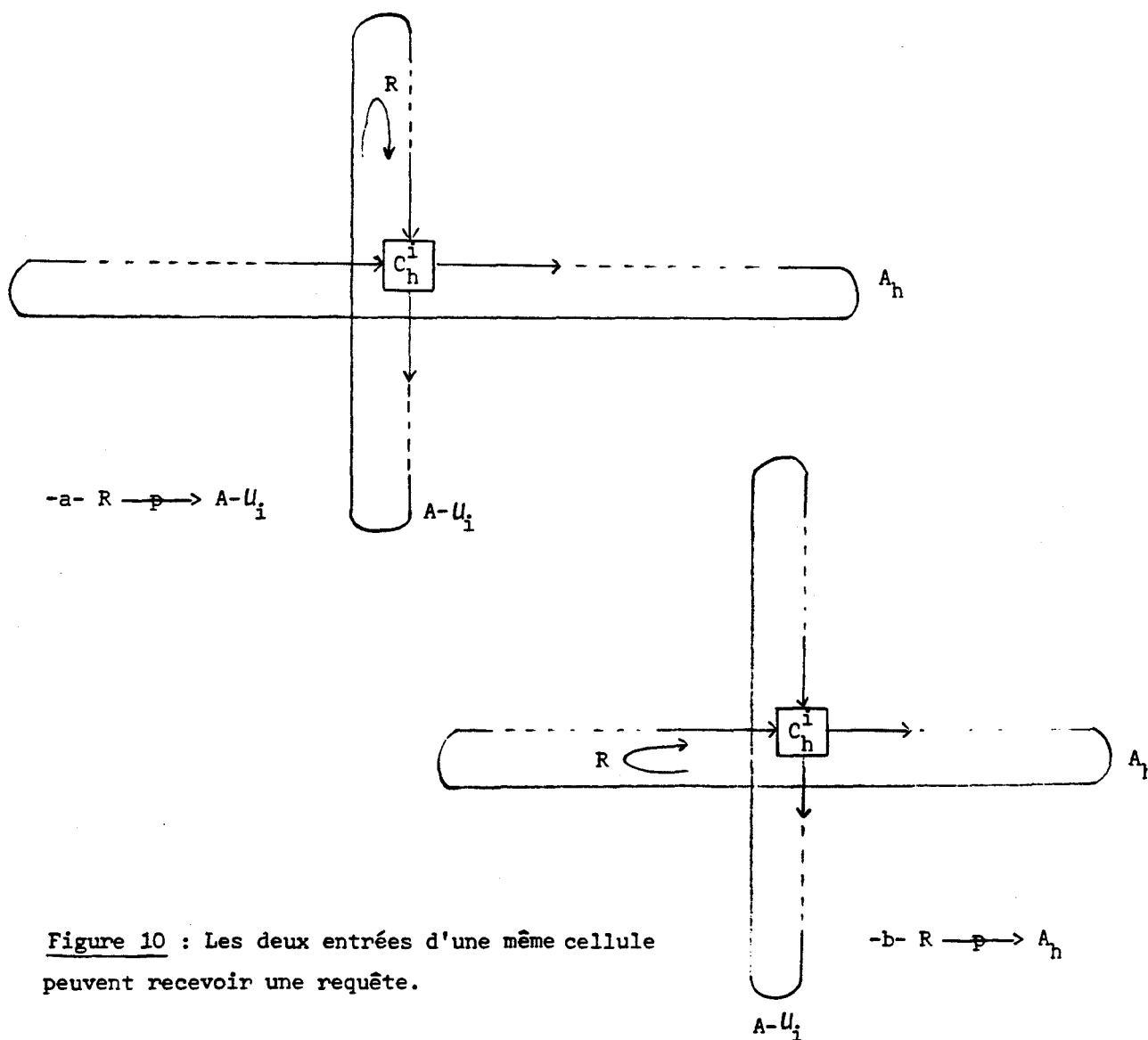


Figure 10 : Les deux entrées d'une même cellule peuvent recevoir une requête.



L'opération d'allocation d'un anneau  $A-U_i$  est assurée par le module de contrôle associé à chaque cellule  $C_h^i$  - figure 11 -. Dès que l'allocation est effective, c'est-à-dire que le jeton circulant dans  $A-U_i$  est capturé, la cellule  $C_h^i$  est commutée et la requête directement émise vers  $U_i$ , sa destination.

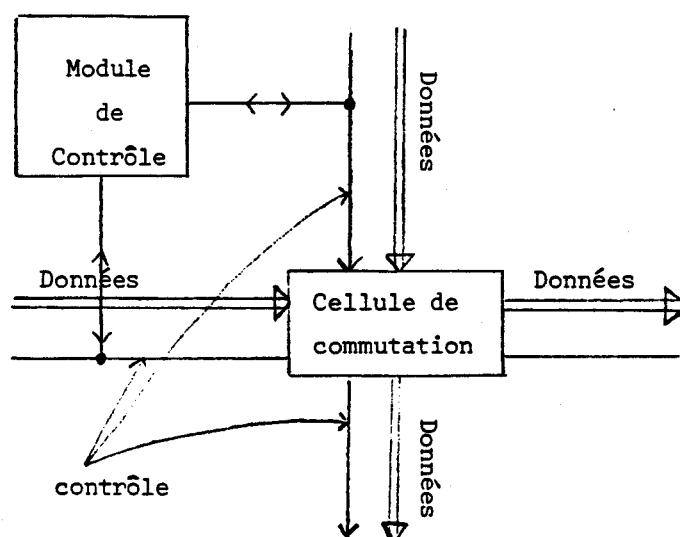


Figure 11 : Module de contrôle d'une cellule.

### Opération de verrouillage des cellules d'un anneau

L'opération de verrouillage peut être déclenchée par l'un quelconque des modules de contrôle. Elle entraîne alors le verrouillage de l'ensemble des cellules de l'anneau  $A-U_i$  auquel appartient le module de contrôle qui a déclenché l'opération. L'objectif de la mise en place de ce mécanisme est double :

- i) Permettre une meilleure utilisation des ressources de communication :

Lorsqu'une requête destinée à une unité  $U_i$ , se présente devant l'une quelconque de ses cellules d'accès et que cette cellule est verrouillée cela signifie que  $U_i$  est impliquée dans une opération de communication et que l'accès à cette unité ne peut être immédiatement obtenu. Un acquittement négatif est alors envoyé à l'unité source de la requête et les ressources acquises sont libérées - figure 12 -.

## ii) Prévenir les situations d'interblocage :

Nous avons vu précédemment qu'une telle situation se produit lorsqu'un couple d'unités  $(U_i, U_j)$ , s'adressent mutuellement. L'unité  $U_i$ , adressant  $U_j$ , procède à l'allocation d'un ANNEAU  $A_h$  et de l'anneau  $A-U_i$ , sa requête, envoyée à travers ces deux anneaux, se présentera devant la cellule  $C_h^j$  et se mettra en attente pour allouer  $A-U_j$ . De la même manière, l'unité  $U_j$  procède à l'allocation d'un ANNEAU  $A_h$ , ( $h' \neq h$ ), et de l'anneau  $A-U_j$  et sa requête, envoyée à travers ces deux anneaux se mettra en attente devant la cellule  $C_h^{i'}$ , pour allouer l'anneau  $-A-U_i$ , d'où interblocage. La figure 13 montre une situation d'interblocage entre les unités  $U_2$  et  $U_4$ . L'interblocage est impossible du fait que les cellules  $C_1^4$  et  $C_2^2$  sont verrouillées, ce qui entraîne l'envoi systématique d'un acquittement négatif vers les deux unités et par conséquent la libération des ressources occupées.

L'interblocage indirect, faisant intervenir plus de deux unités, est résolu de la même façon.

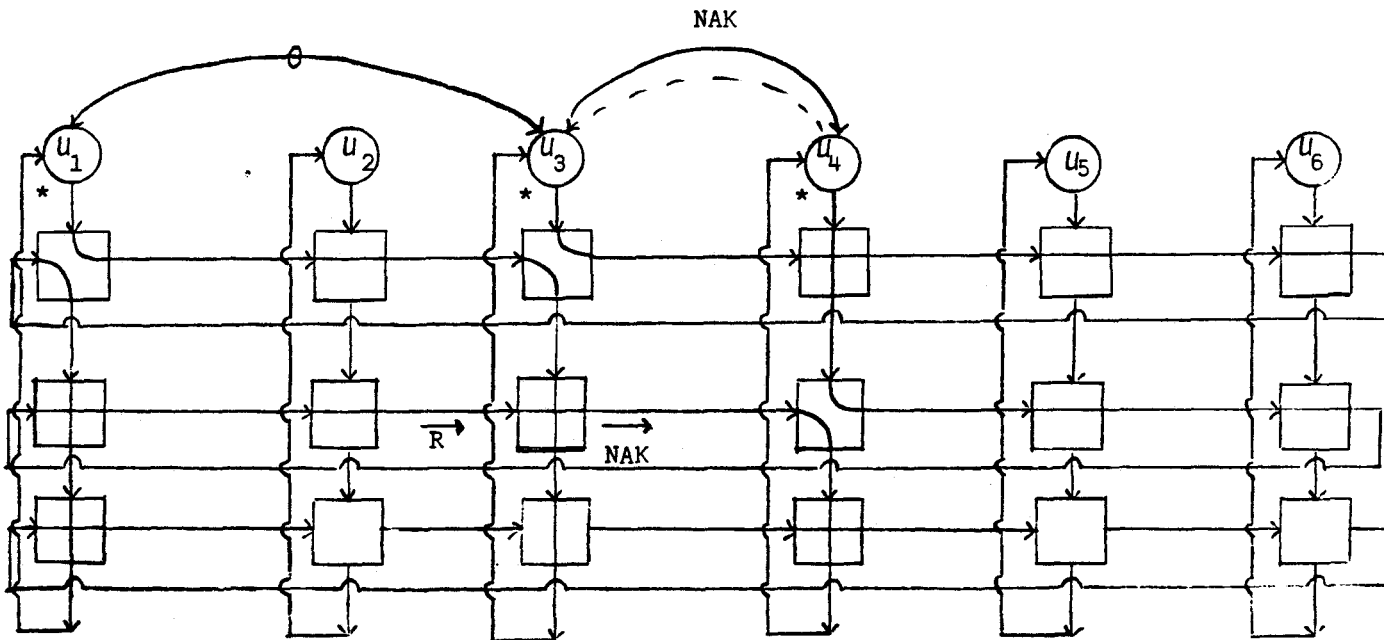


Figure 12 : La requête émise par  $U_4$  vers  $U_3$  est acquittée négativement, puisque  $U_3$  est en communication avec  $U_1$ . Les ANNEAUX dont les cellules sont verrouillées sont marqués d'un \*.

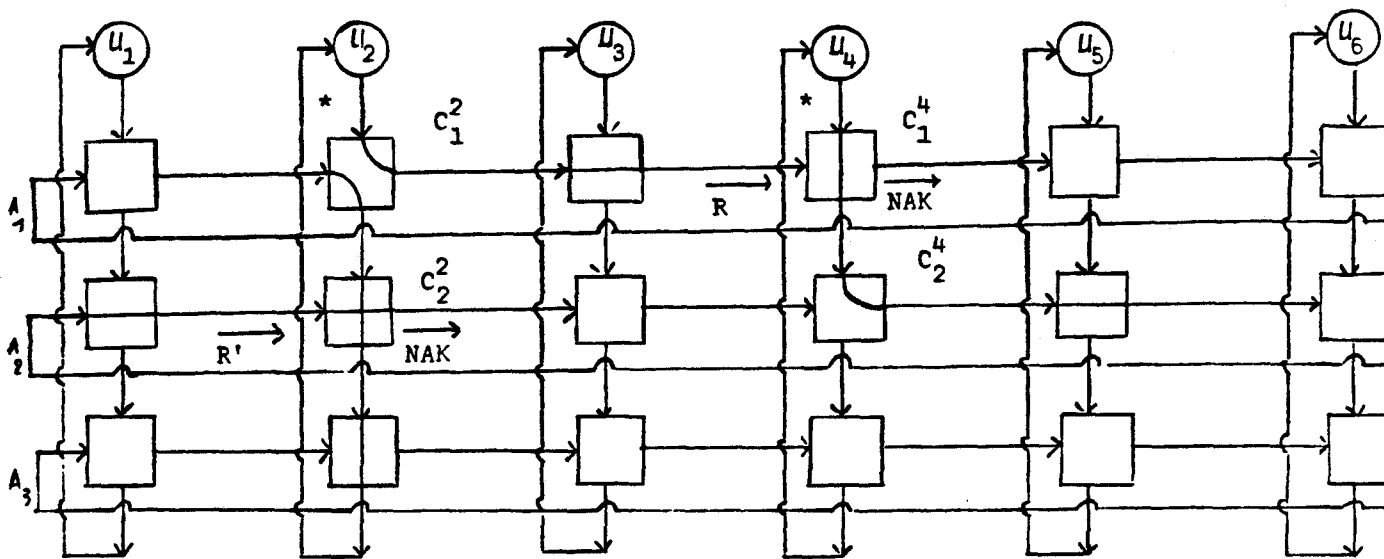


Figure 13 : La requête R occupe les ANNEAUX A- $U_2$ , A- $U_1$  et demande l'allocation de A- $U_4$  - R est émise par  $U_2$  vers  $U_4$  - La requête R' occupe les ANNEAUX A- $U_4$  et A- $U_2$  et demande l'allocation de A- $U_2$  - R' est émise par  $U_4$  vers  $U_2$  -.

#### 4. Conclusion

Une troisième extension du modèle de base, le MULTI-ANNEAU, vient d'être présentée. C'est un outil de communication simple face et sans blocage dans le cas d'une configuration de  $\frac{n}{2}$  ANNEAUX. Des configurations ne comportant pas exactement  $\frac{n}{2}$  ANNEAUX, peuvent, cependant, être considérées. De telles configurations permettent de définir la configuration de l'outil en fonction non seulement du nombre des unités qu'il interconnecte, mais également en fonction de leurs charges de communication. Elles permettent aussi de répondre à des objectifs tels que



i) La tolérance aux pannes :

Si la configuration comporte plus de  $\frac{n}{2}$  ANNEAUX, le réseau reste fonctionnel et avec les mêmes performances, malgré la mise hors service de l'un ou plusieurs de ses ANNEAUX.

ii) Le fonctionnement en mode dégradé :

Dans le cas d'une configuration de  $\frac{n}{2}$  ANNEAU, la mise hors service d'un ou de  $\frac{n}{2} - 1$  ANNEAUX transforme le MULTI-ANNEAU en un outil avec blocage. L'outil continue donc d'être fonctionnel mais avec des performances moindres.

Notons que le protocole d'interconnexion proposé pour le MULTI-ANNEAU peut être repris indépendamment de sa configuration.

\*\*\*\*\*  
\*  
\* CONCLUSION \*  
\*  
\*\*\*\*\*

Dans ce chapitre, nous avons étudié les possibilités d'extension régulières qu'offre le modèle de base. De cette étude il découle que de nombreuses topologies peuvent être définies à partir de l'ANNEAU. Nous nous sommes limités à l'étude de trois d'entre-elles : ARBRE, CELLULAIRE et MULTI-ANNEAU. L'étude de chacune de ces trois topologies a été faite avec la contrainte que les unités communiquent entre-elles par commutation de ligne bidirectionnelle et simultanée.

Le réseau-ARBRE a été entièrement défini. Une mise en oeuvre répartie de son protocole d'interconnexion a été proposée. Cependant, cet outil est spécifique et son utilisation ne peut être envisagée que si des conditions sur la localité des interactions entre les unités attachées à un même sous-ARBRE, sont vérifiées.

Le réseau CELLULAIRE, n'a été que partiellement défini. L'étude de sa capacité fonctionnelle a montré que, d'une part, son utilisation ne peut être envisagée, du moins en commutation de ligne, pour interconnecter des unités fortement couplées et que d'autre part, l'ANNEAU, pris comme cellule de base, n'est pas adaptée à cette topologie et que par conséquent, une redéfinition de la cellule de base s'avère nécessaire.

Le MULTI-ANNEAU, notre dernière extension, a été défini dans le cas d'une configuration sans blocage. Nous avons montré que d'autres configurations de cet outil sont possibles et que certaines de ces configurations autorisent un fonctionnement en mode dégradé ou encore sont tolérantes aux pannes.

\*\*\*\*\*  
\*  
\* CHAPITRE IV \*  
\*  
\*\*\*\*\*

EVALUATION DES OUTILS PROPOSES

## 0. INTRODUCTION

L'objet de ce chapitre est d'évaluer les outils qui ont été définis dans les Chapitres II et III de ce travail. Pour chacun de ces outils deux éléments seront considérés : Performance et complexité. Cette dernière sera mesurée selon la méthode classique d'évaluation de la complexité. Une comparaison avec des outils existants, lorsque cela est possible, est également effectuée.

### 1. L'ANNEAU

L'évaluation de la complexité logique, ainsi que des performances de l'ANNEAU repose sur une implémentation de ses modules de contrôle, selon le schéma de la figure 1. Le circuit logique que représente ce schéma est une implémentation possible du protocole d'interconnexion - cf. II.2.b.

#### 1.a. Performances

Nous déterminons les performances de l'ANNEAU dans deux cas. Dans le premier nous supposons qu'il n'y a pas de conflits d'accès à l'ANNEAU, et que, à tout instant, une connexion est établie ; cela nous permet d'évaluer les performances potentielles. Dans le second, nous tenons compte des conflits d'accès en considérant une configuration particulière que nous précisons plus loin, cela nous permet de déterminer les performances effectives de l'ANNEAU pour cette configuration.

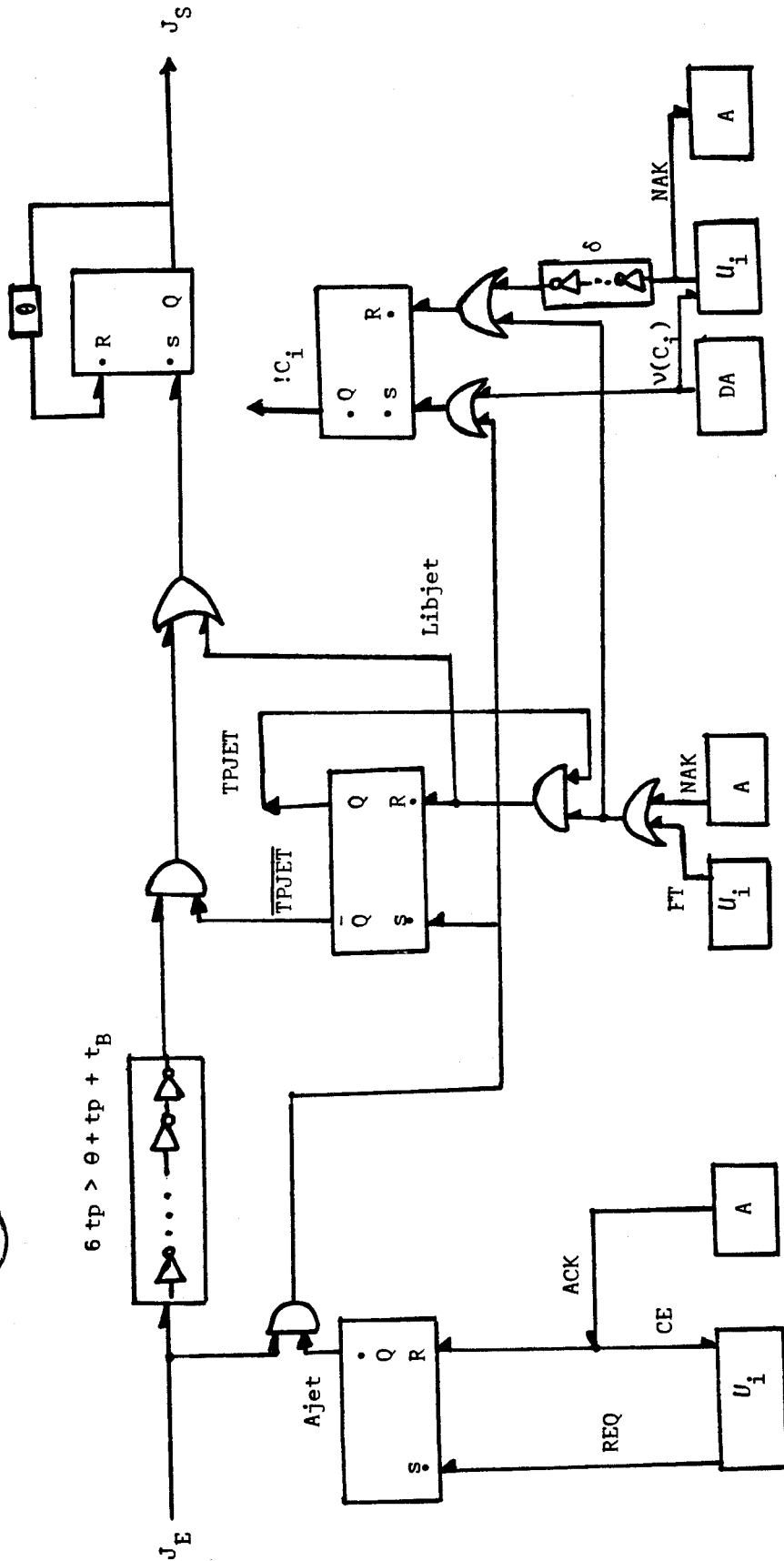
##### 1.a.1. Performances potentielles

Les valeurs temporelles seront exprimées en fonction du temps de traversée d'une porte  $t_p$ .

###### 1.a.1.a. Délai d'accès : $\delta$

Dans le cas idéal, où il n'y a pas de conflit d'accès à l'ANNEAU, seul le temps de propagation du jeton intervient dans le calcul du délai d'accès d'une unité à l'ANNEAU.





A : ANNEAU.  
 tp : temps de traversé d'une porte logique.  
 t<sub>B</sub> : temps de transition d'une Bascule RS (t<sub>B</sub> ≈ 2 tp).  
 DA : Décodeur d'adresse.

Figure 1 : Circuit logique du module de contrôle d'une cellule C<sub>i</sub>.

\* Valeur minimale de  $\delta$  :  $\delta_{\min}$

Le délai d'accès est minimal, chaque fois que le jeton est présent à l'entrée du module de contrôle, à l'"instant" où celle-ci demande sa capture.

$$\delta_{\min} = 6 \text{ tp} \quad [\text{voir figure 1}]$$

\* Valeur maximale de  $\delta$  :  $\delta_{\max}$

Le délai d'accès est maximal, lorsque le jeton quitte le module de contrôle d'une unité, à l'"instant" même où celle-ci réclame sa capture.

$$\delta_{\max} = (10n + 4) \text{ tp}$$

n : étant le nombre d'unités

10 tp est le délai de traversée d'un module de contrôle

4 tp est le temps nécessaire pour capter le jeton lorsque celui-ci est disponible à l'entrée du module de contrôle.

\* Valeur moyenne de  $\delta$  :  $\delta_{\text{moy}}$

C'est le temps nécessaire pour que le jeton traverse la moitié de l'ANNEAU avant d'être capté par le module de contrôle d'une unité :

$$\begin{aligned} \delta_{\text{moy}} &= \left(10 \left\lfloor \frac{n}{2} \right\rfloor + 4\right) \text{ tp} \\ &= (5n + 4) \text{ tp} \end{aligned}$$

Pour une configuration de quatre unités le délai d'accès moyen est de 24 tp et est de 84 tp pour une configuration de 16 unités.

L'utilisation d'une technologie ultra-rapide, est un élément qui contribue à l'obtention de hautes performances. En prenant la nanoseconde comme délai de traversé d'une porte, le délai d'accès moyen pour une configuration de quatre unités est de 24 ns, et est de 84 ns pour une configuration de 16 unités. Ces résultats, comparés à ceux obtenus avec des arbitres asynchrones [SCG 83], nous paraissent tout-à-fait satisfaisants.

Cela démontre, par ailleurs, que la technique du jeton circulant, peut être exploitée efficacement en dehors de son domaine privilégié, celui des réseaux locaux en boucle.

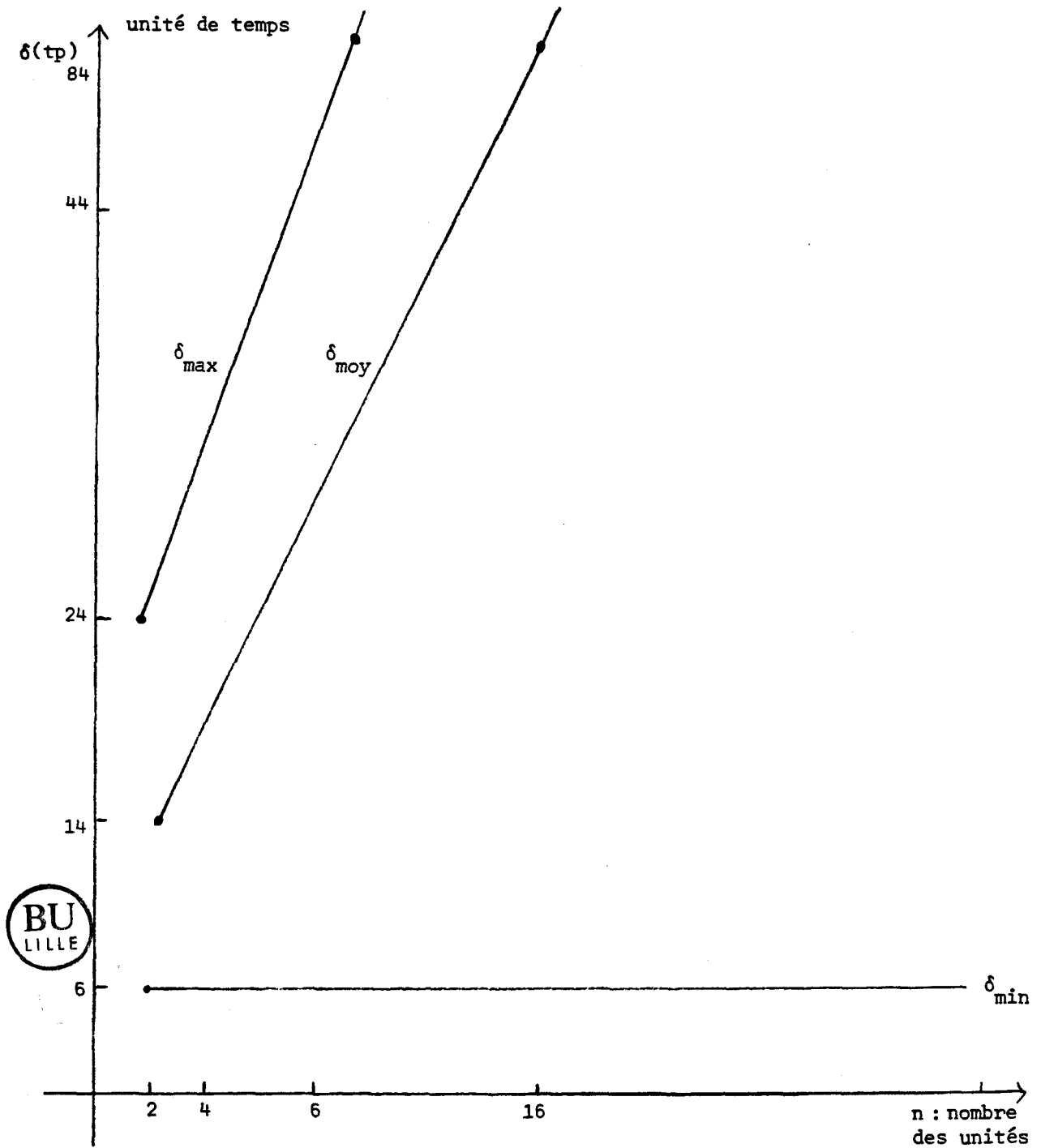


Figure 2 : Les délais d'accès dans le cas idéal : absence de conflit d'accès.

### 1.a.1.b. Débit potentiel

Le débit potentiel de l'ANNEAU,  $D_a^P$ , est fonction du choix du protocole, adopté au niveau "transfert" - cf. II.2.c.3. -. Nous avons opté pour le transfert séquentiel asynchrone.

$$D_a^P = \frac{2}{2 * n * t_p} = \frac{1}{n * t_p}$$

Le débit potentiel décroît de façon exponentielle en fonction du nombre des unités qu'il interconnecte - figure ci-après -. Il varie de 0,0625 mot par unité de temps, pour une configuration de 16 unités à, jusque 0,25 mot par unité de temps pour une configuration de quatre unités.

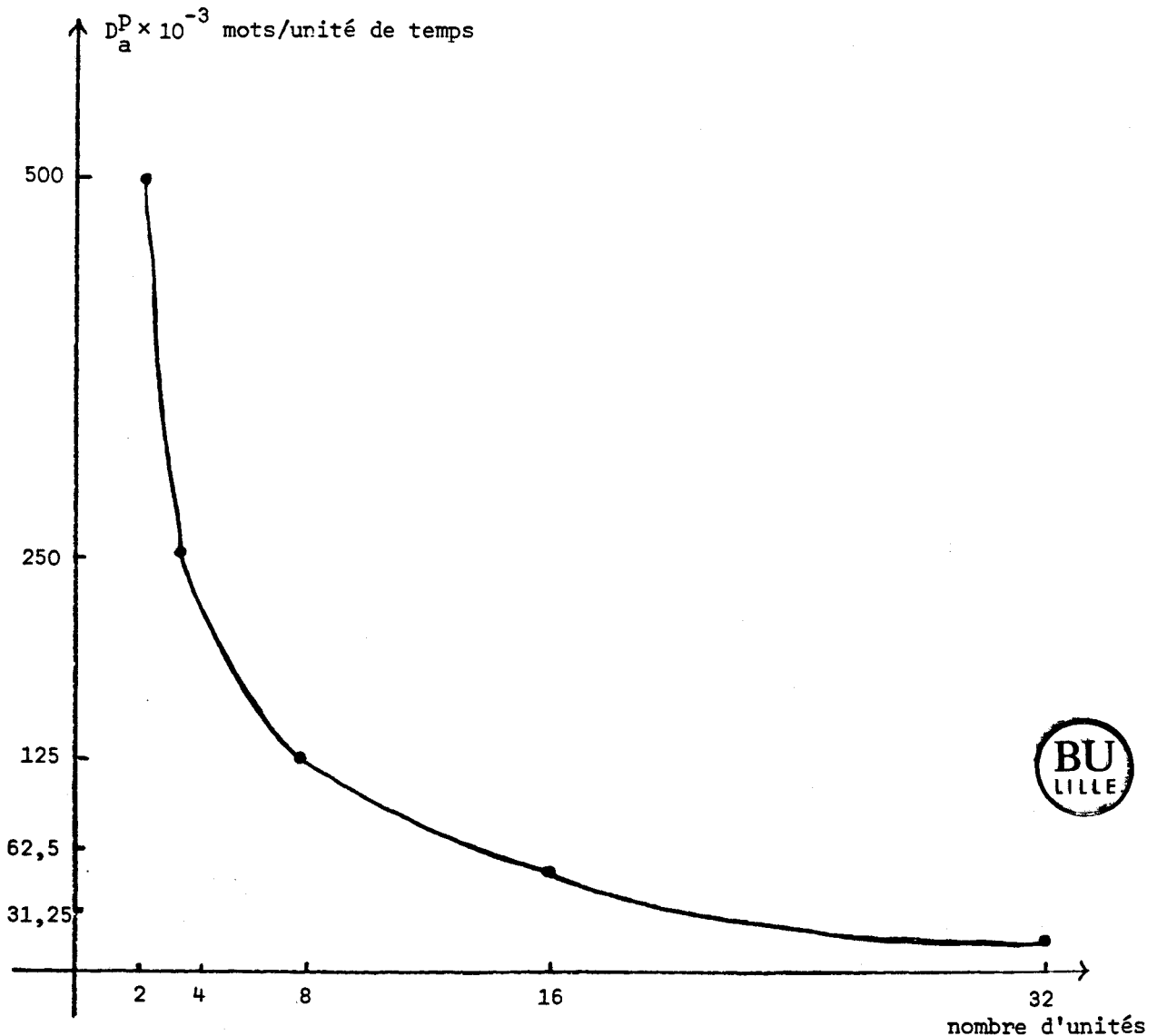


Figure 3 : Le débit potentiel en fonction du nombre d'unités.

### 1.a.2. Performances effectives

Les performances effectives de l'ANNEAU tiennent compte des caractéristiques des unités interconnectées et des conflits d'accès à l'ANNEAU. Nous nous proposons d'évaluer ces performances par simulation. Les configurations effectivement simulées, ont été restreintes à des cas très simples. Cela nous a permis d'établir les équations des performances pour ces configurations et de comparer dans quelle mesure les résultats obtenus par simulation coïncident avec ceux obtenus à partir des équations.

#### 1.a.2.a. Le logiciel de simulation

Il se décompose en trois modules :

1. Un premier module, définit la configuration à simuler :

- i) Les ensembles  $U_a$  des unités actives, et  $U_p$  des unités passives. Les unités actives sont les seules unités admises à émettre des demandes d'accès à l'ANNEAU.
- ii) La loi statistique ou périodique qui régit la fréquence des accès d'une unité.
- iii) Les lois qui régissent la répartition des demandes d'une même unité active  $u_i$ , d'une part entre les deux sous-ensembles  $U_a \setminus \{u_i\}$  et  $U_p$ , et d'autre part entre les unités de chacun de ces deux sous-ensembles.
- iv) Enfin, les lois qui régissent la longueur des transferts entre deux unités actives, ainsi qu'entre une unité active et une unité passive.

2. Le second module simule la configuration définie par le premier et enregistre pour chaque unité, les évènements qui se sont produits au cours de la simulation.

3. Le dernier module évalue, à partir des résultats fournis par le module précédent, les performances réelles de l'ANNEAU. Il évalue, en particulier :

- i)  $\bar{d}$  : la valeur moyenne du délai d'accès.
- ii)  $D_a^m$  : la valeur moyenne du débit dans l'ANNEAU.
- iii)  $D_n$  : le débit normalisé

$$D_n = \frac{D_a^m}{D_a^p}$$

### 1.a.2.b. Les hypothèses de la simulation

1. Les unités interconnectées sont assynchrones et leurs demandes de connexions indépendantes.
2. L'intervalle de temps qui sépare la génération de deux demandes successives, par une même unité, obéit à une loi exponentielle de moyenne  $\lambda$ .
3. Une unité active adresse une seconde unité active avec une probabilité  $p$ . Les unités passives sont donc adressées par une même unité active avec une probabilité  $1-p$ .
4. Les demandes d'une unité active  $u_i$  sont uniformément réparties sur chacun des deux sous-ensembles :  $U_a \setminus \{u_i\}$  et  $U_p$ .

Notons  $\text{Prob}[u_i \rightarrow u_j]$ , la probabilité pour que l'unité  $u_i$  adresse l'unité  $u_j$ , on a donc :

$$\text{Prob}[u_i \rightarrow u_j \mid u_j \in U_a] = \frac{1}{\text{Card}(U_a) - 1}$$

$$\text{Prob}[u_i \rightarrow u_j \mid u_j \in U_p] = \frac{1}{\text{Card}(U_p)}$$

5. Longueur des échanges.

Deux cas pourront être envisagés :

- a) Les échanges sont de longueur fixe.
- b) Les échanges sont de longueur variable, la loi qui régit alors cette longueur est une loi exponentielle de moyenne  $\mu$ .

6. Le protocole d'interconnexion simulé, est supposé être implémenté selon le circuit de la figure 1.

7. Protocole de Transfert.

Les configurations sont simulées dans l'hypothèse d'un transfert séquentiel assynchrone.

8. Réponse d'une unité à une demande de connexion.

. Une unité passive est supposée acquitter positivement, toute requête qui lui est adressée.

. Une unité active peut être dans l'un des deux états suivants, lorsqu'elle reçoit une requête :

. *Al* : Activité locale.

. *Att* : Attente d'établissement d'une liaison.

La réponse d'une unité active, lorsqu'elle est sollicitée par une seconde unité active, dépend du type d'application qu'élabore l'ensemble des unités, de l'interdépendance ou non des tâches qu'elles exécutent et de l'état de leurs ressources locales impliquées par une opération d'échange. Il existe même des cas où la réponse dépendra de l'émetteur de la demande, sa priorité par exemple.

Il existe deux façons de simuler cette réponse :

La première consiste à simuler avec précision l'activité interne de chaque unité et de fixer les conditions d'une réponse positive et négative, à une demande de connexion.

La seconde attache à une réponse positive une probabilité ( $q$ ), et la probabilité  $(1-q)$  à une réponse négative - figure 4.a -. La valeur de cette probabilité caractérisera l'application qu'élaborent l'ensemble des unités. Nous pensons que ces préoccupations relèvent beaucoup plus des caractéristiques du système interconnecté, que de l'outil d'interconnexion, proprement dit. Nous avons alors supposé qu'une unité active acquitte positivement toute demande qui

lui est soumise, et ce quelque soit l'état dans lequel elle se trouve au moment de l'arrivée de la demande - figure 4.b -. Nous estimons d'ailleurs, qu'il est impératif qu'une unité acquitte positivement une demande de connexion lorsqu'elle se trouve dans l'état Attente. Cela permet en effet de rendre impossible la production d'interblocage entre deux ou plusieurs unités, se trouvant toutes dans le même état Attente, adressant chacune l'autre et qui acquittent négativement toutes les demandes qu'elles se soumettent.

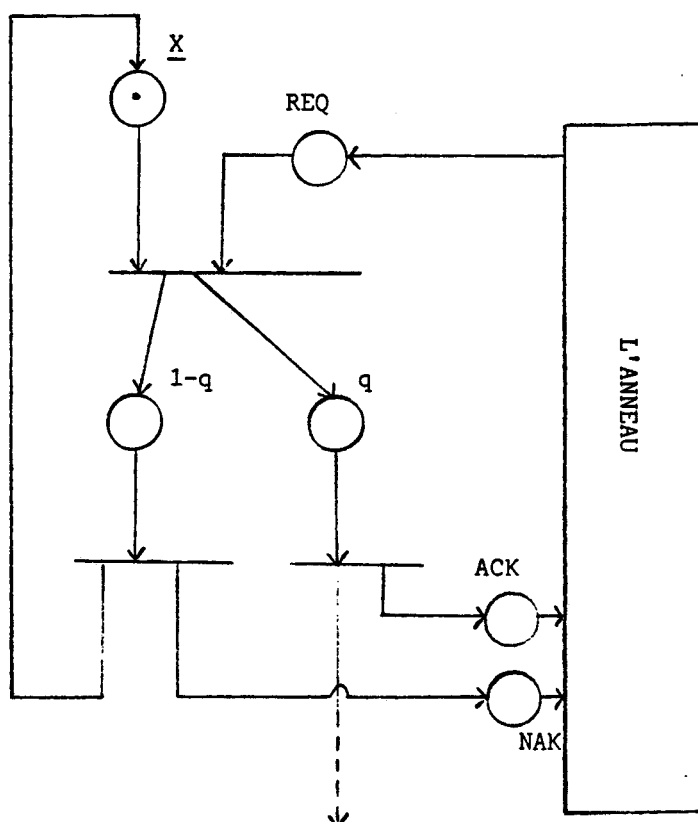


Figure 4.a. : La réponse d'une unité active est régie par une loi de probabilité  $\text{Prob}[\text{ACK}] = q$ ,  $\text{Prob}[\text{NAK}] = 1-q$ .



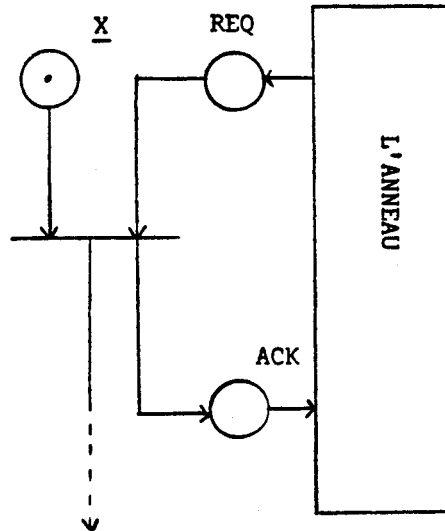


Figure 4.b. : Toute demande est positivement acquittée.

Figure 4 : Réponse d'une unité active à une demande de connexion.

#### 1.a.2.c. Résultats de la simulation

Les configurations que nous avons simulées ont été restreintes au cas où les unités actives adressent exclusivement des unités passives,  $\text{Prob}[U_i \rightarrow U_j \mid U_i \text{ et } U_j \in U_a] = 0$ . Cette restriction, qui n'est nullement due au logiciel de simulation, permet de déterminer aisément les équations qui régissent les performances de l'ANNEAU. Ces équations doivent permettre la validation des performances obtenus par simulation et leur extension à des configurations non simulées.

Débit potentiel d'une unité active :  $D_u^P$ .

$$D_u^P = \frac{2\mu}{\lambda + \frac{2\mu}{D_a^P}}$$

où  $\mu$  : est le nombre moyen de mots envoyés par une unité active à chaque accès.

$2\mu$  : est alors le total de mots envoyés et reçus lors d'un accès.

$\lambda$  : le temps moyen qui sépare l'émission, par une même unité active, de deux demandes d'accès successifs.

$D_a^P$  : le débit potentiel de l'ANNEAU.

Les équations des performances, à savoir le délai d'accès moyen  $\bar{d}$  d'une unité, ainsi que le débit moyen de l'ANNEAU  $D_a^m$ , sont obtenues en distinguant deux cas, selon que le total des débits des unités actives est, ou non, inférieur au débit potentiel de l'ANNEAU.

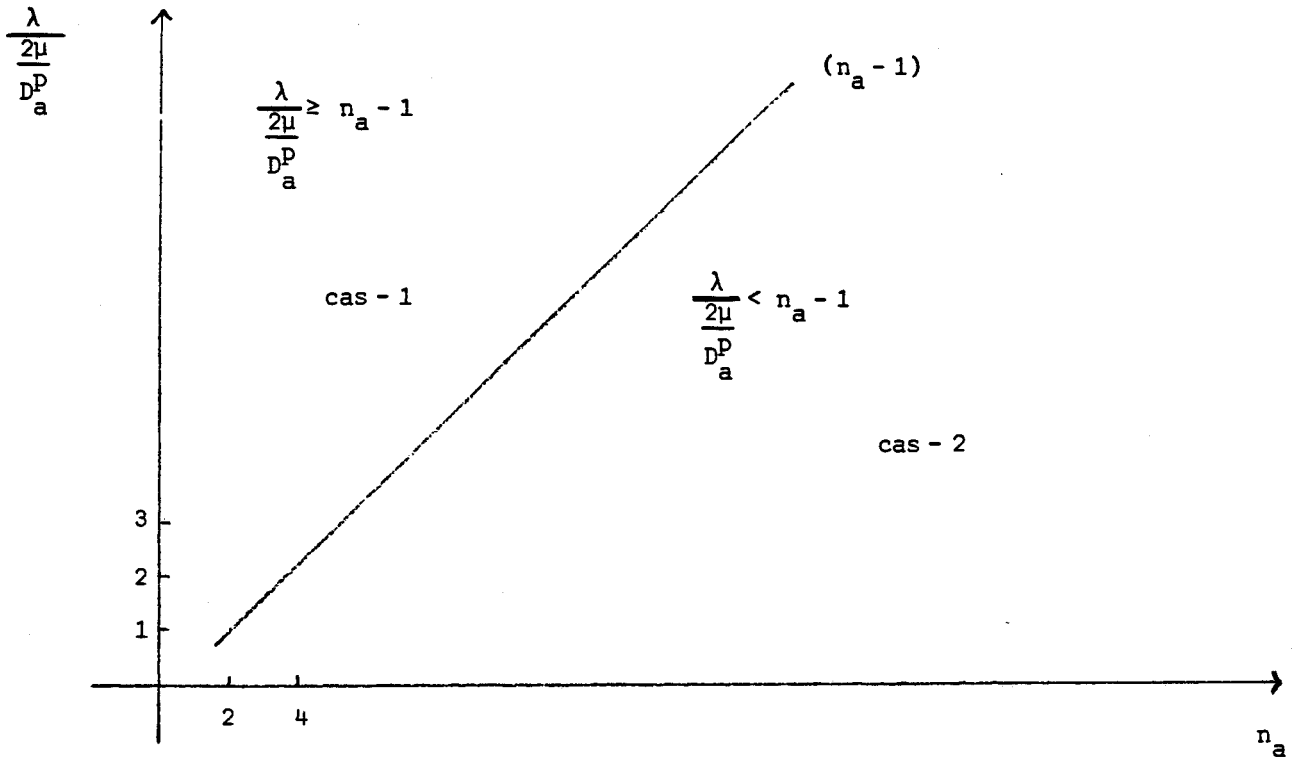


Figure 5

cas - 1 :  $n_a D_u^P \leq D_a^P$ .

Dans ce cas, la charge totale appliquée par l'ensemble des  $n_a$  unités actives est plus faible que la capacité potentielle de transfert de l'ANNEAU. Le délai d'accès moyen  $\bar{d}$  est identique à celui obtenu précédemment en considérant que l'ANNEAU est sans conflit

$$\delta_{\min} = 6 \leq \bar{d} \leq \delta_{\max} = 10n + 4 \quad (1.a)$$

$D_a^m$  : Le débit moyen d'une unité,

$$D_u^m = \frac{2\mu}{\lambda + \bar{d} + \frac{2\mu}{D_a^p}}$$

$$D_a^m = n_a D_u^m$$

Ces performances sont obtenues dans le cas :  $n_a D_u^p \leq D_a^p$

$$n_a D_u^p \leq D_a^p \Leftrightarrow \frac{\lambda}{\frac{2\mu}{D_a^p}} \geq n_a - 1$$

Cette relation exprime la relation que doit vérifier le rapport entre le temps d'activité interne d'une unité ( $\lambda$ ) et le temps de communication  $\left(\frac{2\mu}{D_a^p}\right)$  pour assurer un meilleur délai d'accès (a.1). Quant au débit moyen de l'ANNEAU, il vérifie la double inégalité suivante (2) :

$$\frac{n_a \mu}{n(n_a + 5) + 4} \leq D_a^m \leq D_a^p \quad (1.b)$$

Cas - 2 :  $n_a D_u^p > D_a^p$ .

La charge totale appliquée à l'ANNEAU, par les  $n_a$  unités actives est supérieure à sa capacité potentielle de transfert.

$$\bar{d} = (n_a - 1) \frac{2\mu}{D_a^p} + \delta_{\max} - \lambda$$

Temps de communication  
de  $n_a - 1$  unités actives

$$\bar{d} = (n_a - 1) \frac{2\mu}{D_a^p} + 10n + 4 - \lambda \quad (2.a)$$

$$D_u^m = \frac{2\mu}{\lambda + \bar{d} + \frac{2\mu}{D_a^p}} = \frac{2\mu}{n_a \frac{2\mu}{D_a^p} + 10n + 4}$$

$$D_a^m = n_a D_u^m = \frac{2n_a \mu}{n_a \frac{2\mu}{D_a^p} + 10n + 4} \quad (2.b)$$

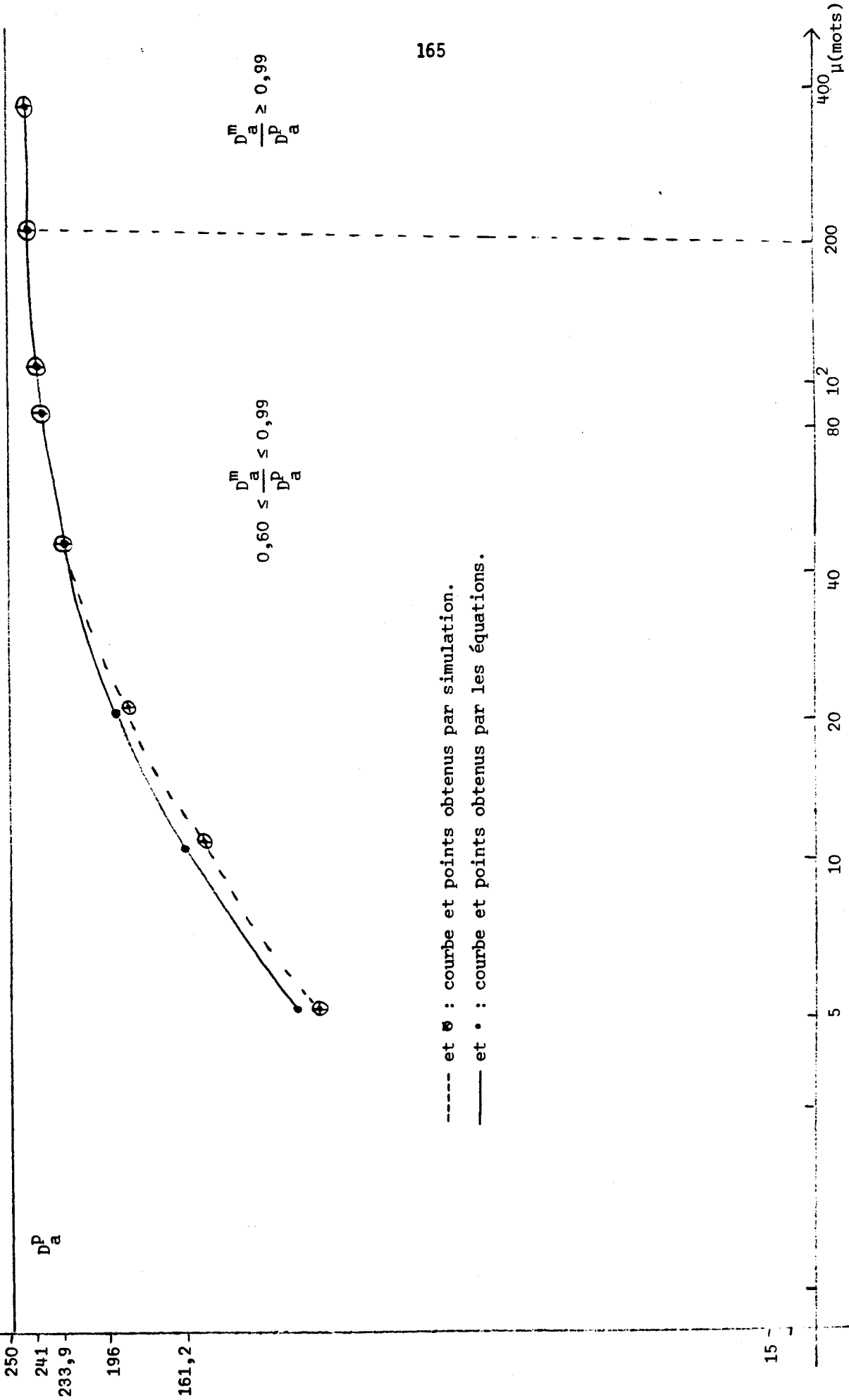
Le débit moyen de l'ANNEAU, lorsque  $\frac{\lambda}{2\mu} < n_a - 1$ , est indépendant de  $\lambda$  (2.b).

Pour ce même cas, le débit moyen croît en fonction de la longueur des échanges,  $\mu$ , et tend asymptotiquement vers le débit potentiel de l'ANNEAU  $D_a^D$ .

Les résultats de la simulation - table 1 et figure ci-après - montrent que le débit moyen est de 60 % du débit potentiel pour des échanges très courts ( $\mu = 5$ ), et atteint les 99 % de ce même débit potentiel pour des échanges longs ( $\mu \geq 200$ ).



$D_m^a \times 10^{-3}$  mot/unité de temps



----- et ⊕ : courbe et points obtenus par simulation.  
—— et • : courbe et points obtenus par les équations.

Figure 6 : Débit moyen de l'anneau en fonction de  $\mu$ .  $n = 2$ ,  $\lambda = 10$  unités de temps

Le délai d'accès d'une unité active - figure 7 -.

- croît linéairement en fonction de  $\mu$  : la longueur des échanges.
- décroît linéairement en fonction de  $\lambda$  : intervalle de temps séparant deux requêtes successives d'une même unité.

Ce délai est de l'ordre de la micro-seconde pour  $t_p = 1$  ns :

$$\begin{cases} \lambda = 10 \text{ ns} \\ \mu < 100 \text{ mots} \end{cases} \Rightarrow \bar{d} < 1 \mu\text{s}.$$

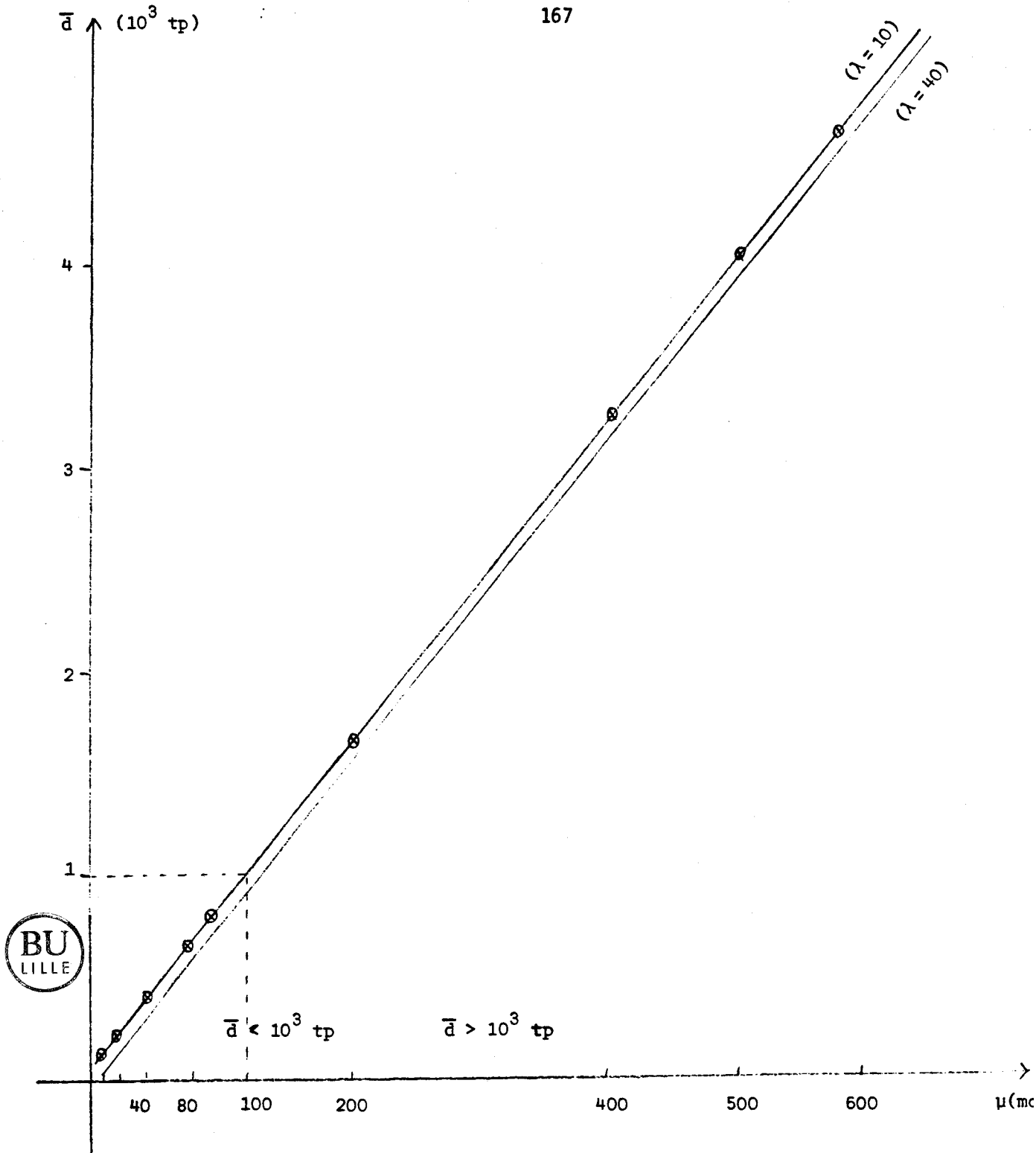


Figure 7 :  $n = 4, n_a = 2$

• : délais d'accès obtenus par simulation avec  $\lambda : 10$ .

Les deux droites correspondent à l'équation (2.a) du délai d'accès pour  $\lambda = 10$ , et  $\lambda = 40$ .

$$n = 4, n_a = 2, \lambda = 10.$$

Echanges fixes,  $p = 0.$

	$\mu(\text{mots})$	$\bar{d}$		$D_a^m$		$D_a^m/I$
		(unité)	(unité = ns)	(mot/unité)	mégamot/sec.	
cas - 1 $\mu \leq 1$	1	52,5	52,5 ns	0,061	61	0,25
cas - 2 $\mu > 1$	5	84	84 ns	0,157	157	0,63
	10	124	0,1 $\mu\text{s}$	0,193	193	0,77
	20	204	0,2 $\mu\text{s}$	0,218	218	0,87
	40	364	0,4 $\mu\text{s}$	0,233	233	0,93
	80	684	0,7 $\mu\text{s}$	0,241	241	0,96
	100	844	0,8 $\mu\text{s}$	0,243	243	0,97
	200	1640	1,6 $\mu\text{s}$	0,247	247	0,98
	400	3240	3,2 $\mu\text{s}$	0,249	249	0,99
	500	4040	4 $\mu\text{s}$	0,249	249	0,99
	600	4840	4,8 $\mu\text{s}$	0,249	249	0,99
700	5640	5,6 $\mu\text{s}$	0,249	249	0,99	

Tableau 1





$n = 4, n_a = 2, \lambda = 40$  unités de temps

Echanges fixes,  $p = 0$ .

	$\mu$ (mots)	$\bar{d}$		$D_a^m$		$D_a^m/D_a^p$
		(unité)	(unité = ns)	(mot/unité)	mégamot/sec.	
cas - 1 $\mu \leq 5$	1	44	44 ns	0,046	46	0,18
cas - 2 $\mu > 5$	10	97,5	97,5 ns	0,189	189	0,75
	20	174	0,1 $\mu$ s	0,217	217	0,86
	40	333	0,3 $\mu$ s	0,233	233	0,93
	80	653	0,6 $\mu$ s	0,241	241	0,96
	100	813	0,8 $\mu$ s	0,243	243	0,97
	200	1610	1,6 $\mu$ s	0,247	247	0,98
	700	5610	5,6 $\mu$ s	0,249	249	0,99

Tableau 2

### 1.a.3. Conclusion



Les configurations que nous avons simulées comportent quatre unités interconnectées par l'ANNEAU, et dont deux sont actives. Le débit moyen obtenu atteint dans ce cas les 0,241 mot par unité de temps, soit environ 99 % du débit potentiel de l'ANNEAU, et ce pour des échanges longs ( $\mu \geq 200$  mots) et une fréquence d'accès très élevée, un accès toutes les 10 unités de temps.

Nous avons évalué le délai d'accès, sans tenir compte des conflits d'accès à l'ANNEAU. Ce délai croît linéairement en fonction du nombre d'unités  $n$ . Avec  $n = 4$  et  $t_p = 1$  ns, ce délai est inférieur à 44 ns. Ceci met en valeur la technique du jeton circulant comme moyen d'accès.

L'évaluation par simulation du délai moyen d'accès est assez élevé. Il est de l'ordre de la microseconde (avec  $n = 4, t_p = 1$  ns).

### 1.b. Compléxité logique

L'ANNEAU comporte deux modules par unité interconnectée : le module de contrôle et le module de commutation.

Le circuit logique - figure 1 -, proposé pour implémenter le module de contrôle comporte une trentaine de portes logiques, environ.

La cellule de commutation est réalisable au moyen de  $2b$  multiplexeurs  $2 \times 1$ ,  $b$  étant le nombre de liaisons de chaque entrée de la cellule. En considérant la compléxité d'un multiplexeur  $2 \times 1$ , équivalente à celle d'une porte logique, le coût d'une cellule est de  $2b$ .

Le coût de l'ANNEAU pour  $n$  unités interconnectées, est donc d'environ  $n(2b + 30)$ . Il croît donc en  $O(n)$ .

$$n = 16, b = 30 \Rightarrow \text{coût} \approx 1504 \text{ portes.}$$

Un circuit ayant cette compléxité peut être facilement intégré dans un seul boîtier. Cependant, le nombre de broches dont il doit disposer, supérieur à 1024, est très élevé. Il est donc indispensable d'envisager une décomposition de l'ANNEAU, à moins qu'il ne soit possible d'intégrer l'ensemble des unités et l'ANNEAU dans le même boîtier.

### 1.c. Comparaison de l'ANNEAU avec le BUS

Le BUS et l'ANNEAU sont des outils de communication équivalents.

- Au niveau de la complexité, elle croît en  $O(n)$  pour les deux outils.
- Au niveau fonctionnel, les deux outils sont de type simple face, et n'autorisent pas de simultanéité. Cependant l'ANNEAU présente l'avantage de permettre un échange bidirectionnel simultané entre ses unités. L'obtention de cette caractéristique au moyen du BUS, nécessite le dédoublement de sa voie et une gestion non triviale des sens de transfert le long des deux voies.

## 2. LES OUTILS EXTENSION

### 2.a. Le réseau ARBRE

La configuration considérée est celle d'un arbre a aire complet, et interconnectant  $n = a^p$  unités.

#### 2.a.1. Performances

L'évaluation des performances d'un réseau consiste en premier lieu, à évaluer sa capacité d'établir des liaisons simultanées entre ses unités. Cette capacité est variable dans le réseau ARBRE, entre 1 - figure 8.a. -, et  $\frac{n}{a}$  - figure 8.b. -.

Une évaluation précise du nombre moyen des liaisons simultanément établies exige l'émission d'hypothèses sur la répartition des demandes envoyées par chaque unité. L'obtention d'un degré élevé de simultanéité passe nécessairement par une organisation des unités, selon leurs degrés d'interaction.

Le délai d'accès est proportionnel au nombre d'ANNEAUX qui constituent une liaison. Le nombre vrai de 1 - figure 8.b. -, dans le cas où la liaison relie deux unités attachées au même ANNEAU du niveau feuille, et atteint  $2[\log_a n - 1] + 1$  ANNEAUX, lorsque la liaison concerne deux unités dont le seul ANNEAU commun est la racine du réseau - figure 8.a. -.

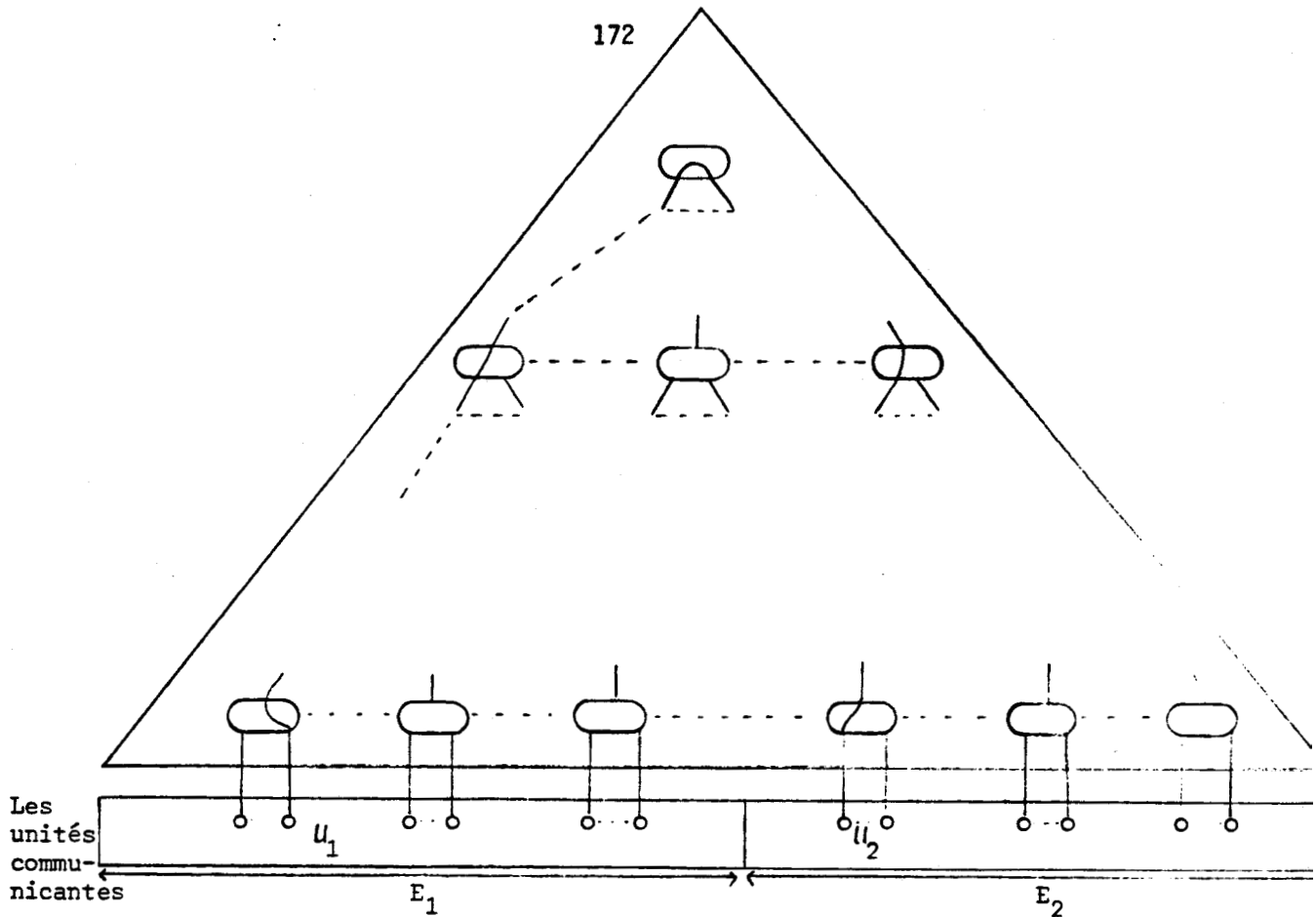


Figure 8.a. : Une seule liaison  $u_1 \leftrightarrow u_2$ , est établie dans le réseau, elle traverse la racine du réseau.

L'ensemble des autres demandes se font entre les unités de  $E_1 \setminus \{u_1\}$  et les unités de  $E_2 \setminus \{u_2\}$ .

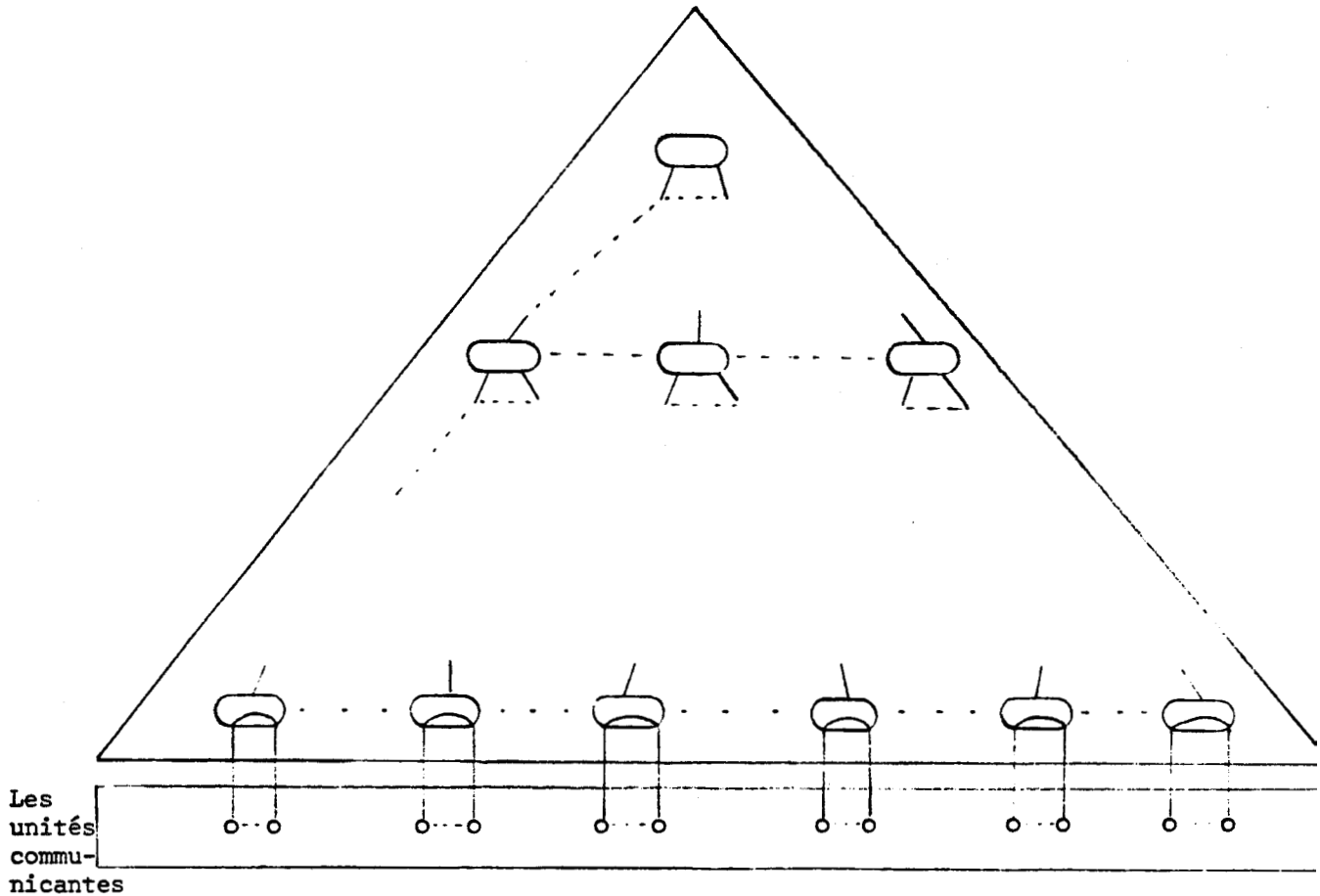


Figure 8.b. :  $\frac{n}{a}$  liaisons sont établies dans le réseau ARBRE.  
Chaque liaison occupe un ANNEAU du niveau feuille.

### 2.a.2. Compléxité



Le nombre de cellules  $2 \times 2$  que comporte le réseau ARBRE est :  $(n-1) \frac{a}{a-1}$

$$\log_a(n)-1 \sum_{i=0} \frac{n}{a^i} = (n-1) \frac{a}{a-1} \leq 2(n-1) \quad (a \geq 2)$$

La mise en oeuvre du protocole d'interconnexion exige deux modules de contrôle par cellule. La compléxité du réseau reste cependant proportionnelle au nombre d'unités qu'il interconnecte, donc équivalente à celle de l'ANNEAU ou encore à celle du BUS.

L'intégration VLSI du réseau ARBRE, si elle est rendue aisée en raison de la régularité de la topologie, exige, au même titre que l'intégration de l'ANNEAU, la décomposition du réseau afin de remédier au problème des broches.

### 2.a.3. Conclusion

Le réseau ARBRE est un outil simple face, avec blocage. Sa complexité croît en  $O(n)$ , de la même manière que pour le BUS et l'ANNEAU. Cependant il présente une capacité fonctionnelle plus élevée. Il autorise en effet, l'établissement, sous certaines conditions sur la répartition des demandes d'accès des unités, jusque  $\frac{n}{a}$  liaisons simultanées.

### 3. CONCLUSION

L'évaluation des performances de l'ANNEAU a montré que son débit effectif peut atteindre 99% de son débit potentiel. Cette évaluation a également permis de mettre en valeur la technique du jeton circulant, pris comme moyen de contrôle d'accès à l'ANNEAU. Comparé au BUS, l'ANNEAU présente l'avantage d'autoriser, chaque fois qu'elle est établie, une communication bidirectionnelle est simultanée.

Le réseau ARBRE a une complexité qui croît en  $O(n)$ . Il autorise un degré de simultanété variable, entre 1 et  $\frac{n}{a}$ ,  $a$  étant l'arité du réseau. La topologie du réseau ARBRE privilégie des systèmes dont les unités sont organisées selon leurs degrés d'interaction.

Le MULTI-ANNEAU, outil de communication sans blocage, est d'une complexité qui croît en  $O(n^2)$ . Comparé au cross-bar il présente l'avantage d'être simple face et d'autoriser des communications bidirectionnelles simultanées.

L'intégration VLSI de chacun de ces outils exige sa décomposition, pour remédier au problème de broches. Leur complexité et performances doivent être réévaluées en conséquence.

### CONCLUSION GENERALE

L'étude que nous avons effectuée a débouché sur la définition d'abord d'un outil de communication, l'ANNEAU, et, ensuite, sur deux extensions de cet outil, le réseau ARBRE et le MULTI-ANNEAU. Pour chacun de ces outils, un protocole d'interconnexion, entièrement réparti, a été proposé.

Les dispositifs proposés, répondent aux exigences d'une intégration VLSI, grâce à la régularité de leurs topologies, et ont les caractéristiques suivantes :

- Ils garantissent la commutation de ligne bidirectionnelle et simultanée,
- assurent l'interconnexion de deux unités quelconques du réseau, ce sont des réseaux simple face et,
- gèrent, de façon simple, les accès asynchrones.

L'ensemble de ces dispositifs représente une gamme d'outils de communication. En effet, en classant ces outils selon leur complexité et performance croissantes, on trouve :

- L'ANNEAU, outil avec blocage ; il autorise l'établissement d'une liaison au plus et a une complexité qui croît en  $O(n)$  ( $n$  étant le nombre des unités).
- Le réseau ARBRE, outil avec blocage ; il autorise un nombre de liaisons variant de 1 à  $\frac{n}{a}$  (ou  $a$  est l'arité de l'arbre), et a une complexité qui croît en  $O(n)$ .
- Le MULTI-ANNEAU, outil sans blocage ; il permet d'établir un nombre maximal de liaisons, jusque  $\frac{n}{2}$  liaisons simultanées, et a une complexité qui croît en  $O(n^2)$ .

Notons cependant, que la mesure de la complexité a été effectuée selon la méthode classique d'évaluation et qui consiste à compter le nombre de commutateurs d'un réseau. Nous estimons qu'une évaluation VLSI, qui calcule la surface d'intégration, est indispensable et qu'elle aboutirait à une réduction de la complexité.

Le travail que nous avons effectué peut donner lieu à des recherches dans au moins les trois directions suivantes :

1. La première complèterait la gamme des outils que nous avons déjà développée. Elle conserverait le modèle de base et l'enrichirait par de nouvelles extensions. L'approche extension par graphe nous semble devoir être reprise et approfondie.
2. La seconde, conserverait toujours la topologie du modèle de base mais l'exploiterait différemment, comme cela a été introduit dans le chapitre II de ce travail.
3. La dernière direction, consiste à reprendre l'ensemble des dispositifs que nous avons proposés et étudier la possibilité de leur utilisation avec la commutation de message et/ou mixte, comme mode transfert.



ANNEXE : ALGORITHME DE ROUTAGE DANS LE RESEAU ARBRE

1. ENONCE DE L'ALGORITHME

Soit  $C_{i,j}^k$  une cellule du réseau ARBRE traversée par la demande  $\underline{D} = d_0 \dots d_n$  et  $D \xrightarrow{p} A_i^j$ .

•  $k = a$

$$\underline{\text{Si}} \ d_0 \dots d_{i-1} = v(C_{i,j}^a) \ \underline{\text{alors}} \quad (R-1) : \begin{cases} n(C_{i,j}^a) = T \\ \underline{\text{et}} \\ D \xrightarrow{p} A_i^j \end{cases}$$

$$\underline{\text{sinon}} \quad (R-2) : \begin{cases} n(C_{i,j}^a) = E \\ \underline{\text{et}} \\ D \xrightarrow{p} A_{i-1}^j \text{ div } a \end{cases}$$

fsi

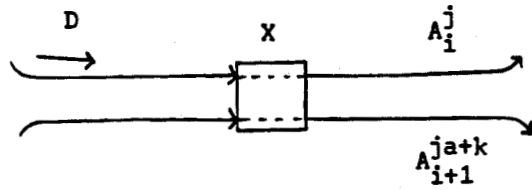
•  $k \neq a$  ( $0 \leq k \leq a-1$ )

$$\underline{\text{Si}} \ d_0 \dots d_i = v(C_{i,j}^k) \ \underline{\text{alors}} \quad (R-3) : \begin{cases} n(C_{i,j}^k) = E \\ \underline{\text{et}} \\ D \xrightarrow{p} A_{i+1}^{ja+k} \end{cases}$$

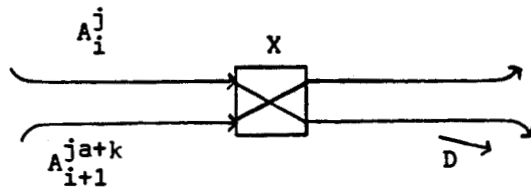
$$\underline{\text{sinon}} \quad (R-4) : \begin{cases} n(C_{i,j}^k) = T \\ \underline{\text{et}} \\ D \xrightarrow{p} A_i^j \end{cases}$$

fsi



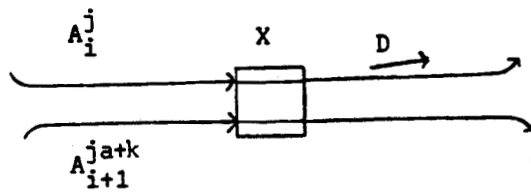


a - Avant traversée de la cellule X par la demande D.



b - Après traversée de la cellule X par D

$$d_0 \dots d_i = x_0 \dots x_i \quad (R-3)$$



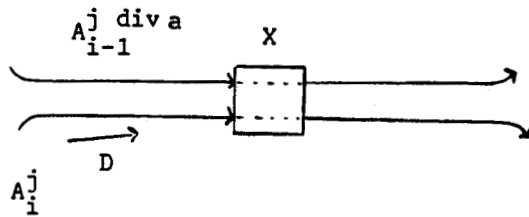
c - Après traversée de la cellule X par D

$$d_0 \dots d_i \neq x_0 \dots x_i \quad (R-4)$$

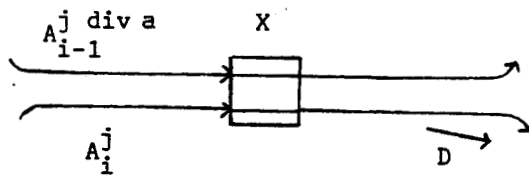
$$. X = x_0 \dots x_i = v(C_{i,j}^k) \text{ avec } k \neq a.$$

La traversée d'une cellule  $C_{i,j}^k$   $k \neq a$  par une demande D.



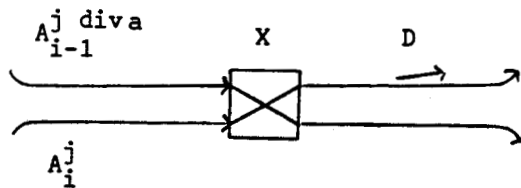


a - Avant la traversée de la cellule X par la demande D.



b - Après traversée de la cellule X par D

$$d_0 \dots d_{i-1} = x_0 \dots x_{i-1} \quad (\text{R-1})$$



c - Après traversée de la cellule X par D

$$d_0 \dots d_{i-1} \neq x_0 \dots x_{i-1} \quad (\text{R-2})$$

.  $X = x_0 \dots x_{i-1} = v(C_{i,j}^a)$ .

. Traversée de la cellule  $C_{i,j}^a$  par D.



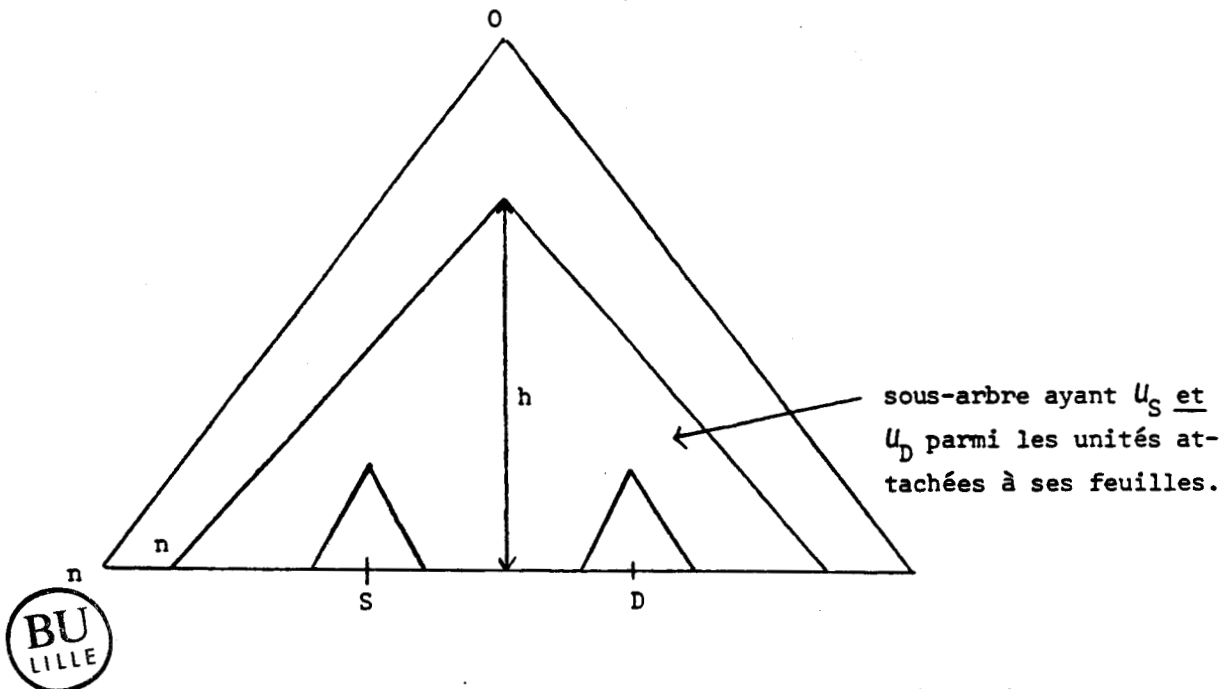
## 2. PREUVE DE L'ALGORITHME

Soient deux unités  $U_S$  et  $U_D$ , attachées respectivement aux deux cellules du niveau feuille, dont les numéros sont S et D

$$S = s_0 \dots s_n$$

$$D = d_0 \dots d_n$$

$U_S$  est l'unité source, et  $U_D$  l'unité destinataire.



La détermination de la hauteur du sous-arbre ayant  $U_S$  et  $U_D$  dans son niveau feuille, est immédiate, d'après la fonction de numérotation.

$$\text{Posons } m = \min_{0 \leq k \leq n} (s_k \neq d_k) \iff \begin{cases} s_0 \dots s_{m-1} = d_0 \dots d_{m-1} \\ \text{et } s_m \neq d_m \end{cases} \quad (2)$$

(la hauteur  $h$ , du sous-arbre est :  $h = n - m + 1$ )

Le lancement d'une demande D dans le réseau, se traduit au départ par deux opérations :

1. L'unité source commute d'abord sa cellule

$$\eta(S) = E$$

2.  $D \xrightarrow{p} A_n^S \text{ div } a$ , la demande D est envoyée dans l'ANNEAU feuille  $A_n^S \text{ div } a$  et auquel l'unité  $U_S$  est attachée

$$\text{donc au départ : (1) } \begin{cases} S \longleftrightarrow S \\ D \xrightarrow{p} A_n^S \text{ div } a \end{cases}$$

La preuve de l'algorithme de routage comporte deux parties, chacune d'elles correspond à une étape de progression de la demande D dans le réseau ARBRE. La première est l'étape de montée de D vers la racine du sous-arbre contenant  $U_S$  et  $U_D$ ; la seconde est l'étape de descente de la demande D depuis cette racine jusque l'unité destinataire.

### Etape 1

$$\forall p : 0, \dots, n-m$$

$$E_1(p) : \begin{cases} S \longleftrightarrow s_0 \dots s_{n-p} \\ D \xrightarrow{p} A_{n-p}^S \text{ div } a^{p+1} \end{cases}$$

Cette étape correspond à la montée dans le réseau ARBRE depuis l'ANNEAU feuille de  $U_S$  ( $p=0, A_n^{s_0 \dots s_{n-1}}$ ), jusqu'à la racine du sous-arbre contenant  $U_S$  et  $U_D$  et de racine  $A_n^{s_0 \dots s_{n-m}} = A_{n-h+1}^{s_0 \dots s_{h-1}}$  obtenu avec  $p = n-m$ .

### Etape 2

$$\text{Pour tout } p : 0, \dots, n-m$$

$$E_2(p) : \begin{cases} S \longleftrightarrow d_0 \dots d_m \dots d_{m+p} \\ \text{et} \\ D \xrightarrow{p} A_{m+p+1}^D \text{ div } a^{n-(m+p)} \end{cases}$$

Cette étape correspond à la descente dans le sous-arbre contenant  $u_S$  et  $u_D$ , jusqu'à atteindre l'ANNEAU  $A_{n+1}^D$  qui contient l'unité  $u_D$ .

La preuve de chacune des deux étapes se fait par récurrence finie sur  $p : 0, \dots, n-m$ .

Preuve de l'étape 1

$$(1) \Rightarrow E_1(0)$$

$$\text{Supposons } E_1(p) : (\text{avec } p < n-m) \begin{cases} S \xleftrightarrow{\phi} s_0 \dots s_{n-p} \\ D \xrightarrow{p} A_{n-p}^S \text{ div } a^{p+1} \end{cases}$$

- Toute cellule  $\mathcal{C}_{n-p, S \text{ div } a^{p+1}}^k$ ,  $k=0, \dots, n-p-1$  de  $A_{n-p}^S \text{ div } a^{p+1}$  est telle que :

$$v(\mathcal{C}_{n-p, S \text{ div } a^{p+1}}^k) = s_0 \dots s_{n-p-1}^k$$

$$(2) \text{ et } p < n-m \Rightarrow s_0 \dots s_{n-p-1}^k \neq d_0 \dots d_{n-p}$$

$$\text{donc d'après (R-4)} \begin{cases} n(\mathcal{C}_{n-p, S \text{ div } a^{p+1}}^k) = T \\ \text{et} \\ D \xrightarrow{p} A_{n-p}^S \text{ div } a^{p+1} \end{cases}$$

- La cellule  $\mathcal{C}_{n-p, S \text{ div } a^{p+1}}^a$  est telle que

$$v(\mathcal{C}_{n-p, S \text{ div } a^{p+1}}^a) = s_0 \dots s_{n-p-1}$$

$$\text{or } p < n-m \Rightarrow m \leq n-p-1, \text{ et } (2) \Rightarrow s_0 \dots s_{n-p-1} \neq d_0 \dots d_{n-p-1}$$

donc d'après (R-2) :

$$\left\{ \begin{array}{l} \eta(C_{n-p, S \operatorname{div} a^{p+1}}^a) = E \\ \text{et} \\ D \xrightarrow{-p} A_{n-(p+1)}^{(S \operatorname{div} a^{p+1}) \operatorname{div} a} = A_{n-(p+1)}^{S \operatorname{div} a^{p+2}} \end{array} \right.$$

d'après  $E_1(p)$   $S \leftrightarrow s_0 \dots s_{n-p}$ .

Les cellules  $s_0 \dots s_{n-p}$  et  $s_0 \dots s_{n-p-1} = v(C_{n-p, S \operatorname{div} a^{p+1}}^a)$  font partie du même ANNEAU  $A_{n-p}^{S \operatorname{div} a^{p+1}}$

donc  $s_0 \dots s_{n-p} \leftrightarrow s_0 \dots s_{n-p-1} \Rightarrow S \leftrightarrow s_0 \dots s_{n-p-1}$

$$\left\{ \begin{array}{l} S \leftrightarrow s_0 \dots s_{n-(p+1)} \\ D \xrightarrow{-p} A_{n-(p+1)}^{S \operatorname{div} a^{p+2}} \end{array} \right. : E_1(p+1)$$

pour  $p = n-m$  on obtient donc :

$$E_1(n-m) \left\{ \begin{array}{l} S \leftrightarrow s_0 \dots s_m \\ D \xrightarrow{-p} A_m^{S \operatorname{div} a^{n-m+1}} \end{array} \right.$$

$A_m^{S \operatorname{div} a^{n-m+1}}$  est l'ANNEAU racine du sous-arbre contenant  $u_S$  et  $u_D$ .

$$\begin{aligned} S \operatorname{div} a^{n-m+1} &= s_0 \dots s_{m-1} = d_0 \dots d_{m-1} \text{ d'après (2)} \\ &= D \operatorname{div} a^{n-m+1} \end{aligned}$$

$$E_1(n-m) : \left\{ \begin{array}{l} S \leftrightarrow d_0 \dots d_{m-1} s_m \\ D \xrightarrow{-p} A_m^D \operatorname{div} a^{n-m+1} \end{array} \right.$$

Traversée des cellules de l'ANNEAU racine par D

- La cellule  $C_{m,D \text{ div } a^{n-m+1}}^a$  de  $A_m^D \text{ div } a^{n-m+1}$  est telle que

$$v(C_{m,D \text{ div } a^{n-m+1}}^a) = d_0 \dots d_{m-1}$$

$$\text{donc d'après (R-1)} \begin{cases} \eta(C_{m,D \text{ div } a^{n-m+1}}^a) = T \\ D \xrightarrow{\mathcal{P}} A_m^D \text{ div } a^{n-m+1} \end{cases}$$

- Toute cellule  $(C_{m,D \text{ div } a^{n-m+1}}^k)_{k=0}^{a-1}$  de  $A_m^D \text{ div } a^{n-m+1}$  est telle que :

$$v(C_{m,D \text{ div } a^{n-m+1}}^k) = d_0 \dots d_{m-1} k$$

$k \neq a$  et  $k \neq s_m$

$$\bullet k \neq d_m \text{ d'après (R-4)} \begin{cases} \eta(C_{m,D \text{ div } a^{n-m+1}}^k) = T \\ D \xrightarrow{\mathcal{P}} A_m^D \text{ div } a^{n-m+1} \end{cases}$$

$$\bullet k = d_m \text{ d'après (R-3)} \begin{cases} \eta(C_{m,D \text{ div } a^{n-m+1}}^{d_m}) = E \\ D \xrightarrow{\mathcal{P}} A_{m+1}^{(D \text{ div } a^{n-m+1})a+d_m} \end{cases}$$

Les cellules  $d_0 \dots d_{m-1} s_m$  et  $d_0 \dots d_m = v(C_{m,D \text{ div } a^{n-m+1}}^{d_n})$  font partie du même ANNEAU.



On a donc  $d_0 \dots d_{m-1} s_m \leftrightarrow d_0 \dots d_m$ , d'après  $E_1(n-m) S \leftrightarrow d_0 \dots d_{m-1} s_m$   
 $(D \operatorname{div} a^{n-m+1}) \cdot a + d_m = D \operatorname{div} a^{n-m}$

$$(4) \begin{cases} S \leftrightarrow d_0 \dots d_m \\ \text{et} \\ D \xrightarrow{P} A_{m+1}^D \operatorname{div} a^{n-m} \end{cases}$$

Preuve de l'étape 2

(4)  $\Rightarrow E_2(0)$

Supposons  $E_2(p)$  avec  $p < n-m$   $\begin{cases} S \leftrightarrow d_0 \dots d_{m+p} \\ \text{et} \\ D \xrightarrow{P} A_{m+p+1}^D \operatorname{div} a^{n-(m+p)} \end{cases}$

- La cellule  $C_{m+p+1, D \operatorname{div} a^{n-(m+p)}}^a$  de  $A_{m+p+1}^D \operatorname{div} a^{n-(m+p)}$  est telle que :

$$v(C_{m+p+1, D \operatorname{div} a^{n-(m+p)}}^a) = d_0 \dots d_{m+p}$$

la traversée par  $D$  de cette cellule a été déjà faite dans  $E_2(p)$   
 puisque  $E_2(p) \Rightarrow S \leftrightarrow d_0 \dots d_{m+p}$ .

- Toute cellule  $C_{m+p+1, D \operatorname{div} a^{n-(m+p)}}^k$  est telle que

$$v(C_{m+p+1, D \operatorname{div} a^{n-(m+p)}}^k) = d_0 \dots d_{m+p} k$$

•  $k \neq d_{m+p+1}$

On a d'après (R-4)  $\begin{cases} \eta(C_{m+p+1, D \operatorname{div} a^{n-(m+p)}}^k) = T \\ D \xrightarrow{P} A_{m+p+1}^D \operatorname{div} a^{n-(m+p)} \end{cases}$

•  $k = d_{m+p+1}$

On a d'après (R-3)  $\begin{cases} \eta(C_{m+p+1, D \operatorname{div} a^{n-(m+p)}}^{d_{m+p+1}}) = E \\ D \xrightarrow{P} A_{m+p+2}^D \operatorname{div} a^{n-(m+p)} \cdot a + d_{m+p+1} \end{cases}$

Les cellules  $d_0 \dots d_{m+p}$  et  $d_0 \dots d_{m+p+1} = v \left( \binom{d_{m+p+1}}{m+p+1, D \operatorname{div} a^{n-(m+p)}} \right)$

appartient au même ANNEAU.

Donc  $d_0 \dots d_{m+p} \leftrightarrow d_0 \dots d_{m+p+1}$ , et puisque  $S \leftrightarrow d_0 \dots d_{m+p}$   
(hypothèse  $E_2(p)$ ).

On a donc  $S \leftrightarrow d_0 \dots d_{m+p+1}$

$$D \operatorname{div} a^{n-(m+p)} \cdot a + d_{m+p+1} = D \operatorname{div} a^{n-(m+p+1)}$$

$$\left\{ \begin{array}{l} S \leftrightarrow d_0 \dots d_{m+p+1} \\ \text{et} \\ D \xrightarrow{p} A_{m+p+2}^{D \operatorname{div} a^{n-(m+p+1)}} \end{array} \right. \Rightarrow E_2(p+1)$$

On obtient enfin pour  $p = n-m$

$$S \leftrightarrow d_0 \dots d_n = D.$$

c.q.f.d.

BIBLIOGRAPHIE

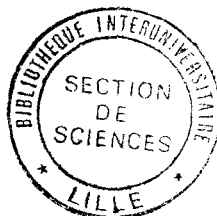
- [ACD 82] : J.M. AYACHE et al.  
*"Conception et validation de protocoles, application au réseau REBUS"*.  
T.S.I, Vol. 1, n° 4, 1982.
- [AUB ] : M. AUGUIN, F. BOERI  
*"Réseaux d'interconnexion et leur commande asynchrone"*.  
1<sup>ère</sup> partie du rapport de l'ATP *"Parallélisme, synchronisation, communication"*.  
LASSY E.R.A, 835 du C.N.R.S, Université de Nice.
- [BAR 68] : G.H. BARNES et al.  
*"The ILLIAC IV computer"*.  
I.T.E.C, Vol. C-17, n° 8, Aug. 1968.
- [BAR 81] : G.H. BARNES  
*"Design and validation of a connection network for many-processor multiprocessor systems"*.  
Computer, Vol. 14, n° 12, Déc. 1981.
- [BAT 80] : K. BATCHER  
*"Architecture of a massively parallel processor"*.  
The 7<sup>th</sup> ann. Symp. on Comp. Arch., pp. 168-173, 1980.
- [BEN 65] : V.E. BENES  
*"Mathematical theory of connecting networks and telephone traffic"*.  
New-York and London, Academic Press, 1965.
- [BCK 72] : L.H. BRANDENBURG et al.  
*"On the addressing problem of loop switching"*.  
The B.S.T.J., Vol. 51, n° 7, Sept. 72.
- [BLG 77] : E. BLOCH et D.J. GALAGE  
*"Component progress : It effect on high speed computer architecture and machine organization"*.  
Proc. of the symp. on high speed computer and Algorithm organization, April 13-15, 1977.

- [BRH 83] : G. BROOMELL et J.R. HEATH  
*"Classification categories and historical development of circuit switching topologies"*.  
Computing survey, Vol. 15, n° 2, June 1983.
- [COR 81] : CORNAFION  
*"Systèmes informatiques réparties"*.  
Editeur Dunod, 1981.
- [DLV 79] : P. DARVAS et al.  
*"A peripheral controller microprocessor for the PDP-11 series minicomputers"*.  
Proc. of the Symp. on Microcomputer and Microprocessor Application, Oct. 1979.
- [EST 83] : P. ESTRAILLER  
*"Modélisation par réseaux de Pétri d'un protocole d'ANNEAU virtuel"*.  
Publication interne n° 1, Jan. 1983, M.A.S.I. Univ. P. et M. CURIE, Institut de Programmation, Paris.
- [FLY 72] : M.J. FLYNN  
*"Some computer organization and their effectiveness"*.  
I.T.E.C, Vol. C-21, n° 9, Sept. 1972, pp. 948-960.
- [FRA 81] : M.A. FRANKLIN  
*"VLSI Performance comparison of Banyan and cross-bar communication networks"*.  
I.T.E.C, Vol. C-30, n° 4, April 1981.
- [FWT 82] : M.A. FRANKLIN et al.  
*"Pin limitation and partitionning VLSI interconnection networks"*.  
I.T.E.C, Vol. C-31, n° 11, Nov. 1982.

- [GKM 83] : A. GOTTLLEB et al.  
*"The NYU-Ultracomputer. Designing an MIMD shared memory parallel computer"*.  
I.T.E.C., Vol. C-23, n° 2, Feb. 1983.
- [GRP 71] : R.L. GRAHAM et H.O. POLLAK  
*"On the addressing problem for loop switching"*.  
The B.S.T.J., Vol. 50, n° 8, Oct. 1971.
- [HOR 77] : S. HOERNER et W. ROEHDER  
*"Efficiency of a multi-microprocessor system with time shared busses"*.  
Euromicro Symp., Amsterdam, Oct. 1977.
- [LAW 75] : D.K. LAWRIE  
*"Access and alignment of data in an array processor"*.  
I.T.E.C., Vol. C-24, pp. 1115-1145, Dec. 1975.
- [MAR 77] : H.G. MARTIN  
*"A discourse on a new super computer PEPE"*.  
Proc. of the symp. on high speed computer and algorithm organization, April 13-15, 1977, pp. 101-111.
- [MAS 81] : R.J. Mc MILLEN et al.  
*"Performance and implementation of 4 x 4 switching nodes in an interconnection network for PASM"*.  
Proc. of the 1981 Int. Conf. on Paral. Process., August 25-28, 1981, pp. 225-234.
- [MEC 80] : C. MEAD and L. CONWAY  
*"Introduction to VLSI systems"*.  
Addison-Wesley publishing company, 1980.
- [MIN 83] : MINIS et MICROS  
*"Le 68000 : Le bus d'adresse et de transfert des données"*.  
Minis et Micros, n° 196, Oct. 1983.

- [NAS 80] : D. NASSIMI et S. SAHNI  
*"A self routing Benes Network"*.  
 The 7<sup>th</sup> Ann. Symp. on Comp. Arch. 1980.
- [OPT 71] : D.C. OFFERMAN et N.T. TSAO-WU  
*"On a class of réarrangeable switching network. Part. I :  
 Control algorithm"*.  
 The B.S.T.J, Vol. 50, n° 1, May-June 1971.
- [PAT 81] : J.H. PATEL  
*"Performance of processor-memory interconnection for multi-  
 processors"*.  
 I.T.E.C, Vol. C-30, n° 10, Oct. 1981.
- [PIE 72] : J.R. PIERCE  
*"Network for block switching of data"*.  
 The B.S.T.J, Vol. 51, n° 6, July-August 1972.
- [SCG 83] : J.P. SECK et al.  
*"Contrôle de l'accès à un bus partagé : les arbitres"*.  
 R.A.I.R.O. Automatiques/Systems Analysis and Control, Vol. 17,  
 n° 4, 1983, pp. 359-403.
- [SIM 81] : H.J. SIEGEL et al.  
*"The multistage cube : a versatile interconnection network"*.  
 Computer, Vol. 14, n° 12, Dec. 1981.
- [SFS 77] : R.J. SWAN et al.  
*"C<sub>m</sub>\* a modular, multi-microprocessor"*.  
 APFIPS Conf. Proc, 1977, pp. 637-645.
- [SUL 77] : H. SULLIVAN  
*"A large scale, homogeneous, fully distributed parallel  
 machine"*.  
 The 4<sup>th</sup> Ann. Symp. on Comp. Arch., March 23-25, 1977.

- [TAH 79] : S. TAHE  
*"Transfert des données dans un calculateur parallèle vectoriel"*.  
Thèse de Docteur-Ingénieur, Université de Rennes I, Juil. 1979.
- [TBB 81] : D. TOMESCU et al.  
*"The circular synchronous bus"*.  
Microprocessing and microprogramming, Vol. 7, 1981, pp. 344-350.
- [WAF 83] : D.F. WANN et al.  
*"Asynchronous and clocked control structures for VLSI based interconnection networks"*.  
I.T.E.C, Vol. C-22, n° 3, March 1983.
- [WAH 84] : B.W. WAH  
*"A comparative study of distributed resource sharing or multiprocessors"*.  
I.T.E.C, Vol. C-33, n° 8, Aug. 1984.
- [WAK 68] : A. WAKSMAN  
*"A permutation network"*.  
J.A.C.M, Vol. 15, n° 1, pp. 159-163, 1968.
- [WUF 80] : C.L. WU et al.  
*"The reverse exchange interconnection network"*.  
I.T.E.C, Vol. C-29, n° 9, Sep. 1980.



## RESUME

Deux classes d'outils de communication ont été développées pour les systèmes fortement couplés : Les multiple-bus et les multi-étages. Chacune de ces deux classes opte pour l'une des deux solutions extrêmes, bus ou cross-bar, et tente d'en réduire les inconvénients.

Notre alternative opte pour la commutation de ligne bidirectionnelle et simultanée, et s'appuie sur une configuration en Anneau de commutateurs  $2 \times 2$  unidirectionnels. Cette configuration nous servira de modèle de base ; une mise en oeuvre asynchrone et répartie de son protocole d'interconnexion est proposée.

Nous poursuivons par l'étude des possibilités d'extension qu'offre le modèle de base. L'objectif étant de concevoir des outils capables de prendre en charge des communications simultanées. Deux outils, le "réseau Arbre", et le "multi-Anneau", sont alors proposés et leurs protocoles d'interconnexion entièrement définis. Une troisième extension : le "réseau cellulaire" montre les limites de notre modèle de base.

Enfin, l'évaluation de ces outils a permis la mise en évidence de leurs avantages par rapport à des outils existants. Ces avantages portent aussi bien sur leur fonctionnalité que sur leur complexité.

## MOTS CLES

Réseaux d'interconnexion, commutation de ligne, Anneau, contrôle asynchrone et réparti, complexité, performance.

