Institutional Repository - Research Portal
Dépôt Institutionnel - Portail de la Recherche
researchportal.unamur.be
University of Namur

THESIS / THÈSE

MASTER EN SCIENCES INFORMATIQUES

Choix d'une réorganisation d'un fichier pour un organisme bancaire

Jacobs, Jean-Louis

Award date:
1982

Awarding institution:
Universite de Namur

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 03. Jul. 2025

FACULTES UNIVERSITAIRES NOTRE-DAME DE LA PAIX (NAMUR).

INSTITUT D'INFORMATIQUE.

CHOIX D'UNE REORGANISATION D'UN FICHIER POUR UN ORGANISME BANCAIRE.

Mémoire présenté par

Jean-Louis JACOBS

en vue de l'obtention
du titre de
LICENCIE et MAITRE en INFORMATIQUE.

Année académique 1981-1982.

REMERCIEMENTS.

Je tiens tout particulièrement à remercier Monsieur Jean-Luc HAINAUT, promoteur de ce mémoire, dont la disponibilité et les conseils judicieux en ont permis la réalisation.

Je remercie également Messieurs les Responsables de la Caisse Générale d'Epargne et de Retraite, et notamment MM. DEMUYNCK, CARDINAEL, VERHAVERT et SERRUYS, sans l'aide desquels ce mémoire n'aurait pu être réalisé.

J'exprime, enfin, toute ma gratitude à ceux qui ont collaboré à l'édition de ce mémoire.

TABLE DES MATIERES.

	Pages
PARTIE I. - PRESENTATION GENERALE.	1 à 4
A. - La C.G.E.R.	1
B. - L'épargne à la C.G.E.R.	1
C. - Le client de la C.G.E.R. en tant qu'épargnant	2
D. - Le pourquoi de ce mémoire	3
E. - La présentation de l'étude effectuée	4
PARTIE II. - ETUDE DE L'EXISTANT.	5 à 22
A. - Historique du fichier C.C.D.	5
B. - Présentation des différents fichiers	6
C. - Etude des relations entre les différents fichiers	14
D. - Présentation des différentes applications	18
PARTIE III. - ANALYSE CONCEPTUELLE.	23 à 37
A. - Introduction	23
B. - Présentation du modèle choisi	24
C. - Présentation du schéma conceptuel	28
D. - Description des entités, relations	29
E. - Expression des contraintes d'intégrité	32
F. - Quantifications	33
G. - Description des applications	34
PARTIE IV. - ORGANISATION ACTUELLE.	38 à 48
A. - Introduction	38
B. - Présentation générale	39
C. - Organisation logique choisie du fichier CCD.	41
D. - Organisation physique choisie du fichier CCD.	42
E. - Applications spécifiques à la situation actuelle	43
F. - Statistiques obtenues sur la situation actuelle	45
PARTIE V. - LES PROBLEMES A RESOUDRE.	49 - 50
A. - Le Temps d'accès	49
B. - Le Format des enregistrements	50
PARTIE VI. - ETUDE DE DIFFERENTES TECHNIQUES.	51 à 69
A. - Introduction	51
B. - Le séquentiel indexé	52
C. - Virtual Storage Access Method (V.S.A.M.)	58
D. - L'accès calculé	64
E. - L'accès calculé avec hashage virtuel	68

Pages

PARTIE VII. - EVALUATION DES DIFFERENTES TECHNIQUES. 70 à 138

A. - Introduction 70
B. - Présentation des critères d'évaluation 70
C. - Présentation des simulations réalisées 72
D. - Evaluation des techniques 79

Solution actuelle 79

A) Evaluation selon les critères retenus 79
B) Evaluation des applications 81
C) Prise de copie triée 82

Le Séquentiel indexé 82

1. Présentation des différents éléments influençant l'évaluation 82
2. Equations analytiques 82
3. Equations de probabilité 84
4. Première répartition 85
   A) Evaluation selon les critères retenus 85
   B) Evaluation des applications 98
   C) Prise de copie triée 100
5. Deuxième répartition 100
6. Comparaison des deux types de répartition 104

Virtual Storage Access Method (V.S.A.M.) 104

1. Présentation des différents éléments influençant l'évaluation 104
2. Equations analytiques 105
3. Détermination de l'overflow 106
4. Première répartition 106
   A) Evaluation selon les critères retenus 106
   B) Evaluation des applications 116
   C) Prise de copie triée 117
5. Deuxième répartition 117
6. Comparaison des deux types de répartition 120

L'Accès Calculé.

1. Présentation des différents éléments influençant l'évaluation 120
2. Equations analytiques 120
3. Etude de la fonction de répartition 121
4. Première répartition 121
   A) Evaluation selon les critères retenus 122
   B) Evaluation des applications 132
   C) Prise de copie triée 133
5. Deuxième répartition 134
6. Comparaison des deux types de répartition 137

PARTIE VIII. - CONCLUSIONS. 139 à 143

Conclusions 139
Tableaux de synthèse 141

BIBLIOGRAPHIE.

PARTIE I. - PRESENTATION GENERALE.

A. - La C.G.E.R.

La Caisse Générale d'Epargne et de Retraite est un organisme financier belge, composé d'un siège central sis à BRUXELLES et de nombreuses agences réparties dans tout le pays.

Comme tout autre organisme du genre, elle travaille sur différents horizons, tels :
- des opérations de crédit, notamment lors de prêts hypothécaires;
- des opérations avec l'étranger;
- des opérations d'épargne, de placements;
- des opérations d'assurances (assurance-vie);
- etc...

Elle dispose d'un système informatique centralisé, doté d'un réseau de télétraitement reliant certaines agences du pays.

B. - L'épargne à la C.G.E.R.

L'étude présente a été réalisée au Centre de Traitement de l'Informatique (C.T.I.) à BRUXELLES et porte uniquement sur l'activité de l'épargne.

Epargner, c'est placer de l'argent pour qu'il en rapporte tout en restant disponible quasi immédiatement. C'est ainsi que la personne qui place son argent à la C.G.E.R. reçoit, à son nom, un livret qui l'identifie en portant un numéro différent des autres livrets. Ce livret peut être conservé à domicile mais il peut, également, être laissé en dépôt dans une des agences du pays.

Pour la C.G.E.R., un livret d'épargne représente un compte; nous parlerons donc, par la suite, indifféremment du livret du client ou de son compte.

Actuellement, le nombre de clients à l'épargne, à la C.G.E.R., s'élève à dix millions.

C. - Le client de la C.G.E.R. en tant qu'épargnant.

Le client peut, à la fois, vouloir épargner, souscrire une assurance-vie , effectuer des opérations avec l'étranger; mais c'est en tant qu'épargnant qu'il nous intéressera principalement dans cette étude.

Dans ce cas précis, le client peut :

a) ouvrir un compte; chaque année, quelque 400.000 nouveaux comptes font leur apparition dans les opérations de la C.G.E.R. Ce chiffre comprend tous les nouveaux clients mais également les anciens clients ouvrant un nouveau compte, auquel cas leur ancien compte sera mis en relation avec le nouveau.
Parmi les nouveaux clients, il y a beaucoup de nouveaux-nés; en effet, lors de la naissance de l'enfant, les parents reçoivent un livret au nom de leur fils (fille), livret au montant initial de 150 francs offerts par la C.G.E.R.

b) cesser son épargne; chaque année, ils sont 300.000 environ à cesser d'être épargnants. Cesser ne veut pas seulement dire avoir un montant nul sur le livret, car il est toujours possible de verser à nouveau de l'argent sur ce livret, mais il faut, en outre, que le livret soit rentré au siège central. Il faut donc que le client décide qu'il ne fera plus d'opérations avec ce livret.
Néanmoins, la C.G.E.R. garde une trace de ce compte durant trente ans.

c) consulter son compte pour en connaître la situation.
Par journée de 8 heures, on dénombre environ 5.000 consultations avec des pointes d'affluence entre 10 heures et 12 heures et entre 14 heures et 15 heures.
On remarque, pendant une période précise de l'année (de la fin janvier à la mi-avril), une consultation particulière : c'est durant cette période que le client vient vérifier si les intérêts auxquels il a droit en tant qu'épargnant, ont été versés sur son livret.
Normalement, les intérêts sont portés en compte dès la mi-janvier aussitôt que le client a effectué une opération financière quelconque via ce livret.

d) effectuer des opérations financières et comptables.

C'est le fait de retirer de l'argent de son compte ou bien celui d'en verser sur ce compte.

On estime à environ 40.000, les opérations financières quotidiennes reprises sur bordereaux et à 4.000, le nombre des opérations comptables mensuelles.

Une opération comptable est celle qui est effectuée directement à partir d'une agence possédant un terminal relié au Centre de BRUXELLES, et disposant du livret du client pour pouvoir en effectuer la mise à jour immédiate.

e) effectuer des opérations non comptables.

Ce sont des opérations sans influence sur le capital, tel le fait de signaler un changement d'adresse, une modification du libellé, etc...

On les évalue entre 500.000 à 1 million par an.

REMARQUE.

Après la présentation de ces opérations, nous pouvons déjà distinguer différents types de comptes - que nous retrouverons dans la présentation de la solution actuelle en partie 4. - à savoir :

- le compte soldé : il s'agit d'un compte au montant nul.
- le compte flaggé : l'ancien compte d'un épargnant est mis en relation avec son nouveau compte.
- le compte actif : c'est celui avec lequel le client effectue différentes opérations

Il existe toutefois un autre type de compte, appelé "compte à terme" qui permet, au client de la C.G.E.R., de placer de l'argent pour un terme déterminé. Ce placement produisant des intérêts, sera également considéré comme compte d'épargne.

D. - Le pourquoi de ce mémoire.

La gestion de l'épargne étant informatisée, tous les comptes des clients sont repris dans un fichier central localisé à

BRUXELLES et c'est donc sur ce fichier que toutes les opérations d'épargne du pays seront répercutées.

La gestion de ce fichier, appelé C.C.D. (comptes clients disques) et l'organisation de celui-ci sont très importantes, par le fait que l'épargne, qui n'est pas la seule activité de la C.G.E.R. à être informatisée, doit forcément s'intégrer dans un ensemble informatique.

C'est pourquoi l'organisation actuelle ne semble plus satisfaire la C.G.E.R. et il m'a été demandé d'étudier des solutions de remplacement envisageables.

E. - La présentation de l'étude effectuée.

En vue d'atteindre le but proposé, on étudiera, dans la deuxième partie du travail, l'existant en termes de fichiers, de relations entre eux et d'applications.

Il en sera fait une étude conceptuelle dans la troisième partie; on développera, dans la quatrième partie du travail, l'organisation actuelle, laquelle sera documentée à l'aide de statistiques.

Dans la partie cinq, les problèmes à résoudre seront exposés. Ensuite, en partie six, on présentera des propositions de solutions qui seront évaluées dans la septième partie du travail, avant de conclure notre étude en huitième partie.

=========

PARTIE II. - ETUDE DE L'EXISTANT.

PLAN

A. - Historique du fichier C.C.D.

B. - Présentation des différents fichiers.
- . fichier C.C.D.
- . fichier historique
- . fichier "termes"
- . fichier Autres Sommes
- . fichier Mouvements
- . fichier Litiges
- . fichier Bordereaux
- . fichiers signalétiques.

C. - Etude des relations entre les différents fichiers.

D. - Présentation des différentes applications.
- . consultation
- . mises à jour
- . opérations comptables et non comptables
- . capitalisation (et recapitalisation)

PARTIE II. - ETUDE DE L'EXISTANT.

A. - Historique du fichier C.C.D.

Avant de parler du développement des différents fichiers, nous pensons qu'il est intéressant de présenter un bref historique de l'évolution du C.C.D. (appelé encore fichier d'épargne).

Ce fichier reprend, comme cité dans la Ière partie du travail, les différents comptes des clients. Un compte est identifié par un numéro. Ce numéro est, notamment, composé d'au moins neuf chiffres décimaux - dont les deux premiers constituent ce que l'on appellera une série - et de caractères de contrôle.

Au début de l'existence du fichier, on attribuait des livrets par province, que l'on distinguait par des numéros de séries différents. Déjà, à cette époque, la C.G.E.R. divisait le BRABANT en deux parties, l'une francophone et l'autre néerlandophone. D'autres numéros de séries étaient attribués pour le service de la Poste et pour les comptes à terme. En effet, comme dit ci-avant, ces comptes produisant des intérêts, se retrouvent également dans le C.C.D.

Dès l'apparition des agences, vers 1958-1959, on attribua, à chacune d'elles, un numéro de série différent. A cette époque, il existait quatre agences : ATH - EISDEN - GAND - NAMUR.

Très vite cependant, les agences se multiplièrent et il fut décidé alors d'allouer un numéro de série pour toutes les agences francophones, un autre pour les agences néerlandophones et un autre également pour les agences du BRABANT.

Actuellement, une nouvelle distinction apparaît dans ce fichier : tous les nouveaux livrets francophones se retrouvent dans une même série et les livrets néerlandophones dans une autre même série.

On dénombre ainsi, dans le fichier actuel, 38 séries différentes.

B. - Présentation des différents fichiers.

Pour chaque fichier, il sera donné, dans la mesure du possible :

. un identifiant;

. certains types de données;

. sa description générale avec, entre autres :

- sa fonction
- le nombre de fois et le moment où il est possible d'y accéder;

Pour certains fichiers :

. une évaluation du nombre d'articles (enregistrements).

FICHIER C.C.D.

1. Identifiant.

C'est le numéro du livret, composé de neuf chiffres décimaux.

2. Types de données.

Bureau d'émission du livret, année de son émission, montant des intérêts bruts ou des intérêts nets, langue d'expression, etc... (type de livret)

NB. - Une liste détaillée sera donnée dans l'étude de la solution actuelle, en partie IV. du présent travail.

3. Description.

Ce fichier reprend tous les livrets d'épargne (c'est-à-dire les comptes) des clients.

Il comporte les différents types de comptes (actifs, soldés, flaggés, à terme) plus un autre type, un compte blanc qui ne possède, comme renseignements, que le numéro, le bureau et l'année d'émission. Nous en verrons son intérêt dans l'étude de la solution actuelle - (partie IV.)

Il ne comprend plus les comptes stoppés.
On accède à ce fichier lors de chaque consultation du client ainsi qu'à l'occasion de toutes ses opérations financières, aussi bien le jour que la nuit, pour les opérations de mise à jour des comptes.

La liste des différentes applications relatives à ce fichier est détaillée dans le point suivant.

4. Nombre d'articles dans le fichier.

On dénombre environ 10.200.000 articles dans ce fichier.

FICHIER HISTORIQUE.

1. Identifiant.

Il s'agit du numéro du livret.

2. Types de données.

---

3. Description.

Lorsqu'un compte est stoppé (solde nul et rentré au siège central) la C.G.E.R. doit le retirer du C.C.D. et, pour satisfaire à la législation en la matière, assurer sa conservation durant trente ans. Ce compte sera gardé dans le fichier historique.

L'ensemble de tous ces comptes stoppés est transféré, une fois par année, dans le fichier historique.

FICHIER "TERMES".

1. Identifiant.

C'est toujours le numéro du livret.

2. Type de données.

Montant du capital.

3. Description.

Ce fichier comprend tous les comptes à terme.

A un compte à terme, correspond un capital. Celui-ci peut produire des intérêts. En ce sens, il intéresse le fichier CCD dans lequel on retrouvera tous les comptes à terme (dans la série 50).

En fait, pour atteindre un compte à terme, il faut passer par le CCD, l'application "terme" ayant été développée après celle

de l'épargne et ayant été basée sur l'existence du fichier CCD. Il faut compter aux environs de 200.000 opérations par an sur le fichier "termes". De plus, chaque mois, tous les comptes à terme arrivant à échéance, sont passés comme mise à jour des comptes termes du fichier CCD, afin de produire les intérêts. On en compte de 80.000 à 90.000 par mois.

4. Nombre d'articles.
Ils sont estimés à environ 800.000 pour ce fichier.

FICHIER "AUTRES SOMMES".

1. Identifiant.
Il s'agit du numéro du livret.

2. Types de données.
. nombre d'opérations effectuées;
. différence de montant entre le livret et le compte.

3. Description.
Lorsqu'un client exécute une opération à partir d'une agence ne disposant pas de terminal relié à l'ordinateur central, la mise à jour du livret peut se faire immédiatement si le client vient avec son livret ou si celui-ci se trouve en dépôt à l'agence.

Par contre, la mise à jour du fichier central se fera plus tard lorsque la modification sera parvenue par la poste, au siège central.

Donc, la situation d'un compte client n'est pas toujours le reflet de la situation réelle du livret.

De même, lors d'un versement effectué au profit du client, le montant est porté à son compte mais ne figure pas immédiatement sur son livret, ce qui donne également une différence entre le livret et le fichier.

Lors de la mise à jour du compte, si l'on constate une différence entre le livret et le compte, la consultation du fichier "Autres Comptes" pourra, peut-être, expliquer cette différence. S'il n'y a plus de problème de concordance, le cas est solutionné; s'il en

subsiste, l'on se trouve dans un cas de litige.

REMARQUE.

Lorsqu'il faut procéder à la mise à jour du livret d'un client, l'agence qui a effectué la dernière opération comptable de ce client, reçoit, du siège central, un avis de mise à jour du livret.

L'agence effectue la mise à jour immédiatement si elle détient ce livret en dépôt et renvoie ensuite l'avis à BRUXELLES. Si elle n'a pas la garde du livret, elle se charge de contacter le client pour l'inviter à se présenter à l'agence, muni du livret, de façon à pouvoir effectuer la mise à jour nécessaire. L'avis de la mise à jour sera renvoyé à BRUXELLES dès que l'opération aura pu être effectuée.

L'avis de mise à jour étant parvenu à BRUXELLES, il peut alors être procédé à la suppression de l'enregistrement correspondant du fichier "autres sommes".

FICHIER "MOUVEMENTS".

1. Identifiant.

Il s'agit du numéro du livret.

2. Types de données.

Numéro des bordereaux, date, nom de l'agence, ...

3. Description.

Ce fichier reprend tous les bordereaux envoyés par les agences ne disposant pas d'un terminal. Ces bordereaux sont encodés au siège central.

Après une première vérification, ce fichier reprend, date après date et agence par agence, ces bordereaux qui sont constitués, en fait, de tous les mouvements des clients effectués à ce type d'agences. Le bordereau sera décrit dans la suite du travail.

Dans ce fichier "Mouvements", passent également les opérations relatives aux comptes à terme (fréquence : une fois par mois).

4. Nombre d'articles.

Ils sont estimés entre 40.000 à 60.000, avec une pointe, par mois, allant de 100.000 à 150.000.

FICHIER "LITIGES".

1. Identification.

Le numéro du livret.

2. Types de données.

---

3. Description.

La nuit, lors de la mise à jour des comptes, des erreurs peuvent être détectées. On ne peut, dès lors, effectuer la modification au fichier ; dans ce cas, l'on place tous les mouvements du bordereau, parmi lesquels se trouve nécessairement le mouvement erroné, dans le fichier "litiges".

Pourquoi tous ces mouvements ? En vue de faciliter la correction par le service qui traitera les erreurs.

Il faut compter entre 6 et 7% du total des mouvements effectués la nuit (soit entre 40.000 à 60.000) avec des pointes de 100.000 à 150.000, lors des mouvements des comptes à terme.

4. Nombre d'articles.

De 6 à 7% du nombre total des mouvements de nuit.

FICHIER "BORDEREAUX".

Remarque. - Contrairement aux autres fichiers décrits jusqu'à présent, il ne s'agit pas d'un fichier que l'on retrouvera sur bande ou disques mais d'un ensemble de fiches reçues des différentes agences.

1. Identifiant.

Le numéro du bordereau et la date à laquelle il a été établi.

2. Types de données.

Le total des sommes versées - le total des sommes retirées - le numéro ce compte (livret) - le solde.

3. Description.

Ces fiches reprennent, jour après jour, et agence par agence, l'ensemble des opérations comptables effectuées par les clients. Etant donné que ces agences ne possèdent pas de terminal, elles signalent les mouvements effectués chez elles par des bordereaux transmis au siège central où chaque mouvement sera encodé dans le fichier "Mouvements".

4. Nombre d'enregistrements.

Pour tous les bordereaux, ces enregistrements constituent une grande partie des 40.000 à 60.000 mouvements.

FICHIER SIGNALETIQUE.

Ce fichier sera décrit d'une manière différente car :

. il est commun à toutes les applications de la C.G.E.R. donc, aussi bien pour la partie "EPARGNE" que pour celles des "COMPTES à TERME" - "EMPRUNTS" - "ASSURANCE-VIE", ...

. il est constitué d'un ensemble de fichiers signalétiques.

C'est ainsi que l'on distingue :

@ un fichier signalétique principal (S.P.)

a) Reprenant un enregistrement par client, exception faite pour les clients qui sont des sociétés à dénomination bilingue, lesquels sont repris en deux enregistrements reliés entre eux (un enregistrement dans chacune des deux langues utilisées par la société).

b) Que peut-on trouver dans les enregistrements de ce fichier ?
- des renseignements signalétiques du client;
- des liens entre deux conjoints, si ceux-ci possèdent un compte en commun;
- des liens entre les deux enregistrements d'une société à dénomination bilingue;
- des liens avec d'autres fichiers.

c) Il existe trois possibilités d'accès à ce fichier :

- soit par les fichiers comptables (un fichier comptable reprend les informations propres à l'application qu'il couvre); parmi eux, se trouve le fichier CCD.
  Pour l'épargne, on sait qu'à un compte correspond un client.
- soit par les nom, prénoms (initiales) et adresse du client.
- soit par le numéro du client au registre national (solution non encore adoptée actuellement).

• un fichier signalétique des numéros (S.N.)

a) Reprenant un enregistrement pour tout client disposant de plus de quatre comptes à la C.G.E.R. Celle-ci lui offre ainsi la possibilité d'avoir huit nouveaux comptes. Si cela s'avère insuffisant, on établira un deuxième enregistrement pour ce client, et ainsi de suite.

b) Que peut-on trouver dans les enregistrements de ce fichier ?
- des liens entre deux enregistrements de S.N.
- des renseignements permettant de retrouver les différents comptes d'un client.

c) Il existe une seule possibilité d'accès à ce fichier, à partir du fichier S.P.

• un fichier signalétique adresse (S.A.) et pour des adresses diverses (S.D.) qui reprend tout ce qui n'est pas conforme à un enregistrement de S.A.

a) Reprenant un enregistrement par client dans S.A. et, éventuellement, un dans S.D.

b) Que peut-on trouver dans l'enregistrement de ces fichiers ?
- des renseignements concernant l'adresse du client pour S.A.
- des renseignements spéciaux concernant le client pour S.D.

c) Il existe deux possibilités d'accéder à ce fichier :
   - à partir du fichier S.P.
   - à partir des fichiers comptables.

- un fichier signalétique mandataire (S.M.)

a) Reprenant des renseignements sur les mandataires, au sujet des comptes ouverts au nom de plusieurs personnes ou de société(s).

b) Que peut-on trouver dans les enregistrements de ce fichier ?
   - des liens avec le fichier S.A.

c) Il existe deux possibilités d'accès à ce fichier :
   - à partir du fichier S.P.
   - à partir des fichiers comptables.

Il faut, enfin, citer :

- un fichier signalétique répertoire (R.P.) qui permet le regroupement de toutes les applications par personne et qui offre une des trois possibilités d'accès à S.P.

- un fichier signalétique numéro national (N.N.)

C. - Etude des relations entre les différents fichiers.

Pour chaque relation, nous présenterons :

1. Le type de relation. Nous prendrons connaissance, ici, des différentes possibilités de relations entre les enregistrements des fichiers concernés.

   Une relation du type 1 - 1 entre deux fichiers (A et B) signifie qu'à un enregistrement du fichier A correspond au maximum un enregistrement du fichier B et vice-versa.

   Une relation du type 1 - N entre deux fichiers (A et B) signifie qu'à un enregistrement du fichier A peut correspondre plus d'un enregistrement du fichier B mais qu'à un enregistrement du fichier B ne correspond qu'un enregistrement du fichier A.

   Une relation du type N - 1 entre deux fichiers (A et B) signifie qu'à un enregistrement du fichier A correspond, au plus, un enregistrement du fichier B et qu'à un enregistrement du fichier B peut correspondre plus d'un enregistrement du fichier A.

   Une relation du type N - M entre deux fichiers (A et B) signifie qu'un enregistrement du fichier A peut être mis en relation avec plusieurs enregistrements du fichier B et qu'un enregistrement du fichier B peut être mis en relation avec plusieurs enregistrements du fichier A.

2. Une notice explicative de cette relation.

\+ + +

Relation CCD - Historique.

1. Type 1 - 1

2. Annuellement, après capitalisation (qui est une application de gestion de l'épargne à la C.G.E.R.), tous les comptes soldés et pour lesquels les livrets sont stoppés, sont stockés dans le fichier "archives" (historique). Chaque compte donne lieu à un enregistrement dans le fichier historique.
   Ces comptes seront conservés durant trente ans, conformément aux dispositions légales régissant cette matière.

Relation CCD - TERME.

1. Type 1 - 1 pour les comptes de la série 50 (°) uniquement.

2. L'application terme est une autre activité de la C.G.E.R. mais, comme déjà cité, les comptes à terme peuvent produire des intérêts; de ce fait, ils se retrouvent dans le fichier CCD.

   Chaque fois que l'on désire accéder aux enregistrements relatifs aux comptes à terme, il faut, au préalable, accéder aux enregistrements relatifs aux comptes à terme dans le fichier CCD où l'on trouvera un lien pour permettre l'accès au fichier terme.

Relation CCD - Mouvements.

1. Type 1 - N

2. Chaque nuit, on déroule le fichier des mouvements et on effectue les mises à jour des comptes sur le fichier CCD, un même compte pouvant subir plusieurs mouvements durant la mise à jour du fichier CCD.
   En cas de virement de compte à compte, il y aura deux opérations au fichier car, à un mouvement, correspond un compte.

Relation CCD - Litiges.

1. Type 1 - 1

2. Tous les matins, les services comptables traitent les cas litigieux et tentent ainsi de rétablir une situation égale sur les comptes du fichier CCD.
   Les litiges proviennent de la mise à jour des comptes qui s'est déroulée la nuit. - (Rappel : de 6 à 7% de litiges pour un chiffre de 40.000 à 60.000 mouvements).

(°) La série 50 est celle qui reprend les comptes à terme.

Relation Mouvements - Litiges.

1. Type 1 - N

2. S'il y a un mouvement erroné, c'est tout le bordereau (c'est-à-dire un ensemble de mouvements pour une date et une agence données) reprenant le mouvement erroné qui est transféré dans le fichier litiges.
   Afin de faciliter la correction comptable, une remarque signale, pour chaque mouvement du bordereau, s'il a été ou non exécuté.

Relation CCD - "Autres sommes".

1. Type 1 - 1

2. Lorsque le montant du livret diffère de celui du compte sur le fichier CCD, après la mise à jour des comptes, on essaie de corriger cette différence en vérifiant s'il n'y a pas, dans le fichier "autres sommes", une opération pour ce numéro de compte.
   Si l'on ne peut rétablir l'égalité, on se trouve alors devant un cas de litige.

Relation CCD - Signalétiques.

Le fichier d'épargne CCD, comme tout autre fichier comptable de la C.G.E.R., peut accéder au signalétique principal (S.P.), aux signalétiques adresse (S.A. ou S.D.) et au signalétique mandataire (S.M.)

CCD - S.P.

1. Type N - M

2. Le cas le plus fréquent est celui où un enregistrement du fichier CCD relate un seul enregistrement du fichier S.P.
   Cependant, il peut se présenter les cas suivants :
   - si la société est à dénomination bilingue, l'enregistrement du fichier CCD relatera deux enregistrements du fichier S.P.
   - si le client possède plusieurs comptes, ils ne seront en relation qu'avec un seul enregistrement du fichier S.P.

CCD - S.A.

1. Type N - 1

2. L'adresse du client n'est reprise qu'une seule fois, même si le client détient plusieurs livrets.

CCD - S.D.

1. Type N - 1

2. Un client n'a un enregistrement dans S.D. que si son adresse ne peut se mettre sous la forme d'un enregistrement de S.A.

CCD - S.M.

1. Type 1 - 1

2. Lorsqu'un client a des mandataires, il existe un enregistrement dans S.M. en liaison avec celui du client dans le fichier comptable.

D. - Présentation des différentes applications.

Pour chaque application, nous trouverons :

1. la description de cette application;
2. sa quantification.

REMARQUE.

Ce paragraphe va détailler, de manière plus précise, les applications propres à la partie EPARGNE de la C.G.E.R. Plus rigoureusement, les applications propres au fichier CCD.

\+ + +

CONSULTATION.

1. Description.

Une consultation ne peut se faire que dans une agence possédant un terminal relié au centre informatique de BRUXELLES.

Lorsque le client se présente à une agence de ce type, il décline, s'il le connaît, son numéro de livret; à défaut, il est possible de retrouver son numéro à partir de ses nom, prénoms (initiales) et date de naissance.

Etant donné que, dans tous les cas, on connaît son numéro de livret, on sait lui donner la situation de son compte (ou, par exemple, vérifier si les intérêts ont été versés sur son compte).

De plus, si le livret est en dépôt dans cette agence ou si le client se présente à l'agence, muni de son livret, il peut prendre connaissance du montant qui y est inscrit.

2. Quantification.

Il faut compter environ 5.000 consultations par journée de 8 heures avec des pointes entre 10 heures et 12 heures, et entre 14 heures et 15 heures.

MISES A JOUR.

On distingue trois types de mise à jour :

- . la mise à jour normale
- . la mise à jour des rejets ou litiges
- . la mise à jour des intérêts.

Mise à jour NORMALE.

1. Description.

Elle reprend :

- . tous les mouvements des bordereaux;
- . parfois, les mouvements relatifs aux comptes à terme (une fois par mois);
- . toute autre application de mise à jour, telle que des corrections automatiques de litiges, lorsque cela est possible.

Il faut savoir cependant que les mouvements d'un bordereau passés un jour J seront mis à jour au jour J + D (D étant égal au délai de la poste + celui de l'encodage).

Comment s'effectue la mise à jour normale ?

Cela se passe en deux étapes :

1°) après l'encodage des bordereaux, il y a un premier contrôle informatique sur des renseignements globaux du bordereau (contrôle date, totaux). Une erreur, à ce niveau, rejette le bordereau dans les litiges.

2°) on effectue alors la mise à jour des comptes. Je rappelle qu'un mouvement, dans le bordereau, est relatif à un compte. Une erreur, à ce niveau, rejette tout le bordereau en litiges (avec une indication pour chaque mouvement du bordereau mentionnant s'il a été exécuté ou non), excepté si l'on sait corriger l'erreur grâce au fichier "autres sommes" (Cfr. fichier "autres sommes").

2. Quantification.

Par nuit, on peut compter entre 40.000 et 60.000 mouvements. Certaines nuits accusent des pointes allant de 100.000 à 150.000 (lors de la mise à jour des comptes à terme).

Mise à jour des REJETS (LITIGES).

1. Description.

La mise à jour des mouvements provoque des rejets (litiges). Ces différents cas sont traités, un à un, pour permettre la correction des comptes du fichier CCD.

Cette mise à jour s'effectue dans la matinée.
Il ne faut pas perdre de vue que la mise à jour a déjà été effectuée sur le livret.

2. Quantification.

De 6 à 7% des mouvements, c'est-à-dire de la mise à jour normale effectuée la nuit.

Mise à jour des INTERETS.

1. Description.

L'objectif du client, quand il épargne, est de percevoir des intérêts. Il est donc normal qu'il s'informe pour savoir si les intérêts de son argent lui ont été versés.

Le versement des intérêts s'effectue après la capitalisation.

Il y a deux types de mise à jour des intérêts :

a) la mise à jour automatique : si, au moment où le client exécute une opération comptable, il est constaté que les intérêts n'ont pas été portés en compte, la mise à jour se fait en même temps.

b) la mise à jour demandée : le client peut consulter son compte pour savoir si les intérêts ont été comptabilisés. Dans la négative, il demande que la mise à jour soit effectuée à ce moment-là.

Il faut encore distinguer le cas où le client demande les intérêts à partir d'une agence qui possède ou non un terminal. Si elle n'en dispose pas, on se retrouve dans le cas d'une mise à jour normale; par contre, si elle possède un terminal, on aura une opération comptable supplémentaire.

2. Quantification.

Tous les comptes actifs produisent des intérêts.

Du type (b) : 4.000.000 sur trois mois dont 25% dans les agences reliées par terminal et 75% qui se feront plus tard.

OPERATIONS COMPTABLES ET NON COMPTABLES.

1. Description.

Une opération comptable est celle qui a une influence sur le capital et/ou sur les intérêts de celui-ci.

Une opération non comptable n'a pas d'influence, ni sur le capital, ni sur les intérêts de celui-ci.

Ces opérations ont souvent lieu à partir des agences disposant d'un terminal.

2. Quantification.

Opérations comptables : environ 4.000 par mois.

Opérations non comptables : il est plus difficile de les évaluer (entre 500.000 et un million par an).

CAPITALISATION ET RECAPITALISATION.

1. Description.

Il s'agit d'une application strictement comptable mais qui, en quelque sorte, est une mise à jour de tous les comptes, se faisant à une période donnée (première quinzaine de décembre).

Pour tous les comptes, on effectue les capitalisation en vue de déterminer leurs intérêts.

D'autre part, il faut signaler que, du 15 décembre jusqu'à la fin de l'année, d'autres opérations peuvent avoir lieu.
De même, les livrets d'épargne pour les écoles disposent d'un régime spécial en ce qui concerne les intérêts qu'ils produisent.
Toute ceci a une répercussion sur les comptes concernés, lesquels doivent être recapitalisés.

Cette nouvelle mise à jour de comptes passera cependant comme une mise à jour normale et, par conséquent, le nombre d'opérations relatif à cette mise à jour s'ajoutera au total des mises à jour normales.

2. Quantification.

Capitalisation : tous les comptes actifs.

Recapitalisation : 10% environ du total des comptes.

Les autres applications sont propres à l'organisation actuelle et seront développées dans une partie ultérieure, à savoir :

- prise de copie triée.
- réorganisation.

PARTIE III. - ANALYSE CONCEPTUELLE.

PLAN.

A. - Introduction.

B. - Présentation du modèle choisi. (E-R)

C. - Présentation du schéma conceptuel.

D. - Description des entités, relations.

E. - Expression des contraintes d'intégrité.

F. - Quantifications.

G. - Description des applications.

PARTIE III. - ANALYSE CONCEPTUELLE.

A. - Introduction.

L'ensemble des conversations et des documents ayant servi à établir l'analyse conceptuelle ne seront pas reproduits. Ils ne seraient, d'ailleurs, que de peu d'intérêt pour le lecteur.

Pour créer le schéma conceptuel que nous présenterons à l'aide du modèle choisi et présenté ci-après, nous nous sommes basés sur les deux premières parties de cette étude.

Nous décrirons sémantiquement les différents composants de notre schéma avant de rappeler certaines quantifications.

Nous terminerons la partie III. de notre travail par l'analyse des différentes applications retenues.

B. - Présentation du modèle choisi.

LE MODELE D'ANALYSE CONCEPTUELLE.

Pour exprimer l'analyse conceptuelle, nous avons retenu le modèle binaire, limité aux concepts suivants : les entités, les relations, les propriétés et contraintes d'intégrité.

Dans un premier temps, nous présenterons ces concepts et nous préciserons, ensuite, les formalismes utilisés pour exprimer notre analyse.

1. Les concepts.

La structure d'une information sera représentée par un modèle élaboré à partir des concepts :

entité	relation	propriété
(type d'entité)	(type de relation)	

a) Entité.

C'est un objet concret ou abstrait du monde réel. L'entité n'existe en tant que telle que par rapport à un individu (ou un groupe d'individus) qui la considère comme un tout et lui confère une existence autonome.

Une entité est caractérisée par un ensemble de propriétés.
(Exemple : un client de la C.G.E.R. en tant qu'épargnant)

b) Type d'entité.

C'est une famille d'entités que l'individu (ou le groupe d'individus) choisit de les concrétiser toutes par le même ensemble de propriétés. (Exemple : CLIENT)

c) Nous déduisons :

Si le client C appartient au type d'entité CLIENT, alors le client C est une réalisation de l'entité CLIENT.

d) Relation.

C'est un ensemble de deux entités dans lequel chacune d'elles assume un rôle déterminé.

Une relation peut posséder des propriétés; elle est contingente à l'existence des entités qu'elle relate.
(Exemple : la relation Possède entre l'entité CLIENT
et l'entité LIVRET D'EPARGNE.)

e) Type de relation.

C'est une famille de relations caractérisées par les mêmes propriétés.
Exemple : POSSEDE (CLIENT - LIVRET D'EPARGNE)

f) Nous déduisons :

Si C est une réalisation de CLIENT,
L est une réalisation de LIVRET D'EPARGNE ,
(C,L) est une réalisation du type POSSEDE,

alors (C, L) est une occurrence de la relation POSSEDE.

g) Propriété.

Une propriété d'une entité ou d'une relation représente une qualité qu'un individu (ou un groupe d'individus) attribue à cette entité ou à cette relation.

L'existence d'une propriété est contingente à l'existence des entités ou de la relation concernée.
(Exemple : numéro du livret d'épargne)

Nous pouvons classer les propriétés comme suit :
. celles qui identifient ou non la relation, l'entité;
. les propriétés obligatoires;
. les propriétés facultatives;
. les propriétés répétitives.

h) Contrainte d'intégrité.

Une contrainte d'intégrité est un prédicat relatif aux données et défini sur un élément du schéma conceptuel qui doit toujours être vrai.

Certaines contraintes seront exprimées à l'aide du modèle tandis que d'autres seront exprimées en français.

Description des contraintes représentées dans le schéma.

1. Contrainte de cardinalité.

Entre deux types d'entités (A et B), nous distinguerons des relations du type :

1 - 1 une occurrence du type d'entité A ne peut être en relation qu'avec une seule entité du type B et réciproquement.

1 - N une occurrence du type A peut être en relation avec plusieurs entités du type B mais une occurrence du type B ne sera en relation qu'avec une seule entité du type A.

N - 1 une occurrence du type B peut être en relation avec plusieurs entités du type A mais une occurrence du type A ne sera en relation qu'avec une seule entité du type B.

N - M une occurrence du type A peut être en relation avec plusieurs entités du type B et une occurrence du type B peut être en relation avec plusieurs entités du type A.

2. Contrainte d'existence.

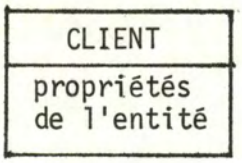
Un type de relation est fort pour un type d'entité si une entité de celui-ci ne peut exister que si elle appartient à au moins une relation de ce type.

Un type de relation est faible pour un type d'entité si l'existence d'une entité de ce type est indépendante de la relation.

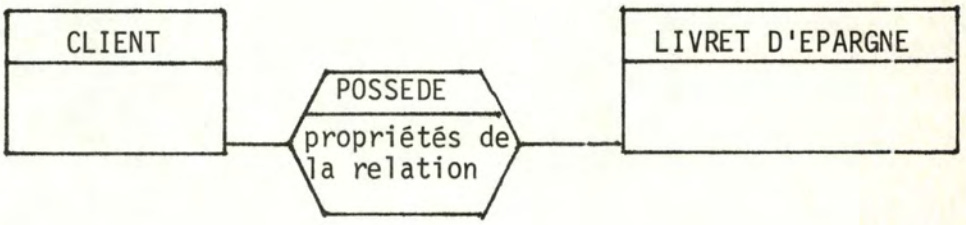
.....

2. Représentation graphique.

Type d'entité



Type de relation



Expression sémantique des entités

Nous décrirons les types d'entité du schéma en prenant une occurrence de chaque type et en donnant sa signification ainsi que celle de ses propriétés.

Expression sémantique et contrainte d'existence et de cardinalité des relations.

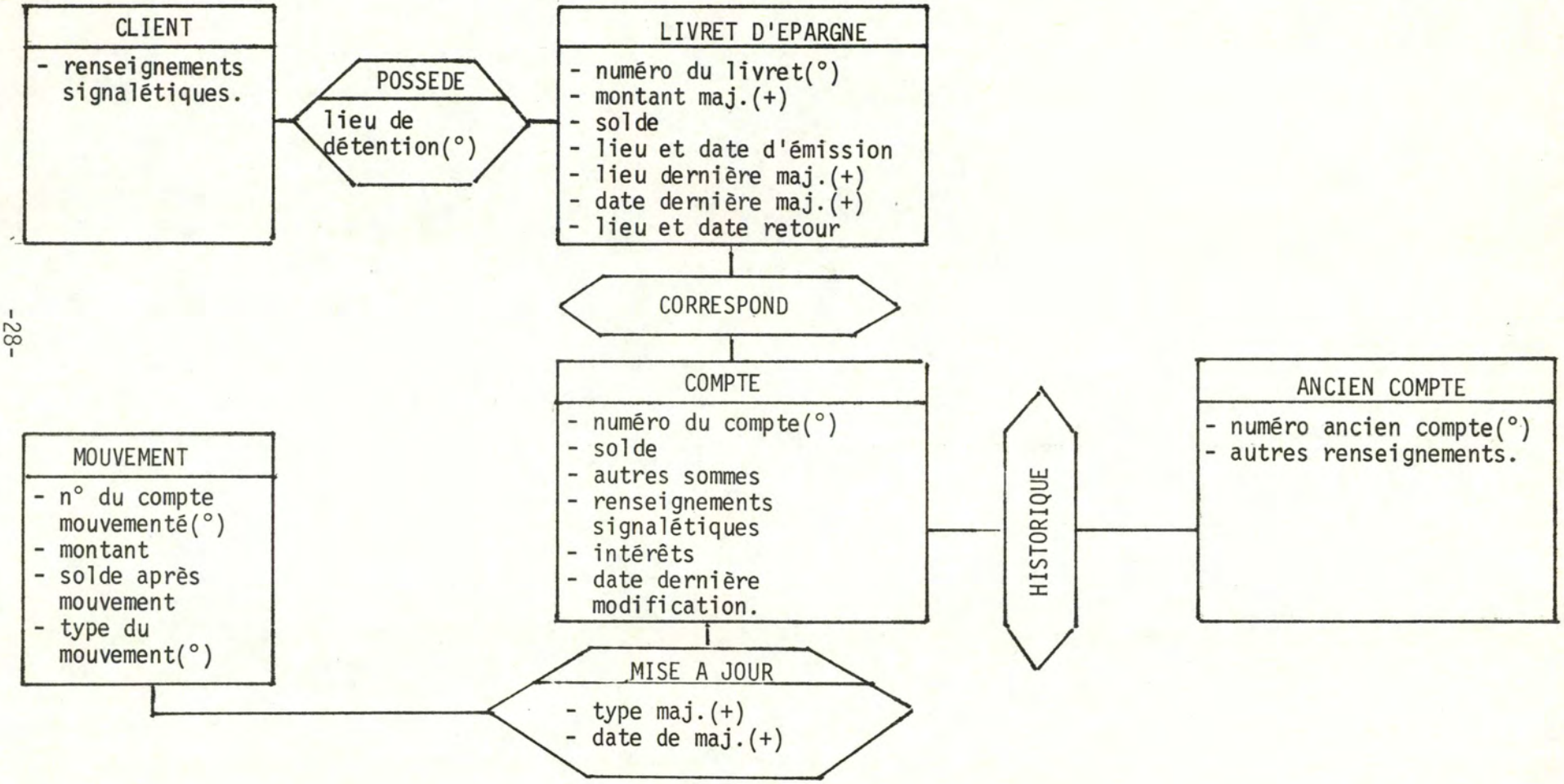
Un type de relation sera représenté par le nom entre parenthèses des types d'entités le composant.

Nous décrirons une occurrence de chaque type de relation en donnant sa signification ainsi que celle de ses propriétés.

Nous exprimerons les contraintes de cardinalité sous la forme décrite plus haut (1-1) (1-N) (N-1) (N-M).

Les contraintes d'existence seront représentées dans l'énoncé de la relation en soulignant le(s) type(s) d'entité(s) pour lequel(lesquels) la relation est forte.

.....



(°) contrainte d'intégrité.

(+) mise à jour

D. - Description des entités, relations.

DESCRIPTION SEMANTIQUE DES ENTITES.

CLIENT

Si C appartient à CLIENT, alors C désigne un client connu de la C.G.E.R. dont les renseignements signalétiques sont représentés par (renseignements signalétiques).

LIVRET D'EPARGNE

Si L appartient à LIVRET D'EPARGNE, alors L désigne un livret d'épargne d'un client C dont :

- le numéro est représenté par (numéro du livret)
- le lieu et la date d'émission sont représentés par (lieu et date d'émission)
- la valeur du dernier montant versé ou retiré par le client C est représentée par (montant maj.)
- le lieu de la dernière opération est représenté par (lieu dernière maj.)
- la date de cette dernière opération est représentée par (date dernière maj.)
- le nouveau solde après la dernière opération est représenté par (solde)
- le lieu et la date de retour sont représentés par (lieu et date de retour).

COMPTE

Si CO appartient à COMPTE, alors CO désigne un capital versé sur un compte du client C dont :

- le numéro est représenté par (numéro du compte)
- les renseignements signalétiques correspondant au client C sont représentés par (renseignements signalétiques)
- les intérêts sont représentés par (intérêts)
- la date de la dernière opération effectuée sur ce compte est représentée par (date dernière modification)
- le solde après la dernière modification est représenté par (solde)
- un montant de correction de solde est représenté par (autres sommes)

MOUVEMENT

Si M appartient à MOUVEMENT, alors M désigne un mouvement effectué sur un compte CO dont :

-le numéro du compte à mouvementer est représenté par (n° du compte mouvementé)

-la valeur du montant versé ou retiré est représentée par (montant)

-le solde calculé pour le compte CO après le mouvement M est représenté par (solde après mouvement)

-le type de mouvement est représenté par (type du mouvement)

ANCIEN COMPTE

Si AC appartient à ANCIEN COMPTE, alors AC désigne un ancien compte d'épargne, conservé par obligation légale, et dont :

-le numéro de compte est représenté par (numéro ancien compte)

-les autres renseignements sont représentés par (autres renseignements).

.....

DESCRIPTION SEMANTIQUE ET CONTRAINTE D'EXISTENCE ET DE CARDINALITE DES RELATIONS.

POSSEDE

(CLIENT - LIVRET D'EPARGNE) CARD. : 1 - N

Si (C,L) appartient à POSSEDE, alors (C,L) désigne un client C qui possède un livret d'épargne L dont le lieu où le livret d'épargne L est retenu, est représenté par (lieu de détention).

CORRESPOND

(LIVRET - COMPTE) CARD. : 1 - 1

Si (L, CO) appartient à CORRESPOND, alors (L, CO) désigne un livret d'épargne L appartenant à un client C et correspondant au compte CO du client C.

HISTORIQUE

(COMPTE - ANCIEN COMPTE) CARD. : 1 - 1

Si (CO, AC) appartient à HISTORIQUE, alors (CO, AC) désigne un ancien compte d'épargne AC correspondant à un compte d'épargne CO qui est conservé par obligation légale.

MISE A JOUR

(MOUVEMENT - COMPTE) CARD. : N - 1

Si (M, CO) appartient à MISE A JOUR, alors (M, CO) désigne un compte CO mis à jour par un mouvement M dont le type de mise à jour est représenté par (type maj.) et dont la date de mise à jour est représentée par (date maj.)

E. - Expression des contraintes d'intégrité.

1. Le numéro du compte CO est représenté par 9 chiffres décimaux.

2. Le livret d'épargne L possède le même numéro de livret (n° livret) que celui qui est repris sur le compte CO correspondant au livret d'épargne L.

3. L'ancien compte AC possède le même numéro d'ancien compte (n° ancien compte) que celui qui est repris sur le compte CO correspondant à l'ancien compte AC.

4. Le mouvement M possède le même numéro de compte mouvementé (n° compte mouvementé) que le compte CO mis à jour par le mouvement M.

5. Le livret d'épargne L d'un client C se trouve, soit chez le client C, soit à la C.G.E.R. (agence ou siège central).

6. Le type de mouvement (type du mouvement) qui peut être appliqué au compte CO est, entre autres :
   - un bordereau, auquel cas il faudra préciser le numéro du bordereau, la date à laquelle il a été établi et l'agence émettrice.
   - un mouvement litigieux;
   - un mouvement correspondant à un compte à terme;
   - un mouvement correspondant à une recapitalisation;
   - un mouvement correspondant à une opération comptable ou non comptable.

7. Le type de mise à jour (type maj.) effectuée par un mouvement M sur un compte CO est :
   - une mise à jour normale, ou
   - une mise à jour des intérêts, ou
   - une mise à jour de litiges, ou
   - une capitalisation, ou
   - une opération comptable, ou
   - une opération non comptable.

F. - Quantifications.

POUR LES ENTITES.

Type entité COMPTE	± 10.200.000
Type entité LIVRET D'EPARGNE	± 10.200.000
Type entité CLIENT	± 10.200.000

Type entité MOUVEMENT

. bordereaux	40.000 à 60.000 par jour
. compte à terme	entre 80.000 et 90.000, une fois/mois.
. litiges	6 à 7% des bordereaux
. opérations comptables	4.000 par mois
. opérations non comptables	entre 500.000 et un million/an.
. recapitalisations	environ 1.000.000 par an.

Type entité ANCIEN COMPTE ± 300.000 par an.

POUR LA RELATION MISE A JOUR.

Mise à jour normale	40.000 à 60.000 par nuit (certaines nuits : de 100.000 à 150.000)
Mise à jour des litiges	6 à 7% de la mise à jour normale
Mise à jour des intérêts	4.000.000, relative à une période de deux à trois mois.
Capitalisation	une fois par an
Opérations comptables	4.000 par mois
Opérations non comptables	entre 500.000 et 1.000.000 par an.

G. - Description des Applications.

1. Application CONSULTATION (y compris consultation des intérêts).

Entrée.

Le numéro du livret ou les renseignements signalétiques du client (nom, initiales du prénom, date de naissance).

Sortie.

L'état du compte correspondant au numéro du livret et/ou état des intérêts du compte.

Fonction.

A partir du numéro du livret, cette application fournit l'état du compte (et/ou celui des intérêts) correspondant au numéro du livret.
A partir des renseignements signalétiques, cette fonction recherche en premier lieu, le numéro du livret correspondant avant de fournir l'état du compte et/ou celui des intérêts, pour le numéro obtenu.

2. Application MISE A JOUR normale.

Entrée.

Type mouvement bordereau.

Sortie.

Compte mis à jour ou mouvement type litigieux.

Fonction.

Cette application se déroule en deux étapes :

. étape 1 - contrôle des renseignements globaux du bordereau lui-même. S'il y a erreur, rejet en litige.

. étape 2 - mise à jour des comptes correspondant aux entrées du bordereau. S'il y a erreur, rejet du bordereau en mouvement litigieux, après avoir mentionné sur chaque entrée du bordereau, si elle a été prise en considération ou non.

3. Application MISE A JOUR - LITIGES.

Entrée.

Mouvement type litigieux.

Sortie.

Compte mis à jour.

Fonction.

Les litiges ayant été traités par un service comptable, cette application actualise le compte correspondant au mouvement litigieux.

4. Application MISE A JOUR - COMPTE TERME.

Entrée.

Mouvement type compte terme.

Sortie.

Compte mis à jour.

Fonction.

Cette application actualise le compte à terme correspondant au mouvement d'entrée.

5. Application MISE A JOUR - INTERETS (à la consultation).

Entrées.

Le numéro du livret ou les renseignements signalétiques du client et/ou le livret.

Sortie.

Compte mis à jour - partie (intérêts).

Fonction.

Cette application (après recherche éventuelle du numéro du livret

sur base des renseignements signalétiques) met à jour les intérêts du compte et, si cela est possible, l'état du solde du livret.

6. Application OPERATION COMPTABLE.

Entrée.

Le livret.

Sortie.

Le livret mis à jour.

Fonction.

Lorsque le client se présente dans une agence disposant d'un terminal relié au centre informatique de la C.G.E.R., cette application met à jour l'état du livret et celui du compte.

7. Application OPERATION NON COMPTABLE.

Entrée.

Le numéro du livret ou les renseignements signalétiques du client.

Sortie.

Les renseignements signalétiques mis à jour.

Fonction.

Sur base du livret ou des renseignements signalétiques (grâce auxquels il est permis d'obtenir le numéro du livret), on met à jour différents renseignements donnés par le client au sujet de son compte.(Ex. changement d'adresse, changement de mandataire)

8. Application CAPITALISATION.

Entrée.

Tous les comptes.

Sortie.

Comptes mis à jour (au point de vue intérêts)

Fonction.

A partir du solde de tous les comptes, cette application calcule les intérêts dus et les inscrit dans la partie (intérêts) du compte.

9. Application RECAPITALISATION.

Entrées.

Mouvement type recapitalisation et compte.

Sortie.

Compte mis à jour (au point de vue intérêts)

Fonction.

Cette application recalcule les intérêts et les inscrit dans la partie (intérêts) du compte correspondant au mouvement.

PARTIE IV. - ORGANISATION ACTUELLE.

PLAN.

A. - Introduction.

B. - Présentation générale.

C. - Organisation logique choisie du fichier CCD.

D. - Organisation physique choisie du fichier CCD.

E. - Applications spécifiques à la situation actuelle.

F. - Statistiques obtenues sur la situation actuelle.

PARTIE IV. - ORGANISATION ACTUELLE.

A. - Introduction.

Cette partie est consacrée à l'étude de la solution existant actuellement.

Nous présenterons une description générale avant de nous attarder plus particulièrement sur le fichier CCD en étudiant son organisation logique et physique.

Nous exposerons, ensuite, des applications spécifiques à la situation existante qui seront suivies d'une série de statistiques recueillies sur le fichier CCD.

Il ne nous est pas permis - pour des raisons bien compréhensibles de discrétion - de présenter les différents programmes travaillant avec le fichier CCD.

.....

B. - Présentation générale.

Etant donné que le schéma conceptuel est tiré de l'existant, nous rappellerons brièvement la situation actuelle. Pour plus de renseignements, nous renvoyons le lecteur aux deux premières parties de cette étude.

En ce qui concerne :

Le LIVRET D'EPARGNE.

Le livret d'épargne est émis et mis à jour par une agence ou par le siège central de la C.G.E.R.

Il se trouve généralement chez le client, mais il peut être en dépôt dans une des agences ou au siège central.

Les autres renseignements du livret sont conservés sur l'enregistrement du fichier CCD correspondant au livret.

Le COMPTE.

Le compte est repris sur le fichier CCD que nous développerons dans la suite.

Les renseignements signalétiques correspondant au client détenteur du compte sont repris dans les différents fichiers signalétiques (signalétique principal, signalétique adresse, signalétique numéro, signalétique mandataire). Les liaisons entre les différents fichiers signalétiques et le fichier CCD sont réalisées à l'aide de pointeurs dans les enregistrements.

En cas de discordance entre le montant du livret et celui du compte, nous rappelons que la différence peut se trouver dans le fichier "autres sommes".

Les autres renseignements du compte sont repris dans l'enregistrement du fichier CCD. Les renseignements complémentaires sur les comptes à terme, repris dans le fichier CCD, se trouvent dans le fichier "Terme", lequel est relaté à partir des enregistrements du CCD et ce, par pointeurs.

Les anciens comptes sont repris dans le fichier historique; ces comptes sont stoppés (leur montant est nul et il n'y a plus d'opérations effectuées sur ces comptes).

Les MOUVEMENTS.

Les mouvements suivants sont repris dans le fichier "Mouvement". Ce sont ceux correspondant :

- à la recapitalisation;
- à un compte à terme;
- aux mises à jour effectuées à partir des bordereaux (y compris celles correspondant à la mise à jour des intérêts).

Les mouvements litigieux sont repris dans le fichier " LITIGES ".

Quant aux opérations comptables et non comptables, ce sont celles prestées par le client, à partir d'une agence.

Seules, les opérations comptables et non comptables ainsi que les mises à jour des litiges se déroulent en ON-LINE, c'est-à-dire pendant la période où l'ordinateur de Bruxelles est relié aux agences par télépressing.

Les autres mouvements se déroulent en OFF-LINE, souvent durant la nuit ou le week-end.

.....

C. - Organisation logique choisie du fichier CCD.

Cette organisation existe depuis environ dix ans. A cette époque, il fut décidé de reprendre le fichier CCD sous forme de fichier séquentiel indexé.

Pour répondre aux activités d'un client vis-à-vis de son compte, il fut décidé de créer quatre types d'enregistrements différents mais de format fixe. C'est ainsi que l'on parle toujours d'enregistrement ou de compte : actif, soldé (montant nul) et flaggé, termes que nous avons déjà cités dans l'introduction initiale du présent travail.

Le quatrième type de compte est le "compte blanc".

Qu'est-ce qu'un compte blanc ?

C'est un compte destiné à un futur client de la C.G.E.R. L'idée retenue est la suivante : on prépare l'arrivée de nouveaux clients en introduisant régulièrement (tous les quinze jours), dans le fichier, de nouveaux numéros de comptes. Ceux-ci correspondent aux livrets envoyés, par le siège central, dans les différentes agences du pays et ce, suivant la demande.

C'est ainsi que, lorsqu'un client demande l'ouverture d'un compte d'épargne, il reçoit un livret dont le numéro est déjà repris dans le fichier CCD.

.....

D. - Organisation physique actuelle du fichier CCD.

La C.G.E.R. travaillant avec du matériel IBM, le fichier CCD sera un séquentiel indexé du type I.S.A.M. Nous reviendrons sur ces concepts I.S.A.M. en partir VI.

En réalité, il y a dix fichiers physiques qui forment, dans leur ensemble, le fichier CCD. Chaque fichier est un séquentiel indexé, c'est-à-dire qu'il possède ses enregistrements, en séquence dans une parties "données", et son index composé du cylinder index et du track index. Il n'existe pas de master index, celui-ci étant remplacé par un index composé de dix enregistrements, chacun représentant la valeur minimum et la valeur maximum des numéros de comptes que l'on trouvera sur un fichier.

Il faut noter aussi qu'il n'existe pas de pistes d'overflow pour les enregistrements qui ne trouveraient pas de place dans la partie "données" et ce, pour la raison expliquée précédemment (partie IV, point C). Pour le même motif, nous ne parlerons pas d'overflow indépendant. On pourra consulter, en annexe, la liste des V.T.O.C. des différents fichiers.

L'enregistrement du fichier, dont on pourra consulter la description actuelle en annexe, possède un format fixe - sa longueur est présentement de 96 positions - et, de plus, est groupé avec d'autres pour former ce que l'on appelle un bloc physique. Celui-ci a constamment évolué suivant l'augmentation des capacités des disques. La taille du bloc physique est déterminée dans le but de perdre le moins de place possible, par piste. Actuellement, un bloc comprend 64 enregistrements et l'on dispose de trois blocs par piste.

.....

E. - Applications spécifiques à la situation actuelle.

Outre les différentes applications présentées jusqu'ici /pour rappel : la consultation - la mise à jour normale - la mise à jour des litiges - la mise à jour des intérêts - les opérations comptables - les opérations non comptables - la capitalisation (et la recapitalisation)_7 , l'organisation actuelle nécessite deux autres applications que nous présentons ci-après.

1. La PRISE de COPIE TRIEE.

Le fichier CCD - on a pu s'en rendre compte - est utilisé, que ce soit en consultation ou en mise à jour, très souvent durant une journée, à partir des agences notamment.

Pour ne pas le surcharger davantage, il a été décidé que plusieurs programmes ne s'exécuteraient pas avec les données de ce fichier mais seulement à partir d'une copie de celles-ci.

Parmi les programmes travaillant sur la copie, il y en a notamment qui produisent des statistiques sur les enregistrements du fichier CCD. Pour obtenir des statistiques fiables, il ne faut pas que la copie soit trop ancienne. C'est pourquoi il fut choisi de prendre une copie du fichier, une fois par semaine et cela, sur bande magnétique.

Etant donné l'organisation actuelle du fichier, la prise de copie triée se résume à une lecture séquentielle d'un fichier séquentiel indexé. Nous reviendrons sur ce point dans les parties suivantes du travail.

2. La REORGANISATION.

Un fichier évolue au cours de son existence, notamment par les insertions. Suivant l'organisation choisie, si l'on ne réorganise pas le fichier, les performances d'accès à un article se détériorent.

C'est ainsi qu'à partir d'un certain moment, il est indispensable de procéder à cette réorganisation.

Normalement, on réorganise le fichier quand le coût d'une réorganisation est moins élevé que celui dû aux pertes de performances d'accès.

Dans la solution actuelle, c'est une autre optique qui a été choisie. On impose une réorganisation tous les quinze jours. Ceci s'explique aisément. En effet, les ajouts d'enregistrements, dans le fichier, se font en bloc tous les quinze jours (voir partie IV, point C), suivant la demande en livrets des agences. En autre temps, il n'y a pas d'autre insertion. On profite donc de ce que l'on insère beaucoup d'enregistrements en une fois pour réorganiser le fichier.

Le procédé est le suivant : prise d'une copie du fichier auquel il est ajouté les nouveaux enregistrements avant de recharger, sur les disques, le fichier ainsi réorganisé.

.....

F. - Statistiques obtenues sur la situation actuelle du fichier CCD.

Dans ce paragraphe, nous rappellerons, dans un premier temps, des statistiques générales sur le fichier CCD, que l'on a déjà présentées dans la partie II. du travail.

Nous présentons ensuite deux statistiques reprenant le nombre d'opérations journalières au fichier sur deux semaines (la 39e et la 40e de l'année 1981).

A l'aide de ces statistiques, nous tirerons des conclusions sur l'évolution du fichier CCD et nous proposerons deux types de répartition du fichier, envisageables dans une nouvelle organisation du fichier.

1. STATISTIQUE GENERALE.

a) Nombre d'articles dans le fichier 10.200.000
Nombre d'insertions par an 400.000
Nombre de suppressions par année 300.000
Nombre de consultations par journée de 8 H. 5.000
Nombre de mises à jour par nuit 40.000 à 60.000
certaines nuits 100.000 à 150.000
Nombre d'opérations comptables, par mois 4.000
Nombre d'opérations non comptables, par an entre 500.000 et 1.000.000

b) Les insertions ne peuvent se faire que dans une partie déterminée du fichier CCD.
Cette partie est représentée par les séries

36, 37	reprenant les comptes nouveaux
50	reprenant les comptes à terme
81, 82, 83	reprenant les agences.

Aucune autre insertion ne peut, désormais, avoir lieu et l'ensemble des insertions se répartit comme suit :

- ± 35% dans les séries 37 et 50
- ± 27% dans la série 36
- ± 0,1% dans les séries 81, 82, 83

Nous rappelons, à nouveau, qu'actuellement, les insertions se font en bloc, tous les quinze jours.

2. STATISTIQUES SUR DEUX SEMAINES.

Un programme, datant de 1977, proposait des résultats sur l'activité des séries 81, 82 et 83 lesquelles, à l'époque, étaient les plus mouvementées. Ce programme n'étant plus disponible, un nouveau programme a été conçu pour me donner des résultats sur l'activité de plusieurs séries.

C'est ainsi qu'après une première exécution, nous obtenons des résultats sur les séries considérées par la C.G.E.R. comme étant les plus mouvementées, à savoir les séries 36, 37, 81, 82 et 83.

Les renseignements obtenus n'étant pas suffisamment éloquents, le programme - après de légères modifications - a été à nouveau exécuté pour produire des résultats sur l'ensemble des séries du fichier.

Ière exécution du programme.

- données : les opérations journalières exécutées sur le fichier CCD. Ces opérations sont constituées des mises à jour et ne comprennent pas les insertions au fichier.

- résultats :

SERIES	Nbre d'OPERATIONS	%
36	47.484	14,95
37	65.563	20,64
81	24.847	7,82
82	14.873	4,88
83	8.070	2,54
TOTAL :	160.837	50,64
séries < 36	136.056	42,83
37 < séries < 81	19.905	6,66
séries > 83	813	0,26
TOTAL :	156.774	49,36
TOTAUX GLOBAUX	317.611	100,-

2ème exécution du programme.

- données reprenant les opérations passées au cours de la 39ème semaine de 1981.

n° de série	NOMBRE	n° de série	NOMBRE
1	13.420	27	493
2	10.432	33	1
3	7.418	36	40.872
4	1.946	37	50.476
5	7.550	38	21
6	9.848	39	20
7	10.039	40	11.093
8	9.767	50	4.623
9	5.445	51	4.554
10	2.865	56	274
11	3.296	58	2
12	6.898	63	1
13	8.813	64	224
14	8	72	543
16	1.252	80	339
18	390	81	23.458
19	469	82	13.414
25	1	83	8.026
26	4.512	TOTAL :	262.803

RECAPITULATIF.

n° de série	NOMBRE	%	% cumulés.
36	96.050	40	40
36	40.872	15,5	55,5
37	50.476	19,2	74,7
T.36, 37	91.348	35	75
37 < S < 81	21.694	8	83
81	23.458	9	92
82	13.414	5	97
83	8.026	3	100
T. 81,82,83	44.298	17	///////
TOTAL GLOBAL	262.803	100%	100%

Observation générale.

50% des opérations se répartissent dans les séries { 36, 37 ; 81, 82, 83.

3. DEDUCTION DES STATISTIQUES.

Il y a deux grandes parties dans le fichier, au point de vue accroissement :

- une partie évolutive, comprenant les séries 36, 37, 80, 81, 82, 83.
- une partie stable, comprenant les autres séries.

Nous observons également deux grandes séries au point de vue accès aux enregistrements, chacune reprenant 50% du taux d'activité, l'une étant constituée des séries 36, 37, 81, 82, 83 et l'autre reprenant les autres séries.

Nous constatons, dès lors, que la partie qui peut s'accroître correspond presque à celle qui comprend 50% du taux d'activité : seule, la série 50 n'est pas reprise dans celle-ci.

4. TYPES DE REPARTITION DU FICHIER.

Nous proposerons d'évaluer deux répartitions différentes par organisation choisie :

. une reprenant tout le fichier CCD
. une autre reprenant deux fichiers,
  -l'un pour la partie évolutive du fichier
    (comprenant les séries 36, 37, 50, 81, 82, 83 pour un total de ± 4.000.000 d'enregistrements)
  -l'autre pour la partie fixe du fichier
    (comprenant les autres séries pour un total de ± 6.200.000 enregistrements).

PARTIE V. - LES PROBLEMES A RESOUDRE.

Nous allons, dans cette cinquième partie, développer de manière plus précise, les problèmes qui ont été exposés.

Nous rappelons que nous travaillons sur la gestion de l'épargne, qui est l'une des applications informatisées de la C.G.E.R. et qui doit s'intégrer dans un ensemble informatisé.

Nous ne nous intéresserons, dorénavant, qu'au fichier CCD pour lequel nous nous efforcerons de proposer des organisations susceptibles de pallier les problèmes présentés ci-après.

Une étude de nouvelle organisation des centres fichiers n'étant pas prévue dans ce mémoire, ils resteront dans leur situation actuelle.

Pour le fichier CCD, nous distinguerons deux problèmes essentiels : le temps d'accès et le format des enregistrements.

A. Le TEMPS D'ACCES.

Comme on a pu le constater dans le début du présent travail, on accède au fichier CCD un très grand nombre de fois durant la journée.
(Rappel :
- 5.000 consultations par jour
- de 40.000 à 60.000 mises à jour par nuit
- 6 à 7% de ce chiffre pour les rejets en litige, par nuit.
- entre 80.000 et 90.000 opérations, une fois par mois, pour les comptes à terme.
- opérations comptables au nombre de 4.000 par mois
- opérations non comptables, entre 500.000 à un million/an
- mise à jour des intérêts, à raison de 4.000.000 sur 3 mois.
- à cela, il faut ajouter les insertions (400.000/an) et les suppressions (300.000/an).

La gestion et l'organisation de ce fichier jouent un rôle prépondérant puisqu'elles influencent fortement le temps d'accès au fichier.

La C.G.E.R. est intéressée de savoir s'il y a possibilité d'améliorer la situation existante, en tenant compte évidemment des applications spécifiques à l'épargne, développées dans la partie précédente de cette étude.

B. Le FORMAT des ENREGISTREMENTS.

Nous avons vu, dans la présentation de la partie IV, que les enregistrements ont un format fixe.

Cette situation pose, en général, deux problèmes, problèmes que l'on retrouve dans le fichier CCD, à savoir :

a) une perte de place pour certains enregistrements qui n'ont pas besoin de toutes les zones définies dans l'article;
b) un manque de place pour d'autres enregistrements que l'on voudrait détailler, chose que l'on n'ose pas faire de crainte d'amplifier le premier problème.

Dans notre étude, ces problèmes sont illustrés dans le fichier CCD notamment par les"enregistrements blancs" pour le premier problème et par la redéfinition de certaines zones d'un enregistrement (ce qui a, pour effet, de compliquer certains programmes d'application) pour le deuxième problème (°).

Une solution à ces problèmes serait de travailler dans un format variable. Nous trouverons, en annexe, une description du nouveau format de l'enregistrement tel qu'il a déjà été prévu par la CGER et qui sera comparée avec la description actuelle.

Toute proposition de nouvelle organisation du fichier CCD devra tenir compte de ces deux problèmes. Pour le problème du format de l'enregistrement, il faudra également savoir si c'est le programmeur système qui devra le gérer ou si la méthode d'organisation le supportera elle-même.

(°) Pour plus de précision pour ces illustrations, nous reportons le lecteur à la partie IV de cette étude : "Organisation actuelle".

PARTIE VI. - ETUDE DE DIFFERENTES TECHNIQUES.

PLAN.

A. - Introduction.

B. - Le séquentiel indexé.

C. - Virtual Storage Access Method (V.S.A.M.).

D. - L'accès calculé.

E. - L'accès calculé avec hashage virtuel.

.....

PARTIE VI. - ETUDE DE DIFFERENTES TECHNIQUES.

A. - Introduction.

Dans cette partie, on va s'intéresser aux différentes techniques que l'on peut proposer comme étant d'éventuelles solutions pouvant remplacer celle qui existe actuellement.

Seules, les trois organisations suivantes ont été retenues :

. le séquentiel indexé (type I.S.A.M.)
. la"Virtual Storage Accès Method"(V.S.A.M.)
. l'accès calculé.

Pour chacune de ces trois organisations, nous présenterons la méthode et le principe de base, à savoir :

. recherche d'un enregistrement;
. ajout d'un enregistrement;
. suppression d'un enregistrement;
. modification de la longueur d'un enregistrement.

Nous rappellerons, chaque fois, les deux types de répartition de fichiers envisageables et nous terminerons par les procédés de lecture d'un fichier, aussi bien de manière aléatoire que de manière séquentielle, ainsi que leur coût (en accès).

Une quatrième proposition (2 méthodes d'accès calculé avec hashage virtuel pour la résolution des collisions) sera brièvement présentée à la fin de cette partie; nous expliquerons le pourquoi du rejet de ces deux méthodes d'organisation.

.....

B. - Le séquentiel indexé.

1. PRESENTATION.

Comme son nom l'indique, un fichier de ce type est un fichier séquentiel avec un ou plusieurs niveaux d'index. Il comprend trois zones distinctes que l'on dénomme : la zone primaire, la zone d'index et la zone de débordement (ou encore zone d'overflow).

Les enregistrements se trouvent, au moment de la création du fichier ou lors d'une réorganisation du fichier, dans la zone primaire et sont, de plus, en séquence sur une clé. La zone primaire peut s'étendre sur plusieurs volumes physiques mais peut aussi être constituée de différentes zones non contiguës.

Il y a plusieurs niveaux d'index qui sont créés par l'Operating System (O.S.) au moment de la création ou de la réorganisation du fichier. Le niveau le plus bas s'appelle le track index; il y en a un pour chaque cylindre de la zone primaire; il se trouve sur la(les) première(s) piste(e) du cylindre qu'il indexe.

Un enregistrement à ce niveau comprend une entrée en zone normale et une en zone d'overflow. Il y a une paire d'entrées par piste en zone primaire dans le cylindre. Une entrée = une adresse + la clé maximum de la piste. La dernière entrée de chaque piste est fictive. S'il reste de la place sur la piste, on peut s'en servir comme d'une zone primaire, excepté lorsque l'enregistrement est de format variable, auquel cas la dernière piste du track index ne peut être complétée par des enregistrements de zone primaire.

Chaque track index peut contenir un enregistrement spécial, dénommé le "cylinder overflow control record".

Le deuxième niveau s'appelle le cylinder index, ses entrées pointant vers le track index. Il y a un cylinder index pour le fichier; il peut se trouver sur un autre volume que le fichier, dans la zone primaire ou dans la zone d'overflow. Il comprend une entrée par cylindre de zone primaire qui est formattée comme pour le track index (1 adresse + 1 clé maximum du cylindre). On ne peut pas retrouver d'enregistrements de zone primaire dans le cylinder index; s'il reste de la place, elle sera réservée pour les additions futures.

Un niveau supérieur d'index est utilisé lorsque le cylinder index est trop important (sup. à 4 pistes). C'est le master index qui comprend une entrée par piste de cylinder index. Il peut se trouver aussi dans une zone indépendante, en zone primaire ou en zone d'overflow. S'il fait encore plus de 4 pistes, on a encore un niveau supérieur et ainsi de suite.

La zone d'overflow peut être de deux types : le cylinder overflow ou l'independent overflow. On peut avoir l'un et/ou l'autre pour un même fichier.

Le cylinder overflow : on réserve un certain nombre de pistes sur chaque cylindre pour les enregistrements qui viendront en overflow (c'est-à-dire qui ne trouveront plus de place dans la zone primaire).
L'avantage du cylinder overflow est le suivant : on se trouve positionné sur le bon cylindre, que l'article soit en zone primaire ou en overflow.
Son inconvénient : la perte de place.

L'independent overflow : la zone d'overflow se trouve dans une zone indépendante du fichier. Si l'avantage de l'I.O. est le gain de place, l'inconvénient en est la perte de temps. En effet, si l'enregistrement ne se trouve pas sur le cylindre, en zone primaire, il faut se déplacer pour atteindre le(s) cylindre(s) d'overflow, ce qui coûte manifestement quelques millisecondes supplémentaires.

2. PRINCIPES DE BASE.

Recherche d'un enregistrement identifié par (a).

. recherche sur master index jusqu'à ce que l'on trouve une entrée avec une clé $\not\geq$ a

. recherche sur cylinder index jusqu'à ce que l'on trouve une entrée avec une clé $\not\geq$ a

en réponse : adresse du cylindre .

. parcours du track index du cylindre jusqu'à ce que l'on trouve une entrée avec une clé $\not\geq$ a

en réponse : adresse de la piste.

. parcours de la piste $\longrightarrow$ enregistrement clé $\not\geq$ a

si clé = a $\Longrightarrow$ recherche fructueuse
si clé $\neq$ a $\Longrightarrow$ recherche infructueuse

REMARQUE.

Le track index me dit si l'élément se trouve en overflow ou non; dans l'affirmative, on parcourt alors uniquement la chaîne d'overflow pour le retrouver. Si on ne l'a pas trouvé lorsqu'on se trouve en fin de chaîne, on a une recherche infructueuse.

Ajout d'un enregistrement.

. recherche de l'endroit d'insertion.

. si cet endroit se trouve dans une piste de zone primaire :
  - insérer l'élément et décaler les suivants, s'il reste de la place sur la piste.
  - dans la négative, envoyer en overflow autant d'enregistrements nécessaires pour insérer le nouveau et modifier l'entrée de l'index.

. si cet endroit se trouve dans la chaîne d'overflow :
  - appliquer le principe d'ajout d'un enregistrement dans une chaîne et modifier l'entrée de l'index.

Supression d'un enregistrement.

L'enregistrement n'est pas physiquement supprimé mais il est marqué. Si, lors d'additions ultérieures d'enregistrements, il devait être expulsé en zone d'overflow, il n'y serait pas réinscrit. Il sera, de toute façon, physiquement supprimé lors d'une réorganisation ultérieure.

Remarque.

Ces principes sont valables pour des enregistrements de longueur fixe ou variable mais, pour les enregistrements de longueur variable, il faut ajouter un nouveau principe : la modification de longueur d'un enregistrement qui serait comparable, selon que l'on augmente ou que l'on diminue la longueur, à l'addition et à la suppression d'un enregistrement.

3. DEUX PROPOSITIONS A ENVISAGER.

1°) Solution actuelle.

On applique les principes développés. Toutefois :

- pas de master index
- pas d'ajouts (sauf lors des réorganisations)

Nous renvoyons le lecteur en partie IV pour une présentation plus précise.

2°) Autre possibilité.

Solution : séquentiel indexé avec :

- enregistrements en longueur variable
- mise à jour des longueurs d'enregistrements
- ajout d'enregistrements
- suppression marquage
- enregistrement dans des zones contiguës
- cylinder overflow et independent overflow.

Pour cette seconde proposition, nous étudierons deux types de répartition choisis, dans la partie IV :

a) un fichier global reprenant tout le fichier CCD;
b) deux fichiers : un pour la partie évolutive du fichier CCD et un autre pour sa partie fixe.

4. PROCEDE DE LECTURE.

Le but de ce paragraphe est de présenter :

1°) le bilan en accès d'une recherche au hasard d'un enregistrement, son procédé de recherche ayant été vu auparavant;
2°) le procédé et le bilan en accès d'une recherche séquentielle de tous les enregistrements du fichier.

a) de manière aléatoire.

Bilan en accès.

. un seek du cylinder index
. un search + un read sur le cylinder index
. un seek du track index
. un search + un read du track index

S'il s'agit d'une entrée en zone primaire :

. search de la piste primaire indiquée (on sait si l'enregistrement est trouvé ou non)

. read de l'enregistrement.

S'il s'agit d'une entrée en overflow :

. search de la piste d'overflow indiqué

. read de l'enregistrement.

si l'enregistrement en bon $\Longrightarrow$ fini.

si l'enregistrement n'est pas bon, recommencer search et read $\rightarrow$ fin de chaîne ,implique enregistrement non trouvé.

b) de manière séquentielle.

Principes.

1. Inialisation.
   Positionner le mécanisme d'accès à la track index du premier cylindre du fichier.
   Recherche et lecture d'une parie d'entrées sur la track index.

2. Lecture de chaque piste de la zone primaire spécifiée.
   Si fin fichier $\rightarrow$ stop.

3. Si entrée overflow n'a pas été modifiée, aller en 5.

   ↓

4. Si non , lire la chaîne d'overflow.

5. Si entrée primaire suivante n'est pas la dernière, aller en 7.

6. Si l'entrée est la dernière $\Longrightarrow$ tous les enregistrements du cylindre ont été lus. Rechercher le cylindre suivant.

7. Rechercher la track index et lire la paire d'entrées suivantes.

8. Aller en 2.

Bilan en accès.

. on a un seek pour chaque cylindre de zone primaire

{ search + read pour chaque enregistrement de données.
{ search et read pour chaque paire d'entrées dans la piste d'index.

. un search + un read peur être évalué à une rotation

. le search + un read du track index peut être évalué à 1/2 rotation car les enregistrements sont plus courts.

.le temps de seek des cylindres est réduit au minimum car ils sont dans une zone continue (option choisie).

c) Remarques.

- Le coût des opérations élémentaires peut être retrouvé à partir de la description du disque employé à la C.G.E.R. (3350 D'IBM) que l'on trouvera en annexe.

- Un seek : temps de déplacement du bras du disque.

  Un search : délai de rotation du disque jusqu'à ce que l'enregistrement se trouve en-dessous de la tête de lecture.

  Un read : lecture de l'enregistrement.

- On considère que le master index se trouve en mémoire centrale : ceci évite le temps d'accès (seek-search-read) au disque.

.....

C. - Virtual Storage Access Method (V.S.A.M.)

1. PRESENTATION.

V.S.A.M. est une méthode d'accès utilisable à partir d'un operating system propre à IBM et utilisé avec des unités de stockage à accès direct pour accélérer la recherche ou le stockage d'un article dans un fichier.

V.S.A.M. comprend trois types de fichiers :

- "Entry sequenced data set" où les enregistrements sont stockés dans le fichier comme ils y sont entrés; on ne peut plus en ajouter.
- "Key sequenced data set" : les enregistrements sont classés suivant une clé et l'on se sert d'un index pour les retrouver. On a la possibilité d'ajouter des enregistrements.
- "Relative record data set" valable uniquement pour des enregistrements à longueur fixe. Chaque enregistrement a un numéro d'identification. On a la possibilité d'insérer des enregistrements.

Par la suite, lorsque l'on parlera d'un fichier V.S.A.M., on entendra "Key sequenced data set" car c'est celui qui répond aux problèmes exposés en partie V. Il faut savoir aussi que tout fichier V.S.A.M. est repris dans un catalogue V.S.A.M. Pour plus de précisions à ce sujet, nous renvoyons le lecteur aux références V.S.A.M. (Bibliographie, 8-9-10)

Un "Key sequenced data set" comprend deux parties : la partie "données" et la partie "index".

a) la partie "données".

Les enregistrements sont stockés dans une zone de longueur variable appelée control interval (C.I.). Mais, pour un fichier donné, le C.I. est toujours d'une même longueur.

Un C.I. est composé :

- de un ou de plusieurs enregistrements;
- éventuellement, d'espaces libres;
- d'informations de contrôle sur les enregistrements du C.I. et sur la place encore libre.

On parlera aussi de control area (C.A.) qui comprend, pour un fichier donné, toujours le même nombre de C.I. Chaque C.A. peut comprendre, parmi ses C.I., un ou plusieurs C.I. d'espace libre, tout ceci étant défini lors de la création du fichier.

b) la partie "index".

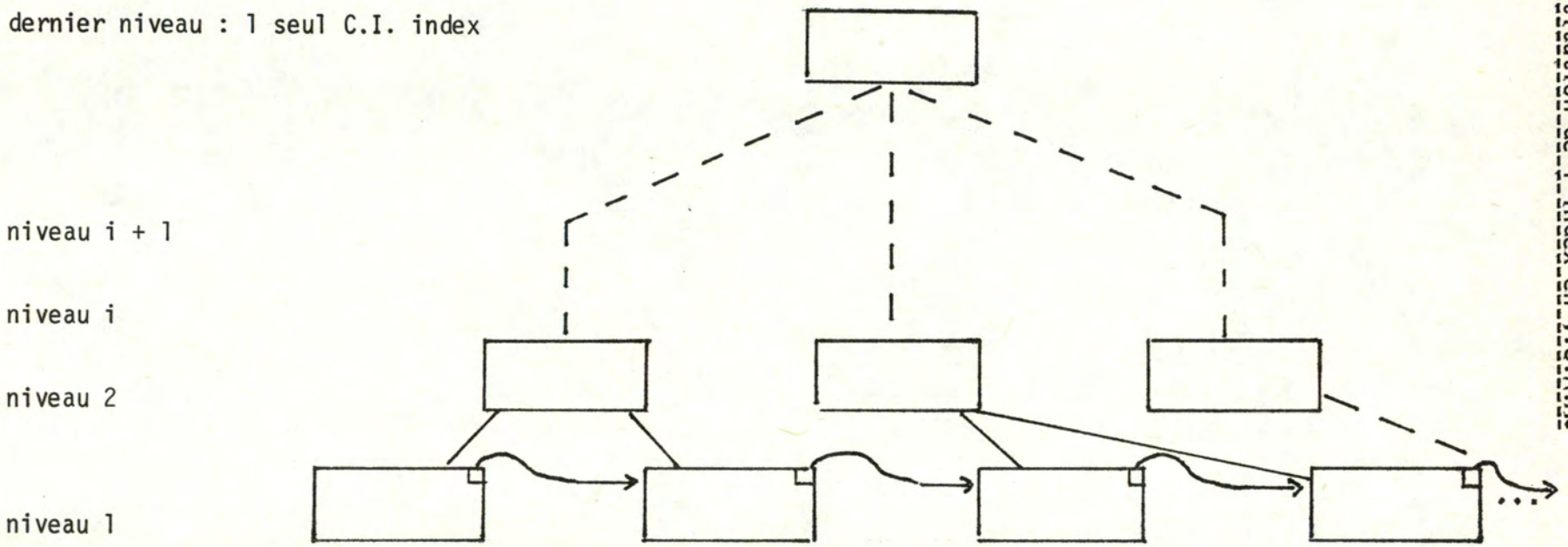
L'index est constitué de plusieurs niveaux. Le niveau le plus bas s'appelle le "sequence set". Un sequence set forme un bloc de plusieurs enregistrements.

Chaque enregistrement du sequence set contrôle une C.A. c'est-à-dire qu'il y a autant de clés dans un enregistrement du sequence set qu'il n'y a de C.I. dans une C.A. (par C.I., on reprend la clé la plus élevée de l'index). Le dernier niveau de l'index ne peut avoir qu'un seul enregistrement.

V.S.A.M. offre un mécanisme de compression de clés pour les enregistrements d'index, ce qui permet un plus grand nombre d'entrées pour un enregistrement d'index. Ce mécanisme est basé sur la "suppression" en avant et en arrière de la clé, des caractères non nécessaires pour distinguer cette clé de ses adjacentes.

On notera, enfin, que pour faciliter le parcours séquentiel, les enregistrements du sequence set sont reliés entre eux alors que, de toute façon, les différents niveaux de l'index sont également reliés entre eux (représentation page suivante).

L'index peut se trouver n'importe où dans le fichier, même dans la zone de données. On le placera dans le milieu du fichier en vue de diminuer le temps de déplacement du bras. De plus, si la possibilité en est offerte, on le recopiera plusieurs fois par piste d'index, en vue de diminuer le délai de rotation.



. Niveau 1 appelé sequence set : un enregistrement par C.A.

. Tous les C.I. index ont la même longueur

. A un niveau i, lorsqu'il y a deux C.I.index $\Rightarrow$ création d'un C.I. index au niveau i + 1

2. PRINCIPES DE BASE.

Recherche d'un enregistrement de clé (a)

. recherche, dans les différents niveaux d'index, avec un mécanisme de reconstitution des clés comprimées. On recherche jusque a la plus haute clé d'un C.I.

. de l'index, on se retrouve dans un C.I. où il y a parcours séquentiel avec les deux possibilités de recherche fructueuse ou de recherche infructueuse.

Ajout d'un enregistrement.

. recherche de l'endroit d'insertion (comme la recherche d'un enregistrement). On se retrouve donc dans un C.I.

. s'il y a de la place libre dans le C.I. :
  - on insère;
  - on effectue la mise à jour de la partie contrôle du C.I.

. si l'on ne dispose pas de place libre dans le C.I., on exécute un SPLIT du C.I. : on sépare le C.I. en deux avec un C.I. d'overflow appartenant à la même control area (C.A.) que le C.I. On mettra alors à jour les informations de contrôle du C.I., du C.I.overflow et l'index.

. si l'on ne sait pas faire le SPLIT de C.I. (plus de place dans le C.I.overflow, par exemple), on réalise alors un SPLIT de C.A. : on coupe la C.A. en deux avec une C.A. d'extension. Ce qui implique la mise à jour des informations de contrôle dans l'index et dans les C.I. des deux C.A.

Suppression d'un enregistrement.

. recherche de l'endroit de suppression comme se fait la recherche d'un enregistrement. On se retrouve dans un C.I.

. on supprime l'élément en réalisant :
  - un tassement
  - une mise à jour des informations de contrôle
  - éventuellement, une mise à jour de l'index.

Remarque.

Ces principes sont valables pour des enregistrements de longueur fixe ou variable mais, pour les enregistrements de longueur variable, il faut ajouter un nouveau principe : la modification de longueur d'un enregistrement.

Ne connaissant pas exactement la manière de travail de V.S.A.M., nous proposons, comme dans les autres techniques, d'appliquer les principes de suppression et d'addition d'enregistrements exposés précédemment.

3. DEUX TYPES DE REPARTITION.

Nous rappelons les deux types de répartition choisis dans la partie IV de cette étude :

1. - un fichier global, reprenant tout le fichier CCD.
2. - deux fichiers, l'un pour la partie évolutive du fichier CCD et l'autre, pour sa partie fixe.

4. PROCEDE DE LECTURE.

Le but de ce paragraphe est de présenter :

1°) le bilan en accès d'une recherche au hasard d'un enregistrement, son procédé de recherche ayant été vu auparavant;

2°) le procédé et le bilan en accès d'une recherche séquentielle de tous les enregistrements du fichier.

a) de manière aléatoire.

Bilan en accès.

. recherche dans l'index; par niveau d'index, il faut un seek
un search
un read.

Remarque.
Un ou plusieurs niveaux peuvent se trouver en mémoire centrale, ce qui supprime l'accès sur disque.

. recherche dans la partie "données" : un seek, un search, un read.

b) de manière séquentielle.

Préliminaires.

Il faut rappeler que les enregistrements du premier niveau d'index (le sequence set) sont chaînés entre eux, ce qui évite un parcours de tous les niveaux d'index pour une lecture séquentielle du fichier.

Principes.

1. Accéder au premier bloc du sequence set (le lire)
2. Pour chaque enregistrement de ce bloc, accéder au C.I. correspondant + lecture du C.I.
3. S'il existe un bloc suivant du sequence set, aller en 4. Si non, aller en 6.
4. Accéder au bloc suivant du sequence set + le lire.
5. Aller en 2.
6. STOP.

Bilan en accès.

Nous obtenons, pour les points 1, 2 et 4, chaque fois un seek
un search
un read.

c) Remarques.

. Le coût des opérations élémentaires peut être retrouvé à partir de la description du disque employé à la C.G.E.R. (3350 d'IBM) que l'on trouvera en annexe.

. Il existe un langage appelé A.M.S. (Acces Method Service) qui, en fait, est un ensemble de fonctions utiles pour la création et la maintenance des fichiers en V.S.A.M. Certaines de ces fonctions peuvent nous procurer des statistiques intéressantes sur l'activité d'un fichier géré par V.S.A.M.

. Beaucoup d'alternatives nous sont proposées pour la création et l'exploitation d'un fichier V.S.A.M. Nous nous limiterons cependant aux principes exposés dans cette partie.
Pour des renseignements complémentaires, le lecteur consultera les références à V.S.A.M. (bibliographie, 8-9-10)

. Un seek : c'est le temps de déplacement du bras du disque.
Un search : il s'agit du délai de rotation du disque jusqu'à ce que l'enregistrement se trouve en-dessous de la tête de lecture.
Un read : c'est la lecture de l'enregistrement.

.....

D. - Accès calculé avec méthodes classiques de résolution de collisions.

Remarque préliminaire.

Deux techniques de résolution de collisions ont été choisies : l'independent overflow et le primary chaining.

1. PRESENTATION.

L'objectif de cette technique d'organisation est de retrouver, très rapidement, un article dans le fichier à partir d'une partie de l'article qui l'identifie, partie qui est appelée : CLE.

On dispose d'un espace pour ranger les articles, espace que l'on découpe logiquement en cases. Chaque case a une adresse physique précise. Une case peut contenir un ou plusieurs articles.

Une fonction, que l'on devra définir, va permettre, à partir de la clé des articles, de répartir ceux-ci dans les différentes cases. Cette fonction s'appelle "fonction de répartition" ou de "hashage". Il en existe de plusieurs types, tels que fonction de division, de folding, de midsquare, etc...
(exemple : fonction de division)

Il y a n cases; chaque article est identifié par une clé que l'on suppose numérique (si non, on peut toujours y arriver). L'adresse de la case où l'article sera rangé est [le reste de la division de la clé par n + 1] si l'on considère que l'espace des cases commence à 1 et qu'il se termine à n.

On se rend compte qu'après un certain nombre d'insertions, certaines cases peuvent être remplies. Que se passe-t-il lorsqu'un nouvel enregistrement doit se retrouver dans une de ces cases ?

On appelle ce phénomène : une collision. Il existe différentes méthodes de résolution d'équation et comme dit plus avant, nous en étudierons deux :

1) L'independent overflow.
   Tous les enregistrements qui, normalement, auraient dû se trouver dans une même case mais qui, par manque de place, n'y sont pas, se retrouvent chaînés dans une zone indépendante du fichier.

2) Le Primary Chaining
   Les enregistrements rejetés sont aussi chaînés mais dans le fichier lui-même.

2. PRINCIPES DE BASE.

Recherche d'un enregistrement.

. Passer la clé par la fonction de répartition, ce qui donne comme résultat l'adresse d'une case en zone primaire.

. Parcourir séquentiellement la case, ce qui donne soit une recherche fructueuse, soit une recherche non fructueuse.

. Si la recherche est infructueuse, parcours de la chaîne d'overflow ce qui donne soit une recherche fructueuse, soit une recherche non fructueuse. Dans ce dernier cas, l'enregistrement est absent du fichier.

Ajout d'un enregistrement.

. Recherche de l'endroit d'intertion, ce qui donne une case en zone primaire.

. S'il y a de la place, insérer l'enregistrement et mettre à jour le compteur indiquant la place libre dans la case.

. S'il n'y a pas de place libre, on insère dans la chaîne d'overflow.

Pour la chaîne d'overflow.

. Il faut mettre à jour le pointeur réalisant le chaînage avec l'élément inséré.
Le dernier pointeur de la chaîne a une valeur fictive pour marquer la fin de celle-ci.

Pour insérer, il faut de la place libre
en independent overflow.

. On peut considérer l'espace libre comme un fichier séquentiel adressable et extensible.

en primary chaining.

1) soit à partir de la case "arrivée" : on parcourt les cases suivantes et on s'arrête à la première où il y a de la place disponible;

2) soit pour toutes les cases, on dispose d'une indication qui nous informe s'il y a de la place disponible dans cette case. On a alors le choix ; on prend :
   - la première case où il y a de la place;
   - la case la plus proche de celle où on se trouve.

Destruction d'un enregistrement.

. Recherche de l'endroit de suppression (comme addition).

. Si l'endroit est dans une case primaire :
  - soit supprimer en le marquant;
  - soit le supprimer physiquement + un tassement de la case et remise éventuelle d'éléments de la chaîne d'overflow en zone primaire.

. Si l'endroit est dans la chaîne d'overflow :
  suppression d'un article dans une chaîne avec remise de l'enregistrement dans l'ensemble des places vides.

Remarques.

Ces principes sont valables pour des enregistrements de longueur fixe ou variable mais, pour les enregistrements de longueur variable, il faut ajouter un nouveau principe : la modification de longueur d'un enregistrement qui serait comparable, selon que l'on augmente ou que l'on diminue la longueur, à l'addition et à la suppression d'un enregistrement.

Dans l'independent overflow, on disposera d'un pointeur par case de zone primaire, qui nous donne l'adresse du premier enregistrement en overflow (s'il en existe un).
Les enregistrements de la chaîne d'overflow seront chaînés entre eux.

Dans le primary chaining, on dispose d'un pointeur par case qui nous donne l'adresse dans la case du premier enregistrement appartenant effectivement à cette case. Ensuite, tous les enregistrements devant appartenir à cette case seront chaînés entre eux bien qu'ils puissent se trouver dans des cases différentes.

3. DEUX TYPES DE REPARTITION.

Nous rappelons les deux types de répartition choisis dans la partie IV de cette étude :

1. - un fichier global reprenant tout le fichier CCD.
2. - deux fichiers, l'un pour la partie évolutive du fichier CCD et l'autre, pour sa partie fixe.

4. PROCEDE DE LECTURE DU FICHIER.

Le but de ce paragraphe est de présenter :

1) le bilan en accès d'une recherche au hasard d'un enregistrement, son procédé de recherche ayant été vu précédemment;

2) le procédé et le bilan en accès de la lecture séquentielle du fichier.

a) de manière aléatoire.

Bilan en accès.

1 seek, 1 search, 1 read. - Coût supplémentaire s'il y a recherche en overflow.

1 seek, 1 search, 1 read par enregistrement lu dans la chaîne.

b) de manière séquentielle.

Préliminaires.

Généralement, la lecture séquentielle d'un fichier organisé en accès calculé, n'est pas demandée sauf pour des situations particulières car les enregistrements sont dans un désordre total. Nous nous intéresserons plus particulièrement au tri du fichier afin d'obtenir les enregistrements en séquence, dans la partie suivante de l'étude.

Principes.

La lecture séquentielle sera évaluée à la lecture du fichier, piste par piste, cylindre par cylindre, avec un déplacement du bras, réduit au minimum.

Bilan en accès.

1 seek, 1 search, 1 read par piste.

REMARQUES.

Le temps des opérations élémentaires peut être retrouvé à partir de la description du disque employé à la C.G.E.R. (3350 d'I.B.M.) que l'on trouvera en annexe.

Un seek : temps de déplacement du bras du disque.
Un search : délai de rotation du disque jusqu'à ce que l'enregistrement se trouve en-dessous de la tête de lecture.
Un read : lecture de l'enregistrement.

E. - Accès calculé avec hashage virtuel pour la résolution de collisions.

Observations.

Cette proposition fait suite aux récentes publications de l'I.N.R.I.A. (°), lors du stage effectué en novembre 1981 et à la thèse du Docteur LITWIN (').

Parmi les différents algorithmes, deux seront introduits avec leurs performances théoriques.

PRINCIPES DU HASHAGE VIRTUEL.

. Principe de l'accès calculé avec :
  - fonction de hashage sur une clé
  - collisions.

. Ce qui change, ce sont les méthodes de résolution de collisions. Lorsqu'une collision arrive, on allonge l'espace d'adressage en faisant éclater en deux, la case où la collision est survenue et on définit une nouvelle fonction de hashage.

. A tout moment, on sait par quelle fonction de hashage une case est adressée.

. Le traitement des collisions peut commencer dès la première collision ou à partir d'un taux d'occupation du fichier, auquel cas les collisions précédentes ont été résolues par les méthodes connues, à savoir : Independent overflow, Primary Chaining, ...)

. Performances recueillies.

		Capacité de la case.	Seuil pour le taux d'occupation
recherche avec succès	1,28 accès	50	0,75
	1,13	50	0,675
recherche sans succès	2,4	50	0,675
	3,57	50	0,75

(°) Institut National de Recherche d'Informatique et d'Automatique.

(') Voir Bibliographie, n° 16, 17.

Un autre algorithme est intéressant, le Trie Hashing.

PRINCIPES DU TRIE ASHING.

. Ce sont les mêmes principes que pour l'algorithme précédent :
  - restriction : on éclate à toutes les collisions.
  - on garde le fichier trié grâce à un index constitué sous forme d'arbre.

. Inconvénient.

La longueur de l'index qui s'accroît très vite et se retrouve rapidement en mémoire secondaire pour des fichiers de plus de un million d'enregistrements.

. Performances recueillies.
sous les conditions suivantes :
  - index en mémoire centrale,
  - recherche (avec ou sans succès) réalisée en un accès.

  → 2 accès, quand il n'y a pas de collision, sinon :
  → 4 accès.

.....

Pourquoi avoir rejeté ces solutions ?

La première car :
. cette technique est optimale pour un fichier très mouvementé, ce qui n'est pas le cas pour le fichier CCD.
. les performances recueillies manquent de précision, notamment quant à l'échantillonnage choisi;
. les données des performances ne sont pas les mêmes (ex. : format des enregistrements fixe au lieu de variable).

La seconde car :
. le fichier CCD nécessiterait un index trop grand, ce qui en diminuerait les performances.
. de plus, la gestion d'un tel algorithme est assez complexe;
. enfin, les performances ne sont encore que théoriques et peu testées en pratique actuellement.

PARTIE VII. - EVALUATION DES DIFFERENTES TECHNIQUES.

PLAN.

A. - Introduction.

B. - Présentation des critères d'évaluation.

C. - Présentation des simulations réalisées.

D. - Evaluation des techniques.

PARTIE VII. - EVALUATION DES DIFFERENTES TECHNIQUES.

A. - Introduction.

L'objectif poursuivi, dans cette partie, est d'évaluer les différentes techniques présentées précédemment. En vue d'atteindre ce but, nous présenterons, dans un premier temps, les critères retenus.

Nous parlerons ensuite des simulations réalisées pour les techniques V.S.A.M. et accès calculé, la technique du séquentiel indexé ayant été étudiée complètement de manière analitique.

Nous développerons, enfin, les évaluations des différentes techniques proposées en commençant toutefois par la solution existant actuellement.

Exception faite pour la solution actuelle, nous procéderons de la manière suivante :

. présentation des éléments à prendre en considération lors de l'évaluation;
. développement des équations analytiques;
. développement des équations de probabilité ou présentation de résultats de simulation;
. présentation des résultats obtenus par rapport aux critères retenus pour les deux types de répartition du fichier CCD choisis;
- comparaison des deux types de répartition du fichier CCD retenus.

B. - Présentation des critères d'évaluation.

Plusieurs critères sont retenus dans l'évaluation des différentes techniques; il s'agit :

a) de la place occupée par le fichier;
b) du temps d'accès aléatoire à un enregistrement du fichier;
c) du temps de lecture séquentielle ou, le cas échéant, d'un tri du(des) fichier(s);
d) du temps de mise à jour d'un enregistrement du fichier ;
e) du temps d'insertion d'un enregistrement dans le fichier ;
f) de la réorganisation du fichier. Il s'agit, ici, de connaître le moment, le coût et la partie du fichier à réorganiser.
g) de l'influence de la technique choisie sur l'existant.

Il nous restera, alors, à évaluer les différentes applications traitées à la C.G.E.R. - Pour rappel, ce sont :

1) la consultation;
2) la mise à jour normale, celle des rejets (litiges) et celle des intérêts;
3) les opérations comptables et les opérations non comptables;
4) la capitalisation et la recapitalisation.

REMARQUE.

Pour l'ensemble des points 2, 3 et 4, il sera établi une moyenne journalière de mises à jour d'enregistrements du fichier.

Enfin, nous terminerons par l'évaluation d'une prise de copie triée du(des) fichier(s) qui est, pour rappel, une application particulière à la solution actuelle mais qui pourrait être envisagée pour toute autre technique retenue.

Rappel des quantifications pour les applications.

. Consultation		5.000 par jour
. Mise à jour normale		40.000 à 60.000 avec pointes de 100.000 à 150.000
	rejets (litiges)	6 à 7% des mises à jour normales
	des intérêts	4 millions en trois mois
. Opérations comptables		4.000 par mois
. Opérations non comptables		entre 500.000 et 1.000.000 par an.

Nous proposerons ainsi une évaluation, par jour, de :

50.000 100.000 150.000 mises à jour

. Capitalisation : tous les comptes.
La recapitalisation concernant +/- un million d'enregistrements, est reprise dans les mises à jour normales.

C. - Présentation des simulations réalisées.

1. INTRODUCTION.

Nous rappelons que deux techniques ont été simulées. Pour chaque simulation, nous présenterons, dans un premier temps, son(ses) objectif(s), ses informations en entrée et en sortie et, enfin, nous développerons ses principes.

Tous les résultats ne seront pas repris dans cette partie du travail mais figureront en annexe. Nous baserons nos évaluations futures sur ces résultats.

2. SIMULATION EN V.S.A.M.

a) Objectif.

Nous savons que V.S.A.M. effectue, lorsqu'il n'y a plus de place disponible dans un C.I., une répartition sur deux C.I.
L'objectif poursuivi, par cette simulation, est d'évaluer le taux d'occupation des C.I. à la suite d'insertions d'enregistrements, les C.I. étant préalablement chargés à un taux d'occupation choisi.

b) Entrées.

Un fichier de 10.000 C.I. chargés à un taux d'occupation de 80%.
Taille des C.I. : 1024 - 2048 - 4096 bytes.
Nombre d'insertions aléatoires : 20.000

- un échantillonnage aléatoire.
- un enregistrement est représenté par un numéro de compte et par sa longueur.

c) Sorties.

Taux d'occupation des C.I. (voir tableau page suivante)

Nombre d'additions d'enregistrements.

Nombre d'additions de C.I.

NB - Le lecteur pourra consulter les listings en annexe.

TABLEAU (résumé) du taux d'occupation des C.I.

Taille C.I.	Nombre d'insertions	% d'occupation
1024	0	80
	400	76,9
	3.600	66,7
	Rem.: Le pourcentage le moins élevé, puis oscillations aux alentours de 68%.	
2048	0	80
	400	82,5
	3.200	67
	Rem. : Le pourcentage le moins élevé, puis oscillations aux alentours de 70%.	
4096	0	80
	400	82,6
	6.000	65,2
	Rem. : Le pourcentage le moins élevé, puis oscillations aux alentours de 73%	

d) Principes.

. Technique d'insertion. - Le C.I. dans lequel l'enregistrement doit être inséré est déterminé par une fonction de randomisation sur l'identifiant de l'enregistrement (le numéro de compte). Fonction choisie : la division.

. Pour chaque C.I., on dispose d'un compteur indiquant la place libre dans le C.I.

. Pour une insertion :

1) recherche du C.I. d'insertion, soit C.I.a
2) s'il y a de la place disponible dans le C.I.a : mise à jour du compteur du C.I. - aller en 4.
3) s'il n'y a pas de place disponible dans le C.I.a, ajouter un C.I.b et répartir en deux l'occupation de C.I.a sur C.I.a et C.I.b en faisant une mise à jour des compteurs.
4) passer à l'insertion suivante.

REMARQUES.

La fonction de randomisation tient compte des ajouts de C.I. pour répartir les enregistrements.

Prise de résultats sur les listings : toutes les 400 insertions.

Calcul du taux d'occupation :

. moyenne du taux d'occupation de l'ensemble des C.I.

. pour un C.I. : $\frac{\text{place occupée}}{\text{place théorique du départ}}$

. . . . . .

3. SIMULATION EN ACCES CALCULE.

Deux algorithmes ont été simulés : l'independent overflow et le primary chaining.

a) Objectifs.

. Calcul du nombre moyen d'accès pour l'insertion et la consultation des enregistrements.
. Calcul du temps d'accès moyen et maximum pour l'insertion et la consultation des enregistrements.
. Calcul de la longueur maximum de la chaîne formée par les éléments ne trouvant pas de place dans leur case initiale.
. Calcul du pourcentage d'enregistrements atteints en 1, 2, 3, ... accès.
. Calcul du taux d'occupation du fichier.

b) Entrées.

. Espace d'adressage variable, suivant les simulations.
. Echantillon de 200.000 éléments :
  - échantillonnage choisi en tenant compte d'une répartition des longueurs des enregistrements et d'une répartition des enregistrements dans les séries (")
  - un enregistrement est représenté par un numéro de compte et par sa longueur.

c) Sorties.

Les objectifs poursuivis sont repris sur les listings que le lecteur pourra consulter, en annexe.

(") Référence au générateur de nombre aléatoire : J-M.NISEN & J.ROULIN. Bibliographie, n° 12.

4. PRINCIPES.

a) Choix de la fonction de répartition.

. Représentons par n le nombre de cases où l'on doit répartir les enregistrements et par x le nombre de chiffres nécessaires pour représenter le nombre de cases où l'on doit répartir les enregistrements.

. Trois fonctions de répartition ont été retenues :

1 - division (reste de la division de la clé numérique de l'enregistrement par n) + 1

2 - midsquare élévation au carré de la clé et prendre les x chiffres centraux. Si le nombre représenté par ces x chiffres est supérieur à n , nous prenons comme résultat le module n de ce nombre.

3 - folding
- . diviser la clé en tranches de x chiffres
- . replier la Ière tranche sur la 2ème et additionner
- . prendre la tranche formée des x derniers chiffres de l'addition et la replier sur la tranche suivante
- . répéter l'opération précédente jusqu'à la dernière tranche.
- . si la dernière tranche obtenue est supérieure à N, nous prenons comme résultat, le module N du nombre formé par cette dernière tranche.

b) Choix de la capacité disque.

. Un des critères retenus est le temps d'accès à l'enregistrement (en consultation ou en insertion). Or, un temps d'accès se compose de trois éléments :
- temps de déplacement du bras du disque (seek)
- temps de délai rotationnel (search)
- temps de lecture (read)

. Si les deux derniers éléments sont représentatifs de la réalité, quelle que soit la simulation réalisée, le premier élément ne l'est pas car nous n'avons que 200.000 enregistrements au lieu de 10.200.000 - ce qui réduit l'espace nécessaire et, par conséquent le temps de déplacement du bras. Pour pallier cet inconvénient, nous avons décidé de réduire la capacité des disques tout en corrigeant le temps de déplacement du bras.

c) Choix de l'overflow.

. 20% de la zone primaire initiale ont été choisis, au départ, répartis sur l'ensemble des cases primaires, pour le primary chaining et des cylindres réservés par disque, pour l'independent overflow.

d) Idée de la simulation.

. Insertion d'un élément et déduction des renseignements en insertion et en consultation.

e) Détails de la simulation.

I.- Pour l'algorithme d'independent overflow, représentons par :

COMPT-PRESENTS	le nombre d'enregistrements présents dans une case de zone primaire.
COMPT-Réels	le nombre d'enregistrements qui devraient se trouver dans une case de zone primaire.
PLACE	la place disponible dans une case de zone primaire.

a) lecture de l'enregistrement suivant à insérer

b) recherche, par la fonction de randomisation, d'une case primaire où l'enregistrement doit être inséré.

c) mise à jour des statistiques :

. COMPT-Réels + 1

. Si place disponible :

mise à jour PLACE

COMPT-PRESENTS + 1

. Sinon, insérer en overflow

se positionner à l'overflow (milieu du disque) et insertion de l'enregistrement.

. Mise à jour du nombre moyen d'accès en insertion et en consultation.

d) Calcul des temps de recherche :

. calcul du temps nécessaire pour le déplacement du bras en vue d'arriver à la case de zone primaire.

. calcul du temps search + read de la case

. calcul du temps passé dans la chaîne d'overflow (longueur chaîne = COMPT Réels - COMPT. PRESENTS)

e) Recommencer pour toute nouvelle insertion.

f) Présentation des calculs.

f_1 - longueur de la chaîne

elle est égale à COMPT Réels - COMPT PRESENTS

f_2 - nombre moyen d'accès

EN INSERTION.

. en zone primaire : 2 (lecture et écriture de l'enregistrement)

. s'il faut aller en overflow : 3 + longueur de la chaîne, à savoir :

1 - lecture zone primaire

. lecture des enregistrements de la chaîne

2 - écriture du dernier enregistrement de la chaîne

3 - écriture du pointeur dans la case réalisant le chaînage.

Remarque : la longueur de la chaîne comprend l'enregistrement à insérer.

EN CONSULTATION.

. en zone primaire : 1

. en overflow : 1 + longueur de la chaîne (celle-ci comprend l'enregistrement à insérer)

f_3 - temps de search.

Il a été tenu compte du temps moyen (8,4 ms.)

Cfr. capacité disque en annexe.

f_4 - temps de transfert.

Capacité case divisée par capacité de transfert.

Cfr. description capacité disque en annexe.

f_5 - temps passé en overflow.

Suivant la longueur de la chaîne, en tenant compte :

. du temps d'arrivée en début de chaîne d'overflow

. du temps de passage d'un élément de la chaîne à un autre (option choisie : équirépartition des éléments chaînés dans l'espace réservé à l'overflow)

. des temps de délai rotationnel et de lecture.

. NB - En insertion, tenir compte de la mise à jour du pointeur, pour réaliser le chaînage.

II. - Pour l'algorithme du primary chaining, représentons par :

COMPT ..... le nombre d'enregistrements par case.
PLACE ..... la place disponible par case.

a) lecture de l'enregistrement suivant à insérer
b) recherche, par la fonction de randomisation, d'une case primaire où l'enregistrement doit être inséré.
c) mise à jour des statistiques :
   . Si place disponible :

mise à jour	PLACE
COMPT	+ 1

   . Sinon, insertion en overflow.
   option choisie : parcours des cases suivantes pour la recherche de place libre et insertion de l'enregistrement.
d) calcul des temps de recherche (id. algorithme ind.overflow)
e) recommencer toute nouvelle insertion.
f) présentation des calculs.

f_1 - longueur de la chaîne
elle est égale à la longueur du parcours des cases effectué pour la recherche de place, dans le cas d'une insertion en overflow.

f_2, f_3, f_4) Cfr. independent overflow, f_2- f_3- f_4

f_5 - temps passé en overflow.
Suivant la longueur de la chaîne, en tenant compte :
. du temps de passage d'un élément de la chaîne à un autre (option choisie : équirépartition des éléments chaînés).
. des temps de délai rotationnel et de lecture
. NB - En insertion, tenir compte de la mise à jour du pointeur pour réaliser le chaînage.

D. - Evaluation des techniques.

Comme présenté dans l'introduction de cette partie, nous allons développer successivement :

- la solution existant actuellement;
- la technique du séquentiel indexé;
- la technique V.S.A.M.;
- la technique de l'accès calculé avec deux algorithmes : l'independent overflow et le primary chaining.

Options choisies.

Répartition sur 200 jours, des insertions d'enregistrements, des opérations comptables et non comptables ainsi que des mises à jour des intérêts.

La semaine est considérée comme une période de cinq jours de travail.

Nous ne nous sommes pas occupés des suppressions d'enregistrements, celles-ci pouvant se faire "en masse" annuellement, lors d'une réorganisation.

SOLUTION ACTUELLE.

A. - EVALUATION SELON LES CRITERES RETENUS.

1. Place occupée par le fichier.

. Rappel : dix fichiers physiques organisés en séquentiel indexé; les neuf premiers possèdent le même nombre d'enregistrements et le dernier en possède le reste.

. Nombre de cylindres :

(183 x 9) + (367 x 1) = 2.014 cylindres répartis sur 4 disques.
Parmi ces cylindres : place réservée à l'index.
Il y a trois pistes par fichier physique.

. Taux d'occupation du fichier.

Actuellement : +/- 10.200.000 enregistrements.
Nombre d'enregistrements possibles dans le fichier :

- les enregistrements sont bloqués (facteur de blocage : 64)
- nombre de blocs par piste : 3
- nombre de pistes par cylindre : 30
- total possible : 64 x 3 x 30 x 2.014 = 11.600.640

Taux d'occupation 10.200.000 / 11.600.640 = +/- 90%.

2. Temps d'accès aléatoire.

. Rappel : pas d'enregistrement en overflow.

. Temps d'accès en ms.

	Minimum	Moyen	Maximum
seek cylinder index	10	25	50
search cylinder index + read	8,8	8,8	8,8
seek trach index	10	25	50
search data index + read	8,8	8,8	8,8
search data	8,4	8,4	8,4
read data	5	5	5
Total (en ms.)	51	81	131

3. Temps de lecture séquentielle du fichier.

. Rappel : les principes ont été développés lors de la présentation générale du séquentiel indexé, en partie VI.

. Un seek pour chaque cylindre de zone primaire.
minimum pour tous les cylindres, sauf le premier (moyen)
seek min. 10 ms.
total 1. 2.014 ms. x 10 = 20.140 ms.

. Un search pour chaque bloc de données + un read de chaque bloc peut être estimé à un délai rotationnel (= 16,7 ms.)
Remarque : il y a 90 blocs par cylindre.
total 2. 16,7 ms. x 90 = 1.503 ms.

. Un search + un read pour chaque paire d'entrées du track index peut être estimé à 1/2 délai rotationnel (= 8,4 ms.)
Remarque : il y a 30 entrées par cylindre (29 + 1 entrée fictive)
total 3. 8,4 ms. x 30 = 252 ms.

. N.B. Le fichier est chargé à 90% mais, par cylindre, il y a toujours le même nombre de blocs que s'il était chargé à 100%; seul, le dernier bloc peut ne pas être complet mais il devra, néanmoins, être lu dans sa totalité.

. Total global
total 1 + (total 2 + total 3) x nombre de cylindres.
20.140 + (1.503 + 252) x 2.014 = 3.554.710 ms.
= 59 minutes 15 secondes
(soit +/- 1 heure)

4. Temps de mise à jour d'un enregistrement.

Une mise à jour { = 1 accès à l'enregistrement avec lecture
= 1 accès pour l'écriture
= 1 relecture pour vérification

peut donc être évaluée à 2 accès aléatoires + 1 lecture (nécessitant un délai rotationnel), ce qui donne :

		Minimum	Moyen	Maximum
2 accès	(en ms)	102	162	262
délai pour lecture (estimé à une toration)		16,7	16,7	16,7
Total	(en ms)	118,7	178,7	278,7

5. Temps d'insertion d'un enregistrement dans le fichier.

Il ne sera pas évalué, car ce n'est pas appliqué dans la solution actuelle (Cfr. présentation solution actuelle, en partie IV.)

6. Réorganisation.

. Quand ? Tous les quinze jours (Cfr. situation actuelle, P. IV.)

. Coût ? Evaluation reçue au mois d'octobre 1981.

Copie du fichier + insertion massive des enregistrements (+ ou - deux heures)

Rechargement sur disques : 1 Heure 26 minutes.

7. Pas d'influence sur l'existant dans ce cas-ci car c'est la solution appliquée.

- - - - - - - - - - - - -

B. - EVALUATION DES APPLICATIONS.

1. Consultation.

. Rappel : 5.000 consultations par jour.
Une consultation = un temps d'accès aléatoire.

	Minimum	Moyen	Maximum
. Temps total (en secondes)	255	405	655

2. Mise à jour (journalière). (')

(en secondes)	Minimum	Moyen	Maximum
. Total pour :			
50.000 opérations	5.935	8.935	13.935
100.000 opérations	11.870	17.870	27.870
150.000 opérations	17.805	26.805	41.805

(') Cfr. remarque technique séquentiel indexé, page

3. Capitalisation (annuelle).

. Coût : une prise de copie du fichier sur bande est estimée à une lecture séquentielle du fichier.

un rechargement du fichier sur les disques est estimé à une lecture séquentielle (le temps de traitement n'étant pas évalué).

. Rappel : une lecture séquentielle = +/- 1 heure (59' 15")

. Total : 2 heures.

C. - PRISE DE COPIE TRIEE.

Elle est évaluée à une lecture séquentielle = 1 heure.

TECHNIQUE I. - LE SEQUENTIEL INDEXE.

1. Présentation des différents éléments influençant l'évaluation.

. Capacité de la piste, exprimée en bytes. Elle est déterminée par le type du disque utilisé. Actuellement, elle est de 19.069 bytes.

. Longueur des enregistrements (LE). Les enregistrements étant décrits dans un format variable, les calculs seront basés sur une longueur moyenne évaluée à 130 bytes en zone primaire et à 140 bytes en zone de débordement (130 + long.pointeur 10)

. Taille des pages (TP). Les enregistrements en zone primaire seront bloqués et ce, afin d'utiliser au mieux la capacité de la piste. Il sera envisagé différentes tailles de pages.

. Taux d'occupation du fichier (TO); en vue d'absorber les insertions, le fichier créé sera rempli à un taux d'occupation variable.

. Place track index (PTi) - Rappel : track index = index de niveau le plus bas. Il faudra évaluer la place occupée par le track index.

. Place de l'overflow (PO), c'est-à-dire le nombre de piste d'overflow que l'on va réserver par cylindre, pour les enregistrements qui ne trouveront plus de place en zone primaire.

2. Equations analytiques.

a) Longueur des pages (TP)
à choisir en fonction de la capacité de la piste.

b) Nombre de pages par piste.

$$NB = \left\lfloor \frac{1 + 19.069 - \text{longueur dernier bloc (page)}}{TP.} \right\rfloor$$

c) Nombre d'enregistrements par piste.

NEP = TO x MB x TP

↳ taux d'occupation.

d) Nombre de pistes nécessaires (NP)

$$\left\lfloor \frac{\text{Nombre maximum d'enregistrements}}{\text{NEP}} \right\rfloor \times 1$$

Remarque : Mêmes principes pour les enregistrements en overflow, excepté qu'ils ne sont pas bloqués.

e) Place track index.

Nombre de pistes du track index : formule (b). Les enregistrements ne sont pas bloqués.

f) Nombre de pistes de track index par cylindre.

$$\left\lfloor \frac{2 \times \text{nbre de pistes disponibles par cylindre} + \textcircled{1}^{\oplus}}{\text{nbre d'entrées du track index par piste} + 2} \right\rfloor_{\oplus\oplus} + 1$$

⊕ entrée fictive à la fin de chaque piste de track index

⊕⊕ car la piste track index n'a pas d'entrée dans cet index (une entrée = entrée normale + entrée overflow)

g) Place cylinder index.

$$\left\lfloor \frac{\text{nbre cylindres} + 1}{\text{entrée index par piste}} \right\rfloor + 1$$

h) Nombre de cylindres nécessaires.

$$NC = \left\lfloor \frac{\text{nombre total de pistes}}{x} \right\rfloor + 1$$

x = 30 - taille track index - taille cylindre overflow

⇓

Nombre de disques :

$$\left\lfloor \frac{NC}{555} \right\rfloor + 1$$

(555 cylindres disponibles par disque (du type 3350.)

Remarque.

Nous nous baserons, pour ces calculs, sur la capacité du disque 3350 que l'on découvrira en annexe.

3. Equations de probabilité.

Pour déterminer l'overflow nécessaire dans cette organisation, nous choisirons un modèle analytique se basant sur les données suivantes :

. nombre de cases (NC) où nous pouvons insérer des enregistrements;

. taille de ces cases (TC);

. nombre d'insertions (N).

Dans le séquentiel indexé, le nombre de cases est représenté par le nombre de pistes et la taille des cases, par la place disponible sur ces pistes.

De ces données, nous tirerons la probabilité suivante :

- Probabilité qu'il y ait K enregistrements dans une case.

Nous appliquons la définition de la loi binomiale que nous rappelons : "la probabilité qu'il y ait un enregistrement exactement K fois dans une case au cours de N insertions indépendantes".

$$P(K) = C_N^K \left(\frac{1}{NC}\right)^K \cdot \left(1 - \frac{1}{NC}\right)^{N-K}$$

Nous déduisons :

1) Nombre de cases ne débordant pas :

$$NC \cdot \sum_{K=0}^{T} P(K)$$

2) Nombre de cases débordant :

$$NC - NC \cdot \sum_{K=0}^{T} P(K)$$

3) Le contenu moyen des cases qui débordent :

$$TC + 1 + \sum_{K = T+2}^{N} P(K) \cdot (K-T)$$

Le calcul de la probabilité se base sur les propriétés suivantes :

1) $P(0) = \left(1 - \frac{1}{NC}\right)^N$

$= \left(1 - \frac{1}{NC}\right)^{NC \cdot \frac{N}{NC}} \longrightarrow = e^{-1}$

$= e^{-\frac{N}{NC}}$

2) $$\frac{P(K+1)}{P(K)} = \frac{C_N^{K+1} \cdot \left(\frac{1}{NC}\right)^{K+1} \cdot \left(1 - \frac{1}{NC}\right)^{N-K-1}}{C_N^{K} \cdot \left(\frac{1}{NC}\right)^{K} \cdot \left(1 - \frac{1}{NC}\right)^{N-K}}$$

$$= \frac{\left(\frac{1}{NC}\right) \cdot (N-K)}{\left(1 - \frac{1}{NC}\right) \cdot (K+1)}$$

D'où :

$$P(K+1) = \frac{(N-K)}{(K+1)} \cdot \frac{\left(\frac{1}{NC}\right)}{\left(1 - \frac{1}{NC}\right)} \cdot P(K)$$

.....

4. Première répartition retenue dans la technique du séquentiel indexé : un fichier pour tout le C.C.D.

A. - EVALUATION SELON LES CRITERES RETENUS.

1. Place occupée par le fichier.

Nombre d'enregistrements	10.200.000
Taille des pages	2 K - 3 K ou 6 K
Taux d'occupation	70, 80 et 90%
Longueur enregistrements	130 en zone primaire 140 en overflow
Nombre d'insertions	400.000 par an

Nombre d'enregistrements par piste :

taux d'occupation (TO) / taille des pages (TP)		70%	80%	90%	(100)% pour mémoire
+/-	2 K	88	100	113	126
+/-	3 K	92	105	118	132
	6 K	98	112	126	141

Nombre de pistes nécessaires :

TP \ TO	70 %	80 %	90 %
+/- 2 K	115.910	102.000	90.266
+/- 3 K	110.870	97.143	86.441
6 K	104.082	91.072	80.953

Place track index :

. Un enregistrement track index = une entrée normale
+ une entrée overflow.

. Taille une entrée = taille clé (9 chiffres ⟶ 4 bytes) + 10 bytes par entrée pour le pointeur total : 2 entrées = 28 bytes.

. Nombre de pistes de track index par cylindre (formule f) : 0,91 à savoir : une piste.

. Rappel : les enregistrements étant en longueur variable, ils ne peuvent se trouver sur la piste d'index.

Place overflow :

. Nous appliquons les principes présentés dans le point 3 (équations de probabilité) pour les parties du fichier qui peuvent évoluer.

. Nous déterminons, en premier lieu, les calculs à effectuer et nous présenterons ensuite les résultats obtenus.

I. Nombre de pistes où nous devons insérer.

La série 36 représente ± 820.000 enregistrements, ce qui donne par taux d'occupation en nombre de pistes, selon le tableau ci-après :

TP \ TO	70%	80%	90%
+/- 2 K	9.319	8.200	7.257
+/- 3 K	8.914	7.810	6.950
6 K	8.368	7.322	6.508

Pour la série 37 (+/- 1.080.000 enregistrements)

TP \ TO	70%	80%	90%
+/- 2 K	12.273	10.800	9.558
+/- 3 K	11.740	10.286	9.153
6 K	11.020	9.643	8.572

Pour la série 50 (+/- 800.000 enregistrements)

TO / TP	70%	80%	90%
+/- 2 K	9.091	8.000	7.080
+/- 3 K	8.696	7.620	6.780
6 K	8.164	7.143	6.350

Pour les séries 81, 82 et 83 (+/- 1.200.000 enregistrements)

TO / TP	70%	80%	90%
+/- 2 K	13.637	12.000	10.620
+/- 3 K	13.044	11.429	10.170
6 K	12.347	10.804	9.604

II. Place disponible par piste en enregistrements.

TO / TP	70%	80%	90%
+/- 2 K	38	26	13
+/- 3 K	40	27	14
6 K	43	29	15

III. Nombre d'insertions.

Pour la série 36	108.000	soit	2.700 par semaine.
Pour la série 37	140.000	soit	3.500 par semaine.
Pour la série 38	140.000	soit	3.500 par semaine.
Pour les séries 81, 82, 83	400	soit	10 par semaine.

IV. - Résultats.

Tableau résumé des résultats obtenus.

Conditions de test : taux d'occupation = 90%
séries choisies : 36, 37, 50.
taille des pages : ± 2 K, ± 3 K, 6 K

ND = nombre de pistes débordant.

EO = nombre moyen d'enregistrements en overflow par piste.

Nombre de semaines écoulées	+/- 2 K		+/- 3 K		6 K	
	ND	EO	ND	EO	ND	EO
Série n° 36.						
10	0	0	0	0	0	0
15	11,15	1,002	4,43	1,0009	1,73	1,0007
20	143,96	1,026	91,24	1,017	69,31	1,015
25	649,20	1,16	477,22	1,12	404,25	1,11
30	1692,98	1,56	1369,02	1,46	1229,02	1,45
35	3112,18	2,33	2685,08	2,18	2487,59	2,2
36	3408,59	2,55	2970,86	2,38	2763,55	2,41
37	3701,80	2,77	3256,79	2,59	3040,04	2,63
38	3989,04	3,005	3539,86	2,82	3313,96	2,88
39	4267,83	3,26	3817,35	3,07	3582,44	3,14
40	4536,05	3,52	4086,78	3,33	3842,95	3,42
Série n° 37.						
10	0	0	0	0	0	0
15	13,16	1,0015	5,42	1,0008	2,34	1,0005
20	169,15	1,023	106,39	1,015	80,30	1,012
25	778,58	1,14	567,74	1,11	477,94	1,097
30	2073,26	1,50	1664,86	1,41	1487,88	1,40
35	3884,28	2,23	3332,27	2,08	3078,46	2,09
36	4269,077	2,43	3700,80	2,27	3433,90	2,29
37	4651,96	2,64	4071,77	2,46	3792,36	2,50
38	5029,25	2,86	4441,29	2,68	4149,82	2,73
39	5397,58	3,10	4805,71	2,92	4502,49	2,98
40	5754,05	3,35	5161,74	3,16	4846,94	3,24
Série n° 50.						
10	1,96	1,0005	0	0	0	0
15	136,72	1,026	86,41	1,017	65,38	1,01
20	897,83	1,25	683,13	1,19	590,97	1,18
25	2522,39	2,01	2133,04	1,88	1957,42	1,88
30	4389,69	3,48	3951,53	2,29	3713,67	3,38
35	5792,31	5,51	5397,42	4,33	5100,94	5,56
36	5993,75	5,96	5610,37	4,79	5301,96	6,05
37	6170,09	6,42	5797,90	5,26	5477,72	6,56
38	6322,96	6,88	5961,32	5,74	5629,67	7,07
39	6454,27	7,35	6102,30	6,23	5759,65	7,59
40	6566,06	7,83	6222,80	6,72	5869,70	8,12

REMARQUE.

A un taux d'occupation de 80%, seule la série n° 50 déborde après 33 semaines pour les tailles de pages de $\pm$ 2 K et de 6 K, après 32 semaines pour la taille de pages de $\pm$ 3 K.

Les résultats, après 40 semaines, sont les suivants :

40	113,5	1,03	158,7	1,04	113,46	1,03

Commentaires du tableau.

1. Vu le nombre trop petit d'insertions pour les séries n° 81, 82 et 83 par rapport à la place restant disponible, nous pouvons affirmer qu'après un an d'insertions d'enregistrements, nous n'aurons aucun débordement.

2. A un taux d'occupation (TO) de 70%, pour les différentes séries testées, nous n'obtenons aucun débordement après un an d'insertions, ce qui implique qu'il ne faut pas réserver de piste de débordement par cylindre de données.

3. A un taux d'occupation (TO) de 80%, nous obtiendrons des débordements pour la série n° 50 après seulement 33 semaines d'insertions. Ce nombre est d'ailleurs très peu élevé (entre 113 et 158 pistes déborderont avec un contenu moyen de débordant égal à $\pm$ 1,033 enregistrement).
   Afin d'éviter une trop grande perte de place réservée à l'overflow par cylindre, nous proposerions, dans ce cas, un overflow indépendant, qui comprendrait aux environs de 165 enregistrements (ce qui correspondrait à quatre pistes), la partie de l'overflow indépendant étant considérée comme un fichier séquentiel adressable et extensible.

4. Si nous admettons que, sur une piste d'overflow par cylindre de données, il y a 47 enregistrements de longueur moyenne égale à 140 bytes, deux pistes d'overflow par cylindre, à un taux d'occupation de 90%, nous satisferaient pour les séries 36, 37, 81, 82 et 83 pour l'ensemble des insertions d'enregistrements sur une période d'un an.
   Par contre, la série n° 50 nécessite, à elle seule, trois pistes d'overflow. Toutefois, deux pistes nous satisferaient si nous arrêtions les insertions d'enregistrements après 33 semaines.
   C'est ainsi que, pour éviter de perdre trop de place, nous proposerions deux pistes d'overflow par cylindre et un overflow indépendant dont la taille pourrait aller jusqu'à +/- 568 pistes (soit 19 cylindres).

^^^^^^^

En résumé.

Par cylindre, nous utiliserons :

30 pistes - le nombre de pistes du track index - le nombre de pistes du cylinder overflow.

A savoir, pour un taux d'occupation de :

70%	30	-	1	-	0	= 29 pistes
80%	30	-	1	-	0	= 29 pistes
90%	30	-	1	-	2	= 27 pistes

Nombre de cylindres nécessaires pour ce fichier :

TP \ TO	70%	80%	90%
+/- 2 K	3.997	3.518	3.344 + 19 ind.overf.
+/- 3 K	3.824	3.350	3.201 + 19 ind.overf.
6 K	3.590	3.141	2.999 + 19 ind.overf.

Nous devons rajouter la place du cylinder index :

= (nombre de cylindres : nombre entrées index par piste) + 1

= 1 entrée du cylinder index = 14 bytes $\Rightarrow$ 68 par piste.

Cela varie entre 44 pistes pour TO = 90%
58 pistes pour TO = 70%
} soit $\pm$ 2 cylindres.

Nous obtenons un master index :

taille 1 entrée = 14 bytes $\Rightarrow$ 68 par piste $\Rightarrow$ 1 piste est suffisante pour le master index.

Place en disques.

TP \ TO	70%	80%	90%
+/- 2 K	7,2 $\longrightarrow$ 8	6,34 $\longrightarrow$ 7	6,06 $\longrightarrow$ 7
+/- 3 K	6,9 $\longrightarrow$ 7	6,04 $\longrightarrow$ 7	5,8 $\longrightarrow$ 6
6 K	6,5 $\longrightarrow$ 7	5,7 $\longrightarrow$ 6	5,44 $\longrightarrow$ 6

2. Temps d'accès aléatoire.

Remarque : le délai rotationnel est toujours pris comme délai moyen.

a) Option choisie.

Master index en mémoire centrale mais par le cylinder index.

b) En zone primaire (en msecondes)

		Minimum	Moyen	Maximum
seek cylinder index		10	25	50
search cylinder index		8,4	8,4	8,4
read cylinder index (bloc de 528 bytes)		0,4	0,4	0,4
seek track index		10	25	50
search track index		8,4	8,4	8,4
read track index (bloc de 528 bytes)		0,4	0,4	0,4
search enregistrement		8,4	8,4	8,4
read enregistrement en zone primaire	+/- 2 K	1,5	1,5	1,5
	+/- 3 K	2,4	2,4	2,4
	6 K	5	5	5
en overflow	par enregistrement lu, il faut compter le temps de lecture augmenté du délai rotationnel. Nous pouvons estimer ce temps à un tour complet de la piste, soit 16,7 ms. (le premier enregistrement de la chaîne = le délai moyen + lecture, soit 8,5ms.)			

c) Ce qui nous donne en zone primaire :

TP	Minimum	Moyen	Maximum
+/- 2 K	47,5	77,5	127,5
+/- 3 K	48,4	78,4	128,4
6 K	51,-	81,-	131,-

d) S'il y a recherche en overflow :

Si TO = 70% pas d'overflow

Si TO = 80% overflow indépendant

		Moyenne
⇒	seek supplémentaire	25
	search	8,4
	read	0,1

en moyenne, nous avons 1,033 enregistrement en overflow après 33 semaines d'insertions.

⇒ coût en overflow 33,5 x 1,033 = 34,61 ms.

Coût total moyen si overflow indépendant : il variera entre
(pour page 2 K) 71,5 + 34,61 = 112,11 ms.
(pour page 6 K) 81 + 34,61 = 115,61 ms.

Ce coût sous réserve que l'on ne réorganise pas avant 33 semaines auquel cas nous n'aurions pas d'overflow.

Si TO = 90% (Nous nous basons sur le tableau de la page 88)
Nous avons réservé deux pistes d'overflow par cylindre + un overflow indépendant nécessaire pour subvenir aux insertions postérieures à la 33e semaine d'insertions.

Avant cette 33e semaine.

Nombre moyen en overflow : varie entre 0 et 5 enregistrements.
⇒ le temps moyen en overflow variera entre

5 → 8,5 x 5 = 42,5 ms.

Après la 33e semaine.

Il faut ajouter, en moyenne, par enregistrement d'overflow indépendant

un temps de seek	25,-	ms.
search	8,4	ms.
read	0,1	ms.
	33,5	ms.

Au maximum, 3 par piste ⟹ 101,5 ms.

e) Tableau résumé des temps moyens à 90%. (°)

Recherche en zone primaire entre 77,5 et 81 ms.
Recherche en overflow :
avant la 33e semaine ..... temps moyen maximum 123,5 ms.
après la 33e semaine ..... temps moyen maximum 182,5 ms.
(°) Nous ne nous intéresserons qu'à ce taux de chargement.

REMARQUE.

Vu sa taille, il a été choisi de ne pas mettre le cylinder index en mémoire centrale.
D'autres solutions seraient envisageables, notamment celle qui prévoierait de mettre, en mémoire centrale, une partie du cylinder index.

3. Lecture séquentielle du fichier.

Nous appliquons les principes détaillés dans la présentation de la technique que nous rappellerons brièvement ci-dessous. Nous posons comme hypothèse que I.S.A.M. ne lit pas les enregistrements vides en fin de piste, bien que le dernier bloc soit lu en entier. Pour faciliter certains calculs, les enregistrements en cylinder overflow sont considérés comme bloqués. Ce ne sera pas le cas pour ceux de l'overflow indépendant.

Tableaux de lecture.

a) Seek pour chaque cylindre de zone primaire qui est minimum (10 ms). (Nombre de cylindres)

TP \ TO	70%	80%	90%
+/- 2 K	3.997	3.518	3.113
+/- 3 K	3.824	3.350	2.981
6 K	3.590	3.141	2.792

b) Search pour chaque bloc de données + un read de chaque bloc. (L'ensemble peut être estimé à un délai rotationnel (16,7 ms.) En ms.

TP \ TO	70%	80%	90%
+/- 2 K	210	240	270
+/- 3 K	150	150	180
6 K	90	90	90

c) Search + un read pour chaque paire d'entrées du track index.
L'ensemble peut être estimé à une demi-rotation (8,4 ms.).
Par cylindre, il y a 27 ou 29 entrées, 1 par piste de zone primaire,
un temps de 226,8 ms. ou de 243,6 ms.

Total des temps de lecture pour le fichier. (En secondes)
Formule employée :

$$(10 \times C_i) + C_i \times (16{,}7 \times B_i) + D_i$$

avec C_i = les valeurs du tableau (a) ci-avant.
B_i = les valeurs du tableau (b) ci-avant.
D_i = soit 226,8 ms., soit 243,6 ms.

TP \ TO	70%	80%	90%
+/- 2 K	15.032	14.993	14.774
+/- 2 K	10.549	9.242	9.667
6 K	6.307	5.518	4.858

A ces temps, il faudrait pouvoir ajouter celui de lecture de l'overflow indépendant. Celui-ci varie de une piste à dix-neuf cylindres, suivant un taux d'occupation de 80% ou de 90%.
Pour rappel : par piste, il y a 47 enregistrements non bloqués.
Lecture d'une piste : 16,7 ms. x 47 = 785 ms.
Lecture de 19 cylindres : 448 secondes.

4. Temps de mise à jour d'un enregistrement.

Nous considérons qu'une mise à jour consiste en :

.un accès à l'enregistrement pour la lecture,

.un accès à l'enregistrement pour l'écriture,

.une relecture pour vérification (évaluée à un délai rotationnel, soit 16,7 ms.)

Nous obtenons :

S'il y a recherche en zone primaire : (en ms.)

TP.	Minimum	Moyen	Maximum
+/- 2 K	111,7	171,7	271,7
+/- 3 K	113,5	173,5	273,5
6 K	118,7	178,7	278,7

S'il y a recherche en zone d'overflow :

Pour TO à 80% - Nous avons un overflow indépendant. Le temps moyen de mise à jour variera entre 240,92 ms. (page ± 2 K) et 247,92 ms. (page 6 K).

Pour TO à 90%.

- avant la 33e semaine ⇒ uniquement cylinder overflow.
  Le temps moyen de mise à jour variera entre 171,7 ms. et 263,7 ms. selon que l'on aura, en moyenne, 0 ou 5 enregistrements en overflow.
- après la 33e semaine :
  . si cylinder overflow : le temps moyen de mise à jour sera le même que ci-avant.
  . si overflow indépendant : le temps moyen de mise à jour variera entre 171,7 ms. et 381,7 ms.

5. Temps d'insertion d'un enregistrement dans le fichier.

Beaucoup de cas distincts peuvent se présenter. Nous nous limiterons à une solution simplifiée.

5.1 - Ajout en zone primaire.

Sera considéré comme temps d'accès + temps de recopiage + temps de déplacement des autres blocs à droite (en moyenne 1/2 piste).

5.2 - Ajout en overflow.

Sera considéré comme le temps de 5.1 augmenté de l'écriture de deux enregistrements en overflow (ceci afin de réaliser le chaînage).

5.3 - Ajout en overflow indépendant.

Ecriture de deux enregistrements (mais ici, il faut les rechercher (⇒ temps seek)) en plus du temps 5.1

Evaluation.

5.1 = en moyenne

temps d'accès	entre 77,5 ms. et 81 ms.
recopiage avec délai	16,4 ms.
temps recopiage 1/2 piste (sans délai)	7,5 ms.
Total :	entre 101,4 ms. et 104,9 ms.

5.2 = délai pour écrire le Ier enregistrement 8,4 ms.
recopiage de deux enregistrements + délai 8,4 ms. + 0,2 ms.
} 17 ms.

Total : entre 118,3 ms. et 121,8 ms.

5.3 = en moyenne

temps seek	25 ms.
temps search	8,4 ms.
temps écriture	0,1 ms.

} x 2 = 67 ms.

Total : entre 144,5 ms. et 148 ms.

6. La Réorganisation.

Préliminaires.

Nous ne pouvons pas, ici, calculer le coût et le moment d'une réorganisation pour toutes les possibilités envisagées. Aussi, nous limiterons nous à un exemple le plus représentatif, à savoir : un taux d'occupation à 90% et une taille de pages de 6 K. Nous rappelons que nous avons choisi de réorganiser le fichier lorsque le total des pertes de temps dues aux insertions est proche ou supérieur au coût de la réorganisation.

Coût d'une réorganisation.

La réorganisation consiste en une lecture du fichier et un rechargement du fichier, celui-ci pouvant être estimé à un temps équivalant une lecture du fichier. Nous obtenons donc un temps de deux lectures séquentielles, à savoir 4.858 sec. x 2 = 9.716 secondes + un temps de lecture de l'overflow indépendant lequel varie, selon la taille.

Quand réorganiser ?

Evaluation moyenne des pertes dues aux enregistrements en cylinder overflow.

Nous allons, pour ce point, nous baser sur les résultats présentés au tableau de la page 88. Nous prendrons, comme nombre moyen d'insertions, la moyenne des trois nombres donnés dans les différents tableaux.
Ex.- Après 35 semaines, le nombre moyen d'enregistrements en débordement, pour TP = 6 K, sera de (2,2 + 2,09 + 5,56) : 3 = 3,28

Nombre d'insertions en overflow après 39 semaines.

série 36	3582,44 x 3,14	= 11.249
série 37	4502,49 x 2,98	= 13.418
série 50	5759,65 x 7,59	= 43.716
		68.383

La probabilité de se retrouver en overflow durant une consultation ou une mise à jour peut être estimée à :

68.383 divisé par le nombre total d'enregistrements, soit 0,0065

Total des pertes subies après 39 semaines.

A - En insertion.

Pour les séries 36 et 37 : toutes peuvent trouver place en cylinder overflow. Total = 24.667

Pour la série 50, seulement 22.184 peuvent aller en cylinder overflow, ce chiffre étant obtenu en multipliant par (47 x 2) le nombre de pistes de cylinder overflow (rappel : 2 pistes par cylindre).

⇒ 21.532 se trouveront en overflow indépendant (ce chiffre représente 31% des insertions).

Coût perdu pour les insertions.

a) en cylinder overflow. (deux écritures)

0,017 sec. x (24.667 + 22.184) = 797 secondes.

b) en overflow indépendant.

0,067 sec. x 21.532 = 1.443 secondes.

B - En consultation.

1) Nombre de consultations après 39 semaines :

5.000 x 5 x 50 = 1.000.000

2) Nombre de consultations en overflow :

975.000 x 0,0065 = 6.338

3) Parmi les 6.338 consultations en overflow, 31% (soit 1.965) se trouvent en overflow indépendant.

Pertes subies par enregistrement.

3.1 - Nombre moyen d'enregistrements par piste en cylinder overflow

$$\frac{\left(3{,}14 + 2{,}98 + (7{,}59 - 31\%)\right)}{3} = 3{,}79$$

Le coût sera conformément à ce que nous avons vu pour la consultation, à savoir :

8,5 ms. x 3,79 = 32,215 ms.

3.2 - Nombre moyen d'enregistrements par piste en overflow indépendant.

31% de 7,59 = 2,35

Le coût sera conformément à ce que nous avons vu pour la consultation, à savoir :

33,5 ms. x 2,35 = 78,725 ms.

4) Coût total des pertes en consultation, après 39 semaines.

4.1 0,032215 sec. x (6.338 - 1.965) = 141 secondes.

4.2 0,078725 sec. x 1.965 = 155 secondes.

296 secondes.

C - En Mise à jour.

Le raisonnement est le même que pour les consultations.

1) Nombre de mises à jour par semaine :

a/ normales + rejet 50.000 x 5 = 250.000 (une semaine sur quatre, on ajoute une moyenne de 85.000 mises à jour, pour les comptes à terme).

b/ opérations comptables et non comptables 1.000 + 18.750 = 19.750

c/ recapitalisation 1.000.000 par an, soit ± 25.000 par semaine.

d/ intérêts 4.000.000 par an, soit ± 100.000 par semaine.

e/ moyenne par semaine 416.000

2) Nombre de mises à jour après 39 semaines :

416.000 x 39 = 16.224.000

3) En overflow, nous en aurons :

16.224.000 x 0,0065 = 105.456 dont 31% en overflow indépendant, soit 32.692

4)Pertes : elles sont, par enregistrement, le double de celles des consultations, soit :

pour le cylinder overflow 4.689 secondes

pour l'overflow indépendant 5.148 secondes

Total des pertes subies après 39 semaines.

Insertion (A) + Consultation (B) + Mises à jour (C).

(797 + 1.443) + (141 + 155) + (4.689 + 5.148) = 12.373 (sec.)

ce qui est supérieur au coût de réorganisation, celui-ci étant égal à 9.716 secondes + un temps de lecture de l'overflow indépendant.

Taille de l'overflow indépendant après 39 semaines : 459 pistes, soit un temps de 361 sec. x 2 = 722 secondes.

Coût de réorganisation : 9.716 sec. + 722 sec. = 10.438 secondes.

Par un raisonnement analogue, nous calculons les pertes après la 38e semaine.

Le total des pertes, après 38 semaines, s'élève à 10.112 secondes pour un temps de réorganisation (lecture de l'overflow indépendant comprise) de 10.306 secondes.

Nous proposerions donc une réorganisation après 38 semaines d'exploitation, conformément à l'option choisie dans les préliminaires, page 95.

7. Influence sur l'existant.

La solution "séquentiel indexé" existe actuellement et elle est donc connue. Elle ne devrait plus poser de problèmes.
Néanmoins, il faudrait, dans les programmes, pouvoir permettre l'insertion et la modification de longueur des enregistrements. Ne connaissant pas ces programmes, nous ne pouvons évaluer l'importance de cette tâche.

==========

B. - EVALUATION DES APPLICATIONS.

Introduction.

. Nous travaillerons comme nous l'avons fait pour la réorganisation, à savoir avec un taux d'occupation (TO) égal à 90% et une taille de page (TP) de 6 K.

. La réorganisation étant proposée après la 38e semaine, nous pouvons avoir à ce moment, 60.675 enregistrements en overflow dont 17.618 en overflow indépendant (mêmes calculs que lors de la réorganisation).

. La probabilité de se retrouver en overflow, aussi bien lors d'une consultation que lors d'une mise à jour, est égale au nombre d'enregistrements en overflow sur le nombre total d'enregistrements, soit 0,0057 %.

. Après 38 semaines, nous avons, en moyenne, 3,54 enregistrements par piste sur le cylinder overflow, ce qui nous procure un coût supplémentaire de 30,09 ms. en consultation et de 60,18 ms. en mise à jour. Par contre, nous obtenons un coût supplémentaire de 68,685 ms. en overflow indépendant pour la consultation et de 137,37 ms. pour la mise à jour.

. Nous nous placerons dans la situation la plus défavorable en ce sens que nous allons considérer que nous sommes en 38e semaine, ce qui nous donnera pour les semaines précédentes, des temps moyens par excès.

1. Consultation.

. Rappel : a) 5.000 consultations journalières.
b) une consultation est un accès aléatoire.

. Nombre de consultations en overflow :
5.000 x 0,005 = 29
Parmi ces 29 consultations, il y en a 9 en overflow indépendant.

. Coût de la consultation journalière :

4.971 x 81 ms.	=	402.651 ms.
20 x (81 + 30,09)ms.	=	2.221,8 ms.
9 x (81 + 68,685) ms.	=	1.347,165 ms.
Total		406.218,965 ms. soit ± 407 secondes.

2. Mise à jour (journalière).

. Rappel : a) nous évaluerons 50.000, 100.000 et 150.000 mises à jour journalières.
b) une mise à jour est constituée par deux accès aléatoires et une lecture de contrôle.

. Nombre de mises à jour en overflow :
50.000 x 0,0057 = 285 (avec 83 en overflow indépendant au maximum).

. Coût moyen pour 50.000 opérations journalières.

49.715 x 0,1737 sec.	=	8.885 sec.
202 x (0,1787 + 0,06018)sec.	=	49 sec.
83 x (0,1787 + 0,13737)sec.	=	27 sec.
Total		8.961 sec. = +/- 2 H.30

. Coût moyen pour 100.000 opérations journalières :
8.961 sec. x 2 = 17.929 secondes = +/- 5 Heures.

. Coût moyen pour 150.000 opérations journalières :
8.961 sec. x 3 = 26.883 secondes = +/- 7 H.30

Remarque.

L'évaluation de 150.000 opérations de mises à jour journalières prend un temps considérable. Nous pourrions envisager une amélioration en triant, au préalable, les mises à jour, ce qui diminuerait le temps moyen de déplacement du bras du disque lors de la mise à jour d'un grand nombre d'enregistrements.
Il ne faut cependant pas perdre de vue que lors d'un rejet en litiges, pour la plupart des mises à jour journalières, c'est tout le bordereau qui doit se retrouver en litiges. Etant triées, les mises à jour ne seraient plus par bordereaux, ce qui nécessiterait un travail supplémen-

taire de reconstitution de bordereaux pour les cas litigieux.
Nous obtiendrons un temps qui approche celui d'une lecture séquentielle du fichier. (Cfr. pages 92-93).

3. Capitalisation.

. Coût : a) une prise de copie du fichier sur bande est estimée à une lecture séquentielle.
b) un rechargement du fichier sur les disques est estimé à une lecture et une écriture séquentielles.

. Nous obtenons un total de deux lectures séquentielles, soit à rappeler : 4.858 sec. x 2 = 9.716 secondes.(+ éventuellement le temps de lecture de l'overflow indépendant variant de 0 à 448 sec. au maximum).

. Le temps de traitement n'est pas évalué et le temps de lecture est donné pour un fichier chargé à 90% et une taille de page de 6 K.

C. - PRISE DE COPIE TRIEE.

Elle est évaluée à une lecture séquentielle, soit 4.858 secondes + éventuellement un temps de lecture de l'overflow indépendant, variant entre 0 et 448 secondes au maximum.

.....

5. Deuxième répartition retenue dans la technique du séquentiel indexé : deux fichiers pour le C.C.D., un pour sa partie évolutive, l'autre pour sa partie fixe.

Introduction.

Nous n'étudierons que les différences par rapport au type de répartition précédent; elles se situent à deux critères, à savoir :
- la place occupée par le fichier;
- la réorganisation;
et sur les applications : consultation, mise à jour et prise de copie triée.

Nous ne développerons plus les calculs, ceux-ci étant les mêmes que dans la technique précédente.

1. Place occupée par les fichiers.

1.1 - Partie fixe.

Rappel : 6.200.000 enregistrements.
taux d'occupation (TO) choisi : 95%
longueur moyenne d'enregistrement : 130 bytes.

Taille du fichier.

	TP = +/- 2 K	+/- 3 K	6 K
Nombre d'enregistrements par piste	119	125	133
Nombre de pistes	52.101	49.600	46.617
Nombre de pistes track index par cylindre	1	1	1
Nombre de cylindres	1.797	1.711	1.608
Place cylinder index	27 pistes } °	26 pistes } °	24 pistes } °
Place master index	1 piste }	1 piste }	1 piste }
Nombre de disques	3,24	3,08	2,9

° Cela représente +/- un cylindre.

1.2 - Partie évolutive.

Rappel : 4.000.000 enregistrements.
taux d'occupation (TO) : 70% - 80% - 90%
longueur d'enregistrement : 130 bytes (140 en overflow).

Nombre de pistes nécessaires.

TP \ TO	70%	80%	90%
+/- 2 K	45.455	40.000	35.399
+/- 3 K	43.479	38.096	33.899
6 K	40.817	35.715	31.747

Nombre de cylindres et de disques (y compris place du track index et du master index)(entre 19 et 25 pistes, soit +/- un cylindre).

TP \ TO	70%	80%	90%
+/- 2 K	1.569 (2,83)	1.381 (2,49)	1.313 + 19 (2,40)
+/- 3 K	1.501 (2,70)	1.315 (2,37)	1.257 + 19 (2,30)
6 K	1.409 (2,54)	1.233 (2,22)	1.177 + 19 (2,15)
			@ 0.I.

@ place occupée par l'overflow indépendant.

1.3 - Place totale du fichier.

à 70%	6 ou 7 disques.
à 80%	6 disques.
à 90%	6 disques.

2. La réorganisation.

Coût.

. Nous ne réorganiserons que la partie évolutive du fichier, à savoir 4.000.000 d'enregistrements. Nous nous placerons dans les mêmes conditions que pour la première technique de répartition, à rappeler : TP = 6 K et TO = 90%.

. Le coût est évalué à deux lectures séquentielles de ce fichier.

. Une lecture séquentielle coûte 2.047 secondes + un temps éventuel de lecture en overflow indépendant.

. Total des deux lectures :
2 x (2.047 sec. + temps de lecture de l'overflow indépendant), soit 4.094 + (2 fois le temps de lecture de l'overflow indépendant).

Quand réorganiser ?

. Total des pertes après :

35 semaines : 3.971 secondes pour un coût de réorganisation = 4.302

36 semaines : 5.393 secondes pour un coût de réorganisation = 4.426

. Nous proposerions donc une réorganisation après 35 semaines (même principe que pour la première technique de répartition).

3. Evaluation des consultations et mises à jour.

Introduction.

Après 35 semaines, nous avons, en moyenne, 3,01 enregistrements par piste sur le cylinder overflow, ce qui nous procure un coût supplémentaire de 25,585 ms. en consultation et de 51,17 ms. en mise à jour. En ce qui concerne l'overflow indépendant, nous avons une dégradation moyenne de 27,938 ms. en consultation et de 55,88 ms. en mise à jour.

Nous nous plaçons dans la situation la plus défavorable en ce sens que nous allons considérer que nous sommes en 35e semaine, ce qui nous donnera pour les semaines précédentes, des temps moyens par excès.

La probabilité de se retrouver en overflow lors d'une consultation ou lors d'une mise à jour est égale au nombre d'enregistrements en overflow divisé par le nombre total d'enregistrements, à savoir, après 35 semaines : 0,0038.

Consultation (journalière)

Nombre de consultations en overflow :

5.000 x 0,0038 = 19 (dont 3 en overflow indépendant)

⇒ coût moyen par excès :

4.981 x 81 ms.	=	403.461 ms.
16 x (81 + 25,585)ms.	=	1.705,36 ms.
3 x (81 + 27,938)ms.	=	326,814 ms.
Total		405.493,174 ms., soit 406 secondes.

Mise à jour (journalière)

Nombre de mises à jour en overflow :

50.000 x 0,0038 = 190 (dont 29 en overflow indépendant)

⇒ coût moyen par excès pour 50.000 opérations :

49.810 x 178,7 ms.	=	8.901.047 ms.
161 x (178,7 + 51,17)ms.	=	37.009,07 ms.
29 x (178,7 + 55,88)ms.	=	6.802,82 ms.
Total		8.944.858,89 ms., soit 8.945 secondes ou +/- 2 H.30

coût moyen par excès pour 100.000 opérations :

8.945 secondes x 2 = 17.890 secondes, soit +/- 5 Heures.

coût moyen par excès pour 150.000 opérations :

8.945 secondes x 3 = 26.835 secondes, soit +/- 7 H.30 (°)

(°) Même remarque que pour le premier type de répartition du fichier.
(Cfr. page 99)

4. Prise de copie triée et capitalisation.

Hypothèse : Nous pouvons lire simultanément sur les deux fichiers afin d'éviter une fusion. Le temps de lecture sera alors la somme des temps de lecture des deux fichiers.

Fichier fixe	2.825 secondes	(TP = 6 K et TO = 95%)
Fichier évolutif	2.047 secondes	(TP = 6 K et TO = 90%)
Total	4.872 secondes + éventuellement de 0 à 448 secondes pour l'overflow indépendant.	

Remarque. - La capitalisation, évaluée à deux lectures séquentielles, coûtera 9.744 secondes + éventuellement un temps de lecture de l'overflow indépendant, variant entre 0 et 448 secondes.

6. Comparaison des types de répartition choisis en séquentiel indexé.

Pour deux raisons, nous donnerions notre préférence à la seconde technique de répartition :

1) elle nécessite moins de place physique que la première (au minimum 215 cylindres en moins);
2) la réorganisation peut se limiter à une partie des enregistrements, ce qui en diminue fortement le coût tout en n'obligeant presque pas une fréquence de réorganisation supplémentaire.
   Fréquence :
   1 après 38 semaine dans la Ière répartition;
   1 après 35 semaine dans la 2ème répartition.

Les temps de consultation et de mise à jour journalière sont, quant à eux, très proches dans les deux répartitions.

Le temps de lecture séquentielle est pratiquement semblable dans les deux répartitions : il ne diffère que de 14 secondes.

TECHNIQUE II. - VIRTUAL STORAGE ACCESS METHOD (V.S.A.M.)

1. Présentation des différents éléments influençant l'évaluation.

- Capacité de la piste, exprimée en bytes. Elle est déterminée par le type du disque utilisé. Actuellement, elle est de 19.069 bytes.
- Longueur des enregistrements (LE). La longueur des enregistrements est une moyenne, à savoir : 130 bytes.
- Taille des control interval (TCI). Nous ne parlons plus de taille de page mais de taille de control interval, pour le regroupement d'enregistrements en bloc afin de bien utiliser l'espace réservé.
- Taille des control area (CA). Une control area (regroupant toujours un même nombre de CI) peut aussi avoir une taille variant entre une piste et un cylindre.
- Taux d'occupation du fichier (TO); en vue d'absorber les insertions futures, le fichier sera chargé à un taux d'occupation variable. En réalité, nous aurons de l'espace libre par CI et aussi par CA.

. Nombre de niveaux d'index.

. Nombre de recopies du niveau d'index , ceci afin de diminuer le délai rotationnel lors de la recherche dans l'index.

. Taille des index (TCi) : elle dépend de la taille des CA.

. Place en mémoire centrale disponible "buffers" pour les index et pour les données.

Remarque. - Dans chaque CI, il faudra tenir compte des informations de contrôle. Il en sera de même pour les control interval d'index (CIi).

2. Equations analytiques.

. Nombre d'enregistrements par CI.

$$A = \frac{\text{TCI - les informations de contrôle}}{\text{LE}}$$

en tenant compte du taux d'occupation des CI.

. Nombre d'enregistrements par CA.

nombre de CI par CA $\quad B = TCA / TCI$

en tenant compte du taux d'occupation

nombre d'enregistrements par CA $\longrightarrow \quad C = A \times B$

. Nombre de cylindres

D = (le nombre d'enregistrements / C) + 1

. Nombre de disques

(D / 555) + 1

. Nombre de niveaux d'index

- Au niveau du sequence set, il y a, dans un enregistrement, autant de clés que de CI par CA. La taille de l'enregistrement en dépend. Elle est fixée au plus petit multiple de 512 bytes qui convient. Il y aura donc D enregistrements à ce niveau si l'on considère une CA de taille 1 cylindre (°)

- Au deuxième niveau, la taille des enregistrements reste la même. Par enregistrement, nous pouvons entrer N clés; si $N \leq D$, il nous faudra plus d'un enregistrement et, par conséquent, un niveau supérieur où l'on appliquera les mêmes principes jusqu'à obtenir un enregistrement au dernier niveau.

(°) Voir remarque en page 106.

3. Détermination de l'overflow.

Lorsque l'on insère dans des control interval de V.S.A.M., c'est la technique de répartition sur deux pages (2 CI) qui est utilisée en cas de défaut de place.

Cette technique a été simulée pour un taux d'occupation initial de 80%. Les résultats obtenus ont été comparés avec d'autres résultats de simulation (cfr. référence - Bibliographie, p.11)

Obtenant des similitudes, nous avons choisi de nous baser sur ces résultats pour d'autres taux d'occupation initiaux.

4. Première répartition dans la technique de V.S.A.M. : un fichier pour tout le C.C.D.

A. - EVALUATION SELON LES CRITERES RETENUS.

1. Place occupée par le fichier.

A. Nombre d'enregistrements 10.200.000
Longueur des reads 130 bytes
Taille des CI 2 K, 4 K, 6 K (°)
Taille des CA 1 cylindre (°) (= 30 pistes)
Longueur enregistrement index 8 bytes (°°)
Taille informations contrôles : fixe 4 bytes + 3 bytes par enregistrement de longueur différente. Par défaut, nous choisirons 3 bytes par enregistrement.

B. Nombre d'enregistrements par CI.

TCI \ TO	70%	80%	90%
2 K	11	12	14
4 K	22	25	28
6 K	32	37	42

C. Nombre de CI pour tout le fichier sur base de 10.200.000 enregistrements.

TCI \ TO	70%	80%	90%
2 K	927.273	850.000	728.572
4 K	463.637	408.000	364.286
6 K	318.750	275.676	242.858

(°) Dimensions proposées dans la référence "Bibliographie, n° 8".

(°°) Grâce au mécanisme de compression de clés (Réf. Bibliographie, n° 8)

D. Nombre de CI possibles par CA.

	par piste	par cylindre = par CA.
2 K	8	240
4 K	4	120
6 K	3	90

E. Place pour l'overflow par control area.

a) Résultats des simulations.

Les simulations réalisées en cours de stage l'ont été avec des tailles de CI (pages) égales à 1, 2 et 4 K. Ces tailles correspondent aux nombres d'enregistrements par page réalisés dans les simulations reçues (10, 20, 40).

Nous allons, dans un premier temps, dresser un tableau avec les résultats obtenus et, dans un deuxième temps, un autre avec des résultats extrapolés. Ces résultats pourront être consultés en annexe.

TO	Taille page en enregistrements	% de pages (CI) ajoutées par rapport au nombre final d'insertions d'enregistrements.
70 %	20	0,071
80 %	10	0,15
	20	0,08
	40	0,04
90 %	20	0,08

TO	Taille CI en K.	% de pages (CI) ajoutées par rapport au nombre final d'insertions d'enregistrements.
80 %	2	0,08
	4	0,04
	6	0,03
70 %	2	0,081
	4	0,035
	6	0,026
90 %	2	0,08
	4	0,04
	6	0,03

b) Nombre de CI d'overflow par CA choisi.

TCI \ TO	70%	80%	90%
2 K	18	20	20
4 K	9	10	10
6 K	7	8	8

F. Nombre de cylindres nécessaires pour le fichier.

TCI \ TO	70%	80%	90%
2 K	4.177	3.864	3.312
4 K	4.177	3.710	3.312
6 K	3.841	3.362	2.962

G. Evaluation de l'index.

. Taille entrée index (toujours un multiple de 512 bytes).
Nombre d'entrées d'index par CA de 1 cylindre.

Si TCI = 2 K 240 x 8 (long.entrée index)= 1920 taille CIi = 2 K
= 4 K 120 x 8 (long.entrée index)= 960 taille CIi = 1 K
= 6 K 90 x 8 (long.entrée index)= 720 taille CIi = 1 K

. Nombre de CIi au sequence set.
Un par control area, donc 1 par cylindre.
Résultats : dans le tableau F.

. Nombre de CIi au deuxième niveau.
Un CIi de 1 K peut contenir 121 entrées de 8 bytes tandis qu'un CIi de 2 K peut en contenir 248 (°).
Ce qui nous donne comme résultats, au deuxième niveau :

TCI \ TO	70%	80%	90%
2 K	17	16	14
4 K	35	31	28
6 K	32	28	25

. Nombre de CIi au 3e niveau : un seul.

. Taille de l'index (en pistes et en cylindres, entre parenthèses), sur l'ensemble des trois niveaux en tenant compte du fait qu'une piste peut contenir en V.S.A.M. 15 blocs de 1 K et 8 de 2 K (°)

TCI \ TO	70%	80%	90%
2 K	525 (18)	486 (17)	416 (14)
4 K	281 (10)	250 (9)	223 (8)
6 K	259 (9)	227 (8)	200 (7)

(°) Cfr. référence "Bibliographie, n° 8".

H. Taille globale du fichier en cylindres si l'on recopie une fois, deux fois ou trois fois l'index.

TCI \ TO	70%			80%			90%		
	1x	2x	3x	1x	2x	3x	1x	2x	3x
2 K	4195	4213	4231	3881	3898	3915	3326	3340	3354
4 K	4187	4197	4207	3719	3728	3737	3320	3328	3336
6 K	3850	3859	3868	3370	3378	3386	2969	2976	2983

I. Taille du fichier en disques si l'on recopie 1, 2 ou 3 fois l'index.

TCI \ TO	70%			80%			90%		
	1x	2x	3x	1x	2x	3x	1x	2x	3x
2 K	7,56	7,60	7,62	6,99	7,02	7,05	5,99	6,02	6,04
4 K	7,54	7,56	7,58	6,70	6,72	6,73	5,98	6,00	6,01
6 K	6,94	6,95	6,97	6,07	6,09	6,10	5,35	5,36	5,37

2. Temps d'accès aléatoire.

. Une recherche d'un enregistrement nécessiterait un accès par niveau d'index + un accès à l'enregistrement, ce qui nous donnerait un total de quatre accès par enregistrement.

. Nous pouvons améliorer ce total en admettant que le 3ème niveau d'index se trouve toujours en mémoire centrale (1 CIi de 1 ou 2 K)

. Nous pourrions envisager le cas où tout le 2ème niveau d'index se trouverait lui aussi en mémoire centrale; il nous faudrait, en plus du "buffer" nécessaire pour le 3e niveau, B "buffers" supplémentaires et, de ce fait, nous n'aurions plus qu'un seul accès à l'index.

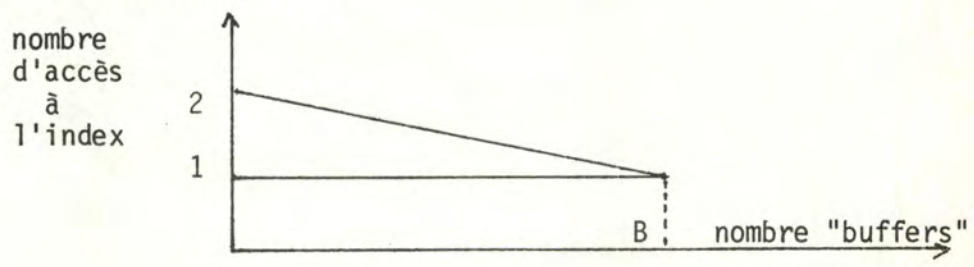
Soit N = le nombre d'accès à l'index

N^2 = le nombre d'enregistrements au niveau 2 de l'index.

Alors :

$$N = \frac{N_i - (B - 1)}{N_i} + \boxed{\frac{N_i^2 - 1}{N_i^2}} \simeq 1 \text{ pour } N_i \text{ grand}$$

Ce qui donne comme figure :



Valeur de B.

$B = N_i + 1$ si nous souhaitons que le niveau 2 soit en mémoire centrale, ce qui nous donne :

TO / TCI	70%	80%	90%
2 K	18	17	15
4 K	36	32	29
6 K	33	29	26

. Temps d'accès à l'index.

Dans le cas où tout le niveau 2 se trouve en mémoire centrale :

1 seek + 1 search + 1 read d'un enregistrement de 1 ou de 2 K

En moyenne, cela nous donne :

25 ms. + 8,4 ms. + 0,8 ms. ou 1,7 ms. $\Rightarrow$ Total = 34,2 ms. ou 35,1 ms.

Dans le cas où le deuxième niveau ne se trouve pas en mémoire centrale, le déplacement maximum est de 18 cylindres; donc, avec un déplacement maximum de 11,23 ms. et minimum de 10 ms. pour un déplacement de 1 cylindre.

Nous nous baserons sur 10,6 ms. pour le déplacement du bras du disque. En moyenne, cela nous donne :

10,5 ms. + 8,4 ms. + 0,8 ms. ou 1,7 ms. $\Rightarrow$ Total = 19,8 ms. ou 20,7 ms.

. Temps d'accès à l'enregistrement.

Il consiste en un seek + un search + un read d'un bloc de 2, 4 ou 6 K.

(Temps en ms.)

		Minimum	Moyen	Maximum
seek		10	25	50
search		8,4	8,4	8,4
read	2 K	1,7	1,7	1,7
	4 K	3,3	3,3	3,3
	6 K	5	5	5

. Temps total de consultation.

1 accès index + 1 accès aux données.

TCI.	Minimum	Moyen	Maximum
2 K	55,2 ms.	70,2 ms.	95,2 ms.
4 K	55,9 ms.	70,9 ms.	95,9 ms.
6 K	57,6 ms.	72,6 ms.	97,6 ms.

2 accès index + 1 accès aux données.

2 K	75,9 ms.	90,9 ms.	115,9 ms.
4 K	75,7 ms.	90,7 ms.	115,7 ms.
6 K	77,4 ms.	92,4 ms.	117,4 ms.

Remarque.

Amélioration possible.

. Placer l'index au milieu des données et recopier l'index 2 ou 3 fois.
Le fichier étant constitué de plusieurs volumes, mettre l'index au milieu d'un disque ne permettrait un gain de temps que pour ce disque et nous n'en tiendrons pas compte (excepté pour la lecture séquentielle du fichier).

. Par contre, la recopie de 2 ou 3 fois l'index, permet un gain de temps :
2 fois en moyenne le temps de search +/- 4,2 ms.
3 fois en moyenne le temps de search +/- 2,8 ms.

. Néanmoins, elle augmente de +/- 1,5 ms. le temps de déplacement du bras du disque pour l'accès au niveau suivant de l'index, puisque ce dernier niveau comportera 36 ou 54 cylindres au lieu des 18 initiaux. Ceci n'est vrai que dans la mesure où il y a deux recherches dans l'index.

Trop de possibilités seraient à envisager et compliqueraient la lecture. Nous ne calcurerons donc pas les gains totaux possibles dans la suite de l'étude. Nous nous limiterons à rappeler, chaque fois, que des améliorations seraient possibles.

3. Lecture séquentielle du fichier.

Pour la lecture séquentielle, il faut rappeler que les enregistrements du premier niveau de l'index (le sequence set) sont chaînés.

. Lecture d'un cylindre (°) - en ms.

TCI \ TO	70%	80%	90%
2 K	3.507	4.008	4.509
4 K	1.503	2.004	2.004
6 K	1.503	1.503	1.503

. Il faut ajouter, à ces temps, le temps d'accès au cylindre et le temps de search, estimés à 25 ms. + 8,4 ms., soit 33,4 ms.

. Il faut également ajouter le temps d'accès à l'index, le temps de search et le temps de read de l'entrée d'index, soit :

17,5 ms. + 8,4 ms. + 0,8 ms. = 26,7 ms. ou
ou 1,7 ms. 27,6 ms.

. Total du temps de lecture. (En secondes)
Le nombre de cylindres de données x par le temps donné ci-dessus.

(°) 1 search + 1 read d'un bloc estimé à 16,7 ms.
Nous posons que V.S.A.M. reconnaît les blocs vides en fin de piste mais ne les lit pas.

TO / TCI	70%	80%	90%
2 K	14.904	15.723	15.136
4 K	6.530	7.658	6.837
6 K	6.004	5.144	4.630

4. Mise à jour d'un enregistrement.

Nous considérons qu'une mise à jour consiste en :

.un accès à l'enregistrement pour la lecture;

.un accès à l'enregistrement pour l'écriture;

.une relecture pour vérification (évaluée à une rotation complète, soit 16,7 ms.)

Nous obtenons alors :

avec 1 accès à l'index et 1 accès aux données :

TCI	Minimum	Moyen	Maximum
2 K	127,1 ms.	157,1 ms.	207,1 ms.
4 K	128,5 ms.	158,5 ms.	208,5 ms.
6 K	131,9 ms.	161,9 ms.	211,9 ms.

avec 2 accès à l'index et 1 accès aux données :

2 K	168,5 ms.	198,5 ms.	248,5 ms.
4 K	168,1 ms.	198,1 ms.	248,1 ms.
6 K	171,5 ms.	201,5 ms.	251,5 ms.

Les améliorations et pertes dues aux variantes possibles de l'index restent d'application.

Remarque. - Mise à jour de l'index.

Il se met à jour lors d'un split de CA ou lors d'un ajout de niveau supplémentaire (phénomènes n'entrant pas en ligne de compte dans notre étude °). Il faut souligner qu'à ce moment, il faut mettre à jour 1, 2 ou 3 fois l'index, suivant qu'il est recopié 1, 2 ou 3 fois. S'il y a recopie 2 fois, cela nécessite 8,4 ms. de délai de rotation et s'il y a recopie 3 fois, le délai de rotation est de 5,57 ms. Ces délais s'ajoutent au temps de mise à jour de l'index.

5. Temps d'insertion d'un enregistrement dans le fichier.

(A) Aussi longtemps qu'il y a de la place dans le CI, le temps d'insertion sera évalué au temps d'accès + le temps de recopiage du CI.

(B) Dans le cas contraire °° :

1) le temps d'insertion sera évalué au temps d'accès

2) + le temps de "split" de CI

3) + le temps de mise à jour de l'index.

° Voir remarque page 114.

°° Nous n'envisagerons pas, ici, l'éclatement d'une control area - Voir remarque, page 114.

Evaluation de (A)

Résultats des tableaux de la page 110 auxquels il faut ajouter 1,7ms., 3,3ms. ou 5ms. pour le recopiage du CI, selon que la taille de celui-ci est égale à 2, 4 ou 6 K.

Evaluation de (B)

1) Temps d'accès (Cfr. tableaux page 110)

2) Temps de "split" de CI.:

. recherche d'un CI $M \times \begin{pmatrix} \\ 8{,}4 \text{ ms.} + \\ \end{pmatrix} \begin{matrix} 1{,}7 \text{ ms. ou} \\ 3{,}3 \text{ ms. ou} \\ 5 \text{ ms.} \end{matrix}$ (°)

. search + écriture de 2 CI.

$2 \times \begin{pmatrix} \\ 8{,}4 \text{ ms.} + \\ \end{pmatrix} \begin{matrix} 1{,}7 \text{ ms. ou} \\ 3{,}3 \text{ ms. ou} \\ 5 \text{ ms.} \end{matrix}$

3) Temps de mise à jour de l'index :

. recherche de l'index (seek) 17,5 ms. en moyenne

. search de l'enregistrement 8,4 ms.

. écriture de l'enregistrement 0,8 ms. ou 1,7 ms. (selon la taille CIi : 1 K ou 2 K)

6. La réorganisation.

Préliminaires.

Nous ne pouvons pas reprendre la réorganisation pour toutes les possibilités envisagées jusqu'ici. Nous nous limiterons à un exemple : celui pour lequel nous possédons le plus de résultats de simulation, à savoir un taux d'occupation de 80% et un TCI égal à 4 K.

Nous rappelons que nous avons choisi de réorganiser le fichier lorsque le total des pertes de temps dues aux insertions est proche ou supérieur au coût de la réorganisation.

Coût de la réorganisation.

Elle consiste en la lecture du fichier et en un rechargement de celui-ci estimé à un temps équivalant une lecture séquentielle. Le temps total sera donc de deux lectures du fichier, soit :

7.658 secondes x 2 = 15.316 secondes.

Quand réorganiser ?

. Au sujet de l'index.

Nous obtiendrons un accès supplémentaire lorsqu'il y aura un 4e niveau d'index. Pour cela, il faudrait 91 entrées supplémentaires au niveau 2

(°) $M \leq N$; N étant le nombre de CI d'overflow pour control area, M sera choisi à la valeur moyenne du nombre de CI d'overflow par CA.

ce qui correspond à 11.132 au niveau 1 (le sequence set) et donc à 11.132 cylindres de données supplémentaires, ce qui ne pourrait pas arriver avant un an.

. Perte de temps à l'insertion des enregistrements.

Nous rappellerons qu'elles sont de deux types :

1) lors de la répartition d'un CI sur 2 CI ("split" de CI)

2) lors de la répartition d'une CA sur 2 CA.("split" de CA)

REMARQUE.

Si les simulations tiennent compte de la répartition des CI, elles ne le font pas quant à la répartition des CA.

Les renseignements obtenus sur V.S.A.M. ne nous permettent pas de nous faire une idée précise quant à la répartition d'une CA sur deux CA. ("Une CA est divisée en deux"). Il faut cependant remarquer que les résultats de simulation de répartition sur deux CI nous procurent un taux d'occupation tendant vers les 70%. Si, dès lors, la CA est divisée en deux, nous allons obtenir un taux d'occupation de 35% (!), ce qui, il faut le reconnaître, n'est pas favorable.

Pour ces raisons, ne voulant pas nous lancer dans des hypothèses qui ne sont pas vérifiées, nous n'étudierons que la répartition des CI sur 2 CI et n'évoquerons plus la répartition des CA.

RAPPEL.

Les insertions ne se font que dans des parties bien précises du fichier CCD, c'est-à-dire dans les séries 36, 37, 50, 81 à 83 et ce, à une fréquence de +/- 108.000 par an pour la série 36
+/- 140.000 par an pour les séries 37 et 50
+/- 400 par an pour les séries 81, 82 et 83.

.....

A un TO = 80% et TCI = 4 K :

la série 36 comprend 32.800 CI soit 299 cylindres ou CA.
la série 37 comprend 43.200 CI soit 393 cylindres ou CA.
la série 50 comprend 32.000 CI soit 291 cylindres ou CA.
les séries 81 à 83
comprennent 48.000 CI soit 441 cylindres ou CA.

L'influence des insertions dans les séries 81 à 83 étant presque nulle, nous ne nous en préoccuperons plus.

Méthode.

Nous nous baserons sur les résultats de simulation et du principe suivant :

Si, pour un fichier de 500 pages, 100 insertions donnent un taux d'occupation de 70%, alors pour un fichier de 1000 pages, après 200 insertions, nous retrouverons le même taux d'occupation.

Ce rapport peut aussi se constater sur les ajouts de CI (un ajout de CI est dû au "split" de CI lorsqu'il n'y a plus de place).

. Analysons les pertes subies après 40 semaines d'insertion.

Dans la série 36, cela revient à 108.000 insertions.
La taille de la série 36 a un rapport de 168, comparée à la taille du fichier simulé.
Les 108.000 insertions correspondent, dès lors, à 643 insertions dans le fichier simulé.
Correspondent à ce nombre d'insertions : 45 additions de CI (pages), soit 22,5% du nombre de pages initiales.
Nous obtenons ainsi 22,5% de 32.800 = 7.380 ajouts de CI dans la série 36.

Par les mêmes principes, nous obtenons :

pour la série 37	9.720 ajouts de CI.
pour la série 50	13.920 ajouts de CI.
Total pour les séries 36, 37, 50 =	31.020 ajouts de CI.

. Temps total des insertions après 40 semaines.

1) Recherche d'une CI avec de la place libre.

Il y a 10 CI d'overflow par control area, ce qui donne en moyenne une recherche dans 5 CI d'overflow.
Coût de cette recherche :

5 x (8,4 ms. + 3,3 ms.) = 58,5 ms.
délai de mise à jour lecture

2) Coût du "split" de CI.

En moyenne (8,4 ms. + 3,3 ms.) x 2 = 23,4 ms.

3) Total des pertes dues aux insertions.

(58,5 ms. + 23,4 ms.) x 31.020 = 2.540.538 ms. soit 2.541 secondes.

. Autres pertes dues aux insertions.

Malgré les insertions, les enregistrements sont toujours atteints en un même nombre d'accès aussi longtemps que le nombre de niveaux d'index ne croît pas, ce qui est le cas comme nous l'avons pu précédemment.

Cette situation offre l'avantage suivant : le temps d'accès et de mise à jour n'est pas détérioré.

. Conclusion.

Après 40 semaines d'exploitation, nous obtenons une perte de 2.541 secondes pour un coût de réorganisation de 15.316 secondes. Il n'y a donc pas nécessité de réorganiser. Néanmoins, nous proposerions une réorganisation annuelle afin de suivre l'évolution physique du fichier.

7. Influence sur l'existant.

V.S.A.M., méthode d'organisation proposée par I.B.M. depuis une dizaine d'années, commence de plus en plus à être utilisée. Il est cependant nécessaire de pouvoir compter sur la présence d'un spécialiste en la matière pour la formation du personnel à cette technique, car même avec la documentation reçue, il n'est pas facile d'optimaliser cette technique.

Au point de vue des programmes existants, V.S.A.M. offre la possibilité de conversion de tous les programmes écrits en I.S.A.M. afin de les rendre compatibles avec V.S.A.M. Il ne devrait donc pas y avoir de problèmes de ce côté.

Comme pour les autres techniques, il faudra que les programmes écrits permettent l'insertion et la modification de longueur des enregistrements.

==========

B. - EVALUATION DES APPLICATIONS.

Introduction.

Pour les applications de consultation et de mises à jour, nous ne reprendrons que les temps moyens donnés auparavant.

Les améliorations et pertes possibles dues à la situation physique de l'index restent valables et ne seront plus reprises ici.

1. Consultation (journalière).

Rappel : a) 5.000 consultations journalières

b) une consultation est un accès aléatoire.

Temps moyen total

TCI	1 accès index + 1 accès données	2 accès index + 1 accès données
2 K	351 sec.	454,5 sec.
4 K	354,5 sec.	453,5 sec.
6 K	363 sec.	462 sec.

2. Mise à jour (journalière).

Rappel : a) nous évaluons à 50.000, 100.000 et 150.000 opérations journalières.

b) une mise à jour est constituée de deux accès à l'enregistrement + d'une recopie pour vérification.

Temps moyen total (en secondes)

TCI	1 accès index + 1 accès données.			2 accès index + 1 accès données.		
	50.000	100.000	150.000	50.000	100.000	150.000
2 K	7.855	15.710	23.565	9.925	19.850	29.775
4 K	7.920	15.840	23.760	9.905	19.810	29.715
6 K	8.095	16.190	24.285	10.075	20.150	30.225

3. Capitalisation.

Elle est estimée à deux lectures séquentielles du fichier (nous n'évaluons pas le temps de traitement). La capitalisation coûtera une fois l'an, soit à rappeler :

2 x 7.658 sec. = 15.316 sec. soit +/- 255 minutes.

C. - PRISE DE COPIE TRIEE.

Elle est évaluée à une lecture séquentielle, soit 7.658 secondes.

.....

5. - Deuxième répartition dans la technique de V.S.A.M. : un fichier pour la partie fixe du C.C.D., un autre pour sa partie évolutive.

Introduction.

Nous n'étudierons que les différences, par rapport à la solution première : elles se situent à plusieurs niveaux, à savoir :

1) la place occupée par le fichier;
2) la réorganisation;
3) la prise de copie triée;
4) la capitalisation.

Nous nous limiterons à réprendre les résultats généraux, les calculs étant identiques.

1. Place occupée par les fichiers.

1.1 - Partie fixe

Rappel : 6.200.000 enregistrements
taux d'occupation choisi : 95%
longueur enregistrement moyen : 130 bytes
control area = 1 cylindre.

Taille du fichier.

TCI =	2 K	4 K	6 K
Nombre d'enregistrements/CI	14	29	44
Nombre CI pour le fichier	442.858	213.794	140.910
Nombre cylindres	1.846	1.782	1.566
Taille index Ier niveau	1.846 CIi	1.782 CIi	1.566 CIi
2me niveau	8 CIi	15 CIi	13 CIi
3me niveau	1 CIi	1 CIi	1 CIi
par recopie de l'index	7 cylind.	4 cylind.	4 cylind.
Total en disques			
si index recopié 1 x	3,34	3,22	2,83
2 x	3,35	3,23	2,84
3 x	3,36	3,23	2,84

1.2 - Partie évolutive.

Rappel : 4.000.000 d'enregistrements.
taux d'occupation : 70, 80, 90%
longueur enregistrement : 130 bytes
control area : un cylindre.

Nombre de CI pour le fichier.

TCI \ TO	70%	80%	90%
2 K	363.637	333.334	285.715
4 K	181.819	160.000	142.858
6 K	125.000	108.109	95.239

Nombre de cylindres nécessaires en tenant compte des cylindres d'overflow. (nous nous basons sur les mêmes résultats de similation)

TCI \ TO	70%	80%	90%
2 K	1.639	1.516	1.299
4 K	1.639	1.455	1.299
6 K	1.507	1.319	1.162

Place index.

Ier niveau : Cfr. tableau ci-dessus. Un CI par cylindre (control area)

2me niveau : nombre de CIi :

TCI \ TO	70%	80%	90%
2 K	7	7	6
4 K	14	13	11
6 K	13	11	10

3me niveau : Un CIi partout.

Taille fichier en cylindres, si on tient compte de la recopie de 1, 2 ou 3 fois l'index.

TCI \ TO	70% 1x	70% 2x	70% 3x	80% 1x	80% 2x	80% 3x	90% 1x	90% 2x	90% 3x
2 K	1646	1653	1660	1523	1530	1537	1305	1311	1317
4 K	1643	1647	1651	1459	1463	1467	1302	1305	1308
6 K	1511	1515	1519	1322	1325	1328	1165	1168	1171

Taille fichier en disques (recopie de l'index 1, 2 ou 3 fois)

TCI \ TO	70% 1x	70% 2x	70% 3x	80% 1x	80% 2x	80% 3x	90% 1x	90% 2x	90% 3x
2 K	2,97	2,98	2,99	2,74	2,76	2,77	2,35	2,36	2,37
4 K	2,96	2,97	2,97	2,63	2,64	2,64	2,35	2,35	2,36
6 K	2,72	2,73	2,74	2,38	2,39	2,39	2,10	2,10	2,11

Total des deux fichiers :

70% entre 6 et 7 disques.
80% entre 6 et 7 disques.
90% entre 5 et 6 disques.

2. La réorganisation.

Nous rappelons ici que nous ne réorganisons que la partie évolutive du fichier, au taux d'occupation de 80% et une taille de CI de 4 K.

Coût de la réorganisation.

Rappel : deux lectures séquentielles du fichier évolutif.

Une lecture séquentielle du fichier évolutif coûte : nous appliquons les mêmes principes que dans la solution première et nous obtenons 3.004 secondes.
Le temps de réorganisation coûtera donc 6.008 secondes.

Quand réorganiser ?

Nous obtenons les mêmes conclusions que dans la solution première, étant donné que les insertions étudiées correspondent à celles que nous avons vues précédemment.

3. La prise de copies triées.

Elle correspond à la lecture séquentielle des deux fichiers, si nous posons l'hypothèse de la lecture simultanée.

Temps de lecture du fichier évolutif	3.004 secondes.
Temps de lecture du fichier fixe	3.679 secondes.
Total	6.683 secondes.

4. La capitalisation.

La capitalisation coûtera, quant à elle :

6.683 secondes x 2 = 13.366 secondes.

.....

6. Comparaison des deux types de répartition choisis en V.S.A.M.

Nous préférerions le 2me type de répartition au premier, pour l'unique raison de gain de place. Au minimum, le deuxième type de répartition demande 238 cylindres de données en moins que le premier type de répartition.

De ce fait, la lecture séquentielle, la réorganisation, la prise de copies triées et la capitalisation prennent moins de temps dans le 2ème type de répartition choisi.

TECHNIQUE III. - L'ACCES CALCULE.

1. Présentation des différents éléments qui influencent l'évaluation.

. Capacité de la piste : pour rappel : 19.069 bytes.

. Longueur des enregistrements (LE) :

l'overflow indépendant - zone primaire : longueur 130 bytes
zone overflow : longueur 130 bytes + 10
(pour le pointeur) = 140 bytes.

primary chaining - pour tout le fichier: longueur 130 bytes + 10
(pour le pointeur) = 140 bytes.

. Taille des pages (TP) :

Rappel : les enregistrements seront bloqués et ce, afin d'utiliser au mieux la capacité de la piste. Il sera envisagé différentes tailles de pages.

. Taux d'occupation du fichier (TO) :

Rappel : en vue d'absorber les insertions, le fichier créé sera rempli avec un certain taux d'occupation variable.

. Place de l'overflow (PO) :

Il s'agit d'évaluer la place nécessaire pour recevoir les enregistrements n'ayant pu trouver de place dans la partie primaire du fichier.

2. Equations analytiques.

a) Taille des pages (TP) choisie en fonction de la capacité par piste.

b) Nombre de pages par piste :

$$NB = \left(\frac{1 + 19.069 - \text{longueur dernier bloc}}{TP} \right)$$

c) Nombre d'enregistrements par piste (NEP) :

$$TO \times TP \times NB$$

d) Nombre de pistes nécessaires :

$$\left(\frac{\text{nombre maximum d'enregistrements}}{\text{nombre d'enregistrements par piste}} \right) + 1$$

e) place overflow : tirée à partir des résultats obtenus dans les simulations.

3. Etude de la fonction de répartition pour les deux algorithmes sur les simulations effectuées.

Trois fonctions ont été retenues : 1) division
2) midsquare
3) folding

Nous allons, sur base des résultats obtenus lors des simulations, classer ces fonctions : A représentera la meilleure
C représentera la moins bonne
B représentera l'intermédiaire.

Les critères retenus pour cette étude, sont ceux qui nous seront nécessaires pour l'évaluation de cette technique et sont repris dans le tableau ci-dessous.

	Fonction 1	Fonction 2	Fonction 3
temps moyen en insertion	A	B	C
temps maximum en insertion	A	B	C
longueur maximum chaîne	A, B	A, B	C
temps moyen en consultation	A	B	C
temps maximum en consultation	A	B	C
nombre moyen accès en insertion	A	B	C
nombre moyen accès en consultat.	A	B	C

De cette brève étude, nous retiendrons que la fonction de division est la meilleure fonction des trois retenues et, dès lors, toutes nos évaluations par rapport à la technique d'accès calculé seront basées sur les résultats obtenus des simulations réalisées à l'aide de cette fonction de décision.

4. Première répartition retenue dans la technique de l'accès calculé : un fichier pour tout le C.C.D.

A. - EVALUATION SELON LES CRITERES RETENUS.

1. Place occupée par le fichier.

. Nous prendrons, ici, des tailles de pages relativement égales à celles réalisées lors des simulations. Nous travaillerons avec des tailles de 4, 5, 6, 8 et 9 K.

. Nombre d'enregistrements : 10.200.000

. Taux d'occupation : 70%, 80%, 90%, 95%.

. Longueur des enregistrements :

overflow indépendant : 130 bytes (140 en overflow)

primary chaining : 140 bytes (°)

. Nombre d'insertions : 400.000

. Nombre d'enregistrements par piste (IO = independent overflow) (PC = primary chaining)

TP \ TO	70%		80%		90%		95%		Pour mém. 100%	
	IO	PC	IO	PC	IO	PC	IO	PC	IO	PC
4 K	86	81	99	92	111	104	117	110	124	116
5 K	77	75	88	86	105	97	111	102	117	108
6 K	98	90	112	103	126	116	133	122	141	129
8 K	86	81	99	92	111	104	117	110	124	116
9 K	98	91	112	104	126	117	133	123	140	130

. Nombre de cylindres nécessaires (pour la zone primaire)

TP \ TO	70%		80%		90%		95%	
	IO	PC	IO	PC	IO	PC	IO	PC
4 K	3954	4198	3435	3696	3064	3270	2742	3091
5 K	4416	4534	3864	3954	3239	3506	3064	3334
6 K	3470	3778	3036	3301	2699	2931	2412	2787
8 K	3954	4198	3435	3696	3064	3270	2742	3091
9 K	3470	3736	3036	3270	2699	2906	2412	2765

. Place occupée par l'overflow.

Nous nous baserons sur les résultats obtenus en simulation pour l'algorithme de l'overflow indépendant.

Pourcentages d'enregistrements nécessitant un accès lors de leur consultation à un taux d'occupation i de la zone primaire :

TP \ i	70%	80%	90%	95%
4 K	99,7	99	96	95
6 K	99,9	99,4	98	97,4
8 K	99,95	99,6	99	97,4

(°) Il faut 10 bytes supplémentaires pour le pointeur de l'enregistrement.

L'overflow va donc varier entre 1 et 5% avec :

1% pour le fichier chargé entre 70% et 80%

5% pour le fichier chargé entre 90% et 95%.

Pour le Primary Chaining, nous nous baserons sur les mêmes chiffres en posant, comme hypothèse, que si la fonction de hashage répartit bien sur N cases, elle répartit aussi sur N + i cases.

5% correspondent à environ :

137 cylindres supplémentaires pour l'Independent Overflow.

154 cylindres supplémentaires pour le Primary Chaining.

. Nombre de disques (en tenant compte de l'overflow)

Taux d'occupation	IO	PC
70%	entre 7 et 8	entre 7 et 8
80%	entre 6 et 7	entre 6 et 7
90%	entre 6 et 7	entre 6 et 7
95%	entre 5 et 6	6

2. Temps d'accès aléatoire.

. Résultats de simulation.

. Nous donnerons les temps moyen et maximum et le nombre moyen d'accès pour consulter un enregistrement.

Independent Overflow.

Pourcentages d'occupation par rapport à la zone totale :	+/- 70%			+/- 80%			+/- 90%			+/- 95%		
TP	Moy.	Max.	Acc.	Moy.	Max.	Acc.	Moy.	Max.	Acc.	Moy.	Max.	Acc.
4K	31,8	338,7	1	32,8	403,4	1,1	35,9	614,2	1,2	41,7	629,2	1,4
6K	33,3	351,8	1	33,7	475,3	1	35,9	636,2	1,1	38,4	759,7	1,2
8K	34,3	262,8	1	34,4	357,3	1	37,5	856,8	1,1	40,4	881,5	1,2

TP	+/- 100%		
	Moy.	Max.	Acc.
4K	46	640,8	1,6
6K	48,2	878,1	1,5
8K	52	1050,1	1,6

Primary Chaining.

Pourcentages d'occupation par rapport à la zone totale :	70%			80%			90%			95%		
TO	Moy.	Max.	Acc.	Moy.	Max.	Acc.	Moy.	Max.	Acc.	Moy.	Max.	Acc.
5K	28,9	75	1	28,9	85	1	29,2	126	1	29,5	189	1,1
6K	29,8	77	1	29,9	77	1	29,9	142	1	30,3	175	1
9K	31,9	65	1	31,9	102	1	32,1	202	1	34,1	1153	1,1

Remarque : lors des simulations, seules ces valeurs ont été testées.

Observation. - Il est normal que le Primary Chaining donne de meilleurs résultats car :

1°) l'overflow normal est déjà acquis lors de la création du fichier, ce qui permet une répartition sur un espace plus grand;

2°) le temps de recherche d'un élément en overflow est plus court par le fait qu'il se trouve souvent dans la case suivant celle où il aurait dû être rangé.

3. Tri du fichier.

Nous rappellerons que nous avons besoin des enregistrements en séquence pour certaines applications (prise de copies triées, réorganisation,...) et, étant donné que les enregistrements sont dans un désordre total, nous devons évaluer un tri du fichier.

a) Evaluation théorique.

. Mémoire centrale disponible pour le tri : 5 mégabytes.

. Longueur moyenne d'un enregistrement : 130 (ou 140) bytes.

. Nous constituons des monotonies (séquences d'enregistrement).

Nombre d'enregistrements par monotonie : 5.000.000 / 130 (ou) 140 soit de l'ordre de 40.000 (ou 37.400).

. Nombre de monotonies :

$$\left(\frac{\text{nombre d'enregistrements}}{\text{taille des monotonies}}\right) + 1$$ Soit de l'ordre de 255 (a) 275 (b)

. Pour trier tout le fichier, il faudra N passages.

$N = 1 + \log_S$ nombre de monotonies

Nous poserons N = 5 c'est-à-dire il faut S = 4 ou 5 suivant (a) ou (b)

. Un passage coûte la lecture et l'écriture de tout le fichier.

Evaluation d'une lecture de 1 disque, piste par piste et cylindre par cylindre (@)

(@) Cfr. description du disque 3350 en annexe.

30 x (15,5 ms. + 15,5 ms.) = ± 960 ms.
nbre pistes temps lecture temps écriture

Par disque 555 x 960 ms. = 532.800 ms. soit 533 secondes.
nbre cylindres

= 8' 53" + temps de déplacement du bras du disque
= 9 minutes.

Le fichier fait, en moyenne, six disques. D'où, temps total de un passage : 9' x 6 = 54 minutes.

. Temps pour les 5 passages : 54' x 5 = 270 minutes.

. A ce temps, il faut ajouter celui du traitement interne. Excepté pour la première phase, on peut considérer que ce temps de traitement est recouvert par le temps de transfert d'entrées/sorties.
Nous ne nous occuperons plus que du temps de tri de la Ière phase.
Ne connaissant pas l'algorithme utilisé, nous nous basons sur l'hypothèse suivante :

- il y a N $\log_2 N$ opérations pour le tri et nous considérons deux transferts et une comparaison par opération.

. Coût des opérations.

(Selon les renseignements obtenus de la C.G.E.R.)
Un transfert de 100 bytes pour un enregistrement se trouvant en mémoire "CASH" coûte ± 2 microsecondes (2 μs). Un transfert de 130 (140) bytes coûterait approximativement 2,5 μs.

Une comparaison coûterait ± 4 μs.

Temps de traitement d'une monotonie :

(2,5 μs x 2 + 4 μs) x 40.000 x $\log_2$40.000 = 5.503.577 μs. (5,5 sec.)
(2,5 μs x 2 + 4 μs) x 37.400 x $\log_2$37.400 = 5.113.207 μs. (5,1 sec.)

Temps de traitement total :
5,5 sec. x 255 = 1.402,5 secondes, soit +/- 23,5 minutes.
5,1 sec. x 275 = 1.402,5 secondes, soit +/- 23,5 minutes.

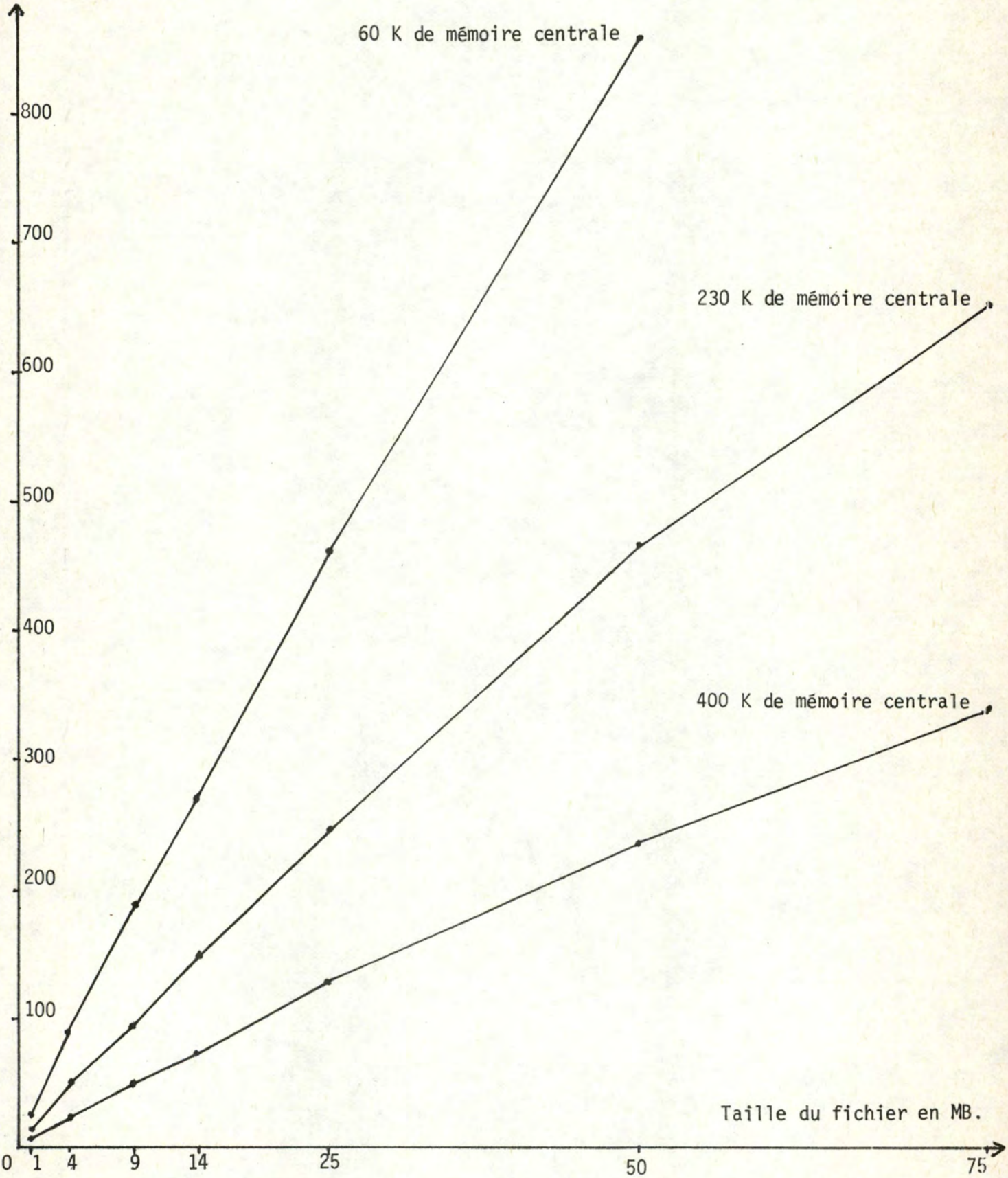
. Temps total du tri.

270 minutes + 23,5 minutes = 293,5 minutes (arrondi à 294 minutes).

.....

b) Evaluation à partir des documents I.B.M.
(Cfr. référence Bibliographie, n° 13)

Temps de tri en secondes.



Observations faites sur le graphe.

A partir d'une taille de fichier de 25 MB, on remarque une progression quasi linéaire du temps de tri avec la taille du fichier.

Déterminons l'équation d'une de ces droites.

On dispose de 400 K de mémoire centrale.

Etablissons l'équation de la droite passant par (25, 127) et (50, 236)

$$y - 127 = \frac{109}{25}(x - 25)$$

$$25\,y - 3175 = 109\,x - 2725$$

$$109\,x - 25\,y + 450 = 0$$

Taille du fichier à trier.

10.200.000 x 130 = de l'ordre de 1300 MB.

si x = 1.300 $\Longrightarrow$ y = 5.686 secondes, soit $\pm$ 95 minutes.

Correction de ce chiffre.

. car le processeur est différent de celui employé pour obtenir ces statistiques.
  Facteur correctif = 0,781 ce qui donne 4.440,766 secondes.

. car les disques ne sont pas les mêmes : il faut augmenter le temps obtenu de 45%, ce qui donne 6.439,11 secondes.
  (+/- 107 minutes)

REMARQUE.

. Les conditions théoriques du tri diffèrent quelque peu de celles nécessaires au tri du fichier CCD, à savoir :
  - longueurs des blocs et des records différentes.
  - le processeur n'est pas multiprogrammé.
  - aucune autre tâche n'utilise le processeur en Independent overflow

. De plus, il est évident que l'extrapolation diminue le degré de fiabilité des chiffres obtenus.
  Nous nous sommes basés sur une approche linéaire correspondant au schéma; nous devons cependant nous méfier car la courbe générale d'un tri est, au mieux, une courbe en N $\log_2$N (N étant le nombre d'enregistrements).

. Pour ces raisons, nous nous baserons sur la valeur théorique du tri.

4. Temps de mise à jour d'un enregistrement.

Rappel : une mise à jour est constituée :

- d'un accès à l'enregistrement avec lecture
- d'un accès à l'enregistrement avec écriture
- d'une relecture pour vérification.

En résumé : 2 accès + (1 lecture (avec délai) = +/- 16,7 ms.)

Nous présentons le temps moyen, le temps maximum n'étant pas représentatif.

Independent Overflow.

Pourcentage d'occupation par rapport à la zone totale				
TP	+/- 70%	+/- 80%	+/- 90%	+/- 95%
4 K	80,3	82,3	88,5	100,1
6 K	83,3	84,1	88,5	93,5
8 K	85,3	85,5	91,7	97,5

Primary Chaining.

Pourcentage d'occupation par rapport à la zone totale				
TP	70%	80%	90%	95%
5 K	74,5	74,5	75,1	75,7
6 K	76,3	76,5	76,5	77,3
9 K	80,5	80,5	80,9	84,9

5. Temps d'insertion d'un enregistrement dans le fichier.

Résultats obtenus par les simulations.

Independent Overflow.

Pourcentage d'occupation par rapport à la zone totale				
TP	+/- 70%	+/- 80%	+/- 90%	+/- 95%
4 K	35,1	48,2	81,4	121,4
6 K	32,9	41,9	68,5	95,2
8 K	33,2	40,7	85,7	103,7

Primary Chaining.

Pourcentage d'occupation par rapport à la zone totale				
TP	70%	80%	90%	95%
5 K	28	29,5	33,2	47,5
6 K	28,9	31,6	36,8	45,2
9 K	30,7	35,5	40,1	83,0

6. La réorganisation.

Coût de la réorganisation.

Il faut redistribuer tous les enregistrements sur un espace plus grand de zone primaire.

La méthode la plus économique de créer (ou de réorganiser) un fichier en accès calculé sur disques, est celle qui consiste à partir d'un fichier séquentiel (comprenant l'ensemble des enregistrements triés sur leur adresse physique) et à réécrire ce fichier sur les disques. Ceci évite un mouvement de va-et-vient du bras du disque, lequel ne se déplace ainsi que dans un seul sens. D'où un gain considérable de temps. Pour cela, il est donc nécessaire :

a) de lire tout le fichier;

b) de trier le fichier sur la nouvelle adresse des enregistrements;

c) de réécrire le fichier.

Le point (b) a déjà été évalué. Il ne nous reste donc que les points (a) et (c) à évaluer.

Une lecture ou une écriture peut être évaluée à :

a) taille de bloc :

4 K	nombre de blocs par piste	120
5 K	"	90
6 K	"	90
8 K	"	60
9 K	"	60

b) temps de lecture par cylindre : nombre de blocs (délai + écriture)

Ce qui donne, approximativement, par cylindre, un temps compris entre 60 x (16,7 ms.) et 270 x(16,7 ms.) 1 et 4,5 sec.

Par disque, on obtiendra entre 555 et 2.497,5 secondes.

En moyenne, pour six disques : 3.330 et 14.985 secondes.

Coût total d'une réorganisation.

2 fois entre 3.330 et 14.985 secondes + 294 minutes du tri = entre 405 et 543,5 minutes.

Exemple : pour TP = 6 K, nous obtenons un total de 460,5 minutes.

Quand réorganiser ?

Il serait trop long de reprendre, pour chaque cas envisagé, l'évaluation faite ci-dessous. Nous ne considérerons qu'un seul exemple : celui du fichier chargé à 90% comme point de départ avec TP = 6 K. Nous souhaitons donc garder le fichier proche des 90% d'occupation.

Nous rappelons que nous réorganisons lorsque le coût de la réorganisation est proche ou supérieur au montant des pertes dues à l'insertion des enregistrements.

Independent Overflow.

La réorganisation, pour ce cas bien précis, peut être évaluée à :

(294 minutes de tri + 166,5 minutes de lecture et écriture)

ce qui donne un total de 460,5 minutes. (Cfr. coût en page précédente)

	à 90% d'occupation	à 95% d'occupation
une consultation coûte	35,9 ms.	38,4 ms.
une mise à jour coûte	88,5 ms.	100,1 ms.
une insertion coûte	68,5 ms.	95,1 ms.

Les 5% de différence d'occupation comprennent entre 510.000 et 530.000 nouveaux enregistrements.

(Rappelons que nous comptons +/- 400.000 insertions par an, ce qui correspond à +/- 10.000 insertions par semaine).

Calcul de la perte due aux insertions de la Ière semaine.

(approximation linéaire)

a) Consultation $\dfrac{38{,}4\text{ms.} - 35{,}9\text{ms.}}{530.000 \times 10.000} = 0{,}05$ ms.

en moyenne, par consultation, nous perdons 0,05 ms.

b) La mise à jour $\dfrac{100{,}1\text{ ms.} - 88{,}5\text{ms.}}{530.000 \times 10.000} = 0{,}22$ ms.

c) L'insertion $\dfrac{95{,}1\text{ ms.} - 68{,}5\text{ms.}}{530.000 \times 10.000} = 0{,}50$ ms.

Total des pertes sur la semaine.

. En consultation 5.000 par jour 25.000 par semaine.

. En mise à jour
- 250.000 mises à jour normales + rejets (en moyenne 1 semaine sur 4 = 335.000)
- 1.000 + 18.750 = 19.750 pour les opérations comptables et non comptables.
- 25.000 pour la recapitalisation.
- 100.000 pour la mise à jour des intérêts (4.000.000 par an)

Total (moyenne de 4 semaines) = 416.000 mises à jour.

. En insertion 8.000

. Coût - en consultation 1.250 ms.
- en mise à jour 91.520 ms.
- en insertion 5.000 ms.

Total : 97.750 ms. soit +/- 1'40.

Les autres semaines : nous admettons ici une approche linéaire.

Total des pertes après 1 semaine 100 sec.
après 2 semaines 100 sec. + 200 sec. = 300 sec.
après 22 semaines +/- 421,7 minutes
après 23 semaines +/- 460 minutes.

Nous réorganiserions donc après la 22e ou la 23e semaine, c'est-à-dire lorsque le coût de réorganisation est proche ou supérieur au montant des pertes subies.

Primary Chaining.

Nous prendrons les mêmes hypothèses que pour l'Independent Overflow et nous effectuerons les mêmes calculs.

Coût des Opérations.

	90%	95%	Perte due aux insertions
consultation	29,1	30,3	0,023 ms.
mise à jour	76,5	77,3	0,015 ms.
insertion	36,8	45,2	0,16 ms.

Perte de la Ière semaine = consultation 575 ms.)
mise à jour 6.240 ms.) ce qui donne :
insertion 1.600 ms.) ± 9 secondes.

Si nous admettons toujours l'approche linéaire des pertes suivant les semaines, nous n'arriverons pas, avant un an, avec une somme de pertes supérieure au coût de réorganisation. Néanmoins, nous pensons qu'une réorganisation annuelle est nécessaire.

7. Influence sur l'existant.

Elle portera sur plusieurs éléments, à savoir :

1. La gestion du fichier à charge du programmeur-système. Par gestion, il faut entendre :
   a) la création du fichier en accès calculé qui, nous le rappelons, est chargé à partir d'un fichier séquentiel comprenant tous les enregistrements triés sur leur adresse physique;
   b) l'accès à un enregistrement;
   c) l'insertion d'enregistrements;
   d) la suppression d'enregistrements;
   e) la mise à jour de longueur des enregistrements;
   f) la gestion de l'overflow.

Pour plus de précisions, nous renvoyons le lecteur aux principes de l'accès calculé décrit, en partie VI. Il pourra, ainsi, se faire une idée sur les programmes à réaliser.

2. Les programmes existants doivent être revus car il l'existe pas de routines permettant les accès à un fichier organisé initialement en séquentiel indexé et réorganisé, par la suite, en accès calculé. Le programmeur doit disposer de primitives lui facilitant la tâche. Parmi ces primitives, doivent se trouver celles pour :
.la lecture d'un enregistrement (accès);
.l'écriture d'un enregistrement (insertion);
.la réécriture d'un enregistrement (mise à jour);
.la suppression d'un enregistrement.

3. Enfin, le programmeur doit aussi disposer d'utilitaires pour travailler en accès calculé. Parmi ceux-ci, nous citerons :
.un utilitaire de tri;
.un utilitaire de fusion de fichiers.

B. - EVALUATION DES APPLICATIONS.

1. **Consultation.**

Rappel : 5.000 consultations.
une consultation : un temps d'accès aléatoire.

Chiffres de base : Cfr. tableau de pages 123 et 124.
(Nous ne reprendrons que les temps moyens).

Independent Overflow. (en secondes)

TP. \ Pourcentages d'occupation par rapport à la zone totale :	+/- 70%	+/- 80%	+/- 90%	+/- 95%
4 K	159	164	179,5	208,5
6 K	166,5	168,5	179,5	192
8 K	171,5	172	187,5	202

Primary Chaining. (en secondes)

TP. \ Pourcentages d'occupation par rapport à la zone totale :	70%	80%	90%	95%
5 K	144,5	144,5	146	147,5
6 K	149	149,5	149,9	151,5
9 K	159,5	159,5	160,5	170,5

2. **La mise à jour.**

Rappel : nous évaluerons 50.000, 100.000 et 150.000 opérations.
temps de base : ce sont ceux repris en page 128.

Résultats (en secondes).

Independent Overflow.

TP. \ Pourcentages d'occupation par rapport à la zone totale :	+/- 70%	+/- 80%	+/- 90%	+/- 95%
a) 50.000 opérations				
4 K	4.015	4.115	4.425	5.005
6 K	4.165	4.205	4.425	4.675
8 K	4.265	4.275	4.585	4.875
b) 100.000 opérations				
4 K	8.030	8.230	8.850	10.010
6 K	8.330	8.410	8.850	9.350
8 K	8.530	8.550	9.170	9.750
c) 150.000 opérations				
4 K	12.045	12.345	13.275	15.015
6 K	12.495	12.615	13.275	14.025
8 K	12.795	12.825	13.755	14.625

Primary Chaining.

TP. \ Pourcentages d'occupation par rapport à la zone totale :	70%	80%	90%	95%
a) 50.000 opérations				
5 K	3.725	3.725	3.755	3.785
6 K	3.815	3.825	3.825	3.865
9 K	4.025	4.025	4.045	4.245
b) 100.000 opérations				
5 K	7.450	7.450	7.510	7.570
6 K	7.630	7.650	7.650	7.730
9 K	8.050	8.050	8.090	8.490
c) 150.000 opérations				
5 K	11.175	11.175	11.265	11.355
6 K	11.445	11.475	11.475	11.595
9 K	12.075	12.075	12.135	12.735

3. Capitalisation.

Coût : une prise de copie du fichier.
un rechargement du fichier sur le(s) disque(s), le temps de traitement n'étant pas évalué.

Remarque. - Pour le rechargement du fichier, nous travaillons comme pour la réorganisation, à savoir tri + réécriture du fichier.

Le coût de la capitalisation est égal à celui d'une réorganisation, à rappeler : entre 405 et 543,5 minutes.

C. - PRISE DE COPIE TRIEE.

Nous renvoyons à la page 124 où ce point a été développé.

5. - Deuxième répartition retenue dans la technique de l'accès calculé : un fichier pour la partie fixe du CCD, un autre pour sa partie évolutive.

Nous étudierons, sous ce point, les différences avec la solution première. Elles se situent sur deux critères, à savoir :

1. la place occupée par les fichiers;
2. la réorganisation;

et sur une application particulière à la C.G.E.R. : la prise de copies triées.

Nous posons, comme hypothèse, que les temps simulés pour la solution première restent les même pour cette solution.

1. Place occupée par les fichiers.

1.1 - Etude de la partie fixe.

Rappel : nombre d'enregistrements 6.200.000
taux d'occupation choisi 95%
longueur enregistrement : pour I-O 130 bytes
140 pour overflow
pour P-C 140 bytes

Nombre d'enregistrements par piste.

TP / %	4K		5K		6K		8K		9K	
	I-O	P-C	I-O	P-C	I-O	P-C	I-O	P-C	I-O	P-C
95	117	110	111	102	133	122	119	110	133	123

Nombre de cylindres nécessaires.

TP / %	4K		5K		6K		8K		9K	
	I-O	P-C	I-O	P-C	I-O	P-C	I-O	P-C	I-O	P-C
95	1767	1879	1862	2027	1554	1694	1737	1879	1554	1681

Il faut prévoir néanmoins un overflow car, lors de la création du fichier avec la fonction de hashage, il y aura des cases qui déborderont. Nous nous baserons sur un overflow de 5% (Cfr. 1ère technique de répartition.)

Place de la partie fixe en disques.

TP / %	4K		5K		6K		8K		9K	
	I-O	P-C	I-O	P-C	I-O	P-C	I-O	P-C	I-O	P-C
95	3,34	3,55	3,55	3,83	2,94	3,20	3,28	3,55	2,94	3,18

1.2 - Etude de la partie évolutive.

Comprenant 4.000.000 d'enregistrements.

Mêmes principes de calcul que pour la solution I.

Nombre de cylindres nécessaires.

Pourcentages d'occupation par rapport à la zone totale :								
TP	70% I-O	70% P-C	80% I-O	80% P-C	90% I-O	90% P-C	95% I-O	95% P-C
4 K	1.551	1.647	1.347	1.450	1.202	1.283	1.140	1.213
5 K	1.732	1.778	1.516	1.551	1.270	1.375	1.202	1.308
6 K	1.361	1.482	1.191	1.295	1.059	1.150	1.003	1.093
8 K	1.551	1.647	1.347	1.450	1.202	1.283	1.140	1.213
9 K	1.361	1.466	1.191	1.283	1.059	1.140	1.003	1.085

Ajout de l'overflow : 2% pour taux d'occupation à 70% et à 80%
5% pour taux d'occupation à 90% et à 95%

(Cfr. solution première de la technique)

Total en disques.

Taux d'occupation	Independent Overflow	Primary Chaining
70%	entre 2,6 et 3,2	entre 2,8 et 3,3
80%	entre 2,2 et 2,8	entre 2,4 et 2,9
90%	entre 2,1 et 2,5	entre 2,2 et 2,7
95%	entre 1,9 et 2,7	entre 2,1 et 2,5

Total général pour les deux fichiers.

Taux d'occupation	Independent Overflow	Primary Chaining
70%	entre 6-7 disques	entre 6-8 disques
80%	entre 6-7 disques	entre 6-7 disques
90%	entre 6-7 disques	entre 6-7 disques
95%	entre 5-7 disques	entre 6-7 disques

2. La réorganisation.

Coût de la réorganisation.

Dans ce cas, nous ne réorganisons que la partie évolutive du fichier, c'est-à-dire les 4.400.000 enregistrements.

. Coût du tri (cfr. tri en solution première)
au point de vue théorique (118 minutes)
équation (43 minutes)

. Coût de lecture (écriture)
En moyenne, nous considérons que la partie évolutive comprend trois disques.
Tri (au point de vue théorique) = 118 minutes.
Lecture : 1,5 sec. (par cylindre) x 555 (nbre cylindres) x 3 (nbre disques) = 2.497,5 secondes ou 41,625 minutes

Temps total (2 lectures + tri) = 201,25 minutes.

Quand réorganiser ?

Nous adoptons ici les mêmes principes que ceux appliqués à la solution première.

. Différence : nous appliquons la même approche linéaire seulement sur la partie évolutive du fichier.

. Nous nous baserons sur les mêmes temps que ceux donnés dans la première solution.

Pour l'algorithme Independent Overflow.

1. Calcul des pertes subies lors de la Ière semaine.

   par consultation (38,4ms.-35,9ms.)/220.000 x 10.000 = 0,11 ms.
   par mise à jour (100,1ms.-88,6ms.)/220.000 x 10.000 = 0,53 ms.
   par insertion (95,1ms.-68,5ms.)/220.000 x 10.000 = 1,21 ms.

   NB : 220.000 = 5% du nombre d'enregistrements de la partie évolutive.
   10.000 = le nombre d'insertions par semaine.

   Total pour la semaine :

consultations	2.750 ms.)	
mises à jour	220.480 ms.)	Total = 235.330 ms.
insertions	12.100 ms.)	soit +/- 235 secondes.

2. Les autres semaines, en admettant l'approche linéaire des pertes :

   après la 2e semaine 235 sec. + 470 sec. = 705 secondes.
   après la 9e semaine ..... 176,25 minutes.
   après la 10e semaine ..... 215,40 minutes.

   Nous proposerions donc une réorganisation toutes les 9 semaines, en rappelant que nous réorganisons lorsque le coût de réorganisation est proche ou supérieur au montant des pertes dues à l'insertion des enregistrements.

Pour l'algorithme Primary Chaining.

1. Calcul des pertes subies lors de la Ière semaine.

par consultation	0,05 ms.	⇒ perte la 1ère semaine	=	1.250 ms.
par mise à jour	0,04 ms.	" "		16.640 ms.
par insertion	0,38 ms.	" "		3.800 ms.
		(soit 22 sec.)		21.690 ms.

2. Perte les autres semaines.

   Toujours la même approche linéaire, qui nous amènerait,
   après la 31e semaine, à une perte totale de 181,87 minutes.
   après la 32e semaine, ... 193,60 minutes.
   après la 33e semaine, ... 205,70 minutes.

   Nous proposerions donc, dans ce cas, une réorgnisation après 32 semaines.

III. - PRISE DE COPIES TRIEES.

Ici, il s'agit de trier les deux fichiers et d'en réaliser la fusion. Celle-ci ne servirait que si nous souhaitions obtenir un seul fichier séquentiel trié pour les deux parties.

Tri des deux fichiers.

Nous appliquons les mêmes principes que pour la solution première.

1. Partie fixe

6.200.000 enregistrements	3,5 disques en moyenne
au point de vue théorique	140 minutes.
au point de vue équation	63 minutes.

2. Partie évolutive

4.000.000 enregistrements	3 disques en moyenne
au point de vue théorique	118 minutes.
au point de vue équation	43 minutes.

3. Total des tris

au point de vue théorique	258 minutes.
au point de vue équation	106 minutes.

4. Coût de la fusion des deux fichiers

Cela équivaut donc à la lecture des deux fichiers, soit +/- 6,5 disques à 9 minutes par disque = +/- 58,5 minutes.

5. Total tris et fusion

au point de vue téorique 258 min. + 58,5 min. = 316,5 minutes.

.....

6. - Comparaison des types de répartition choisis en accès calculé.

A. - Comparaison Independent Overflow - Primary Chaining.

Cette comparaison se fera selon deux critères : l'espace utilisé et le temps d'accès à un enregistrement.

Au point de vue de l'espace utilisé, l'Independent Overflow obtient de meilleurs résultats dus au fait que les enregistrements sont moins grands en moyenne car ils n'ont un pointeur que lorsqu'ils sont en overflow. De plus, cet overflow peut être nul au départ et s'accroître au fur et à mesure des besoins, ce qui n'est pas le cas pour le Primary Chaining où l'overflow fait partie du fichier de base, donc dès la création du fichier.

En ce qui concerne le temps d'accès, le Primary Chaining apparaît meilleur (entre 6 et 10 ms. de gain par accès). Cela provient du fait que les enregistrements qui sont en overflow se trouvent souvent dans une case assez proche

de celle où ils devraient se trouver normalement, réduisant par le fait même de beaucoup le temps nécessaire au déplacement du bras du disque pour se diriger d'une case à une autre.

Ce temps d'accès impliquant un temps de mise à jour meilleur dans le Primary Chaining - celui-ci ayant, en outre, le temps d'insertion le plus favorable - nous pouvons constater que l'on réorganisera un peu moins souvent que lorsque l'algorithme Independent Overflow est utilisé (une fois par an pour le Primary Chaining et deux fois par an pour l'Independent Overflow.)

B. - Comparaison des deux répartitions proposées.

Nous ne pouvons pas affirmer, à ce niveau, qu'une solution soit préférable à l'autre. En effet, nous obtenons des chiffres assez semblables en ce qui concerne l'espace utilisé en nombre de disques bien qu'en nombre de cylindres, il y ait un avantage pour la deuxième répartition, avantage qui devrait se chiffre aux environs de 300 à 400 cylindres.

Les différences essentielles se situent dans le coût de la réorganisation : pour un TP. de 6 K et un TO. de 90%, 201,25 minutes pour la répartition 2 contre 460,5 minutes pour la répartition 1. Mais la fréquence de réorganisation est différente : une à deux fois par an pour la première répartition et de deux à cinq fois par an pour la deuxième.

Enfin, la prise de copie triée est assez semblable dans les deux cas mais, lorsqu'il y a, dans la 2ème répartition, la fusion des deux fichiers triés, la prise de copie triée est moins bonne.

Pour rappel (au point de vue théorique) :

Ière solution	294 minutes.
2ème solution	258 minutes.
2ème solution + fusion	316,5 minutes.

==========

PARTIE VIII. - CONCLUSIONS.

Le but de ce travail était de trouver une technique d'organisation du fichier des livrets d'épargne de la C.G.E.R. Nous en avons étudié plusieurs, afin de permettre le choix sur base de comparaisons.

La méthode de travail que nous avons suivie se rapproche beaucoup, quoique différente, de celle reçue à l'Institut d'Informatique. Nous avons étudié, dans un premier temps, l'organisation de l'épargne à la C.G.E.R. et nous avons donné une description de l'existant informatique. Ces deux points nous ont permis de réaliser notre analyse conceptuelle avant de proposer des techniques de résolution. Parmi celles-ci, nous rappelons la solution existant actuellement.

Le choix des autres techniques a été retenu à la suite des discussions avec les responsables de la C.G.E.R. qui avaient exprimé le souhait d'obtenir des renseignements sur une organisation en accès calculé et sur une organisation du type V.S.A.M. (cette technique étant de plus en plus utilisée à la C.G.E.R.). La technique du séquentiel indexé a été étudiée pour permettre la comparaison avec le séquentiel indexé repris dans la solution actuelle.

Dans chaque technique proposée, nous avons évoqué différentes possibilités mais il est évident que nous n'avons pas pu, pour chaque cas, entrer dans tous les détails. Nous avons alors posé nos limites d'études. C'est ainsi que nous avons signalé les améliorations possibles dans chaque technique.

Nous avons donc constitué un ensemble de références nécessaires avant une autre étude plus détaillée, une fois le choix de l'organisation du fichier effectué. Ce choix reste du domaine de la C.G.E.R. mais nous ne voudrions pas terminer cette étude sans proposer nos idées retenues, à la suite de ce travail.

Suivant deux objectifs possibles, nous proposerions l'une ou l'autre technique. C'est ainsi que :
- si la C.G.E.R. désire améliorer la situation actuelle sans apporter un profond changement à l'existant, nous accorderions notre préférence à la technique de V.S.A.M. en posant, comme réserve, une étude approfondie sur la répartition des control area;

- si, par contre, l'objectif de la C.G.E.R. est de gagner du temps, nous proposerions la technique de l'accès calculé (bien que le tri hebdomadaire nécessaire prenne un temps considérable). Nous rappelons toutefois que cette technique nécessite une modification de l'existant non négligeable par rapport aux autres techniques.

Par ces choix, nous ne voulons pas favoriser une technique plutôt qu'une autre mais simplement en faire remarquer les points importants.

Nous terminerons cette dernière partie par la présentation de tableaux reprenant les principaux résultats dans les différentes techniques.

==========

TABLEAUX DE SYNTHESE.

1. Tailles du fichier en disques.

Taux d'occupation ⟶	70%	80%	90%
Séquentiel indexé (Ière répartition)	7-8	6-7	6-7
(2ème répartition)	6-7	6	6
V.S.A.M. (Ière répartition)	7-8	7	6
(2ème répartition)	6-7	6-7	5-6
Accès calculé			
Indep. Overflow (Ière répartition)	7-8	6-7	6-7
Primary Chaining(Ière répartition)	7-8	6-7	6-7
Indep. Overflow (2ème répartition)	6-7	6-7	6-7
Primary Chaining(2ème répartition)	6-8	6-7	6-7

2. Temps moyens d'accès à un enregistrement.

Pour TO = 90% et TP = 6 K

Séquentiel indexé :
- a) en zone primaire : 81 ms.
- b) en overflow - avant 33 semaines : au maximum : 123,5 ms.
- b) en overflow - après 33 semaines : au maximum : 182,5 ms.

V.S.A.M. :
- a) un accès index + un accès données : 72,6 ms.
- b) deux accès index + un accès données : 92,4 ms.

Accès calculé :
- a) Independent Overflow : 35,9 ms.
- b) Primary Chaining : 29,9 ms.

3. Temps moyens de mise à jour d'un enregistrement.

Pour TO = 90% et TP = 6 K

Séquentiel indexé :
- a) en zone primaire : 178,7 ms.
- b) en overflow - avant 33 semaines : entre 178,7 et 263,7 ms.
- b) en overflow - après 33 semaines : entre 178,7 et 381,7 ms.

V.S.A.M. :
- a) un accès index + un accès données : 161,4 ms.
- b) deux accès index + un accès données : 201,5 ms.

Accès calculé :
- a) Independent Overflow : 88,5 ms.
- b) Primary Chaining : 76,5 ms.

4. Temps moyens d'insertion d'un enregistrement.

Pour TO = 90% et TP = 6 K

Séquentiel indexé :
- a) en zone primaire — 104,9 ms.
- b) en overflow :
  - en cylinder overflow — entre 118,3 et 121,8 ms.
  - en independent overflow — entre 144,5 et 148,- ms.

V.S.A.M. :
- a) 1 accès index + 1 accès données — entre 72,6 et 154,5 ms.
- b) 2 accès index + 1 accès données — entre 92,4 et 174,3 ms.

Accès calculé :
- a) Independent Overflow — 68,5 ms.
- b) Primary Chaining — 36,8 ms.

5. Temps de lecture séquentielle.

Pour TO = 90% et TP = 6 K

Séquentiel indexé :
- (Ière répartition) — entre 4858 et 5282 sec.(entre 81 et 88 min.)
- (2ème répartition) ⊕ — entre 4872 et 5296 sec.(entre 81 et 88 min.)

V.S.A.M. :
- (Ière répartition) — 4630 sec. (+/- 77 min.)
- (2ème répartition) ⊕ — 4264 sec. (+/- 71 min.)

Accès calculé :
- (Ière répartition) — 294 min.
- (2ème répartition) — 258 min.
- id. si fusion des deux fichiers triés — 316,5 min.

⊕ Lecture simultanée sur les deux fichiers.

6. Réorganisation.

Séquentiel indexé :
- (Ière répartition) coût — 10.306 secondes (TP = 6K, TO = 90%)
  fréquence — après 38 semaines.
- (2ème répartition) coût — 4.302 secondes (TP = 6K, TO = 90%)
  fréquence — après 35 semaines.

V.S.A.M. :
- (Ière répartition) coût — 15.316 secondes (TP = 4K, TO = 80%)
  fréquence — 1 fois par an.
- (2ème répartition) coût — 6.008 secondes (TP = 4K, TO = 80%)
  fréquence — 1 fois par an.

Accès calculé :
- (Ière répartition) coût — 460,50 minutes (TP = 6K, TO = 90%)
  ou — 27.630 secondes
  fréquence — après 22 semaines (I-O)
  1 fois par an (P-C)
- (2ème répartition) coût — 201,25 minutes (TP = 6K, TO = 90%)
  ou — 12.075 secondes
  fréquence — après 9 semaines (I-O)
  après 32 semaines (P-C)

7. Influence sur l'existant.

Séquentiel indexé	faible.
V.S.A.M.	peu importante.
Accès calculé	importante.

==========

B I B L I O G R A P H I E .

1. C.G.E.R. — Analyse fonctionnelle du projet : "Epargne 80" Caisse Générale d'Epargne et de Retraite, Bruxelles - 1979.

2. BODART F. — Analyse conceptuelle-fonctionnelle. Notes de cours , 1980.

3. HAINAUT J-L. — Modèles conceptuels. Lecture notes, publications. Institut d'Informatique, NAMUR - 1980.

4. HAINAUT J-L. — Fichiers et banques de données. Notes de cours, 1980.

5. LESUISSE R. — Analyse générale des fichiers. Notes de cours, 2ème graduat. I.E.S.N. - 1978.

6. REILAND J. — Description des ordinateurs. Notes de cours, 2ème graduat. I.E.S.N. - 1978.

7. I.B.M. — Introduction to I.B.M. Direct-Access Storage Devices and Organization Methods. Mars 1979.

8. I.B.M. — V.S.A.M. Primer and Reference G.320 - 5774 Août 1979.

9. I.B.M. — O.S./V.S./Virtual Storage Access Method (V.S.A.M.) Programmer's Guide VS 1 Release 7 - VS 2 Release 3.8 Juillet 1978.

10. I.B.M. — Introduction to Virtual Storage in System/370 Student text. Février 1973.

11. HAINAUT J-L. — Résultats de simulations. Répartition sur deux pages. Institut d'Informatique, NAMUR - 1981.

12. NISEN J-M. & ROULIN J. — Un système d'aide à la simulation à événements discrets en FORTRAN. Institut d'Informatique, NAMUR - 1981.

13. I.B.M. — O.S./V.S. Sort / Merge Programmer's Guide Program Number 5740 - SM1 - Release 5 Mars 1981.

14. I.B.M. — O.S./V.S. Sort / Merge Prgrammer's Guide Program Number 5740 - SM1 - Release 6 Juillet 1979.

15. I.B.M. — Data Facility Data Set Services User's Guide and Reference Program Number 5740 - UT 3 - Release 1-1 Juin 1981.

16. LITWIN W. — Hachage virtuel : une nouvelle technique d'adressage de mémoires. Paris, 1979.

17. LITWIN W. — Trie Hashing. I.N.R.I.A. 78 - 153 (MAP-I-014) Le Chesnay (France) - 1981.

FACULTES UNIVERSITAIRES NOTRE-DAME DE LA PAIX (NAMUR)

INSTITUT D'INFORMATIQUE

CHOIX D'UNE REORGANISATION D'UN FICHIER POUR UN ORGANISME BANCAIRE

(annexes)

Mémoire présenté par

Jean-Louis JACOBS

en vue de l'obtention
du titre de

LICENCIE et MAITRE en INFORMATIQUE.

Année académique 1981-1982.

PLAN des ANNEXES.

A. - Description de l'enregistrement du fichier CCD.

B. - Description de l'enregistrement en format variable du fichier CCC, en comparaison avec l'enregistrement existant.

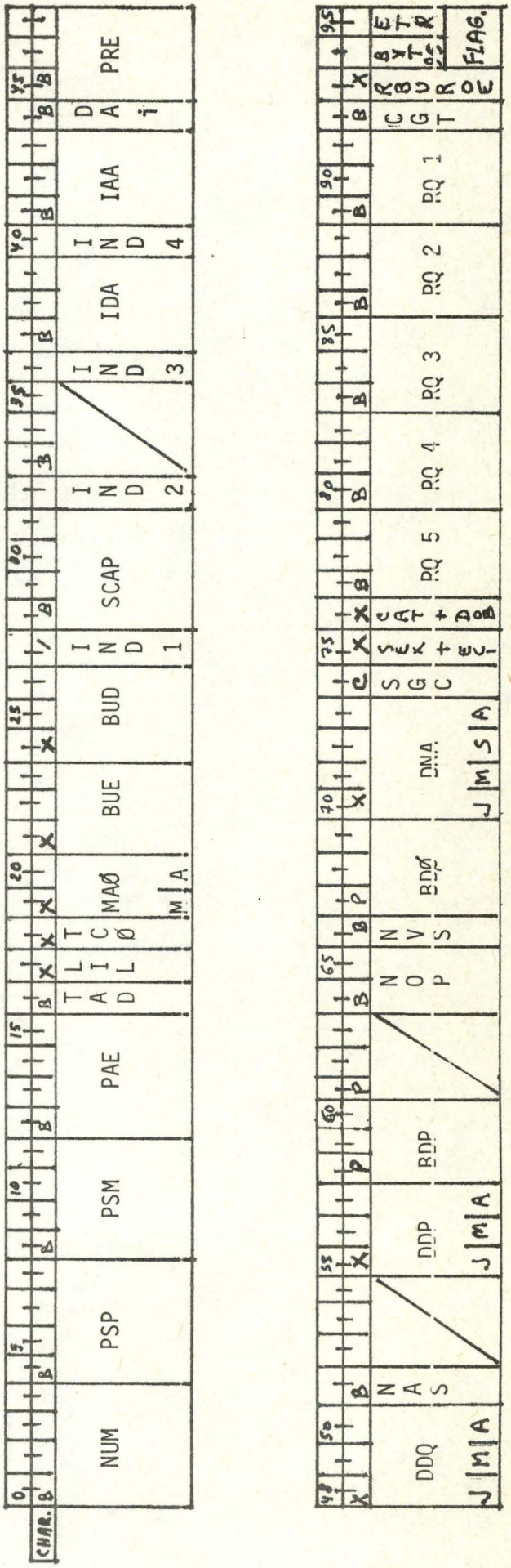
C. - Description du disque 3350 d'IBM.

D. - Listes des V.T.O.C. du fichier CCD.

E. - Résultats simulations :
- V.S.A.M.
- répartition sur deux pages
- accès calculé Independant Overflow
- accès calculé Primary Chaining.

A. - Description de l'enregistrement du fichier CCD.

1. SCHEMA.



2. DESCRIPTION DES ZONES.

NUM	Numéro de compte
PSP	Pointeur signalétique
PSM	Nombre de mandataires
PAE	Pointeur adresse extrait
TAD	Liaison titulaire - adresse
LIL	Langue d'impression du livret sauf en terme : rôle linguistique du compte (RLC)
TCØ	Liaison compte - titulaire
MAØ	Mois et année d'ouverture
BUE	Bureau d'émission
BUD	Bureau de domiciliation
IND(x)	(X) prend les valeurs de 1 à 4 - Cfr. page 3.
SCAP	Solde en capital
IDA	Intérêts nets de la dernière année
IAA	Intérêts nets des années antérieures, non inscrits
DAi	Dernière année inscription des intérêts
PRE	Précompte retenu sur IDA
DDØ	Date de la dernière opération
NAS	Nombre d'autres sommes à inscrire
DDP	Date de la dernière pièce
BDP	Bureau de la dernière pièce
N Ø P	Nombre d'opérations comptables
NVS	Nombre de versements scolaires
BDØ	Bureau de la dernière opération
DNA	Date de naissance
SGC	Service de gestion du compte
SEX	1/2 byte = sexe titulaire pour les sociétés = TCØ
ECi	1/2 byte = état civil du titulaire
CAT	--
DØB	Origine du livret
RQx	Intérêts bruts de l'année - x
CGT	Nombre contrôle du NUM
RBUROE	Régime linguistique du BUE (cf. p.4

FLAG	Zone flag. (= high-value si enregistrement flaggé)	
BYT 95	Cf. description page 4	
ETR	Catégorie étranger (si bit ETR ON dans IND 1)	

.....

IND.1	DUP	I	Duplicata
	OPP	I	Opposition
	CAR	I	Carnet de rentes
	ETR	I	Compte etranger (IBLC)
	SOL	I	Compte soldé
	DSO	I	Dernière opération avec diff.solde
	RECAPIT.		Précapitalisation (avant capt. off line)
	VALIDITE SEX-TEØ	I	ON si SEX-TEØ Ø K. (août 79)

IND.2 _____

IND.3	LDI	I	Livret en dépôt à l'institution
	LDA	I	Livret en dépôt aux archives (V-77)
	PREPSOL	I	Préparation remboursement total (mis on 986)(mis off 600)
	PNF	I	Procuration F
	RNN	I	Procuration N
	INI	I	Intérêts inscrits
	EXIT	I	Livret à renvoyer
	ØPER	I	Une opération depuis dernière capitalisation.

IND.4	CHN	I	Chèque de naissance
	LOI	I	Loi 10.2.1900
	PRO	I	Procuration (série 40)
	RES	I	Réservé jusqu'à 21 ans
	FIC	I	Fiche-dossier subsiste

IND.4	DOS	I	Dossier liquidation
(suite)	ACC	I	Accusé réception pour A.S.
	E.47	I	Demande attestation envoyée

DECOMPOSITION EN BITE DE BYT 95.

CPH — Compte cédant P.M.
CES — Compte cédant escomptes
} exclusifs.

SNCI — Compte cédant géré par SNCI

compte out put d'une normalisation.

ZONE RBUROE.

hexa ' x y ' avec :

(x = rôle linguistique BUE
)
(y = base (0 à 9) du calcul SGC.
)

REMARQUE : Les noms de zones ont été définis dans la présentation de l'enregistrement actuel.

NOMS	Types records Projet 80.	Changements NOMS.	LONGUEUR	Longueur prochaine.
PRE	3 4	PDA	3 bytes	8 chiffres
NUM	1 2 3 4		9 ch. (4 bytes)	9 chiffres
PSP	2 3 4		9 ch. (4 bytes)	9 chiffres
PSM	2 3 4		9 ch. (4 bytes)	9 chiffres
TAD	2 3 4		1 byte	1 byte
LIL	1 2 3 4		1 byte	1 byte
TCO	2 3 4		1 byte	1 byte
MAO	2 3 4		2 bytes	6 chiffres
BUE	1 2 3 4		3 bytes	5 chiffres
SUD	2 3 4		3 bytes	5 chiffres
SCAP	2 3 4	CAP	4 bytes	9 chiffres
IDA	3 4		3 bytes	9 chiffres
IAA	3 4		3 bytes	9 chiffres
DAI	3 4		1 byte	2 chiffres
ODO	2 3 4		3 bytes	6 chiffres
NAS	3 4		1 byte	3 chiffres
ODP	2 3 4	DOB	3 bytes	6 chiffres
BOP	2 3 4	BOB	3 bytes	5 chiffres
NOP	4	NRT	2 bytes	1 chiffre

NOMS	Types records Projet 80.	Changements NOMS.	LONGUEUR	Longueur prochaine.
PAE	2 3 4		4 bytes	5 chiffres
NUS	3 4	LES	1 byte	1 bit
BDO	?		3 bytes	-
DNA	2 3 4		4 bytes	8 chiffres
SGC	1 2 3 4		1 byte	1 chiffre ou 1 lettre
SEX	2 3 4		1 byte	1 chiffre
ECI	2 3 4	ECIC		1 chiffre
CAT	2 3 4		1 byte	1 chiffre
DOB	?			-
RQI i = 1 → 5	2 3 4	IRQI i = 1 → 5	3 bytes par i	8 chiffres par i
CGT	1 2 3 4	NC	1 byte	2 chiffres
RBUROE	1 2 3 4	RLEUE	1 byte	1 chiffre
BYT 95	(2 3 4)		1 byte	5 chiffres
ETR	4	CETR	1 byte	
FLAG	(2 3 4)		2 bytes	-

NOMS	Types records Projet 80.	Changements NOMS.	LONGUEUR	Longueur prochaine.
	Zones reprises sous le nom des différents ind. : ↙ Toutes ces zones ne font qu'un bit lorsqu'il y a une modification de nom, des renseignements complémentaires ajoutés → zone un peu plus longue.			
DUP	2 3 4		1 bit	2 chiffres
OPP	2 3 4		1 bit	1 chiffre
CAR	3 4		1 bit	1 bit
ETR	4	CETR	1 bit	5 chiffres
SOL	2 3 4		1 bit	1 bit
DSO	3 4		1 bit	9 chiffres
PRECAPIT	3 4	LAC	1 bit	1 bit
LDI	2 3 4	LDB	1 bit	1 bit
LDA	2 3 4		1 bit	1 bit
PREPSOL	2 3 4		1 bit	1 bit
PNF	?		1 bit	-
RNN	?		1 bit	-
INI	3 4		1 bit	1 bit
EXIT	2 3 4		1 bit	1 bit
OPER	?		1 bit	-

LONGUEUR	TYPES RECORDS	CHANGEMENTS	REMARQUES.
1 bit	2 3 4	ANNUL	compte annulé (ex. Le guichetier renverse de l'encre sur un livret qui devient inutilisable.)
1 bit	2 3 4	NOR	
9 chiffres	2 3 4	TNOR	
6 chiffres	3 4	DVH	date dernière opération fichier historique
6 chiffres	3 4	OOSC	
10 chiffres	3 4	IOSC	
9 chiffres	3 4	COSC	
9 chiffres	3 4	PDSC	
1 bit	3 4	PBS	plus bas solde à vérifier
1 bit	3 4	SRC	pour les sociétés (savoir qui peut retirer de l'argent, par ex.)
2 chiffres	3 4	NM8	remboursement total; c'est une fiche. on souhaite pouvoir les comptabiliser.
1 bit	3 4	LES	déduit de NVS pour savoir si c'est une abonnement scolaire c.à.d. une opération faite par l'intermédiaire des écoles.
1 bit	3 4	LEA	en ordre de prélèvement à la fin du mois de X..... Frs. du compte pour mettre sur le livret, si cela est possible.
1 bit	3 4	CED	prélèvement pour prêt hypothécaire
1 bit	3 4	SES	renseignements pour les comptes ES

NOMS	Types records Projet 80.	Changements NOMS	LONGUEUR	Longueur prochaine
CHN	3 4		1 bit	1 bit
LOI	2 3 4		1 bit	1 bit
PRO	2 3 4		1 bit	1 chiffre
RES	2 3 4	TLR	1 bit	1 bit
FIC	2 3 4	INFO	1 bit	1 bit
DOS	2 3 4	DOSLO	1 bit	1 bit
ACC	2 3 4		1 bit	1 bit
147	(2 3 4)		1 bit	1 bit

NOUVELLES ZONES POUR LE PROJET 80.

LONGUEUR	TYPES RECORDS	CHANGEMENTS	REMARQUES.
1 bit	2 3 4	TLR	indication spéciale prendre 1 dossier
3 chiffres	2 3 4	CRES	
1 bit	2 3 4	PASCOR	on n'envoie pas de correspondance à ce client
1 chiffre	2 3 4	ATT	ancien Bit 26 dans la description des feuilles
2 chiffres	2 3 4	MCF	
4 chiffres	2 3 4	ONLH	n° dernière ligne de la dernière opération du fichier historique.

LONGUEUR	TYPES RECORDS	CHANGEMENTS	REMARQUES.
8 chiffres	4	INCA	ancien RIN (plus repris sur dernier fichier mais s'y trouvant toujours.
8 chiffres	4	PICA	précompte sur RIN (vont disparaître fin d'année)
2 chiffres	4	NATT	compl.au bit 26 (ATT) (Nombre de membres de la société.)
4 chiffres	4	VGAR	propositions qui sont déjà rejetées avant l'exécution du projet 80.
6 chiffres	4	NGAR	
1 bit	4	APRE	attestation délivrée pour le précompte
1 chiffre	1 2 3 4	TCCD	type record (1 → 4)

REMARQUES.

Nous distinguons quatre types d'enregistrements pour le nouveau fichier CCD, à savoir :

. type 1	record blanc	longueur + ou - 17 bytes
. type 2	record non actif (soldé + flaggé)	longueur + ou - 82 bytes
. type 3	actif	longueur +/- 134 bytes
. type 4	actif spécial	longueur +/- 156 bytes

La longueur moyenne sera fixée à 130 bytes.

C. - Description du disque 3350 d'IBM.

Speed

Seek time*
- Cylinder to cylinder . . . . . 10 ms
- Average . . . . . 25 ms
- Maximum . . . . . 50 ms

Data rate . . . . . 1198 KB/second

Latency
- Minimum . . . . . 0 ms
- Average . . . . . 8.4 ms
- Maximum . . . . . 16.7 ms

*The fixed heads on Models A2F, B2F and C2F provide up to 1.144 million bytes of zero-seek-time storage per drive.

Capacity

	Native Mode	Compatibility Modes	
		3330-1	3330-11
Logical volumes per drive	1	2	1
Cylinders per drive	555 (plus 5 alternates)	404 per logical volume (plus 7 alternates)	808 (plus 7 alternates)
Tracks per cylinder	30	19	19
Tracks per drive	16,650 (plus 150 alternates)	7,676 per logical volume (plus 133 alternates)	15,352 (plus 133 alternates)
Track capacity (bytes)	19,069	13,030	13,030
Cylinder capacity (bytes)	572,070	247,570	247,570
Drive capacity (bytes)	317.5 million	100 million per logical volume	200 million

Capacity Table – Native Mode

Bytes per Record				Number of Records		
Without Keys		With Keys		Per	Per	Per
Min.	Max.	Min.	Max.	Track	Cylinder	Drive
9443	19069	9361	18987	1	30	16650
6234	9442	6152	9360	2	60	33300
4629	6233	4547	6151	3	90	49950
3666	4628	3584	4546	4	120	66600
3025	3665	2943	3583	5	150	83250
2566	3024	2484	2942	6	180	99900
2222	2565	2140	2483	7	210	116550
1955	2221	1873	2139	8	240	133200
1741	1954	1659	1872	9	270	149850
1566	1740	1484	1658	10	300	166500
1420	1565	1338	1483	11	330	183150
1297	1419	1215	1337	12	360	199800
1191	1296	1109	1214	13	390	216450
1099	1190	1017	1108	14	420	233100
1019	1098	937	1016	15	450	249750
948	1018	866	936	16	480	266400
885	947	803	865	17	510	283050
829	884	747	802	18	540	299700
778	828	696	746	19	570	316350
732	777	650	695	20	600	333000
691	731	609	649	21	630	349650
653	690	571	608	22	660	366300
618	652	536	570	23	690	382950
586	617	504	535	24	720	399600
556	585	474	503	25	750	416250
529	555	447	473	26	780	432900
503	528	421	446	27	810	449550
479	502	397	420	28	840	466200
457	478	375	396	29	870	482850
437	456	355	374	30	900	499500
417	436	335	354	31	930	516150
399	416	317	334	32	960	532800
382	398	300	316	33	990	549450
366	381	284	299	34	1020	566100
350	365	268	283	35	1050	582750
336	349	254	267	36	1080	599400
322	335	240	253	37	1110	616050
309	321	227	239	38	1140	632700
297	308	215	226	39	1170	649350
285	296	203	214	40	1200	666000
274	284	192	202	41	1230	682650
263	273	181	191	42	1260	699300
253	262	171	180	43	1290	715950
243	252	161	170	44	1320	732600
234	242	152	160	45	1350	749250
225	233	143	151	46	1380	765900
217	224	135	142	47	1410	782550
208	216	126	134	48	1440	799200
201	207	119	125	49	1470	815850
193	200	111	118	50	1500	832500

Capacity Table (cont'd.)

Bytes per Record				Number of Records		
Without Keys		With Keys		Per	Per	Per
Min.	Max.	Min.	Max.	Track	Cylinder	Drive
186	192	104	110	51	1530	849150
179	185	97	103	52	1560	865800
172	178	90	96	53	1590	882450
166	171	84	89	54	1620	899100
159	165	77	83	55	1650	915750
153	158	71	76	56	1680	932400
147	152	65	70	57	1710	949050
142	146	60	64	58	1740	965700
136	141	54	59	59	1770	982350
131	135	49	53	60	1800	999000
126	130	44	48	61	1830	1015650
121	125	39	43	62	1860	1032300
116	120	34	38	63	1890	1048950
112	115	30	33	64	1920	1065600
107	111	25	29	65	1950	1082250
103	106	21	24	66	1980	1098900
99	102	17	20	67	2010	1115550
95	98	13	16	68	2040	1132200
91	94	9	12	69	2070	1148850
87	90	5	8	70	2100	1165500
83	86	4	4	71	2130	1182150
79	82			72	2160	1198800
76	78			73	2190	1215450
72	75			74	2220	1232100
69	71			75	2250	1248750
66	68			76	2280	1265400
62	65			77	2310	1282050
59	61			78	2340	1298700
56	58			79	2370	1315350
53	55			80	2400	1332000
50	52			81	2430	1348650
47	49			82	2460	1365300
45	46			83	2490	1381950
42	44			84	2520	1398600
39	41			85	2550	1415250
37	38			86	2580	1431900
34	36			87	2610	1448550
32	33			88	2640	1465200
29	31			89	2670	1481850
27	28			90	2700	1498500
25	26			91	2730	1515150
23	24			92	2760	1531800
20	22			93	2790	1548450
18	19			94	2820	1565100
16	17			95	2850	1581750
14	15			96	2880	1598400
12	13			97	2910	1615050
10	11			98	2940	1631700
8	9			99	2970	1648350
6	7			100	3000	1665000
4	5			101	3030	1681650
2	3			102	3060	1698300
1	1			103	3090	1714950

D. - LISTE DES V.T.O.C. du fichier CCD.

Explication des renseignements les plus importants des contenus des V.T.O.C.

DATA SET NAME	Le nom du fichier sur le disque.
DSORG	Organisation du fichier IS → séquentiel indexé.
RECFM	Format de l'enregistrement FB → fixe bloqué.
BLKSIZE	Longueur des blocs en bytes
LRECL	Longueur de l'enregistrement en bytes
KEYLEN	Longueur de la clé en bytes
EXTENTS	NO : numéro de l'extent LOW : adresse début de l'extent HIGH : adresse fin de l'extent
CYLAD	Adresse du cylinder index.
ADLIN	Adresse du plus bas niveau du master index.
ADHIN	Adresse du plus haut niveau du master index.
NOBYT	Nombre de bytes pour le niveau le plus élevé d'index.
NOTRK	Nombre de pistes pour le master index.
LTRAD	Adresse dernière entrée de la track index.
LCYAD	Adresse dernière entrée du cylinder index.
LPRAD	Adresse dernier enregistrement en zone primaire.
NOLEV	Nombre de niveaux d'index.
CYLOV	Nombre de pistes pour le cylinder overflow.
LMSAD	Dernière entrée du master index.

.....

SYSTEMS SUPPORT UTILITIES---IEHLIST

DATE: 81.287 TIME: 11.55.00
CONTENTS OF VTOC ON VOL FFFFF4

```
----------DATA SET NAME---------- ID  SER NO  SEQ NO  CREDT  EXPDT  REFDT  NO EXT  DSORG  RECFM OPTCD BLKSIZE
MD.P10000.CCD01                   1  FFFFF4       1  34876  00000  28781      2     IS     FB     12    6144

          LRECL  KEYLEN  INITIAL ALLOC  2ND ALLOC/LAST BLK PTR(T-R-L)   USED PDS BYTES  FMT 2 OR 3(C-H-R)/DSCB(C-H-R)
             96      4     ABSTR                0                                        0   1   3     0   1   4

          EXTENTS  NO  LOW(C-H)   HIGH(C-H)    NO  LOW(C-H)   HIGH(C-H)
                    0     1   0        1  29     1      3   0      185  29

     2MIND(M-B-C-H)/3MIND(M-B-C-H)/L2MEN(C-H-R)/L3MIN(C-H-R)/CYLAD(M-B-C-H)/ADLIN(M-B-C-H)/ADHIN(M-B-C-H)/NOBYT/ NOTRK
       0 0   0   0    0 0   0   0    0   0   0    0   0   0    2 0   1  27    0 0   0   0    2 0   1  27   2576     0

     LTRAD(C-H-R)/LCYAD(C-H-R)/LMSAD(C-H-R)/LPRAD(M-B-C-H-R) /NULEV /CYLOV/ TAGDT/    PRCTR  / OVRCT/ RORG1/PTRDS(C-H-R)
     180   0  47    1  29  42    0   0   0    1 0 180  24   1     1       0          990001           0      0
                              ----UNABLE TO CALCULATE EMPTY SPACE.
```

```
                    SYSTEMS SUPPORT UTILITIES---IEHLIST

DATE: 81.287  TIME: 11.55.01
              CONTENTS OF VTOC ON VOL FFFFF4

---------------DATA SET NAME---------------  ID  SER NO  SEQ NO  CREDT  EXPDT  REFDT  NO EXT  DSORG  RECFM OPTCD BLKSIZE
MD.P10000.CCD02                               1  FFFFF4      1   34876  00000  28781     2     IS     FB     12     6144

          LRECL  KEYLEN  INITIAL ALLOC  2ND ALLOC/LAST BLK PTR(T-R-L)   USED PDS BYTES  FMT 2 OR 3(C-H-R)/DSCB(C-H-R)
             96      4     ABSTR                0                                          0   1   5     0   1   6

          EXTENTS  NO  LOW(C-H)   HIGH(C-H)    NO  LOW(C-H)   HIGH(C-H)
                    0     2   0       2  29     1   186   0     368  29

     2MIND(M-B-C-H)/3MIND(M-B-C-H)/L2MEN(C-H-R)/L3MIN(C-H-R)/CYLAD(M-B-C-H)/ADLIN(M-B-C-H)/ADHIN(M-B-C-H)/NOBYT/ NOTRK
       0 0   0   0    0 0   0   0    0   0   0    0   0   0    2 0   2  27    0 0   0   0    2 0   2  27   2576     0

     LTRAD(C-H-R)/LCYAD(C-H-R)/LMSAD(C-H-R)/LPRAD(M-B-C-H-R) /NOLEV /CYLOV/ TAGDT/   PRCTR  / OVRCT/ RORG1/PTRDS(C-H-R)
     363   0  47    2  29  42    0   0   0    1 0 363  24   1    1      0         990001          0      0
                              ----UNABLE TO CALCULATE EMPTY SPACE.
```

```
                    SYSTEMS SUPPORT UTILITIES---IEHLIST

DATE: 81.287  TIME: 11.55.03
              CONTENTS OF VTOC ON VOL FFFFF4

---------------DATA SET NAME---------------  ID  SER NO  SEQ NO  CREDT  EXPDT  REFDT  NO EXT  DSORG  RECFM OPTCD BLKSIZE
MD.P10000.CCD03                               1  FFFFF4      1   19677  00000  28781     1     IS    FB     12     6144

          LRECL  KEYLEN  INITIAL ALLOC  2ND ALLOC/LAST BLK PTR(T-R-L)   USED PDS BYTES  FMT 2 OR 3(C-H-R)/DSCB(C-H-R)
             96      4     ABSTR                0                                          0   1   7     0   1   8

          EXTENTS  NO  LOW(C-H)   HIGH(C-H)
                    0   369   0     551  29

     2MIND(M-B-C-H)/3MIND(M-B-C-H)/L2MEN(C-H-R)/L3MIN(C-H-R)/CYLAD(M-B-C-H)/ADLIN(M-B-C-H)/ADHIN(M-B-C-H)/NOBYT/ NOTRK
       0 0   0   0    0 0   0   0    0   0   0    0   0   0    1 0 551  27    0 0   0   0    1 0 551  27   2562     0

     LTRAD(C-H-R)/LCYAD(C-H-R)/LMSAD(C-H-R)/LPRAD(M-B-C-H-R) /NOLEV /CYLOV/ TAGDT/   PRCTR  / OVRCT/ RORG1/PTRDS(C-H-R)
     546   0  47  551  29  42    0   0   0    1 0 546  24   1    1      0         990001          0      0
                              ----UNABLE TO CALCULATE EMPTY SPACE.
```

```
                    SYSTEMS SUPPORT UTILITIES---IEHLIST

DATE: 81.287  TIME: 11.55.05
              CONTENTS OF VTOC ON VOL FFFFF5

---------------DATA SET NAME---------------  ID  SER NO  SEQ NO  CREDT  EXPDT  REFDT  NO EXT  DSORG  RECFM OPTCD BLKSIZE
MD.P10000.CCD04                               1  FFFFF5      1   24877  00000  28781     2     IS     FB     12     6144

          LRECL  KEYLEN  INITIAL ALLOC  2ND ALLOC/LAST BLK PTR(T-R-L)   USED PDS BYTES  FMT 2 OR 3(C-H-R)/DSCB(C-H-R)
             96      4     ABSTR                0                                                0   1   3     0   1   4

          EXTENTS  NO  LOW(C-H)   HIGH(C-H)    NO  LOW(C-H)   HIGH(C-H)
                    0    1   0       1  29     1     3   0     185  29

     2MIND(M-B-C-H)/3MIND(M-B-C-H)/L2MEN(C-H-R)/L3MIN(C-H-R)/CYLAD(M-B-C-H)/ADLIN(M-B-C-H)/ADHIN(M-B-C-H)/NOBYT/ NOTRK
       0 0   0   0    0 0   0   0    0   0   0    0   0   0    2 0   1  27    0 0   0   0    2 0   1  27   2576     0

     LTRAD(C-H-R)/LCYAD(C-H-R)/LMSAD(C-H-R)/LPRAD(M-B-C-H-R) /NOLEV /CYLOV/ TAGDT/   PRCTR  / OVRCT/ RORG1/PTRDS(C-H-R)
     180   0  47    1  29  42    0   0   0    1 0 180  24   1    1      0          990001          0      0
                                  ----UNABLE TO CALCULATE EMPTY SPACE.
```

SYSTEMS SUPPORT UTILITIES---IEHLIST

DATE: 81.287 TIME: 11.55.07
CONTENTS OF VTOC ON VOL FFFFF5

```
----------DATA SET NAME---------- ID  SER NO  SEQ NO  CREDT  EXPDT  REFDT  NO EXT  DSORG  RECFM OPTCD BLKSIZE
MD.P10000.CCD05                   1  FFFFF5       1  24877  00000  28781       2     IS     FB     12    6144

          LRECL  KEYLEN  INITIAL ALLOC  2ND ALLOC/LAST BLK PTR(T-R-L)   USED PDS BYTES  FMT 2 OR 3(C-H-R)/DSCB(C-H-R)
             96       4     ABSTR              0                                          0   1   5     0   1   6

          EXTENTS  NO  LOW(C-H)   HIGH(C-H)    NO  LOW(C-H)   HIGH(C-H)
                    0     2   0        2  29     1   180   0      360  29

     2MIND(M-B-C-H)/3MIND(M-B-C-H)/L2MEN(C-H-R)/L3MIN(C-H-R)/CYLAD(M-B-C-H)/ADLIN(M-B-C-H)/ADHIN(M-B-C-H)/NOBYT/ NOTRK
       0 0   0   0    0 0   0   0    0   0   0    0   0   0    2 0   2  27    0 0   0   0    2 0   2  27   2576     0

     LTRAD(C-H-R)/LCYAD(C-H-R)/LMSAD(C-H-R)/LPRAD(M-B-C-H-R) /NOLEV /CYLOV/ TAGDT/    PRCTR  / OVRCT/ RORG1/PTRDS(C-H-R)
     363   0  47    2  29  42    0   0   0    1 0 363  24    1     1       0          990001          0      0
                              ----UNABLE TO CALCULATE EMPTY SPACE.
```

```
                    SYSTEMS SUPPORT UTILITIES---IEHLIST

DATE: 81.287  TIME: 11.55.14
              CONTENTS OF VTOC ON VOL FFFFF3

---------------DATA SET NAME---------------  ID  SER NO  SEQ NO  CREDT  EXPDT  REFDT  NO EXT  DSORG  RECFM OPTCD BLKSIZE
MD.P10000.CCD06                               1  FFFFF3      1   32380  00000  28781     1     IS     FB     12     6144

          LRECL  KEYLEN  INITIAL ALLOC  2ND ALLOC/LAST BLK PTR(T-R-L)   USED PDS BYTES  FMT 2 OR 3(C-H-R)/DSCB(C-H-R)
             96      4     ABSTR              0                                          0   1   3     0   1   4

          EXTENTS  NO  LOW(C-H)   HIGH(C-H)
                    0     2   0       2  29

     2MIND(M-B-C-H)/3MIND(M-B-C-H)/L2MEN(C-H-R)/L3MIN(C-H-R)/CYLAD(M-B-C-H)/ADLIN(M-B-C-H)/ADHIN(M-B-C-H)/NOBYT/ NOTRK
       0 0   0   0    0 0   0   0    0   0   0    0   0   0    2 0   2  27    0 0   0   0    2 0   2  27   2576     0

     LTRAD(C-H-R)/LCYAD(C-H-R)/LMSAD(C-H-R)/LPRAD(M-B-C-H-R) /NOLEV /CYLOV/ TAGDT/   PRCTR  / OVRCT/ RORG1/PTRDS(C-H-R)
     546   0  47    2  29  42    0   0   0    1 0 546  24   1    1      0          990001          0      0
                              ----UNABLE TO CALCULATE EMPTY SPACE.
```

SYSTEMS SUPPORT UTILITIES---IEHLIST

DATE: 81.287 TIME: 11.55.10
CONTENTS OF VTOC ON VOL FFFFF6

```
----------DATA SET NAME---------- ID  SER NO  SEQ NO  CREDT  EXPDT  REFDT  NO EXT  DSORG  RECFM OPTCD BLKSIZE
MD.P10000.CCD07                   1  FFFFF6       1  34876  00000  28781      2     IS     FB     12     6144

          LRECL  KEYLEN  INITIAL ALLOC  2ND ALLOC/LAST BLK PTR(T-R-L)   USED PDS BYTES  FMT 2 OR 3(C-H-R)/DSCB(C-H-R)
             96       4     ABSTR              0                                        0   1   3     0   1   4

          EXTENTS  NO  LOW(C-H)    HIGH(C-H)     NO  LOW(C-H)    HIGH(C-H)
                    0     1   0         1  29      1     3   0       185  29

     2MIND(M-B-C-H)/3MIND(M-B-C-H)/L2MEN(C-H-R)/L3MIN(C-H-R)/CYLAD(M-B-C-H)/ADLIN(M-B-C-H)/ADHIN(M-B-C-H)/NOBYT/ NOTRK
       0 0   0   0    0 0   0   0    0   0   0    0   0   0    2 0   1  27    0 0   0   0    2 0   1  27   2576     0

     LTRAD(C-H-R)/LCYAD(C-H-R)/LMSAD(C-H-R)/LPRAD(M-B-C-H-R) /NOLEV /CYLOV/ TAGDT/    PRCTR  / OVRCT/ RORG1/PTRDS(C-H-R)
     180   0  47    1  29  42    0   0   0    1 0 180  24    1     1       0          990001          0       0
                              ----UNABLE TO CALCULATE EMPTY SPACE.
```

```
                    SYSTEMS SUPPORT UTILITIES---IEHLIST

DATE: 81.287  TIME: 11.55.11
              CONTENTS OF VTOC ON VOL FFFFF6

--------------DATA SET NAME--------------- ID  SER NO  SEQ NO  CREDT  EXPDT  REFDT  NO EXT  DSORG  RECFM OPTCD BLKSIZE
MD.P10000.CCD08                             1  FFFFF6      1   34876  00000  28781     2     IS     FB     12     6144

          LRECL  KEYLEN  INITIAL ALLOC  2ND ALLOC/LAST BLK PTR(T-R-L)   USED PDS BYTES  FMT 2 OR 3(C-H-R)/DSCB(C-H-R)
             96      4     ABSTR                0                                          0   1   5     0   1   6

          EXTENTS  NO  LOW(C-H)   HIGH(C-H)    NO  LOW(C-H)   HIGH(C-H)
                    0     2   0       2  29     1   186   0     368  29

    2MIND(M-B-C-H)/3MIND(M-B-C-H)/L2MEN(C-H-R)/L3MIN(C-H-R)/CYLAD(M-B-C-H)/ADLIN(M-B-C-H)/ADHIN(M-B-C-H)/NOBYT/ NOTRK
      0 0   0   0    0 0   0   0    0   0   0    0   0   0    2 0   2  27    0 0   0   0    2 0   2  27   2576     0

    LTRAD(C-H-R)/LCYAD(C-H-R)/LMSAD(C-H-R)/LPRAD(M-B-C-H-R) /NOLEV /CYLOV/ TAGDT/   PRCTR  / OVRCT/ RORG1/PTRDS(C-H-R)
    363   0  47    2  29  42    0   0   0    1 0 363  24   1    1      0          990001          0      0
                          ----UNABLE TO CALCULATE EMPTY SPACE.
```

```
                    SYSTEMS SUPPORT UTILITIES---IEHLIST

DATE: 81.287  TIME: 11.55.12
              CONTENTS OF VTOC ON VOL FFFFF7

---------------DATA SET NAME---------------  ID  SER NO  SEQ NO  CREDT  EXPDT  REFDT  NO EXT  DSORG  RECFM OPTCD BLKSIZE
MD.P10000.CCD09                               1  FFFFF7      1   34876  00000  28781     2     1S    FB     12     6144

          LRECL  KEYLEN  INITIAL ALLOC  2ND ALLOC/LAST BLK PTR(T-R-L)   USED PDS BYTES  FMT 2 OR 3(C-H-R)/DSCB(C-H-R)
             96      4     ABSTR                0                                          0   1   3     0   1   4

          EXTENTS  NO  LOW(C-H)   HIGH(C-H)    NO  LOW(C-H)   HIGH(C-H)
                    0     1   0       1  29     1     3   0     185  29

     2MIND(M-B-C-H)/3MIND(M-B-C-H)/L2MEN(C-H-R)/L3MIN(C-H-R)/CYLAD(M-B-C-H)/ADLIN(M-B-C-H)/ADHIN(M-B-C-H)/NOBYT/ NOTRK
       0 0   0   0    0 0   0   0    0   0   0    0   0   0    2 0   1  27    0 0   0   0    2 0   1  27   2576     0

     LTRAD(C-H-R)/LCYAD(C-H-R)/LMSAD(C-H-R)/LPRAD(M-B-C-H-R) /NOLEV /CYLOV/ TAGDT/   PRCTR  / OVRCT/ RORG1/PTRDS(C-H-R)
     180   0  47    1  29  42    0   0   0    1 0 180  24   1    1      0          990001          0      0
                              ----UNABLE TO CALCULATE EMPTY SPACE.
```

```
                    SYSTEMS SUPPORT UTILITIES---IEHLIST

DATE: 81.287  TIME: 11.55.13
              CONTENTS OF VTOC ON VOL FFFFF7

--------------DATA SET NAME--------------- ID  SER NO  SEQ NO  CREDT  EXPDT  REFDT  NO EXT  DSORG  RECFM OPTCD BLKSIZE
MD.P10000.CCD10                             1  FFFFF7      1   34876  00000  28781     2     IS    FB     12     6144

          LRECL  KEYLEN  INITIAL ALLOC  2ND ALLOC/LAST BLK PTR(T-R-L)   USED PDS BYTES  FMT 2 OR 3(C-H-R)/DSCB(C-H-R)
             96      4     ABSTR                0                                          0   1   5     0   1   6

          EXTENTS  NO  LOW(C-H)   HIGH(C-H)    NO  LOW(C-H)   HIGH(C-H)
                    0     2   0       2  29     1   186   0     552  29

    2MIND(M-B-C-H)/3MIND(M-B-C-H)/L2MEN(C-H-R)/L3MIN(C-H-R)/CYLAD(M-B-C-H)/ADLIN(M-B-C-H)/ADHIN(M-B-C-H)/NOBYT/ NOTRK
      0 0   0   0    0 0   0   0    0   0   0    0   0   0    2 0   2  24    0 0   0   0    2 0   2  24   5152     0

    LTRAD(C-H-R)/LCYAD(C-H-R)/LMSAD(C-H-R)/LPRAD(M-B-C-H-R) /NOLEV /CYLOV/ TAGDT/   PRCTR  / OVRCT/ RORG1/PTRDS(C-H-R)
    481   0  55    2  28  24    0   0   0    1 0 481  28   3    1      0        1647875          0      0
                          ----UNABLE TO CALCULATE EMPTY SPACE.
```

SIMULATION ECLATEMENT SUR 2 PAGES
NBRE DE SIMULATIONS:5
PAGES CHARGEES @ 80 % (EN BYTES) CI=1024
RECORDS VARIABLES
NBRE DE RECORDS DE DEPART:10000

OPERATIONS	ADDIT-ART	ADDIT-PAGE	NBRE-ART	NBRE-PAGES	%OCC
400.	400.	76	10400.	576	76.9
800.	800.	175	10800.	675	71.5
1200.	1200.	258	11200.	758	69.1
1600.	1600.	336	11600.	836	67.8
2000.	2000.	409	12000.	909	66.9
2400.	2400.	468	12400.	968	67.0
2800.	2800.	534	12800.	1034	66.7
3200.	3200.	594	13200.	1094	66.6
3600.	3600.	649	13600.	1149	66.7
4000.	4000.	703	14000.	1203	67.0
4400.	4400.	746	14400.	1246	67.4
4800.	4800.	800	14800.	1300	67.5
5200.	5200.	850	15200.	1350	67.7
5600.	5600.	897	15600.	1397	68.0
6000.	6000.	954	16000.	1454	68.0
6400.	6400.	1004	16400.	1504	67.8
6800.	6800.	1057	16800.	1557	68.0
7200.	7200.	1110	17200.	1610	68.2
7600.	7600.	1163	17600.	1663	68.3
8000.	8000.	1218	18000.	1718	68.5
8400.	8400.	1280	18400.	1780	68.3
8800.	8800.	1342	18800.	1842	68.3
9200.	9200.	1408	19200.	1908	68.3
9600.	9600.	1467	19600.	1967	68.4
10000.	10000.	1531	20000.	2031	68.5
10400.	10400.	1600	20400.	2100	68.3
10800.	10800.	1658	20800.	2158	68.5
11200.	11200.	1725	21200.	2225	68.4
11600.	11600.	1791	21600.	2291	68.4
12000.	12000.	1856	22000.	2356	68.3
12400.	12400.	1916	22400.	2416	68.4
12800.	12800.	1991	22800.	2491	68.1
13200.	13200.	2051	23200.	2551	68.2
13600.	13600.	2118	23600.	2618	68.1
14000.	14000.	2177	24000.	2677	68.3
14400.	14400.	2245	24400.	2745	68.2
14800.	14800.	2311	24800.	2811	68.2
15200.	15200.	2381	25200.	2881	68.0
15600.	15600.	2444	25600.	2944	68.1
16000.	16000.	2507	26000.	3007	68.0
16400.	16400.	2550	26400.	3050	67.9
16800.	16800.	2593	26800.	3093	68.0
17200.	17200.	2650	27200.	3150	67.9
17600.	17600.	2695	27600.	3195	67.9
18000.	18000.	2737	28000.	3237	67.8
18400.	18400.	2768	28400.	3268	67.8
18800.	18800.	2824	28800.	3324	67.9
19200.	19200.	2874	29200.	3374	68.3
19600.	19600.	2942	29600.	3442	68.1
20000.	20000.	3001	30000.	3501	68.3

SIMULATION ECLATEMENT SUR 2 PAGES
NBRE DE SIMULATIONS:5
PAGES CHARGEES @ 80 % (EN BYTES) CI=2048
RECORDS VARIABLES
NBRE DE RECORDS DE DEPART:10000

OPERATIONS	ADDIT-ART	ADDIT-PAGE	NBRE-ART	NBRE-PAGES	%OCC
400.	400.	15	10400.	515	82.5
800.	800.	69	10800.	569	78.9
1200.	1200.	134	11200.	634	74.2
1600.	1600.	200	11600.	700	70.3
2000.	2000.	248	12000.	748	68.8
2400.	2400.	290	12400.	790	67.9
2800.	2800.	326	12800.	826	67.6
3200.	3200.	365	13200.	865	67.0
3600.	3600.	391	13600.	891	67.3
4000.	4000.	422	14000.	922	67.4
4400.	4400.	441	14400.	941	68.0
4800.	4800.	464	14800.	964	68.5
5200.	5200.	485	15200.	985	69.1
5600.	5600.	509	15600.	1009	69.3
6000.	6000.	534	16000.	1034	69.7
6400.	6400.	556	16400.	1056	70.0
6800.	6800.	579	16800.	1079	70.5
7200.	7200.	602	17200.	1102	70.9
7600.	7600.	629	17600.	1129	71.1
8000.	8000.	656	18000.	1156	71.4
8400.	8400.	684	18400.	1184	71.5
8800.	8800.	717	18800.	1217	71.6
9200.	9200.	751	19200.	1251	71.5
9600.	9600.	786	19600.	1286	71.4
10000.	10000.	825	20000.	1325	71.1
10400.	10400.	858	20400.	1358	71.1
10800.	10800.	893	20800.	1393	71.1
11200.	11200.	925	21200.	1425	71.1
11600.	11600.	963	21600.	1463	70.9
12000.	12000.	996	22000.	1496	71.0
12400.	12400.	1032	22400.	1532	70.8
12800.	12800.	1068	22800.	1568	70.7
13200.	13200.	1104	23200.	1604	70.6
13600.	13600.	1137	23600.	1637	70.6
14000.	14000.	1174	24000.	1674	70.5
14400.	14400.	1218	24400.	1718	70.1
14800.	14800.	1254	24800.	1754	70.0
15200.	15200.	1290	25200.	1790	70.0
15600.	15600.	1325	25600.	1825	69.9
16000.	16000.	1362	26000.	1862	69.7
16400.	16400.	1380	26400.	1880	69.8
16800.	16800.	1402	26800.	1902	69.9
17200.	17200.	1430	27200.	1930	69.9
17600.	17600.	1454	27600.	1954	69.9
18000.	18000.	1471	28000.	1971	70.0
18400.	18400.	1488	28400.	1988	70.0
18800.	18800.	1518	28800.	2018	70.1
19200.	19200.	1549	29200.	2049	70.2
19600.	19600.	1582	29600.	2082	70.2
20000.	20000.	1613	30000.	2113	70.3

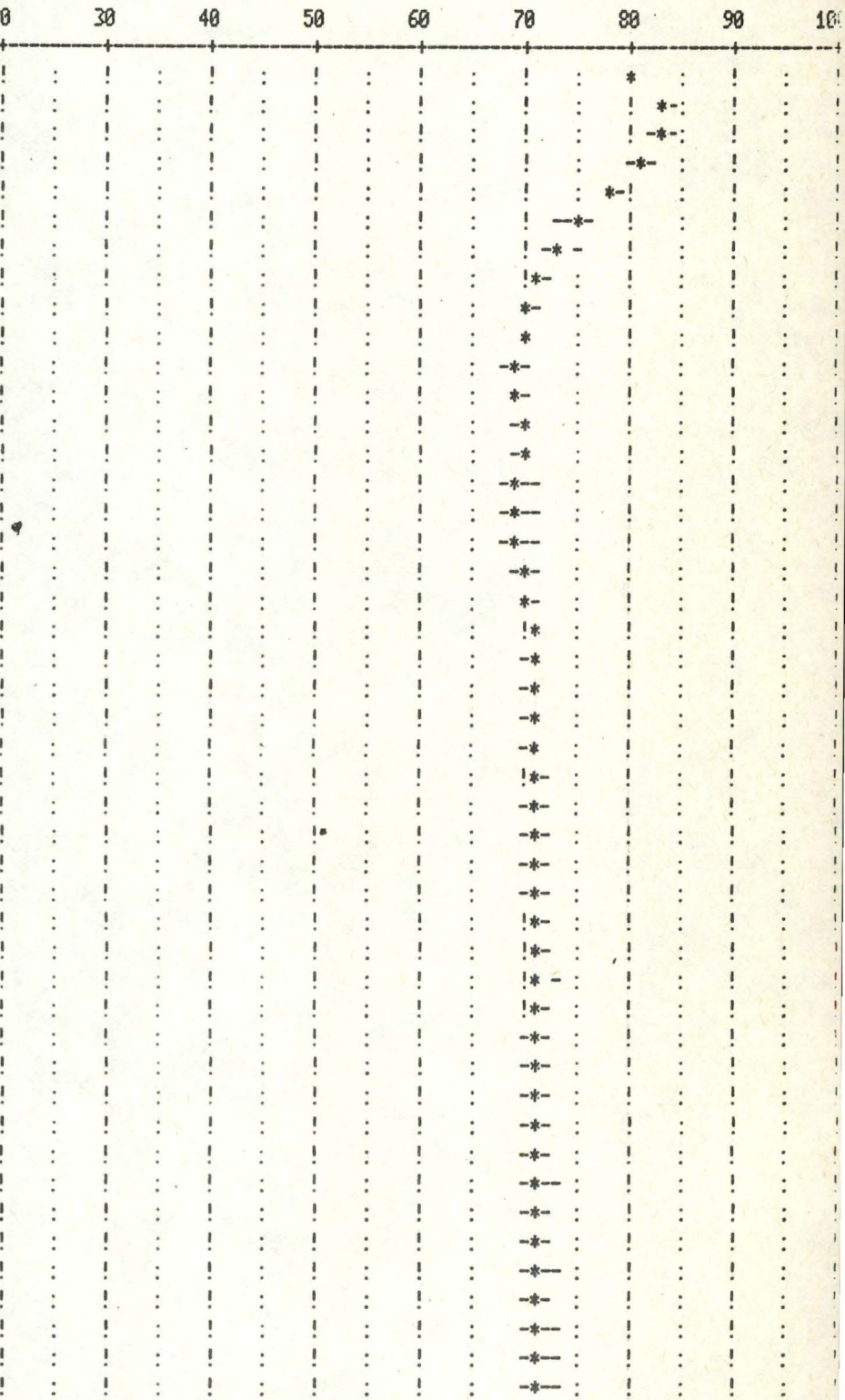
SIMULATION ECLATEMENT SUR 2 PAGES
NBRE DE SIMULATIONS:5
PAGES CHARGEES @ 80 % (EN BYTES) CI=4096
RECORDS VARIABLES
NBRE DE RECORDS DE DEPART:10000

OPERATIONS	ADDIT-ART	ADDIT-PAGE	NBRE-ART	NBRE-PAGES	%OCC
400.	400.	0	10400.	500	82.6
800.	800.	1	10800.	501	85.2
1200.	1200.	8	11200.	508	86.5
1600.	1600.	31	11600.	531	85.2
2000.	2000.	66	12000.	566	82.0
2400.	2400.	101	12400.	601	79.0
2800.	2800.	140	12800.	640	75.9
3200.	3200.	186	13200.	686	72.4
3600.	3600.	224	13600.	724	69.9
4000.	4000.	253	14000.	753	68.6
4400.	4400.	278	14400.	778	67.6
4800.	4800.	305	14800.	805	66.6
5200.	5200.	328	15200.	828	66.0
5600.	5600.	346	15600.	846	65.7
6000.	6000.	369	16000.	869	65.2
6400.	6400.	381	16400.	881	65.4
6800.	6800.	400	16800.	900	65.2
7200.	7200.	413	17200.	913	65.4
7600.	7600.	425	17600.	925	65.8
8000.	8000.	437	18000.	937	66.2
8400.	8400.	448	18400.	948	66.7
8800.	8800.	458	18800.	958	67.3
9200.	9200.	466	19200.	966	68.0
9600.	9600.	475	19600.	975	68.8
10000.	10000.	484	20000.	984	69.4
10400.	10400.	495	20400.	995	69.9
10800.	10800.	506	20800.	1006	70.4
11200.	11200.	517	21200.	1017	70.9
11600.	11600.	529	21600.	1029	71.4
12000.	12000.	536	22000.	1036	72.1
12400.	12400.	547	22400.	1047	72.6
12800.	12800.	556	22800.	1056	73.2
13200.	13200.	568	23200.	1068	73.5
13600.	13600.	579	23600.	1079	73.9
14000.	14000.	592	24000.	1092	74.2
14400.	14400.	605	24400.	1105	74.5
14800.	14800.	620	24800.	1120	74.7
15200.	15200.	638	25200.	1138	74.5
15600.	15600.	653	25600.	1153	74.7
16000.	16000.	670	26000.	1170	74.6
16400.	16400.	680	26400.	1180	74.6
16800.	16800.	693	26800.	1193	74.6
17200.	17200.	708	27200.	1208	74.4
17600.	17600.	720	27600.	1220	74.4
18000.	18000.	735	28000.	1235	74.1
18400.	18400.	744	28400.	1244	74.1
18800.	18800.	759	28800.	1259	74.2
19200.	19200.	783	29200.	1283	73.7
19600.	19600.	805	29600.	1305	73.4
20000.	20000.	829	30000.	1329	73.0

GESTION DE L'ESPACE: ADDITION AVEC DEBORDEMENT : REPARTITION SUR 2 PAGES
SUPPRESSION : RESTITUTION DES PAGES VIDES
CONDITIONS D'EXPLOITATION : 100 % D'ADDITIONS
PARAMETRES PHYSIQUES : PAGES DE 10 ARTICLES EN LONGUEUR FIXE

DEPART : 250 PAGES CHARGEES A 80 %, SOIT 2000 ARTICLES
NOMBRE DE SIMULATIONS : 5

OPER	ADD-A	SUP-A	ADDP	SUPP	ART	PAGE	%OCC
0	0	0	0	0	2000	250	80.0
100	100	0	3	0	2100	253	82.9
200	200	0	15	0	2200	265	83.2
300	300	0	34	0	2300	284	81.0
400	400	0	57	0	2400	307	78.3
500	500	0	85	0	2500	335	74.7
600	600	0	106	0	2600	356	73.0
700	700	0	128	0	2700	378	71.4
800	800	0	148	0	2800	398	70.3
900	900	0	163	0	2900	413	70.2
1000	1000	0	182	0	3000	432	69.4
1100	1100	0	197	0	3100	447	69.4
1200	1200	0	209	0	3200	459	69.8
1300	1300	0	223	0	3300	473	69.7
1400	1400	0	240	0	3400	490	69.4
1500	1500	0	254	0	3500	504	69.5
1600	1600	0	268	0	3600	518	69.5
1700	1700	0	279	0	3700	529	70.0
1800	1800	0	290	0	3800	540	70.4
1900	1900	0	301	0	3900	551	70.8
2000	2000	0	315	0	4000	565	70.8
2100	2100	0	328	0	4100	578	71.0
2200	2200	0	344	0	4200	594	70.7
2300	2300	0	357	0	4300	607	70.8
2400	2400	0	370	0	4400	620	70.9
2500	2500	0	385	0	4500	635	70.9
2600	2600	0	398	0	4600	648	71.0
2700	2700	0	411	0	4700	661	71.1
2800	2800	0	426	0	4800	676	71.1
2900	2900	0	438	0	4900	688	71.3
3000	3000	0	450	0	5000	700	71.4
3100	3100	0	464	0	5100	714	71.4
3200	3200	0	479	0	5200	729	71.3
3300	3300	0	496	0	5300	746	71.1
3400	3400	0	509	0	5400	759	71.2
3500	3500	0	524	0	5500	774	71.1
3600	3600	0	538	0	5600	788	71.1
3700	3700	0	552	0	5700	802	71.1
3800	3800	0	567	0	5800	817	71.0
3900	3900	0	581	0	5900	831	71.0
4000	4000	0	594	0	6000	844	71.1
4100	4100	0	607	0	6100	857	71.2
4200	4200	0	622	0	6200	872	71.1
4300	4300	0	636	0	6300	886	71.2
4400	4400	0	648	0	6400	898	71.3
4500	4500	0	662	0	6500	912	71.3

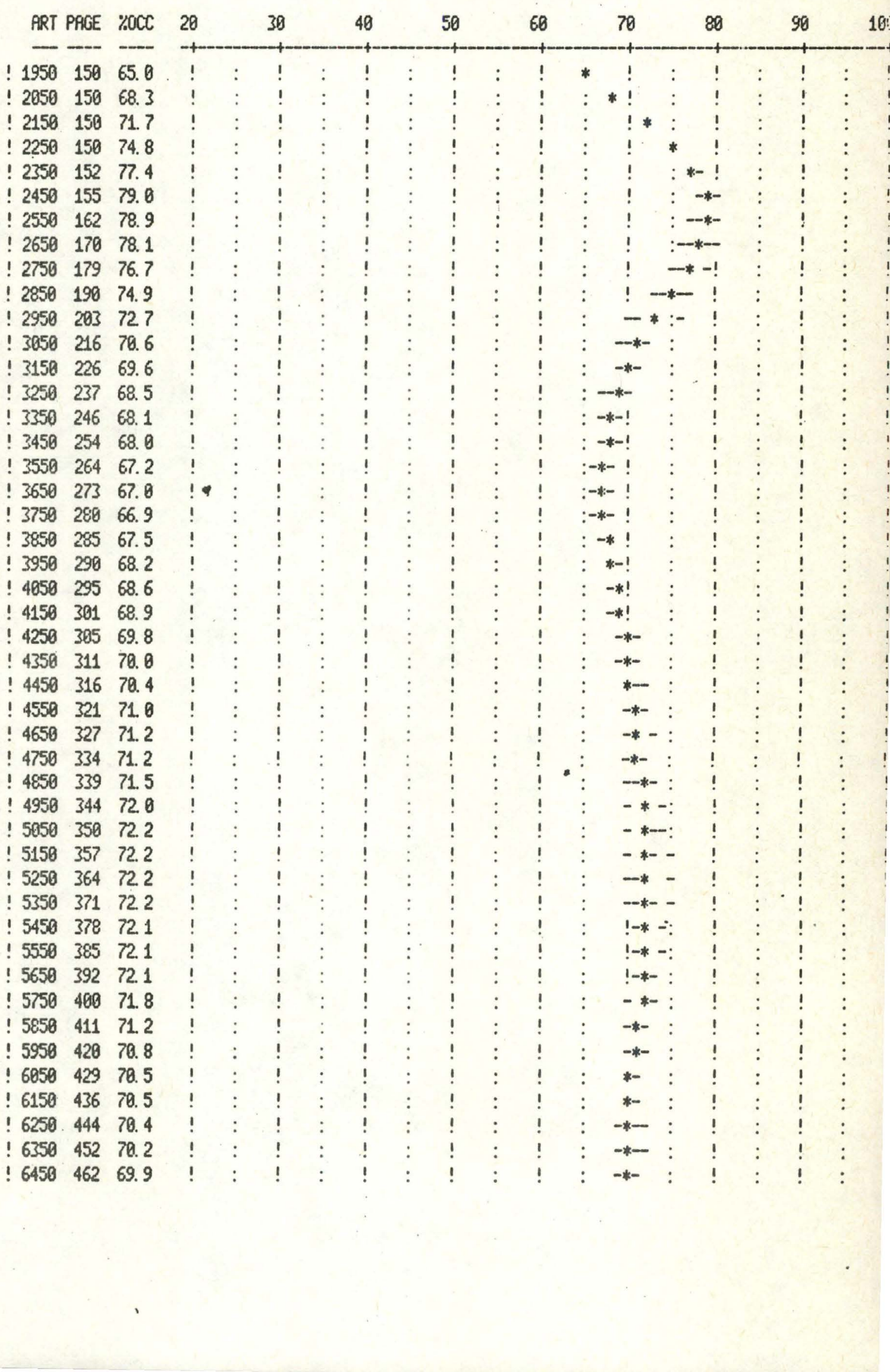


OVERFLOW IN 6245READY

GESTION DE L'ESPACE: ADDITION AVEC DEBORDEMENT : REPARTITION SUR 2 PAGES
SUPPRESSION : RESTITUTION DES PAGES VIDES
CONDITIONS D'EXPLOITATION : 100 % D'ADDITIONS
PARAMETRES PHYSIQUES : PAGES DE 20 ARTICLES EN LONGUEUR FIXE

DEPART : 150 PAGES CHARGEES A 66.6667 %, SOIT 1950 ARTICLES
NOMBRE DE SIMULATIONS : 5

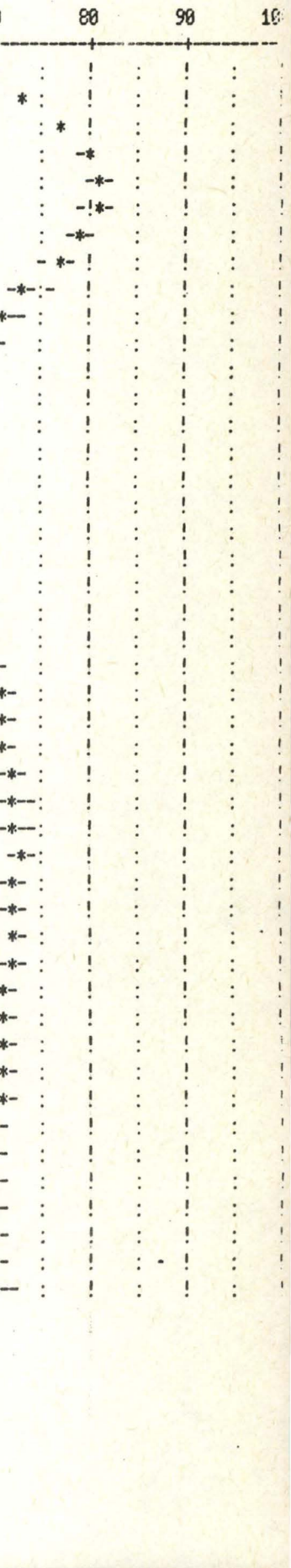
OPER	ADD-A	SUP-A	ADDP	SUPP	ART	PAGE	%OCC
0	0	0	0	0	1950	150	65.0
100	100	0	0	0	2050	150	68.3
200	200	0	0	0	2150	150	71.7
300	300	0	0	0	2250	150	74.8
400	400	0	2	0	2350	152	77.4
500	500	0	5	0	2450	155	79.0
600	600	0	12	0	2550	162	78.9
700	700	0	20	0	2650	170	78.1
800	800	0	29	0	2750	179	76.7
900	900	0	40	0	2850	190	74.9
1000	1000	0	53	0	2950	203	72.7
1100	1100	0	66	0	3050	216	70.6
1200	1200	0	76	0	3150	226	69.6
1300	1300	0	87	0	3250	237	68.5
1400	1400	0	96	0	3350	246	68.1
1500	1500	0	104	0	3450	254	68.0
1600	1600	0	114	0	3550	264	67.2
1700	1700	0	123	0	3650	273	67.0
1800	1800	0	130	0	3750	280	66.9
1900	1900	0	135	0	3850	285	67.5
2000	2000	0	140	0	3950	290	68.2
2100	2100	0	145	0	4050	295	68.6
2200	2200	0	151	0	4150	301	68.9
2300	2300	0	155	0	4250	305	69.8
2400	2400	0	161	0	4350	311	70.0
2500	2500	0	166	0	4450	316	70.4
2600	2600	0	171	0	4550	321	71.0
2700	2700	0	177	0	4650	327	71.2
2800	2800	0	184	0	4750	334	71.2
2900	2900	0	189	0	4850	339	71.5
3000	3000	0	194	0	4950	344	72.0
3100	3100	0	200	0	5050	350	72.2
3200	3200	0	207	0	5150	357	72.2
3300	3300	0	214	0	5250	364	72.2
3400	3400	0	221	0	5350	371	72.2
3500	3500	0	228	0	5450	378	72.1
3600	3600	0	235	0	5550	385	72.1
3700	3700	0	242	0	5650	392	72.1
3800	3800	0	250	0	5750	400	71.8
3900	3900	0	261	0	5850	411	71.2
4000	4000	0	270	0	5950	420	70.8
4100	4100	0	279	0	6050	429	70.5
4200	4200	0	286	0	6150	436	70.5
4300	4300	0	294	0	6250	444	70.4
4400	4400	0	302	0	6350	452	70.2
4500	4500	0	312	0	6450	462	69.9



GESTION DE L'ESPACE: ADDITION AVEC DEBORDEMENT : REPARTITION SUR 2 PAGES
SUPPRESSION : RESTITUTION DES PAGES VIDES
CONDITIONS D'EXPLOITATION : 100 % D'ADDITIONS
PARAMETRES PHYSIQUES : PAGES DE 20 ARTICLES EN LONGUEUR FIXE

DEPART : 143 PAGES CHARGEES A 70 %, SOIT 2002 ARTICLES
NOMBRE DE SIMULATIONS : 5

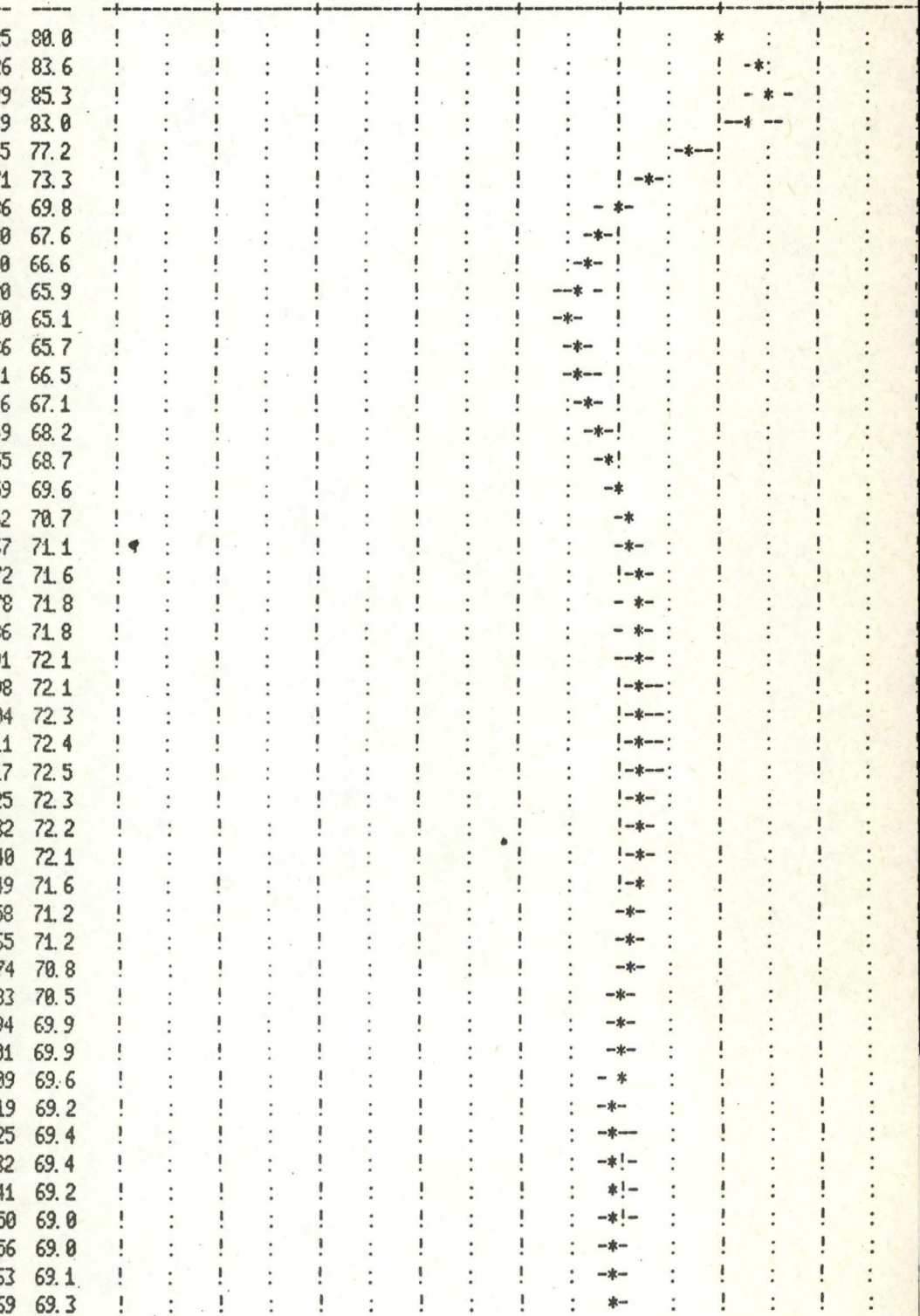
OPER	ADD-A	SUP-A	ADDP	SUPP	ART	PAGE	%OCC
0	0	0	0	0	2002	143	70.0
100	100	0	0	0	2102	143	73.5
200	200	0	0	0	2202	143	77.0
300	300	0	1	0	2302	144	79.7
400	400	0	5	0	2402	148	81.0
500	500	0	12	0	2502	155	80.9
600	600	0	22	0	2602	165	78.7
700	700	0	34	0	2702	177	76.5
800	800	0	48	0	2802	191	73.5
900	900	0	60	0	2902	203	71.4
1000	1000	0	73	0	3002	216	69.6
1100	1100	0	84	0	3102	227	68.3
1200	1200	0	96	0	3202	239	67.0
1300	1300	0	106	0	3302	249	66.4
1400	1400	0	112	0	3402	255	66.6
1500	1500	0	120	0	3502	263	66.7
1600	1600	0	126	0	3602	269	66.9
1700	1700	0	131	0	3702	274	67.6
1800	1800	0	137	0	3802	280	67.9
1900	1900	0	141	0	3902	284	68.6
2000	2000	0	147	0	4002	290	69.0
2100	2100	0	153	0	4102	296	69.3
2200	2200	0	157	0	4202	300	70.0
2300	2300	0	161	0	4302	304	70.7
2400	2400	0	166	0	4402	309	71.2
2500	2500	0	172	0	4502	315	71.5
2600	2600	0	177	0	4602	320	71.9
2700	2700	0	183	0	4702	326	72.2
2800	2800	0	189	0	4802	332	72.3
2900	2900	0	194	0	4902	337	72.7
3000	3000	0	205	0	5002	348	71.8
3100	3100	0	212	0	5102	355	71.8
3200	3200	0	221	0	5202	364	71.5
3300	3300	0	228	0	5302	371	71.5
3400	3400	0	236	0	5402	379	71.4
3500	3500	0	244	0	5502	387	71.2
3600	3600	0	251	0	5602	394	71.0
3700	3700	0	259	0	5702	402	70.9
3800	3800	0	267	0	5802	410	70.7
3900	3900	0	276	0	5902	419	70.5
4000	4000	0	283	0	6002	426	70.4
4100	4100	0	291	0	6102	434	70.4
4200	4200	0	299	0	6202	442	70.2
4300	4300	0	307	0	6302	450	70.0
4400	4400	0	312	0	6402	455	70.4
4500	4500	0	318	0	6502	461	70.5



GESTION DE L'ESPACE: ADDITION AVEC DEBORDEMENT : REPARTITION SUR 2 PAGES
SUPPRESSION : RESTITUTION DES PAGES VIDES
CONDITIONS D'EXPLOITATION : 100 % D'ADDITIONS
PARAMETRES PHYSIQUES : PAGES DE 20 ARTICLES EN LONGUEUR FIXE

DEPART : 125 PAGES CHARGEES A 80 %, SOIT 2000 ARTICLES
NOMBRE DE SIMULATIONS : 5

OPER	ADD-A	SUP-A	ADDP	SUPP	ART	PAGE	%OCC
0	0	0	0	0	2000	125	80.0
100	100	0	1	0	2100	126	83.6
200	200	0	4	0	2200	129	85.3
300	300	0	14	0	2300	139	83.0
400	400	0	30	0	2400	155	77.2
500	500	0	46	0	2500	171	73.3
600	600	0	61	0	2600	186	69.8
700	700	0	75	0	2700	200	67.6
800	800	0	85	0	2800	210	66.6
900	900	0	95	0	2900	220	65.9
1000	1000	0	105	0	3000	230	65.1
1100	1100	0	111	0	3100	236	65.7
1200	1200	0	116	0	3200	241	66.5
1300	1300	0	121	0	3300	246	67.1
1400	1400	0	124	0	3400	249	68.2
1500	1500	0	130	0	3500	255	68.7
1600	1600	0	134	0	3600	259	69.6
1700	1700	0	137	0	3700	262	70.7
1800	1800	0	142	0	3800	267	71.1
1900	1900	0	147	0	3900	272	71.6
2000	2000	0	153	0	4000	278	71.8
2100	2100	0	161	0	4100	286	71.8
2200	2200	0	166	0	4200	291	72.1
2300	2300	0	173	0	4300	298	72.1
2400	2400	0	179	0	4400	304	72.3
2500	2500	0	186	0	4500	311	72.4
2600	2600	0	192	0	4600	317	72.5
2700	2700	0	200	0	4700	325	72.3
2800	2800	0	207	0	4800	332	72.2
2900	2900	0	215	0	4900	340	72.1
3000	3000	0	224	0	5000	349	71.6
3100	3100	0	233	0	5100	358	71.2
3200	3200	0	240	0	5200	365	71.2
3300	3300	0	249	0	5300	374	70.8
3400	3400	0	258	0	5400	383	70.5
3500	3500	0	269	0	5500	394	69.9
3600	3600	0	276	0	5600	401	69.9
3700	3700	0	284	0	5700	409	69.6
3800	3800	0	294	0	5800	419	69.2
3900	3900	0	300	0	5900	425	69.4
4000	4000	0	307	0	6000	432	69.4
4100	4100	0	316	0	6100	441	69.2
4200	4200	0	325	0	6200	450	69.0
4300	4300	0	331	0	6300	456	69.0
4400	4400	0	338	0	6400	463	69.1
4500	4500	0	344	0	6500	469	69.3



GESTION DE L'ESPACE: ADDITION AVEC DEBORDEMENT : REPARTITION SUR 2 PAGES
SUPPRESSION : RESTITUTION DES PAGES VIDES

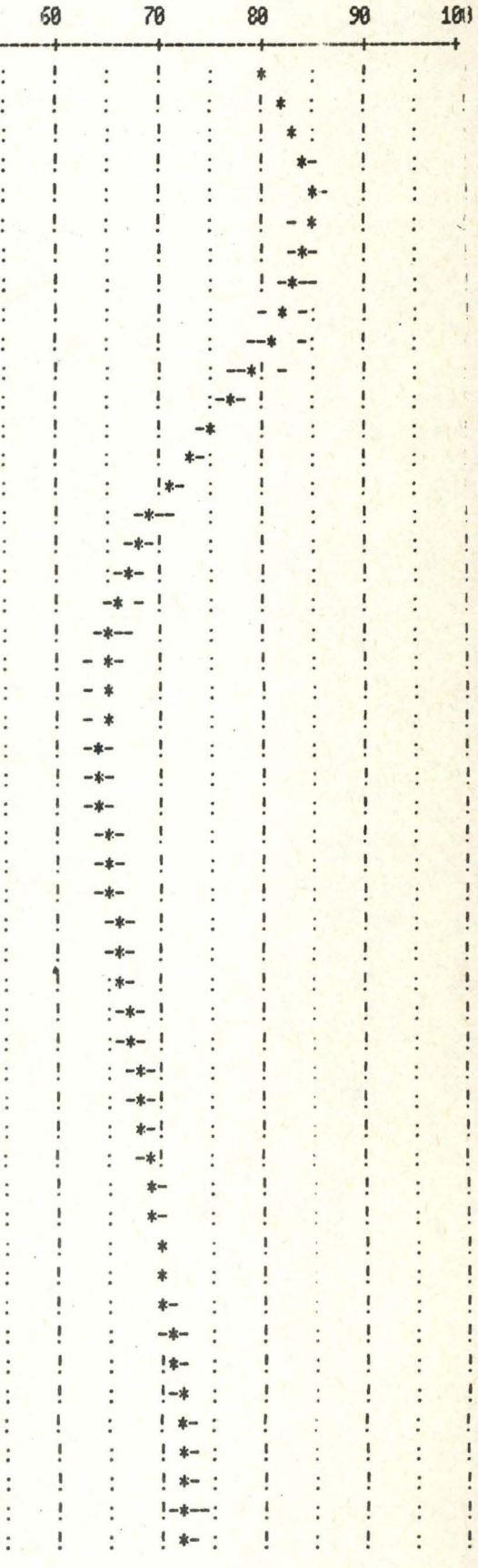
CONDITIONS D'EXPLOITATION : 100 % D'ADDITIONS

PARAMETRES PHYSIQUES : PAGES DE 20 ARTICLES EN LONGUEUR FIXE

DEPART : 200 PAGES CHARGEES A 80 %, SOIT 3200 ARTICLES

NOMBRE DE SIMULATIONS : 5

OPER	ADD-A	SUP-A	ADDP	SUPP	ART	PAGE	%OCC
0	0	0	0	0	3200	200	80.0
64	64	0	0	0	3264	200	81.6
128	128	0	0	0	3328	200	83.2
192	192	0	1	0	3392	201	84.5
256	256	0	2	0	3456	202	85.4
320	320	0	8	0	3520	208	84.7
384	384	0	13	0	3584	213	84.2
448	448	0	19	0	3648	219	83.5
512	512	0	26	0	3712	226	82.0
576	576	0	34	0	3776	234	80.7
640	640	0	45	0	3840	245	78.5
704	704	0	54	0	3904	254	76.9
768	768	0	66	0	3968	266	74.7
832	832	0	76	0	4032	276	73.1
896	896	0	87	0	4096	287	71.3
960	960	0	100	0	4160	300	69.4
1024	1024	0	111	0	4224	311	67.8
1088	1088	0	120	0	4288	320	67.0
1152	1152	0	129	0	4352	329	66.2
1216	1216	0	138	0	4416	338	65.3
1280	1280	0	146	0	4480	346	64.8
1344	1344	0	152	0	4544	352	64.5
1408	1408	0	157	0	4608	357	64.5
1472	1472	0	164	0	4672	364	64.1
1536	1536	0	169	0	4736	369	64.1
1600	1600	0	173	0	4800	373	64.4
1664	1664	0	175	0	4864	375	64.9
1728	1728	0	178	0	4928	378	65.2
1792	1792	0	183	0	4992	383	65.2
1856	1856	0	185	0	5056	385	65.7
1920	1920	0	188	0	5120	388	66.1
1984	1984	0	191	0	5184	391	66.4
2048	2048	0	193	0	5248	393	66.8
2112	2112	0	194	0	5312	394	67.3
2176	2176	0	197	0	5376	397	67.7
2240	2240	0	199	0	5440	399	68.1
2304	2304	0	202	0	5504	402	68.4
2368	2368	0	205	0	5568	405	68.7
2432	2432	0	207	0	5632	407	69.2
2496	2496	0	211	0	5696	411	69.4
2560	2560	0	213	0	5760	413	69.8
2624	2624	0	216	0	5824	416	70.0
2688	2688	0	218	0	5888	418	70.4
2752	2752	0	220	0	5952	420	70.8
2816	2816	0	223	0	6016	423	71.1
2880	2880	0	225	0	6080	425	71.5
2944	2944	0	227	0	6144	427	72.0
3008	3008	0	230	0	6208	430	72.1
3072	3072	0	234	0	6272	434	72.3
3136	3136	0	238	0	6336	438	72.3
3200	3200	0	242	0	6400	442	72.4



GESTION DE L'ESPACE: ADDITION AVEC DEBORDEMENT : REPARTITION SUR 2 PAGES
SUPPRESSION : RESTITUTION DES PAGES VIDES

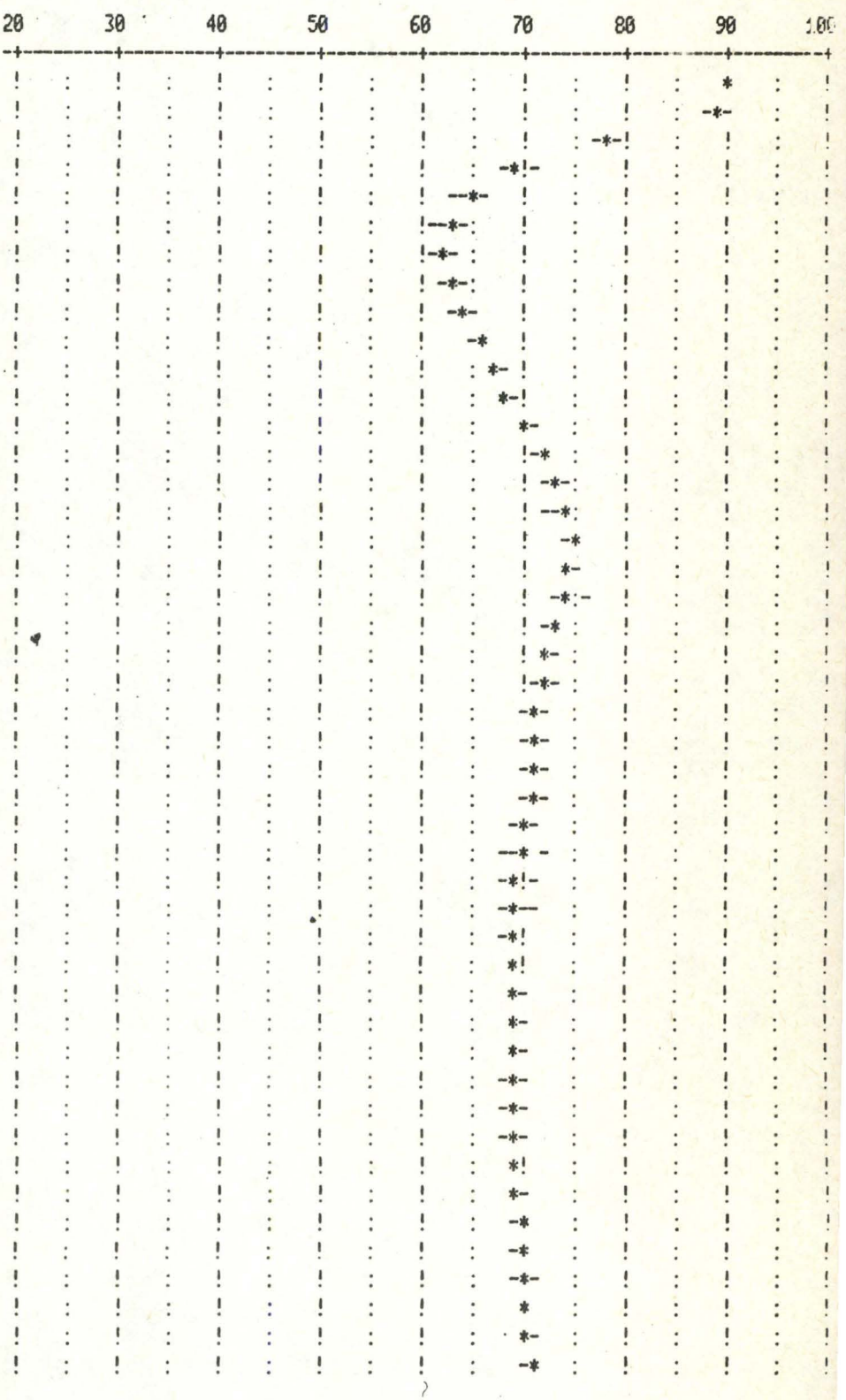
CONDITIONS D'EXPLOITATION : 100 % D'ADDITIONS

PARAMETRES PHYSIQUES : PAGES DE 20 ARTICLES EN LONGUEUR FIXE

DEPART : 111 PAGES CHARGEES A 90.0901 %, SOIT 1998 ARTICLES

NOMBRE DE SIMULATIONS : 5

OPER	ADD-A	SUP-A	ADDP	SUPP	ART	PAGE	%OCC
0	0	0	0	0	1998	111	90.0
100	100	0	7	0	2098	118	88.9
200	200	0	30	0	2198	141	77.7
300	300	0	55	0	2298	166	69.3
400	400	0	75	0	2398	186	64.6
500	500	0	89	0	2498	200	62.6
600	600	0	97	0	2598	208	62.3
700	700	0	103	0	2698	214	63.0
800	800	0	107	0	2798	218	64.2
900	900	0	110	0	2898	221	65.5
1000	1000	0	112	0	2998	223	67.3
1100	1100	0	115	0	3098	226	68.5
1200	1200	0	116	0	3198	227	70.4
1300	1300	0	119	0	3298	230	71.6
1400	1400	0	122	0	3398	233	72.9
1500	1500	0	127	0	3498	238	73.6
1600	1600	0	130	0	3598	241	74.6
1700	1700	0	138	0	3698	249	74.3
1800	1800	0	146	0	3798	257	74.0
1900	1900	0	156	0	3898	267	73.1
2000	2000	0	166	0	3998	277	72.2
2100	2100	0	174	0	4098	285	72.0
2200	2200	0	183	0	4198	294	71.4
2300	2300	0	191	0	4298	302	71.2
2400	2400	0	199	0	4398	310	70.9
2500	2500	0	208	0	4498	319	70.6
2600	2600	0	216	0	4598	327	70.4
2700	2700	0	227	0	4698	338	69.6
2800	2800	0	236	0	4798	347	69.2
2900	2900	0	243	0	4898	354	69.1
3000	3000	0	251	0	4998	362	69.0
3100	3100	0	257	0	5098	368	69.2
3200	3200	0	264	0	5198	375	69.3
3300	3300	0	270	0	5298	381	69.5
3400	3400	0	279	0	5398	390	69.3
3500	3500	0	286	0	5498	397	69.2
3600	3600	0	294	0	5598	405	69.2
3700	3700	0	302	0	5698	413	69.0
3800	3800	0	310	0	5798	421	68.9
3900	3900	0	315	0	5898	426	69.3
4000	4000	0	320	0	5998	431	69.6
4100	4100	0	327	0	6098	438	69.6
4200	4200	0	332	0	6198	443	69.9
4300	4300	0	338	0	6298	449	70.1
4400	4400	0	345	0	6398	456	70.2
4500	4500	0	349	0	6498	460	70.6



GESTION DE L'ESPACE: ADDITION AVEC DEBORDEMENT : REPARTITION SUR 2 PAGES
SUPPRESSION : RESTITUTION DES PAGES VIDES

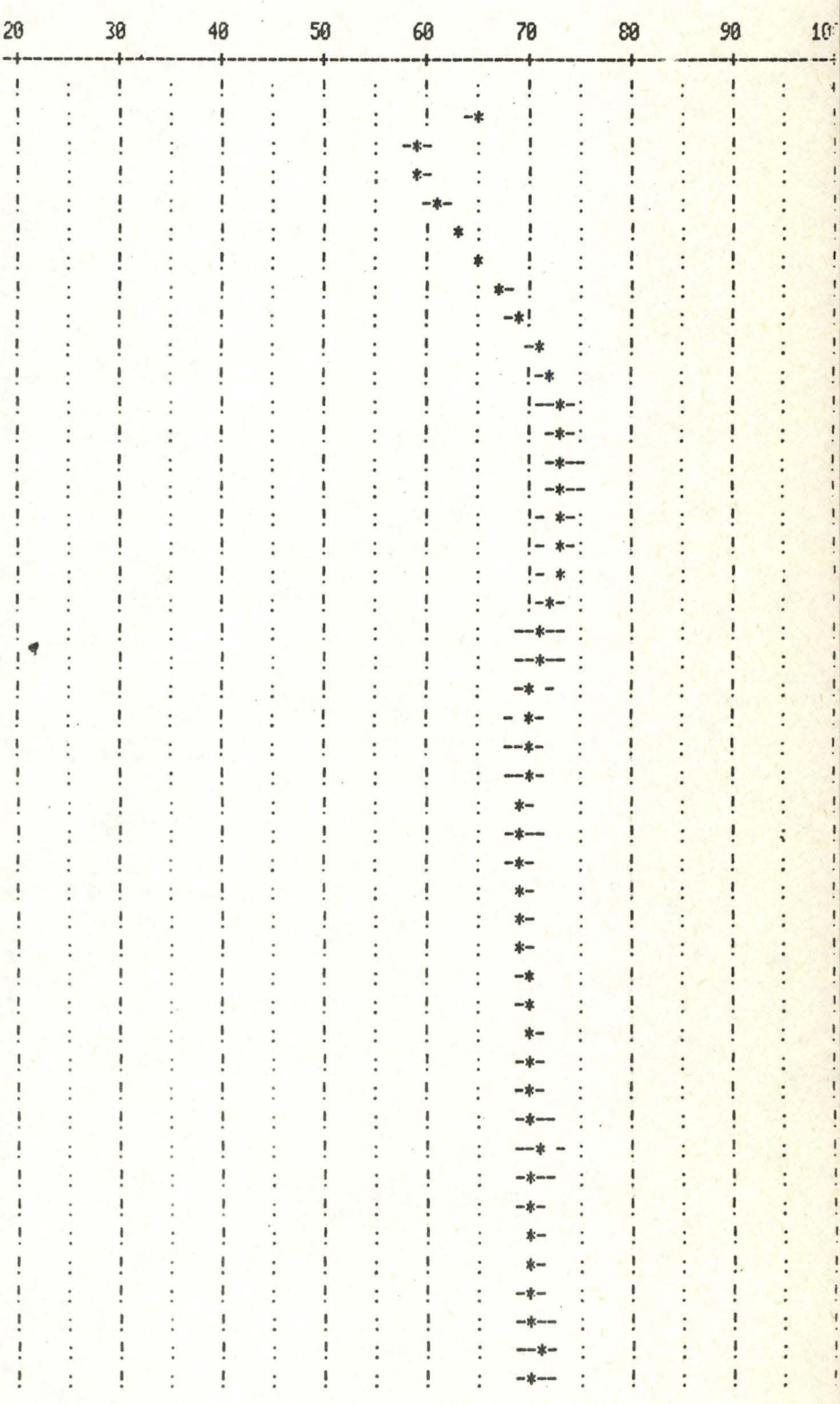
CONDITIONS D'EXPLOITATION : 100 % D'ADDITIONS

PARAMETRES PHYSIQUES : PAGES DE 20 ARTICLES EN LONGUEUR FIXE

DEPART : 100 PAGES CHARGEES A 100 %. SOIT 2000 ARTICLES

NOMBRE DE SIMULATIONS : 5

OPER	ADD-A	SUP-A	ADDP	SUPP	ART	PAGE	%OCC
0	0	0	0	0	2000	100	100.0
100	100	0	62	0	2100	162	64.7
200	200	0	86	0	2200	186	59.1
300	300	0	94	0	2300	194	59.2
400	400	0	98	0	2400	198	60.7
500	500	0	99	0	2500	199	62.8
600	600	0	100	0	2600	200	64.9
700	700	0	101	0	2700	201	67.2
800	800	0	103	0	2800	203	69.0
900	900	0	105	0	2900	205	70.6
1000	1000	0	109	0	3000	209	71.9
1100	1100	0	113	0	3100	213	72.7
1200	1200	0	119	0	3200	219	73.2
1300	1300	0	125	0	3300	225	73.3
1400	1400	0	132	0	3400	232	73.4
1500	1500	0	140	0	3500	240	72.9
1600	1600	0	147	0	3600	247	72.8
1700	1700	0	155	0	3700	255	72.6
1800	1800	0	164	0	3800	264	71.9
1900	1900	0	173	0	3900	273	71.4
2000	2000	0	182	0	4000	282	71.0
2100	2100	0	191	0	4100	291	70.4
2200	2200	0	199	0	4200	299	70.2
2300	2300	0	208	0	4300	308	69.8
2400	2400	0	216	0	4400	316	69.7
2500	2500	0	224	0	4500	324	69.4
2600	2600	0	232	0	4600	332	69.3
2700	2700	0	240	0	4700	340	69.2
2800	2800	0	246	0	4800	346	69.4
2900	2900	0	253	0	4900	353	69.4
3000	3000	0	260	0	5000	360	69.4
3100	3100	0	266	0	5100	366	69.6
3200	3200	0	273	0	5200	373	69.7
3300	3300	0	278	0	5300	378	70.0
3400	3400	0	285	0	5400	385	70.2
3500	3500	0	290	0	5500	390	70.4
3600	3600	0	297	0	5600	397	70.5
3700	3700	0	304	0	5700	404	70.6
3800	3800	0	313	0	5800	413	70.3
3900	3900	0	320	0	5900	420	70.2
4000	4000	0	327	0	6000	427	70.3
4100	4100	0	333	0	6100	433	70.4
4200	4200	0	342	0	6200	442	70.2
4300	4300	0	348	0	6300	448	70.3
4400	4400	0	354	0	6400	454	70.5
4500	4500	0	362	0	6500	462	70.4

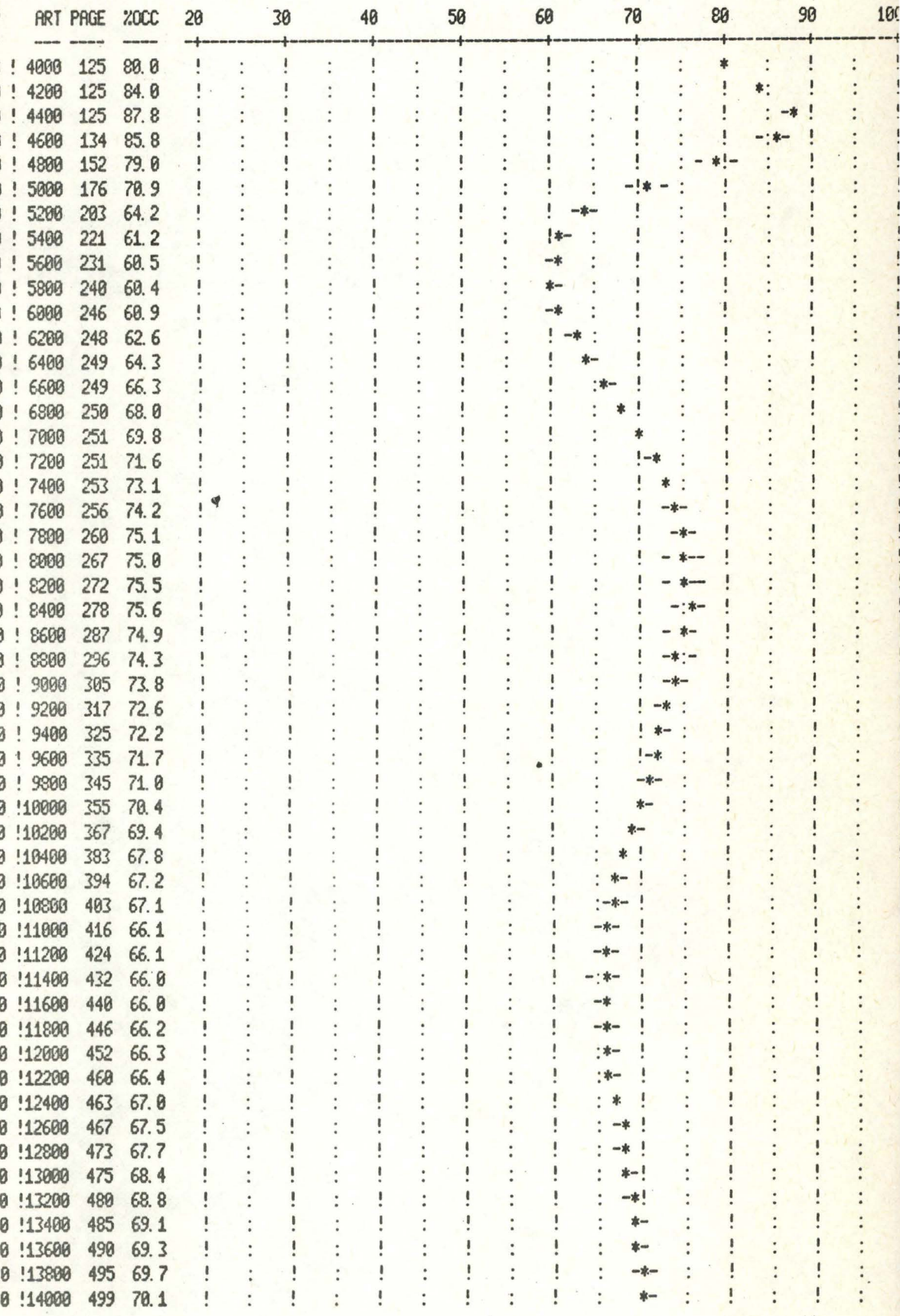


OVERFLOW IN 6245READY

GESTION DE L'ESPACE: ADDITION AVEC DEBORDEMENT : REPARTITION SUR 2 PAGES
SUPPRESSION : RESTITUTION DES PAGES VIDES
CONDITIONS D'EXPLOITATION : 100 % D'ADDITIONS
PARAMETRES PHYSIQUES : PAGES DE 40 ARTICLES EN LONGUEUR FIXE

DEPART : 125 PAGES CHARGEES A 80 %, SOIT 4000 ARTICLES
NOMBRE DE SIMULATIONS : 3

OPER	ADD-A	SUP-A	ADDP	SUPP	ART	PAGE	%OCC
0	0	0	0	0	4000	125	80.0
200	200	0	0	0	4200	125	84.0
400	400	0	0	0	4400	125	87.8
600	600	0	9	0	4600	134	85.8
800	800	0	27	0	4800	152	79.0
1000	1000	0	51	0	5000	176	70.9
1200	1200	0	78	0	5200	203	64.2
1400	1400	0	96	0	5400	221	61.2
1600	1600	0	106	0	5600	231	60.5
1800	1800	0	115	0	5800	240	60.4
2000	2000	0	121	0	6000	246	60.9
2200	2200	0	123	0	6200	248	62.6
2400	2400	0	124	0	6400	249	64.3
2600	2600	0	124	0	6600	249	66.3
2800	2800	0	125	0	6800	250	68.0
3000	3000	0	126	0	7000	251	69.8
3200	3200	0	126	0	7200	251	71.6
3400	3400	0	128	0	7400	253	73.1
3600	3600	0	131	0	7600	256	74.2
3800	3800	0	135	0	7800	260	75.1
4000	4000	0	142	0	8000	267	75.0
4200	4200	0	147	0	8200	272	75.5
4400	4400	0	153	0	8400	278	75.6
4600	4600	0	162	0	8600	287	74.9
4800	4800	0	171	0	8800	296	74.3
5000	5000	0	180	0	9000	305	73.8
5200	5200	0	192	0	9200	317	72.6
5400	5400	0	200	0	9400	325	72.2
5600	5600	0	210	0	9600	335	71.7
5800	5800	0	220	0	9800	345	71.0
6000	6000	0	230	0	10000	355	70.4
6200	6200	0	242	0	10200	367	69.4
6400	6400	0	258	0	10400	383	67.8
6600	6600	0	269	0	10600	394	67.2
6800	6800	0	278	0	10800	403	67.1
7000	7000	0	291	0	11000	416	66.1
7200	7200	0	299	0	11200	424	66.1
7400	7400	0	307	0	11400	432	66.0
7600	7600	0	315	0	11600	440	66.0
7800	7800	0	321	0	11800	446	66.2
8000	8000	0	327	0	12000	452	66.3
8200	8200	0	335	0	12200	460	66.4
8400	8400	0	338	0	12400	463	67.0
8600	8600	0	342	0	12600	467	67.5
8800	8800	0	348	0	12800	473	67.7
9000	9000	0	350	0	13000	475	68.4
9200	9200	0	355	0	13200	480	68.8
9400	9400	0	360	0	13400	485	69.1
9600	9600	0	365	0	13600	490	69.3
9800	9800	0	370	0	13800	495	69.7
10000	10000	0	374	0	14000	499	70.1



EXPLICATION DE L'OUTPUT.

NBRECASES:	NBRE CASES DE LA ZONE PRIMAIRE
BSIZE-BYTES:	LONGUEUR MOYENNE D'UNE CASE (EN BYTES) (LONG. MOYENNE D'UN RECORD = 130 BYTES)
NBRE RECORDS INSERES:	A PARTIR DE 0 RECORD ON EN INSERE LE NOMBRE INDIQUE
FONCTION:	FONCTION DE HASHAGE 1= FONCTION DE DIVISION 2= FONCTION DE FOLDING 3= FONCTION DE MIDSQUARRE
N:	NOMBRE DE RECORDS DEJA INSERES
%N:	POURCENTAGE DE RECORDS DEJA INSERES
%ZPR:	POURCENTAGE DE PLACE OCCUPEE EN ZONE PRIMAIRE
%ZTOT:	POURCENTAGE DE PLACE OCCPUPEE PAR RAPPORT A LA ZONE TOTALE DU FICHIER(ZONE PRIMAIRE THEORIQUE + ZONE D'OVERFLOW)

I N S E R T I O N

A-M:	NBRE ACCES MOYEN
E-T:	ECART-TYPE DE A-M
%I (I=0,2,4,5):	% DE RECORDS NECESSITANT I ACCES SUPPLEMENTAIRES
TP-MOY:	TEMPS MOYEN D'ACCES A UN RECORD PAR PERIODE
E-T:	ECART-TYPE DE TP-MOY
MAX:	TEMPS MAX. D'ACCES A UN RECORD PAR PERIODE
MAX-CH:	LONGUEUR MAXIMUN DE LA CHAINE DES ELEMENTS N'AYANT PAS TROUVE DE PLACE DANS LA CASE PRIMAIRE

C O N S U L T A T I O N

A-M:	NBRE D'ACCES MOYEN
E-T:	ECART-TYPE DE A-M
%I (I=1,2,3,4):	POURCENTAGE D'ARTICLES NECESSITANT I ACCES
TP-MOY:	TEMPS MOYEN D'ACCES A UN RECORD (TP-MOY EST UN TEMPS GLOBAL)
E-T:	ECART-TYPE DE TP-MOY
MAX:	TEMPS D'ACCES MAXIMUN A UN RECORD DEPUIS LE DEBUT DES OPERATIONS

INDEPENDENT OVERFLOW ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS INSERES	FONCTION
6000	3900.	200000	1

				INSERTION										CONSULTATION								
N	%N	%ZPR	%ZTOT	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	5	5	2.0	0.0	100.0	0.0	0.0	0.0	33.2	10.4	61.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.2	10.4	61.6
20000	10	11	11	2.0	0.0	100.0	0.0	0.0	0.0	32.2	10.3	61.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	32.7	10.3	61.6
30000	15	16	16	2.0	0.0	100.0	0.0	0.0	0.0	31.8	10.2	61.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	32.4	10.2	61.6
40000	20	22	22	2.0	0.0	100.0	0.0	0.0	0.0	32.1	10.3	61.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	32.2	10.5	61.6
50000	25	27	27	2.0	0.0	100.0	0.0	0.0	0.0	32.2	10.2	61.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	32.1	10.8	61.6
60000	30	33	33	2.0	0.0	100.0	0.0	0.0	0.0	32.1	10.3	61.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	32.0	10.9	61.6
70000	35	38	38	2.0	0.0	100.0	0.0	0.0	0.0	32.4	10.2	61.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	32.0	11.1	61.6
80000	40	44	44	2.0	0.0	100.0	0.0	0.0	0.0	31.6	10.2	91.3	1.	1.0	0.0	100.0	0.0	0.0	0.0	31.9	11.1	61.6
90000	45	50	49	2.0	0.0	100.0	0.0	0.0	0.0	31.5	10.7	164.5	3.	1.0	0.0	100.0	0.0	0.0	0.0	31.8	11.2	126.3
100000	50	55	55	2.0	0.1	100.0	0.0	0.0	0.0	32.0	12.1	216.1	4.	1.0	0.0	100.0	0.0	0.0	0.0	31.7	11.3	177.9
110000	55	61	61	2.0	0.1	99.8	0.0	0.0	0.0	33.1	16.4	244.2	6.	1.0	0.1	99.8	0.1	0.0	0.0	31.7	11.7	206.1
120000	60	66	66	2.0	0.2	99.6	0.1	0.0	0.0	33.4	23.9	335.5	12.	1.0	0.2	99.6	0.1	0.1	0.1	31.7	12.5	297.4
130000	65	72	71	2.0	0.4	99.3	0.2	0.1	0.0	35.1	30.4	371.8	14.	1.0	0.3	99.3	0.3	0.2	0.1	31.8	13.8	338.7
140000	70	77	76	2.0	0.5	98.7	0.3	0.1	0.1	40.4	40.7	436.6	17.	1.0	0.4	98.7	0.5	0.3	0.2	32.1	15.9	403.4
150000	75	83	81	2.1	0.7	97.8	0.5	0.2	0.2	48.2	51.2	376.8	14.	1.1	0.5	97.8	0.7	0.5	0.3	32.8	18.7	403.4
160000	80	88	86	2.1	0.9	96.7	0.7	0.4	0.3	59.6	64.0	478.2	18.	1.1	0.7	96.7	1.0	0.7	0.5	33.9	22.6	445.0
170000	85	94	90	2.2	1.2	95.0	1.0	0.6	0.4	81.4	87.1	647.4	27.	1.2	1.0	95.0	1.3	1.0	0.8	35.9	29.2	614.2
180000	90	100	93	2.3	1.5	93.1	1.3	0.8	0.6	95.1	95.8	576.1	23.	1.3	1.3	93.1	1.6	1.3	1.0	38.4	35.6	614.2
190000	95	105	96	2.5	1.9	90.9	1.5	1.0	0.8	121.4	109.3	662.4	27.	1.4	1.6	90.9	1.9	1.5	1.3	41.7	43.3	629.2
200000	100	111	98	2.7	2.3	88.2	1.8	1.3	1.1	152.6	119.8	674.0	28.	1.6	2.0	88.2	2.1	1.8	1.5	46.0	52.1	640.8

INDEPENDENT OVERFLOW ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS	INSERES	FONCTION
4500	5200.	200000		1

				INSERTION									CONSULTATION								
N	%N	%ZPR	%ZTOT	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	5	5	2.0	0.0	100.0	0.0	0.0	0.0	34.2	10.4	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	34.2	10.4	62.6
20000	10	11	11	2.0	0.0	100.0	0.0	0.0	0.0	33.1	10.3	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.7	10.2	62.6
30000	15	16	16	2.0	0.0	100.0	0.0	0.0	0.0	32.8	10.2	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.4	10.1	62.6
40000	20	22	22	2.0	0.0	100.0	0.0	0.0	0.0	33.3	10.4	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.2	10.5	62.6
50000	25	27	27	2.0	0.0	100.0	0.0	0.0	0.0	33.3	10.2	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.1	10.7	62.6
60000	30	33	33	2.0	0.0	100.0	0.0	0.0	0.0	33.1	10.3	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.0	10.9	62.6
70000	35	38	38	2.0	0.0	100.0	0.0	0.0	0.0	33.5	10.3	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.0	11.0	62.6
80000	40	44	44	2.0	0.0	100.0	0.0	0.0	0.0	32.7	10.3	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	32.9	11.1	62.6
90000	45	50	49	2.0	0.0	100.0	0.0	0.0	0.0	32.5	10.5	188.5	4.	1.0	0.0	100.0	0.0	0.0	0.0	32.8	11.2	148.2
100000	50	55	55	2.0	0.0	100.0	0.0	0.0	0.0	32.8	10.9	271.4	7.	1.0	0.0	100.0	0.0	0.0	0.0	32.7	11.3	231.1
110000	55	61	61	2.0	0.1	100.0	0.0	0.0	0.0	33.2	11.5	299.0	8.	1.0	0.1	100.0	0.0	0.0	0.0	32.7	11.4	258.7
120000	60	66	66	2.0	0.1	99.9	0.0	0.0	0.0	32.6	17.8	354.3	10.	1.0	0.1	99.9	0.1	0.0	0.0	32.6	11.8	314.0
130000	65	72	72	2.0	0.2	99.6	0.1	0.0	0.0	33.6	24.6	354.3	11.	1.0	0.2	99.6	0.1	0.1	0.0	32.5	12.6	314.0
140000	70	77	77	2.0	0.4	99.3	0.2	0.1	0.0	37.9	35.7	367.2	13.	1.0	0.3	99.3	0.2	0.2	0.1	32.7	14.3	331.9
150000	75	83	82	2.1	0.6	98.7	0.3	0.1	0.1	44.9	50.5	457.8	17.	1.0	0.5	98.7	0.4	0.3	0.2	33.2	17.5	422.5
160000	80	88	87	2.1	0.8	97.7	0.5	0.2	0.2	55.9	64.1	540.7	20.	1.1	0.7	97.7	0.6	0.5	0.3	34.1	21.8	505.4
170000	85	94	91	2.2	1.1	96.5	0.7	0.4	0.3	72.7	86.2	593.6	23.	1.1	1.0	96.5	0.8	0.7	0.5	35.8	28.2	558.3
180000	90	100	95	2.3	1.5	94.8	0.9	0.6	0.4	94.4	102.3	646.5	26.	1.2	1.3	94.8	1.1	0.9	0.7	38.2	35.7	611.2
190000	95	105	97	2.4	1.9	92.5	1.2	0.8	0.7	128.2	122.0	784.7	31.	1.4	1.7	92.5	1.4	1.2	1.0	41.9	45.3	749.4
200000	100	111	99	2.7	2.4	89.8	1.4	1.0	0.9	166.8	136.7	820.0	33.	1.6	2.2	89.8	1.6	1.4	1.2	46.9	56.5	784.7

INDEPENDENT OVERFLOW ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS	INSERES	FONCTION
3600	6500.	200000		1

				INSERTION									CONSULTATION								
N	%N	%ZPR	%ZTOT	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	5	5	2.0	0.0	100.0	0.0	0.0	0.0	35.0	10.2	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.0	10.2	63.7
20000	10	11	11	2.0	0.0	100.0	0.0	0.0	0.0	34.2	10.3	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	34.6	10.1	63.7
30000	15	16	16	2.0	0.0	100.0	0.0	0.0	0.0	33.9	10.2	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	34.4	10.0	63.7
40000	20	22	22	2.0	0.0	100.0	0.0	0.0	0.0	34.0	10.1	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	34.1	10.5	63.7
50000	25	27	27	2.0	0.0	100.0	0.0	0.0	0.0	34.3	10.2	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	34.0	10.8	63.7
60000	30	33	33	2.0	0.0	100.0	0.0	0.0	0.0	34.1	10.3	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.9	11.0	63.7
70000	35	38	38	2.0	0.0	100.0	0.0	0.0	0.0	34.5	10.2	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.9	11.2	63.7
80000	40	44	44	2.0	0.0	100.0	0.0	0.0	0.0	33.6	10.2	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.8	11.3	63.7
90000	45	50	50	2.0	0.0	100.0	0.0	0.0	0.0	33.5	10.3	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.7	11.3	63.7
100000	50	55	55	2.0	0.0	100.0	0.0	0.0	0.0	33.8	10.4	114.8	1.	1.0	0.0	100.0	0.0	0.0	0.0	33.6	11.4	67.4
110000	55	61	61	2.0	0.0	100.0	0.0	0.0	0.0	34.0	10.9	195.9	4.	1.0	0.0	100.0	0.0	0.0	0.0	33.6	11.5	153.5
120000	60	66	66	2.0	0.1	99.9	0.0	0.0	0.0	32.8	15.3	368.1	10.	1.0	0.1	99.9	0.0	0.0	0.0	33.4	11.8	325.7
130000	65	72	72	2.0	0.2	99.8	0.0	0.0	0.0	32.9	20.1	389.2	13.	1.0	0.2	99.8	0.0	0.0	0.0	33.3	12.3	351.8
140000	70	77	77	2.0	0.3	99.6	0.1	0.0	0.0	35.8	29.7	474.0	17.	1.0	0.2	99.6	0.1	0.1	0.0	33.3	13.5	436.6
150000	75	83	82	2.0	0.5	99.2	0.2	0.1	0.0	41.9	43.9	512.7	18.	1.0	0.4	99.2	0.2	0.2	0.1	33.7	16.0	475.3
160000	80	88	87	2.1	0.7	98.5	0.3	0.2	0.1	52.4	61.1	550.1	20.	1.1	0.6	98.5	0.4	0.3	0.2	34.5	20.2	512.7
170000	85	94	92	2.1	1.0	97.4	0.5	0.3	0.2	68.5	84.3	673.6	25.	1.1	0.9	97.4	0.6	0.5	0.4	35.9	26.7	636.2
180000	90	100	95	2.2	1.4	95.8	0.7	0.4	0.3	95.2	107.8	797.1	30.	1.2	1.2	95.8	0.8	0.7	0.6	38.4	35.4	759.7
190000	95	105	98	2.4	1.9	93.5	1.0	0.7	0.6	136.5	133.0	834.4	32.	1.3	1.7	93.5	1.1	1.0	0.8	42.5	47.1	797.1
200000	100	111	**	2.6	2.5	90.8	1.2	0.9	0.8	182.6	150.7	915.5	35.	1.5	2.2	90.8	1.3	1.2	1.0	48.2	60.5	878.1

INDEPENDENT OVERFLOW ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS INSERES	FONCTION
3000	7800.	200000	1

				INSERTION										CONSULTATION								
N	%N	%ZPR	%ZTOT	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	5	5	2.0	0.0	100.0	0.0	0.0	0.0	36.3	10.4	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	36.3	10.4	64.8
20000	10	11	11	2.0	0.0	100.0	0.0	0.0	0.0	35.3	10.3	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.8	10.2	64.8
30000	15	16	16	2.0	0.0	100.0	0.0	0.0	0.0	35.1	10.3	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.5	10.2	64.8
40000	20	22	22	2.0	0.0	100.0	0.0	0.0	0.0	35.4	10.4	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.3	10.8	64.8
50000	25	27	27	2.0	0.0	100.0	0.0	0.0	0.0	35.4	10.3	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.2	11.1	64.8
60000	30	33	33	2.0	0.0	100.0	0.0	0.0	0.0	35.1	10.3	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.0	11.3	64.8
70000	35	38	38	2.0	0.0	100.0	0.0	0.0	0.0	35.6	10.3	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.0	11.5	64.8
80000	40	44	44	2.0	0.0	100.0	0.0	0.0	0.0	34.8	10.3	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	34.9	11.6	64.8
90000	45	50	50	2.0	0.0	100.0	0.0	0.0	0.0	34.5	10.3	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	34.8	11.7	64.8
100000	50	55	55	2.0	0.0	100.0	0.0	0.0	0.0	34.7	10.3	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	34.7	11.8	64.8
110000	55	61	61	2.0	0.0	100.0	0.0	0.0	0.0	34.9	10.3	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	34.6	11.8	64.8
120000	60	66	66	2.0	0.1	100.0	0.0	0.0	0.0	33.3	11.8	228.1	5.	1.0	0.0	100.0	0.0	0.0	0.0	34.5	11.9	183.5
130000	65	72	72	2.0	0.1	99.9	0.0	0.0	0.0	33.2	16.2	307.3	8.	1.0	0.1	99.9	0.0	0.0	0.0	34.3	12.2	262.8
140000	70	77	77	2.0	0.2	99.8	0.0	0.0	0.0	35.4	25.5	366.9	10.	1.0	0.2	99.8	0.1	0.0	0.0	34.2	13.0	322.3
150000	75	83	82	2.0	0.4	99.4	0.1	0.1	0.0	40.7	39.6	401.9	12.	1.0	0.3	99.4	0.2	0.1	0.1	34.4	15.1	357.3
160000	80	88	87	2.1	0.6	98.8	0.2	0.1	0.1	52.7	62.9	584.4	20.	1.0	0.5	98.8	0.3	0.2	0.2	35.2	19.8	544.9
170000	85	94	92	2.2	1.2	97.4	0.4	0.3	0.2	85.7	112.1	896.3	33.	1.1	1.1	97.4	0.5	0.4	0.3	37.5	31.8	856.8
180000	90	100	96	2.3	1.6	95.9	0.6	0.4	0.3	103.7	127.5	921.1	34.	1.2	1.5	95.9	0.7	0.6	0.5	40.4	42.4	881.5
190000	95	105	98	2.5	2.2	93.7	0.8	0.6	0.5	150.4	159.2	995.3	37.	1.4	2.0	93.7	0.9	0.8	0.7	45.1	56.4	955.8
200000	100	111	**	2.7	2.9	91.0	1.0	0.8	0.7	208.8	183.4	1089.6	40.	1.6	2.7	91.0	1.1	1.0	0.9	52.0	73.1	1050.1

INDEPENDENT OVERFLOW ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS INSERES	FONCTION
6000	3900.	200000	2

				INSERTION										CONSULTATION								
N	%N	%ZPR	%ZTOT	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	5	5	2.0	0.0	100.0	0.0	0.0	0.0	31.2	10.4	61.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	31.2	10.4	61.6
20000	10	11	11	2.0	0.0	100.0	0.0	0.0	0.0	30.6	10.4	61.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	30.9	10.4	61.6
30000	15	16	16	2.0	0.0	100.0	0.0	0.0	0.0	30.7	10.5	61.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	30.8	10.4	61.6
40000	20	22	22	2.0	0.0	100.0	0.0	0.0	0.0	31.0	10.6	61.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	30.8	10.6	61.6
50000	25	27	27	2.0	0.0	100.0	0.0	0.0	0.0	30.7	10.5	61.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	30.7	10.9	61.6
60000	30	33	33	2.0	0.0	100.0	0.0	0.0	0.0	30.0	10.5	61.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	30.5	11.1	61.6
70000	35	38	38	2.0	0.0	100.0	0.0	0.0	0.0	31.2	10.7	106.3	1.	1.0	0.0	100.0	0.0	0.0	0.0	30.5	11.2	63.2
80000	40	44	44	2.0	0.0	100.0	0.0	0.0	0.0	30.6	11.5	184.5	4.	1.0	0.0	100.0	0.0	0.0	0.0	30.4	11.4	146.3
90000	45	50	49	2.0	0.1	99.9	0.0	0.0	0.0	30.7	14.1	250.8	7.	1.0	0.1	99.9	0.0	0.0	0.0	30.3	11.6	212.6
100000	50	55	55	2.0	0.2	99.7	0.1	0.0	0.0	32.8	22.1	330.5	10.	1.0	0.2	99.7	0.1	0.1	0.0	30.4	12.5	292.4
110000	55	61	60	2.0	0.4	99.1	0.2	0.1	0.0	37.8	33.6	370.5	12.	1.0	0.3	99.1	0.3	0.2	0.1	30.8	14.4	332.4
120000	60	66	65	2.1	0.6	98.2	0.4	0.2	0.1	43.9	47.9	366.8	14.	1.1	0.5	98.2	0.5	0.4	0.3	31.5	17.7	333.7
130000	65	72	70	2.1	0.8	97.0	0.6	0.4	0.2	53.3	61.5	376.8	14.	1.1	0.7	97.0	0.8	0.6	0.5	32.6	22.2	343.7
140000	70	77	75	2.2	1.1	95.3	0.9	0.6	0.4	69.6	77.8	431.6	17.	1.2	1.0	95.3	1.1	0.9	0.7	34.4	28.2	398.4
150000	75	83	78	2.3	1.5	93.3	1.2	0.8	0.6	87.5	93.0	561.1	23.	1.3	1.3	93.3	1.4	1.2	1.0	37.0	35.2	527.9
160000	80	88	82	2.5	1.9	90.9	1.5	1.0	0.8	111.2	108.1	587.6	24.	1.4	1.6	90.9	1.7	1.5	1.2	40.6	43.5	554.5
170000	85	94	85	2.7	2.3	88.5	1.7	1.3	1.1	135.7	124.3	669.0	28.	1.6	2.0	88.5	1.9	1.7	1.4	45.0	53.0	635.8
180000	90	100	87	2.9	2.8	85.8	1.9	1.5	1.3	167.6	140.3	781.8	33.	1.8	2.5	85.8	2.1	1.9	1.6	50.6	63.8	748.7
190000	95	105	89	3.2	3.4	83.0	2.0	1.6	1.5	201.7	149.7	889.7	38.	2.1	3.1	83.0	2.2	2.0	1.8	57.2	75.2	856.6
200000	100	111	91	3.6	3.9	80.2	2.1	1.8	1.6	233.1	163.2	932.9	40.	2.4	3.6	80.2	2.2	2.1	1.9	64.7	87.1	899.7

INDEPENDENT OVERFLOW ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS	INSERES	FONCTION
4500	5200.	200000		2

				INSERTION								CONSULTATION									
N	%N	%ZPR	%ZTOT	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	5	5	2.0	0.0	100.0	0.0	0.0	0.0	33.8	10.3	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.8	10.3	62.6
20000	10	11	11	2.0	0.0	100.0	0.0	0.0	0.0	33.2	10.5	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.5	10.3	62.6
30000	15	16	16	2.0	0.0	100.0	0.0	0.0	0.0	33.2	10.4	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.4	10.2	62.6
40000	20	22	22	2.0	0.0	100.0	0.0	0.0	0.0	33.6	10.6	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.3	10.6	62.6
50000	25	27	27	2.0	0.0	100.0	0.0	0.0	0.0	33.4	10.6	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.2	10.9	62.6
60000	30	33	33	2.0	0.0	100.0	0.0	0.0	0.0	32.9	10.6	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.0	11.1	62.6
70000	35	38	38	2.0	0.0	100.0	0.0	0.0	0.0	33.9	10.7	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	33.1	11.3	62.6
80000	40	44	44	2.0	0.0	100.0	0.0	0.0	0.0	33.1	10.7	150.6	2.	1.0	0.0	100.0	0.0	0.0	0.0	33.0	11.4	100.3
90000	45	50	49	2.0	0.1	99.9	0.0	0.0	0.0	33.2	13.1	251.1	5.	1.0	0.1	99.9	0.0	0.0	0.0	32.9	11.5	210.8
100000	50	55	55	2.0	0.2	99.8	0.0	0.0	0.0	35.4	22.5	349.3	10.	1.0	0.2	99.8	0.1	0.0	0.0	33.0	12.5	309.0
110000	55	61	60	2.0	0.4	99.4	0.1	0.1	0.0	39.3	34.9	389.3	13.	1.0	0.3	99.4	0.2	0.1	0.1	33.4	14.8	351.9
120000	60	66	66	2.1	0.7	98.7	0.2	0.2	0.1	46.3	55.5	518.1	19.	1.1	0.6	98.7	0.3	0.2	0.2	34.1	19.6	482.8
130000	65	72	71	2.1	0.9	97.9	0.3	0.2	0.2	52.6	68.3	508.1	19.	1.1	0.8	97.9	0.4	0.3	0.3	35.1	24.9	482.8
140000	70	77	75	2.2	1.3	96.7	0.5	0.4	0.3	66.7	88.6	545.7	20.	1.2	1.1	96.7	0.6	0.5	0.4	36.8	32.0	510.4
150000	75	83	80	2.3	1.7	95.4	0.6	0.5	0.4	81.9	108.3	646.2	25.	1.3	1.5	95.4	0.7	0.6	0.5	39.2	40.3	611.0
160000	80	88	84	2.4	2.1	94.0	0.7	0.6	0.5	98.8	128.7	794.7	31.	1.4	1.9	94.0	0.8	0.7	0.6	42.2	49.8	759.4
170000	85	94	88	2.6	2.6	92.4	0.8	0.7	0.6	125.5	155.2	887.9	36.	1.5	2.4	92.4	0.9	0.8	0.7	46.3	61.5	852.6
180000	90	100	91	2.8	3.2	90.7	0.9	0.7	0.6	148.4	179.1	1038.7	42.	1.7	3.0	90.7	1.0	0.9	0.8	51.1	74.0	1003.5
190000	95	105	94	3.1	3.8	88.7	1.0	0.8	0.7	176.2	198.4	1084.0	44.	2.0	3.6	88.7	1.2	1.0	0.9	56.7	87.0	1048.7
200000	100	111	96	3.4	4.5	86.5	1.2	1.0	0.9	215.0	222.0	1101.7	45.	2.3	4.2	86.5	1.4	1.2	1.1	63.6	101.4	1066.4

INDEPENDENT OVERFLOW ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS	INSERES	FONCTION
3600	6500.	200000		2

				INSERTION									CONSULTATION								
N	%N	%ZPR	%ZTOT	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	5	5	2.0	0.0	100.0	0.0	0.0	0.0	36.1	10.5	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	36.1	10.5	63.7
20000	10	11	11	2.0	0.0	100.0	0.0	0.0	0.0	35.5	10.6	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.8	10.4	63.7
30000	15	16	16	2.0	0.0	100.0	0.0	0.0	0.0	35.7	10.5	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.7	10.4	63.7
40000	20	22	22	2.0	0.0	100.0	0.0	0.0	0.0	35.8	10.5	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.6	10.9	63.7
50000	25	27	27	2.0	0.0	100.0	0.0	0.0	0.0	35.5	10.5	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.4	11.2	63.7
60000	30	33	33	2.0	0.0	100.0	0.0	0.0	0.0	35.3	10.7	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.3	11.4	63.7
70000	35	38	38	2.0	0.0	100.0	0.0	0.0	0.0	35.9	10.5	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.3	11.6	63.7
80000	40	44	44	2.0	0.0	100.0	0.0	0.0	0.0	35.4	10.6	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.2	11.7	63.7
90000	45	50	50	2.0	0.0	100.0	0.0	0.0	0.0	35.3	10.7	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	35.1	11.8	63.7
100000	50	55	55	2.0	0.0	100.0	0.0	0.0	0.0	35.4	10.7	134.8	1.	1.0	0.0	100.0	0.0	0.0	0.0	35.1	11.8	87.4
110000	55	61	61	2.0	0.1	99.9	0.0	0.0	0.0	36.1	14.3	310.7	8.	1.0	0.1	99.9	0.0	0.0	0.0	35.1	12.0	268.3
120000	60	66	66	2.0	0.2	99.8	0.0	0.0	0.0	36.2	20.4	295.7	8.	1.0	0.1	99.8	0.1	0.0	0.0	35.1	12.6	268.3
130000	65	72	71	2.0	0.3	99.5	0.1	0.0	0.0	39.2	32.4	403.1	11.	1.0	0.2	99.5	0.2	0.1	0.1	35.2	14.2	360.7
140000	70	77	77	2.0	0.5	99.0	0.2	0.1	0.1	46.6	47.8	424.2	15.	1.0	0.4	99.0	0.3	0.2	0.1	35.7	17.2	386.8
150000	75	83	82	2.1	0.8	98.1	0.4	0.2	0.2	58.4	69.0	517.7	19.	1.1	0.7	98.1	0.5	0.4	0.3	36.8	22.5	480.3
160000	80	88	86	2.2	1.2	96.7	0.5	0.4	0.3	79.4	95.0	659.9	24.	1.1	1.0	96.7	0.7	0.5	0.4	38.8	30.7	622.5
170000	85	94	90	2.3	1.6	95.1	0.7	0.5	0.4	104.4	120.5	768.4	29.	1.3	1.4	95.1	0.9	0.7	0.6	41.8	41.0	731.0
180000	90	100	93	2.5	2.2	92.9	0.9	0.7	0.6	137.3	143.8	825.7	31.	1.4	1.9	92.9	1.1	0.9	0.8	46.1	53.1	788.4
190000	95	105	95	2.7	2.8	90.3	1.1	0.9	0.8	186.9	169.3	906.8	34.	1.6	2.6	90.3	1.3	1.1	1.0	52.2	67.9	869.4
200000	100	111	97	3.1	3.5	87.4	1.3	1.1	1.0	239.0	189.3	944.2	36.	2.0	3.3	87.4	1.4	1.3	1.1	60.2	84.6	906.8

INDEPENDENT OVERFLOW ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS	INSERES	FONCTION
3000	7800.	200000		2

				INSERTION									CONSULTATION									
N	%N	%ZPR	%ZTOT	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	5	5	2.0	0.0	100.0	0.0	0.0	0.0	31.6	10.0	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	31.6	10.0	64.8
20000	10	11	11	2.0	0.0	100.0	0.0	0.0	0.0	31.1	10.3	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	31.4	10.1	64.8
30000	15	16	16	2.0	0.0	100.0	0.0	0.0	0.0	31.0	10.0	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	31.2	10.0	64.8
40000	20	22	22	2.0	0.0	100.0	0.0	0.0	0.0	31.5	10.3	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	31.2	10.4	64.8
50000	25	27	27	2.0	0.0	100.0	0.0	0.0	0.0	31.1	10.2	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	31.0	10.7	64.8
60000	30	33	33	2.0	0.0	100.0	0.0	0.0	0.0	30.1	9.8	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	30.7	10.9	64.8
70000	35	38	38	2.0	0.0	100.0	0.0	0.0	0.0	31.5	10.3	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	30.7	11.1	64.8
80000	40	44	44	2.0	0.0	100.0	0.0	0.0	0.0	30.8	10.1	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	30.7	11.2	64.8
90000	45	50	50	2.0	0.0	100.0	0.0	0.0	0.0	29.9	9.8	64.8	0.	1.0	0.0	100.0	0.0	0.0	0.0	30.5	11.2	64.8
100000	50	55	55	2.0	0.1	100.0	0.0	0.0	0.0	30.6	11.8	277.8	6.	1.0	0.0	100.0	0.0	0.0	0.0	30.4	11.4	233.3
110000	55	61	61	2.0	0.1	99.9	0.0	0.0	0.0	31.3	16.0	347.1	9.	1.0	0.1	99.9	0.0	0.0	0.0	30.4	11.8	302.6
120000	60	66	66	2.0	0.3	99.7	0.0	0.0	0.0	32.4	31.1	500.2	17.	1.0	0.3	99.7	0.1	0.0	0.0	30.4	13.6	460.6
130000	65	72	71	2.0	0.5	99.4	0.1	0.1	0.0	36.2	45.4	554.7	19.	1.0	0.4	99.4	0.1	0.1	0.1	30.6	16.8	515.2
140000	70	77	77	2.1	0.7	98.7	0.2	0.1	0.1	45.9	64.3	554.7	19.	1.1	0.6	98.7	0.2	0.2	0.2	31.4	21.8	515.2
150000	75	83	81	2.1	1.1	97.7	0.3	0.2	0.2	63.9	94.1	767.5	28.	1.1	1.0	97.7	0.4	0.3	0.3	33.0	30.3	728.0
160000	80	88	86	2.3	1.6	96.4	0.5	0.3	0.3	88.0	123.7	822.0	30.	1.2	1.4	96.4	0.5	0.5	0.4	35.8	41.4	782.5
170000	85	94	90	2.4	2.2	95.0	0.5	0.4	0.4	120.9	164.1	1099.4	41.	1.4	2.0	95.0	0.6	0.5	0.5	40.0	56.6	1059.8
180000	90	100	93	2.6	3.0	93.2	0.6	0.5	0.5	156.9	204.2	1366.7	52.	1.6	2.8	93.2	0.7	0.6	0.6	45.6	74.5	1327.2
190000	95	105	96	2.9	3.7	91.0	0.8	0.6	0.6	202.1	232.3	1446.0	55.	1.8	3.5	91.0	0.9	0.8	0.7	52.7	93.1	1406.5
200000	100	111	98	3.3	4.6	88.6	0.9	0.8	0.7	254.8	263.4	1515.3	58.	2.2	4.3	88.6	1.0	0.9	0.8	61.6	113.3	1475.7

INDEPENDENT OVERFLOW ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS	INSERES	FONCTION
4800	5200.	200000		3

				INSERTION									CONSULTATION									
N	%N	%ZPR	%ZTOT	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	5	5	2.0	0.0	100.0	0.0	0.0	0.0	36.8	10.1	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	36.8	10.1	62.6
20000	10	10	10	2.0	0.0	100.0	0.0	0.0	0.0	36.4	10.0	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	36.6	9.9	62.6
30000	15	15	15	2.0	0.0	100.0	0.0	0.0	0.0	37.0	10.1	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	36.7	9.9	62.6
40000	20	20	20	2.0	0.0	100.0	0.0	0.0	0.0	36.7	10.1	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	36.5	10.4	62.6
50000	25	26	26	2.0	0.0	100.0	0.0	0.0	0.0	36.8	10.2	62.6	0.	1.0	0.0	100.0	0.0	0.0	0.0	36.5	10.7	62.6
60000	30	31	31	2.0	0.0	100.0	0.0	0.0	0.0	36.7	10.3	160.8	3.	1.0	0.0	100.0	0.0	0.0	0.0	36.4	10.9	120.6
70000	35	36	36	2.0	0.1	100.0	0.0	0.0	0.0	36.9	12.3	304.0	8.	1.0	0.1	100.0	0.0	0.0	0.0	36.4	11.2	263.7
80000	40	41	41	2.0	0.2	99.9	0.0	0.0	0.0	37.5	18.1	437.5	15.	1.0	0.2	99.9	0.0	0.0	0.0	36.4	12.1	402.2
90000	45	46	46	2.0	0.4	99.7	0.0	0.0	0.0	39.2	26.6	545.7	20.	1.0	0.3	99.7	0.0	0.0	0.0	36.6	14.0	510.4
100000	50	52	51	2.0	0.5	99.4	0.1	0.0	0.0	42.2	37.1	643.9	25.	1.0	0.5	99.4	0.1	0.1	0.1	36.9	16.7	608.6
110000	55	57	56	2.1	0.8	98.9	0.2	0.1	0.1	49.3	56.5	729.4	29.	1.1	0.7	98.9	0.2	0.2	0.1	37.7	21.9	694.2
120000	60	62	61	2.1	1.2	98.1	0.3	0.2	0.2	58.9	76.3	825.0	33.	1.1	1.1	98.1	0.3	0.3	0.2	39.1	28.9	789.7
130000	65	67	66	2.2	1.5	97.2	0.4	0.3	0.2	68.6	94.6	933.2	38.	1.2	1.4	97.2	0.4	0.4	0.3	40.9	36.6	897.9
140000	70	72	70	2.3	1.9	96.1	0.5	0.3	0.3	81.3	114.0	998.5	40.	1.3	1.8	96.1	0.5	0.5	0.4	43.2	45.4	963.2
150000	75	78	75	2.4	2.4	94.8	0.6	0.4	0.4	96.5	133.8	1033.7	42.	1.4	2.2	94.8	0.7	0.6	0.5	46.1	55.0	998.5
160000	80	83	78	2.6	2.8	93.3	0.7	0.5	0.5	113.4	153.6	1169.6	48.	1.5	2.7	93.3	0.8	0.7	0.7	49.6	65.2	1134.3
170000	85	88	82	2.8	3.4	91.7	0.8	0.6	0.6	141.1	187.0	1277.8	53.	1.7	3.2	91.7	0.9	0.8	0.8	54.2	78.2	1242.5
180000	90	93	85	3.0	4.1	89.9	1.0	0.8	0.7	169.4	212.5	1396.0	58.	1.9	3.9	89.9	1.1	1.0	0.9	59.7	92.3	1360.7
190000	95	98	88	3.3	4.8	87.8	1.1	0.9	0.8	205.2	237.0	1619.7	67.	2.2	4.5	87.8	1.2	1.1	1.0	66.3	107.3	1584.4
200000	100	104	90	3.7	5.5	85.6	1.2	1.0	0.9	240.5	262.3	1667.6	70.	2.5	5.2	85.6	1.3	1.2	1.1	73.9	123.3	1632.4

INDEPENDENT OVERFLOW ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS	INSERES	FONCTION
3600	6500.	200000		3

				INSERTION										CONSULTATION								
N	%N	%ZPR	%ZTOT	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	5	5	2.0	0.0	100.0	0.0	0.0	0.0	37.5	10.0	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	37.5	10.0	63.7
20000	10	11	11	2.0	0.0	100.0	0.0	0.0	0.0	37.2	9.9	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	37.3	9.8	63.7
30000	15	16	16	2.0	0.0	100.0	0.0	0.0	0.0	37.5	10.0	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	37.3	9.9	63.7
40000	20	22	22	2.0	0.0	100.0	0.0	0.0	0.0	37.3	10.1	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	37.2	10.5	63.7
50000	25	27	27	2.0	0.0	100.0	0.0	0.0	0.0	37.0	9.9	63.7	0.	1.0	0.0	100.0	0.0	0.0	0.0	37.0	10.8	63.7
60000	30	33	33	2.0	0.0	100.0	0.0	0.0	0.0	37.4	10.2	148.5	2.	1.0	0.0	100.0	0.0	0.0	0.0	37.0	11.0	106.1
70000	35	38	38	2.0	0.0	100.0	0.0	0.0	0.0	37.3	10.3	229.6	5.	1.0	0.0	100.0	0.0	0.0	0.0	36.9	11.2	187.2
80000	40	44	44	2.0	0.1	100.0	0.0	0.0	0.0	37.8	13.3	383.1	10.	1.0	0.1	100.0	0.0	0.0	0.0	36.9	11.5	340.7
90000	45	50	49	2.0	0.2	99.9	0.0	0.0	0.0	38.6	19.1	378.1	11.	1.0	0.2	99.9	0.0	0.0	0.0	37.0	12.4	340.7
100000	50	55	55	2.0	0.4	99.7	0.0	0.0	0.0	40.9	30.8	512.7	18.	1.0	0.3	99.7	0.1	0.0	0.0	37.2	14.5	475.3
110000	55	61	60	2.0	0.7	99.2	0.1	0.1	0.1	47.8	51.5	636.2	23.	1.0	0.6	99.2	0.2	0.1	0.1	37.9	19.4	598.8
120000	60	66	65	2.1	1.0	98.3	0.2	0.2	0.1	60.5	78.7	834.4	32.	1.1	0.9	98.3	0.3	0.2	0.2	39.3	27.6	797.1
130000	65	72	70	2.2	1.5	97.2	0.4	0.3	0.2	76.1	103.8	840.7	32.	1.2	1.4	97.2	0.4	0.4	0.3	41.6	37.5	803.4
140000	70	77	75	2.3	2.0	95.8	0.5	0.4	0.3	95.1	130.8	1024.0	40.	1.3	1.9	95.8	0.6	0.5	0.4	44.7	49.2	986.6
150000	75	83	79	2.5	2.6	94.1	0.6	0.5	0.4	121.1	160.9	1171.2	46.	1.5	2.4	94.1	0.7	0.6	0.6	48.9	62.9	1133.8
160000	80	88	83	2.7	3.3	92.2	0.8	0.6	0.5	150.0	192.0	1289.7	51.	1.7	3.1	92.2	0.8	0.8	0.7	54.3	78.2	1252.3
170000	85	94	86	3.0	4.1	90.2	0.9	0.7	0.6	185.1	230.1	1569.1	63.	1.9	3.8	90.2	0.9	0.9	0.8	61.0	96.1	1531.7
180000	90	100	89	3.4	4.9	88.0	1.0	0.8	0.7	229.9	259.9	1716.3	69.	2.3	4.6	88.0	1.0	1.0	0.9	69.3	115.1	1678.9
190000	95	105	91	3.8	5.8	85.6	1.1	0.9	0.8	279.2	291.2	1758.7	71.	2.7	5.5	85.6	1.1	1.1	1.0	79.1	135.5	1721.3
200000	100	111	93	4.3	6.7	83.1	1.1	1.0	0.9	333.9	327.5	1924.6	78.	3.1	6.5	83.1	1.2	1.1	1.1	90.4	157.7	1887.2

EXPLICATION DE L'OUTPUT.

NBRECASES:	NBRE CASES DE LA ZONE PRIMAIRE
BSIZE-BYTES:	LONGUEUR MOYENNE D'UNE CASE (EN BYTES) (LONG. MOYENNE D'UN RECORD = 130 BYTES)
NBRE RECORDS INSERES:	A PARTIR DE 0 RECORD ON EN INSERE LE NOMBRE INDIQUE
FONCTION:	FONCTION DE HASHAGE 1= FONCTION DE DIVISION 2= FONCTION DE FOLDING 3= FONCTION DE MIDSQUARRE
N:	NOMBRE DE RECORDS DEJA INSERES
%N:	POURCENTAGE DE RECORDS DEJA INSERES
%ZPR:	POURCENTAGE DE PLACE OCCUPEE PAR RAPPORT A LA ZONE TOTALE DU FICHIER

I N S E R T I O N

A-M:	NBRE ACCES MOYEN
E-T:	ECART-TYPE DE A-M
%I (I=0,2,4,5):	% DE RECORDS NECESSITANT I ACCES SUPPLEMENTAIRES
TP-MOY:	TEMPS MOYEN D'ACCES A UN RECORD PAR PERIODE
E-T:	ECART-TYPE DE TP-MOY
MAX:	TEMPS MAX. D'ACCES A UN RECORD PAR PERIODE
MAX-CH:	LONGUEUR MAXIMUN DE LA CHAINE DES ELEMENTS N'AYANT PAS TROUVE DE PLACE DANS LEUR CASE INITIALE DE DEPART

C O N S U L T A T I O N

A-M:	NBRE D'ACCES MOYEN
E-T:	ECART-TYPE DE A-M
%I (I=1,2,3,4):	POURCENTAGE D'ARTICLES NECESSITANT I ACCES
TP-MOY:	TEMPS MOYEN D'ACCES A UN RECORD (TP-MOY EST UN TEMPS GLOBAL)
E-T:	ECART-TYPE DE TP-MOY
MAX:	TEMPS D'ACCES MAXIMUN A UN RECORD DEPUIS LE DEBUT DES OPERATIONS

PRIMARY CHAINING ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS INSERES	FONCTION
5250	5200.	200000	1

			INSERTION									CONSULTATION								
N	%N	%ZPR	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	4	2.0	0.0	100.0	0.0	0.0	0.0	30.8	8.7	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	30.8	8.7	62.
20000	10	9	2.0	0.0	100.0	0.0	0.0	0.0	30.0	8.3	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	30.4	8.4	62.
30000	15	14	2.0	0.0	100.0	0.0	0.0	0.0	29.8	8.2	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	30.2	8.1	62.
40000	20	19	2.0	0.0	100.0	0.0	0.0	0.0	30.3	8.4	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	30.1	8.3	62.
50000	25	23	2.0	0.0	100.0	0.0	0.0	0.0	30.3	8.5	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	30.0	8.7	62.
60000	30	28	2.0	0.0	100.0	0.0	0.0	0.0	29.8	8.5	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	29.9	8.9	62.
70000	35	33	2.0	0.0	100.0	0.0	0.0	0.0	30.6	8.6	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	29.9	9.1	62.
80000	40	38	2.0	0.0	100.0	0.0	0.0	0.0	29.9	8.5	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	29.8	9.2	62.
90000	45	42	2.0	0.0	100.0	0.0	0.0	0.0	29.5	8.4	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	29.7	9.3	62.
100000	50	47	2.0	0.0	100.0	0.0	0.0	0.0	29.7	8.4	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	29.6	9.4	62.
110000	55	52	2.0	0.0	100.0	0.0	0.0	0.0	29.8	8.3	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	29.6	9.4	62.
120000	60	57	2.0	0.0	100.0	0.0	0.0	0.0	28.3	7.8	85.6	1	1.0	0.0	100.0	0.0	0.0	0.0	29.4	9.4	62.
130000	65	61	2.0	0.0	100.0	0.0	0.0	0.0	27.6	7.6	85.6	1	1.0	0.0	100.0	0.0	0.0	0.0	29.2	9.4	62.
140000	70	66	2.0	0.1	99.9	0.0	0.0	0.0	27.8	8.2	100.6	1	1.0	0.0	99.9	0.1	0.0	0.0	29.1	9.4	75.
150000	75	71	2.0	0.1	99.9	0.0	0.0	0.0	28.0	8.6	93.2	2	1.0	0.0	99.9	0.1	0.0	0.0	28.9	9.4	75.
160000	80	76	2.0	0.1	99.7	0.0	0.0	0.0	29.5	10.6	100.6	2	1.0	0.1	99.7	0.3	0.0	0.0	28.9	9.4	75.
170000	85	80	2.0	0.1	99.5	0.0	0.0	0.0	31.3	11.6	120.6	2	1.0	0.1	99.5	0.5	0.0	0.0	28.9	9.4	85.
180000	90	85	2.0	0.2	98.9	0.1	0.0	0.0	33.2	16.2	128.2	4	1.0	0.1	98.9	1.0	0.1	0.0	29.0	9.5	97.
190000	95	90	2.0	0.3	97.9	0.3	0.0	0.0	37.7	21.1	152.8	7	1.0	0.2	97.9	1.7	0.3	0.1	29.1	9.8	126.
200000	100	95	2.1	0.5	96.2	0.7	0.1	0.0	47.5	29.8	215.8	12	1.1	0.3	96.2	2.6	0.7	0.3	29.5	10.7	189.

PRIMARY CHAINING ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS INSERES	FONCTION
3000	9100.	200000	1

			INSERTION									CONSULTATION								
N	%N	%ZPR	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	4	2.0	0.0	100.0	0.0	0.0	0.0	33.8	8.3	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	33.8	8.3	65.
20000	10	9	2.0	0.0	100.0	0.0	0.0	0.0	33.2	8.3	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	33.5	8.2	65.
30000	15	14	2.0	0.0	100.0	0.0	0.0	0.0	32.9	8.1	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	33.3	8.1	65.
40000	20	19	2.0	0.0	100.0	0.0	0.0	0.0	33.3	8.3	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	33.1	8.7	65.
50000	25	23	2.0	0.0	100.0	0.0	0.0	0.0	33.3	8.3	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	33.0	9.2	65.
60000	30	28	2.0	0.0	100.0	0.0	0.0	0.0	32.9	8.3	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	32.8	9.5	65.
70000	35	33	2.0	0.0	100.0	0.0	0.0	0.0	33.6	8.3	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	32.8	9.7	65.
80000	40	38	2.0	0.0	100.0	0.0	0.0	0.0	33.0	8.4	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	32.8	9.8	65.
90000	45	42	2.0	0.0	100.0	0.0	0.0	0.0	32.6	8.3	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	32.6	10.0	65.
100000	50	47	2.0	0.0	100.0	0.0	0.0	0.0	32.8	8.2	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	32.6	10.0	65.
110000	55	52	2.0	0.0	100.0	0.0	0.0	0.0	32.9	8.2	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	32.5	10.1	65.
120000	60	57	2.0	0.0	100.0	0.0	0.0	0.0	31.4	7.6	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	32.4	10.1	65.
130000	65	61	2.0	0.0	100.0	0.0	0.0	0.0	30.7	7.1	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	32.2	10.1	65.
140000	70	66	2.0	0.0	100.0	0.0	0.0	0.0	30.6	7.2	65.8	0	1.0	0.0	100.0	0.0	0.0	0.0	32.0	10.1	65.
150000	75	71	2.0	0.0	100.0	0.0	0.0	0.0	30.7	7.3	98.3	1	1.0	0.0	100.0	0.0	0.0	0.0	31.9	10.1	65.
160000	80	76	2.0	0.0	100.0	0.0	0.0	0.0	31.8	8.4	108.3	1	1.0	0.0	100.0	0.0	0.0	0.0	31.8	10.1	76.
170000	85	80	2.0	0.1	99.8	0.0	0.0	0.0	35.5	13.5	139.1	2	1.0	0.0	99.8	0.1	0.0	0.0	31.9	10.3	102.
180000	90	85	2.0	0.1	99.6	0.0	0.0	0.0	34.2	14.6	168.2	5	1.0	0.1	99.6	0.4	0.0	0.0	31.9	10.4	134.
190000	95	90	2.0	0.3	98.9	0.2	0.0	0.0	40.1	24.9	235.9	8	1.0	0.2	98.9	0.9	0.2	0.0	32.1	10.9	202.
200000	100	95	2.2	2.0	97.5	0.3	0.1	0.0	83.0	141.3	1184.9	70	1.1	1.9	97.5	1.4	0.3	0.1	34.1	32.9	1153.

PRIMARY CHAINING ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS INSERES	FONCTION
4200	6500.	200000	1

			INSERTION								CONSULTATION								
N	%N	%ZPR	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	4	2.0	0.0	100.0	0.0	0.0	0.0	31.8	8.4	63.7	0	1.0	0.0	100.0	0.0	0.0	0.0	31.8	8.4	63.
20000	10	9	2.0	0.0	100.0	0.0	0.0	0.0	30.9	8.0	63.7	0	1.0	0.0	100.0	0.0	0.0	0.0	31.3	8.1	63.
30000	15	14	2.0	0.0	100.0	0.0	0.0	0.0	30.7	7.9	63.7	0	1.0	0.0	100.0	0.0	0.0	0.0	31.1	7.9	63.
40000	20	19	2.0	0.0	100.0	0.0	0.0	0.0	31.3	8.3	63.7	0	1.0	0.0	100.0	0.0	0.0	0.0	31.1	8.3	63.
50000	25	23	2.0	0.0	100.0	0.0	0.0	0.0	31.2	8.3	63.7	0	1.0	0.0	100.0	0.0	0.0	0.0	31.0	8.7	63.
60000	30	28	2.0	0.0	100.0	0.0	0.0	0.0	30.8	8.3	63.7	0	1.0	0.0	100.0	0.0	0.0	0.0	30.8	9.0	63.
70000	35	33	2.0	0.0	100.0	0.0	0.0	0.0	31.6	8.5	63.7	0	1.0	0.0	100.0	0.0	0.0	0.0	30.8	9.2	63.
80000	40	38	2.0	0.0	100.0	0.0	0.0	0.0	30.8	8.1	63.7	0	1.0	0.0	100.0	0.0	0.0	0.0	30.7	9.3	63.
90000	45	42	2.0	0.0	100.0	0.0	0.0	0.0	30.4	8.1	63.7	0	1.0	0.0	100.0	0.0	0.0	0.0	30.6	9.4	63.
100000	50	47	2.0	0.0	100.0	0.0	0.0	0.0	30.7	8.2	63.7	0	1.0	0.0	100.0	0.0	0.0	0.0	30.6	9.4	63.
110000	55	52	2.0	0.0	100.0	0.0	0.0	0.0	30.8	8.3	89.8	1	1.0	0.0	100.0	0.0	0.0	0.0	30.5	9.5	63.
120000	60	57	2.0	0.0	100.0	0.0	0.0	0.0	29.2	7.6	94.8	1	1.0	0.0	100.0	0.0	0.0	0.0	30.3	9.5	63.
130000	65	61	2.0	0.0	100.0	0.0	0.0	0.0	28.6	7.5	84.8	1	1.0	0.0	100.0	0.0	0.0	0.0	30.2	9.5	63.
140000	70	66	2.0	0.0	100.0	0.0	0.0	0.0	28.7	7.9	89.8	1	1.0	0.0	100.0	0.0	0.0	0.0	30.0	9.5	63.
150000	75	71	2.0	0.1	99.9	0.0	0.0	0.0	28.9	8.2	108.5	2	1.0	0.0	99.9	0.1	0.0	0.0	29.9	9.5	77.
160000	80	76	2.0	0.1	99.8	0.0	0.0	0.0	30.1	9.5	113.5	2	1.0	0.0	99.8	0.2	0.0	0.0	29.8	9.5	86.
170000	85	80	2.0	0.1	99.7	0.0	0.0	0.0	31.6	10.6	118.5	2	1.0	0.1	99.7	0.3	0.0	0.0	29.8	9.6	86.
180000	90	85	2.0	0.2	99.3	0.0	0.0	0.0	33.0	14.9	124.8	3	1.0	0.1	99.3	0.7	0.0	0.0	29.8	9.6	87.
190000	95	90	2.0	0.3	98.5	0.2	0.0	0.0	36.8	20.0	171.6	6	1.0	0.2	98.5	1.2	0.2	0.0	29.9	9.9	142.
200000	100	95	2.1	0.4	97.1	0.5	0.1	0.0	45.2	28.2	204.6	10	1.0	0.3	97.1	2.1	0.5	0.2	30.3	10.5	175.

PRIMARY CHAINING ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS INSERES	FONCTION
2750	10400.	200000	1

			INSERTION									CONSULTATION								
N	%N	%ZPR	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	4	2.0	0.0	100.0	0.0	0.0	0.0	34.9	8.5	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	34.9	8.5	66.
20000	10	9	2.0	0.0	100.0	0.0	0.0	0.0	34.0	8.0	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	34.4	8.1	66.
30000	15	13	2.0	0.0	100.0	0.0	0.0	0.0	33.8	7.9	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	34.2	7.9	66.
40000	20	18	2.0	0.0	100.0	0.0	0.0	0.0	34.5	8.4	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	34.1	8.7	66.
50000	25	22	2.0	0.0	100.0	0.0	0.0	0.0	34.4	8.4	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	34.0	9.2	66.
60000	30	27	2.0	0.0	100.0	0.0	0.0	0.0	33.9	8.3	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	33.8	9.5	66.
70000	35	31	2.0	0.0	100.0	0.0	0.0	0.0	34.7	8.4	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	33.8	9.8	66.
80000	40	36	2.0	0.0	100.0	0.0	0.0	0.0	33.9	8.3	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	33.7	9.9	66.
90000	45	40	2.0	0.0	100.0	0.0	0.0	0.0	33.6	8.2	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	33.6	10.0	66.
100000	50	45	2.0	0.0	100.0	0.0	0.0	0.0	33.7	8.1	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	33.5	10.1	66.
110000	55	50	2.0	0.0	100.0	0.0	0.0	0.0	33.9	8.3	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	33.5	10.2	66.
120000	60	54	2.0	0.0	100.0	0.0	0.0	0.0	32.4	7.7	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	33.3	10.2	66.
130000	65	59	2.0	0.0	100.0	0.0	0.0	0.0	31.7	7.3	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	33.1	10.2	66.
140000	70	63	2.0	0.0	100.0	0.0	0.0	0.0	31.8	7.3	66.9	0	1.0	0.0	100.0	0.0	0.0	0.0	33.0	10.2	66.
150000	75	68	2.0	0.0	100.0	0.0	0.0	0.0	31.8	7.4	87.5	1	1.0	0.0	100.0	0.0	0.0	0.0	32.8	10.2	66.
160000	80	72	2.0	0.0	100.0	0.0	0.0	0.0	32.7	8.1	102.5	1	1.0	0.0	100.0	0.0	0.0	0.0	32.8	10.2	68.
170000	85	77	2.0	0.0	100.0	0.0	0.0	0.0	34.5	9.4	112.5	1	1.0	0.0	100.0	0.0	0.0	0.0	32.8	10.3	78.
180000	90	81	2.0	0.1	99.9	0.0	0.0	0.0	33.1	9.6	117.5	2	1.0	0.0	99.9	0.1	0.0	0.0	32.7	10.3	83.
190000	95	86	2.0	0.1	99.7	0.0	0.0	0.0	34.6	13.0	132.5	3	1.0	0.1	99.7	0.3	0.0	0.0	32.7	10.4	92.
200000	100	90	2.0	0.2	99.2	0.1	0.0	0.0	37.9	19.2	162.0	5	1.0	0.1	99.2	0.7	0.1	0.0	32.8	10.5	126.

PRIMARY CHAINING ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS INSERES	FONCTION
5500	5200.	200000	2

			INSERTION									CONSULTATION								
N	%N	%ZPR	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	4	2.0	0.0	100.0	0.0	0.0	0.0	34.5	9.1	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	34.5	9.1	62.
20000	10	9	2.0	0.0	100.0	0.0	0.0	0.0	34.2	9.3	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	34.4	9.1	62.
30000	15	13	2.0	0.0	100.0	0.0	0.0	0.0	34.4	9.3	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	34.4	9.0	62.
40000	20	18	2.0	0.0	100.0	0.0	0.0	0.0	34.3	9.2	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	34.2	9.5	62.
50000	25	22	2.0	0.0	100.0	0.0	0.0	0.0	34.3	9.4	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	34.1	9.8	62.
60000	30	27	2.0	0.0	100.0	0.0	0.0	0.0	34.1	9.5	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	34.0	10.1	62.
70000	35	31	2.0	0.0	100.0	0.0	0.0	0.0	34.5	9.4	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	34.0	10.2	62.
80000	40	36	2.0	0.0	100.0	0.0	0.0	0.0	34.4	9.4	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	34.0	10.3	62.
90000	45	40	2.0	0.0	100.0	0.0	0.0	0.0	34.2	9.5	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	33.9	10.4	62.
100000	50	45	2.0	0.0	100.0	0.0	0.0	0.0	34.1	9.5	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	33.9	10.5	62.
110000	55	50	2.0	0.0	100.0	0.0	0.0	0.0	34.2	9.4	75.6	1	1.0	0.0	100.0	0.0	0.0	0.0	33.8	10.5	62.
120000	60	54	2.0	0.0	100.0	0.0	0.0	0.0	33.7	10.2	90.6	1	1.0	0.0	100.0	0.0	0.0	0.0	33.8	10.6	65.
130000	65	59	2.0	0.0	99.9	0.0	0.0	0.0	33.7	10.6	100.6	2	1.0	0.0	99.9	0.0	0.0	0.0	33.7	10.7	70.
140000	70	63	2.0	0.1	99.9	0.0	0.0	0.0	34.3	11.3	95.6	2	1.0	0.0	99.9	0.1	0.0	0.0	33.7	10.8	70.
150000	75	68	2.0	0.1	99.8	0.0	0.0	0.0	34.6	12.0	115.6	2	1.0	0.1	99.8	0.2	0.0	0.0	33.7	10.9	82.
160000	80	72	2.0	0.2	99.5	0.0	0.0	0.0	35.8	13.6	126.0	4	1.0	0.1	99.5	0.5	0.0	0.0	33.7	11.0	98.
170000	85	77	2.0	0.2	99.1	0.1	0.0	0.0	37.1	14.9	126.0	4	1.0	0.1	99.1	0.8	0.1	0.0	33.8	11.1	98.
180000	90	81	2.0	0.3	98.2	0.3	0.0	0.0	42.1	22.3	187.8	9	1.0	0.2	98.2	1.4	0.3	0.1	33.9	11.4	161.
190000	95	86	2.1	0.6	96.7	0.6	0.1	0.1	53.0	36.1	331.3	17	1.1	0.4	96.7	2.2	0.6	0.2	34.4	12.9	302.
200000	100	90	2.2	1.4	94.3	1.0	0.3	0.2	89.5	89.4	853.2	37	1.2	1.2	94.3	2.9	1.0	0.5	36.4	23.5	805.

PRIMARY CHAINING ALGORITHM

NBRECASES	BSIZE-BYTES	NBRE RECORDS INSERES	FONCTION
5500	5200.	200000	3

			INSERTION									CONSULTATION								
N	%N	%ZPR	A-M	E-T	%0	%2	%4	%5	TP-MOY	E-T	MAX	MAX-CH	A-M	E-T	%1	%2	%3	%4	TP-MOY	E-T	MAX
10000	5	4	2.0	0.0	100.0	0.0	0.0	0.0	35.2	8.3	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	35.2	8.3	62.
20000	10	9	2.0	0.0	100.0	0.0	0.0	0.0	35.1	8.3	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	35.1	8.1	62.
30000	15	13	2.0	0.0	100.0	0.0	0.0	0.0	35.1	8.3	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	35.1	8.0	62.
40000	20	18	2.0	0.0	100.0	0.0	0.0	0.0	35.2	8.4	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	35.0	8.6	62.
50000	25	22	2.0	0.0	100.0	0.0	0.0	0.0	35.1	8.5	62.6	0	1.0	0.0	100.0	0.0	0.0	0.0	34.9	9.0	62.
60000	30	27	2.0	0.0	100.0	0.0	0.0	0.0	35.1	8.5	80.6	1	1.0	0.0	100.0	0.0	0.0	0.0	34.8	9.2	62.
70000	35	31	2.0	0.0	99.9	0.0	0.0	0.0	35.2	8.7	80.6	1	1.0	0.0	99.9	0.0	0.0	0.0	34.8	9.4	62.
80000	40	36	2.0	0.1	99.9	0.0	0.0	0.0	35.4	9.1	110.6	2	1.0	0.0	99.9	0.1	0.0	0.0	34.7	9.5	75.
90000	45	40	2.0	0.1	99.6	0.0	0.0	0.0	35.9	10.6	115.6	3	1.0	0.1	99.6	0.4	0.0	0.0	34.7	9.6	85.
100000	50	45	2.0	0.2	99.2	0.0	0.0	0.0	36.9	12.6	119.2	3	1.0	0.1	99.2	0.7	0.0	0.0	34.8	9.7	90.
110000	55	50	2.0	0.3	98.5	0.1	0.0	0.0	39.0	15.2	134.2	4	1.0	0.1	98.5	1.3	0.1	0.0	34.9	9.9	105.
120000	60	54	2.1	0.3	97.5	0.3	0.0	0.0	41.0	17.7	138.1	5	1.0	0.2	97.5	2.1	0.3	0.0	35.0	10.2	110.
130000	65	59	2.1	0.4	96.6	0.5	0.0	0.0	41.8	18.9	165.3	8	1.0	0.2	96.6	2.8	0.5	0.1	35.2	10.4	138.
140000	70	63	2.1	0.5	95.3	0.8	0.1	0.0	45.8	22.8	201.3	10	1.1	0.3	95.3	3.7	0.8	0.2	35.5	10.9	174.
150000	75	68	2.1	0.6	93.8	1.1	0.1	0.0	48.8	25.5	236.2	12	1.1	0.4	93.8	4.6	1.1	0.3	35.9	11.4	209.
160000	80	72	2.2	0.8	91.9	1.6	0.2	0.1	54.7	31.1	266.3	14	1.1	0.5	91.9	5.5	1.6	0.5	36.4	12.5	239.
170000	85	77	2.3	0.9	90.1	2.1	0.4	0.2	59.1	35.1	279.5	15	1.2	0.7	90.1	6.3	2.1	0.8	37.1	13.9	252.
180000	90	81	2.4	1.1	87.7	2.6	0.6	0.4	72.3	44.5	347.2	20	1.2	0.9	87.7	7.0	2.6	1.1	38.2	16.3	320.
190000	95	86	2.5	1.5	84.9	3.1	0.9	0.6	93.7	64.1	470.0	24	1.4	1.3	84.9	7.6	3.1	1.5	40.1	21.9	441.
200000	100	90	2.8	2.4	81.9	3.4	1.2	0.8	150.4	125.0	856.0	47	1.6	2.2	81.9	7.8	3.4	1.8	44.5	38.9	826.

BUMP
0 0 3 2 2 1 3 7 1
*FM B16/1982/13/2