

PUC

Series: Monografias em Ciência da Computação

Nº 16/78

O SUPORTE BÁSICO DO SISTEMA HYADES DE GERÊNCIA
DE BANCO DE DADOS

- VOLUME 2 -

O PACOTE DE AVALIAÇÃO DE EFICIÊNCIA

por

Sônia M. M. Passos
Roberto P. Vasques

Departamento de Informática

Pontificia Universidade Católica do Rio de Janeiro

Rua Marquês de São Vicente 225 — ZC 19

Rio de Janeiro — Brasil

DEPARTAMENTO DE INFORMÁTICA
SETOR DE DOCUMENTAÇÃO
E INFORMAÇÃO

Série: Monografias em Ciência da Computação
Nº 16/78

Editor da Série: Michael F. Challis

Agosto, 1978

O SUPORTE BÁSICO DO SISTEMA HYADES DE GERÊNCIA
DE BANCO DE DADOS *

- VOLUME 2 -

O PACOTE DE AVALIAÇÃO DE EFICIÊNCIA

por

Sônia M.M. Passos

Roberto P. Vasques

M 3862

SETOR DE DOCUMENTAÇÃO E INFORMAÇÃO	
CÓDIGO / REGISTRO 4401	DATA 10 7 78
DEPT.º DE INFORMÁTICA	

* Trabalho patrocinado pelo CNPq TC 2222.0060/77

Para obter cópias, dirija-se a:

Rosane T. L. Castilho
Chefe, Setor de Documentação e Informação
Deptº de Informática - PUC/RJ
Rua Marquês de São Vicente, 209 - Gávea
22.453 - Rio de Janeiro - RJ - Brasil.

SUMÁRIO

APRESENTAÇÃO..... iii

VOLUME 1 - O PACOTE DE GERÊNCIA DE BANCO DE DADOS

PARTE 1

1. INTRODUÇÃO.....	1
1.1. Apresentação do Problema.....	1
1.2. Objetivo do Trabalho.....	2
1.3. Desenvolvimento.....	3
1.4. Organização.....	3
2. MANUAL DO SISTEMA.....	5
2.1. Concepção do Sistema.....	5
2.2. Arquivos do Sistema.....	9
2.3. Procedimentos do Sistema.....	33
2.4. Fluxo do Sistema.....	41
2.5. Padronização dos Módulos.....	46
2.6. Especificação dos Primitivos e Subrotinas do Sistema.....	53

PARTE 2

3. MANUAL DO USUÁRIO.....	108
3.1. Funções do Sistema.....	108
3.2. Necessidades de Hardware e Software.....	111
3.3. Arquivos do Sistema.....	112
3.4. Fluxo do Sistema.....	120
3.5. Especificação dos Primitivos a Nível de Usuário.....	121
3.6. Códigos de Retorno dos Primitivos.....	148
3.7. Restrições de Implementação.....	149
4. MANUAL DE OPERAÇÃO.....	151
4.1. Execução do Sistema.....	151
4.2. Exemplo de Uso do Suporte.....	154
4.3. Manutenção do Suporte.....	157
REFERÊNCIAS BIBLIOGRÁFICAS.....	187

VOLUME 2 - O PACOTE DE AVALIAÇÃO DE EFICIÊNCIA

1. INTRODUÇÃO.....	1
1.1. Apresentação do Problema.....	1
1.2. Desenvolvimento do Trabalho.....	1
1.3. Objetivo do Trabalho.....	2
1.4. Organização.....	4
2. DESENVOLVIMENTO E IMPLANTAÇÃO DO SISTEMA DE BANCO DE DADOS.....	6
2.1. Considerações Gerais.....	6
2.2. Arquivos.....	6
2.3. Procedimentos.....	9
3. DESENVOLVIMENTO DO SISTEMA DE AVALIAÇÃO DE EFICIÊNCIA...14	
3.1. O que entendemos por Avaliação de Eficiência de Sistema de Banco de Dados.....	14
3.2. Concepção.....	16
3.3. Considerações Gerais.....	16
3.4. Parâmetros a Analisar.....	18
3.5. Arquivos.....	19
3.6. Programas.....	29
3.7. Subrotinas.....	41
4. MANUAL DE UTILIZAÇÃO DO SISTEMA DE AVALIAÇÃO DE EFICIÊNCIA.....	53
4.1. Programa de Distribuições.....	53
4.2. Programa de Frequência.....	61
4.3. Programa de Alocação de Área.....	64
5. ESTUDO DE UM CASO.....	67
5.1. Relatórios de Distribuições.....	67
5.2. Relatórios de Frequência.....	74
5.3. Relatórios de Alocação de Espaço.....	77
REFERÊNCIAS BIBLIOGRÁFICAS.....	82

O SUPORTE BÁSICO DO SISTEMA HYADES DE GERÊNCIA
DE BANCO DE DADOS

- VOLUME 2 -

O PACOTE DE AVALIAÇÃO DE EFICIÊNCIA

Este volume foi elaborado, como tese de
mestrado, por Sônia Maria Miguel Passos.

RESUMO

Este trabalho é um Sistema de Avaliação de Eficiência (SAE) de um Sistema de Banco de Dados (SBD). A análise dos resultados fornecidos pelo SAE dá ao usuário subsídios que permitem adaptar o SBD às suas necessidades e obter uma maior eficiência; esta análise fornece também suporte para modificações implementadas em futuras versões do SBD.

O SAE é um sistema definido especificamente sobre o SBD que é o trabalho de tese de Roberto Pires Vasques e foi apresentado na mesma época deste trabalho.

ABSTRACT

This work is a System of Performance Evaluation (SAE) of a System of Data Base (SBD). The analysis of the results get of SAE gives to the users elements that allow to adapt the SBD to his necessities and to get a grather efficiency; this analysis also give us a support to modification into versions futures of SBD.

This SAE is a system defined specifically about the SBD, which is the work of tese of ROBERTO PIRES VASQUES and was presented at same time of this work.

INTRODUÇÃO

1.1 Apresentação do problema

A utilização de sistemas integrados de informação através da metodologia de Banco de Dados, vem substituindo o uso de arquivos tradicionais.

Os sistemas de Banco de Dados mais usados atualmente, são pacotes prontos fornecidos por firmas especializadas (TOTAL, IMS, ADABAS, etc.). Entretanto, estes pacotes nem sempre atendem a características particulares do usuário e não permitem, em geral, que o usuário faça modificações.

Devido a esses problemas, desenvolvemos e implementamos um suporte de Sistema de Banco de Dados que pretende atender as características particulares de cada usuário. Este sistema pode ser usado diretamente, ou pode servir como base a outro Sistema de Definição e Manipulação de Dados, escolhido de acordo com a necessidade específica do usuário. Esse outro sistema pode ser a definição de uma linguagem de alto nível onde são definidas funções Macros que resolvem o problema daquele usuário particular.

Para que possamos avaliar o desempenho do sistema de Banco de Dados, quando utilizado por diferentes usuários, desenvolvemos e implantamos um sistema de Avaliação de Eficiência.

1.2 Desenvolvimento do trabalho

O desenvolvimento e implementação do sistema de Banco de Dados foi realizado pelo autor e por ROBERTO PIRES VASQUES, no

período de fevereiro de 1977 a janeiro de 1978, englobando de finição, estruturação, codificação, testes e integração dos módulos do sistema, sendo que quarenta por cento (40%) da codificação foi realizada por 5 (cinco) alunos de pós-graduação do Departamento de Informática da PUC/RJ.

A partir de outubro de 1977 foi desenvolvido e implementado o Sistema de Avaliação da Eficiência do Sistema de Banco de Dados.

Tanto o Sistema de Banco de Dados, quanto o de Avaliação de Eficiência foram desenvolvidos conscientes de seu uso como protótipo. A utilização do Sistema de Banco de Dados e mesmo resultados da avaliação de eficiência deverão resultar em versões mais otimizadas de ambos os sistemas.

A documentação referente ao Sistema de Banco de Dados consta da tese de ROBERTO PIRES VASQUES e a avaliação de eficiência do Sistema de Banco de Dados é o principal objeto desta tese.

1.3 Objetivo do trabalho

Desenvolver e implementar um sistema de Banco de Dados, com a finalidade de ser um suporte básico, para 3 (três) principais modelos de Banco de Dados e dotar este sistema de recursos que permita avaliar a sua eficiência. Com esta finalidade foi desenvolvido o Sistema de Avaliação de Eficiência.

O Sistema de Avaliação de Eficiência visa permitir ao usuário avaliar a eficiência do seu Banco de Dados, e com base nos resultados da análise adaptar o sistema às suas necessidades, bem como dar suporte a futuras versões do sistema de Base de Dados.

Como as ferramentas de medida de eficiência foram estabelecidas a posteriori, elas sofreram limitações impostas pela estrutura do sistema e dificuldades de acesso a pontos de medida interessantes.

Os Sistemas de Avaliação de Eficiência e de Banco de Dados são distintos, mas não independentes. Podemos classificar as medições em 3 (três) tipos em relação ao sistema de Banco de Dados.

- medições que necessitam de informações obtidas durante a execução do Sistema de Banco de Dados; se o usuário desejar efetuar alguma destas avaliações deverá indicar este fato, em parâmetro previamente definido, quando ativar o sistema de Banco de Dados.
p. ex. frequência de uso e tempo gasto por cada primitivo (função básica).

- medições que atuam sobre os arquivos de Banco de Dados, mas que podem ser solicitadas a qualquer momento pelo usuário.
p. ex. características dos dados armazenados no Banco de Dados.

- medições que orientam o usuário no dimensionamento da área a ser ocupada pelos seus arquivos, este tipo de medição pode ser utilizado pelo usuário antes mesmo de ter sido criado o Banco de Dados.

Os resultados destas medições serão discutidos de modo a orientar o usuário em que pontos o Sistema de Banco de Dados deve ser alterado para que se obtenha uma maior eficiência e melhor se adapte as suas necessidades.

1.4 Organização da Tese

O presente trabalho será organizado em 6 (seis) capítulos, a saber:

Capítulo 1, a Introdução que descrevo, apresenta a tese de forma geral, é composta dos seguintes itens: apresentação do problema, desenvolvimento do trabalho, objetivo do trabalho e a organização da tese.

o Capítulo 2, que explica o Desenvolvimento e Implementação do Sistema de Banco de Dados. Neste Capítulo são apresentadas as características básicas do Sistema de Banco de Dados, não são tratados detalhes de definição do sistema e restrições de utilização, pois a descrição do sistema em detalhe, se encontra na tese de ROBERTO PIRES VASQUES. É composto dos seguintes itens: considerações gerais, arquivos e procedimentos.

No Capítulo 3, temos o Desenvolvimento do Sistema de Eficiência. Neste capítulo estão descritos os critérios adotados para avaliação de eficiência e o Sistema de Avaliação de Eficiência. É composto dos seguintes itens: o que entendemos por avaliação de eficiência, concepção, os parâmetros analisados, descrição dos arquivos, programas, variáveis externas e subrotinas do Sistema de Avaliação de Eficiência.

No Capítulo 4, apresentamos o Manual de Utilização do Sistema de Avaliação de Eficiência, onde estão descritos de que forma pode-se utilizar o sistema, tanto no que se refere a ativar o sistema, quanto a análise dos resultados oriundos do sistema. Estão descritos, para cada programa do Sistema de Avaliação de Eficiência, o preenchimento ou obtenção da entrada e como fazer a análise dos resultados.

No Capítulo 5, apresentamos o estudo de um caso. Neste capítulo é feita a avaliação de eficiência do Banco de Dados em estudo. É composto das saídas do Sistema de Avaliação de Eficiência e da respectiva Análise de Eficiência.

Como anexos temos as listagens dos programas e das subrotinas do Sistema de Avaliação de Eficiência e os Relatórios Resumo dos Arquivos de Sistema de Banco de Dados utilizados no estudo do caso.

2. DESENVOLVIMENTO E IMPLEMENTAÇÃO DO SISTEMA DE BANCO DE DADOS

2.1 Considerações Gerais

O sistema desenvolvido, teve como características básicas, a modularidade e a flexibilidade, visando sua adaptação a situações e sistemas operacionais. Assim sendo, o sistema foi desenvolvido de forma estruturada de modo a que seja fácil ao usuário adicionar ou modificar o que for conveniente.

Esse sistema inclui mecanismos que tornam possível seu uso como suporte para implementação de modelos do tipo hierárquico, relacional ou de redes.

A codificação do sistema é em PL/1-F e utiliza a organização sequencial indexada, por restrição do software disponível na PUC/RJ.

Esta implementação permite que diversos usuários consultem o Banco de Dados concorrentemente; mas durante atualizações apenas um usuário deve estar utilizando o Banco de Dados.

Da estrutura básica do Sistema de Banco de Dados constam Arquivos e Procedimentos.

2.2 Arquivos

O Sistema de Banco de Dados contém quatro (4) arquivos. Podemos classificar estes arquivos quanto a sua natureza em:

- Arquivos fundamentais: neste grupo classificamos os arquivos de dados, elos e inversões.
- Dicionário de dados: que é o arquivo de estruturas.

Estes quatro arquivos podem ser de dois tipos para o sistema:

- Permanentes: são os arquivos que constituem a base do sistema.
- Temporários: são arquivos criados durante uma execução do sistema, mas que são liberados ao final desta execução.

Tendo em vista a necessidade de se criar arquivos em tempo de execução do sistema, a solução adotada foi a criação de arquivos compostos de segmentos. Um segmento para o sistema é um arquivo para o usuário.

Para evitar ambiguidade chamamos de:

- Segmento: arquivo criado pelo usuário.
- Arquivo: arquivo do sistema, e que contém todos os segmentos de determinada natureza e tipo.

A cada segmento criado pelo usuário, o Sistema de Banco de Dados atribui um número, este número é chamado de identificador do segmento.

O primeiro registro de cada segmento é chamado mestre do segmento e contém informações gerais sobre os registros do segmento.

2.2.1 Arquivos de Dados: Contém as informações propriamente ditas. Os registros do arquivo de dados são identificados por uma chave gerada pelo sistema, que chamamos de identificador do registro.

O primeiro registro deste arquivo é o mestre do Banco de Dados e contém informações sobre todos os arquivos permanentes do sistema.

2.2.2 Arquivos de Elos: Contém ligações entre registros de 2 (dois) diferentes segmentos de dados. Os registros de elos contêm juntos os identificadores dos registros dos segmentos de dados que se relacionam.

A cada registro de elo inserido pelo usuário, o sistema insere automaticamente o registro do elo inverso.

Chamamos de classe de elos ao conjunto de registros de elos com o mesmo identificador de registro do primeiro segmento de dados.

Cada classe de elos é precedida por um registro chamado mestre de classe de elos e que serve para identificar uma classe de elos.

Os segmentos de elos são necessários se o modelo adotado for o hierárquico ou o de redes.

2.2.3 Arquivos de Inversões: Contém os vários valores que um conjunto de itens apresenta em um segmento de dados e para cada um desses valores, os identificadores dos registros, em que o valor aparece no segmento de dados.

Os registros que constam de um segmento de inversão são resultados de uma classificação de um conjunto de itens do segmento de dados, que deu origem ao segmento de inversão.

Chamamos de classe de inversão ao conjunto de registros com o mesmo valor para um conjunto de itens.

Cada classe de inversão é precedida por um registro chamado mestre de classe de inversão e que serve para identificar uma classe de inversão.

2.2.4 Arquivos de Estruturas:- Este arquivo funciona como um dicionário de dados, que contém informações sobre os campos dos registros dos arquivos de dados do Sistema.

Este arquivo não é manuseado pelo usuário, a criação, consulta e deleção de seus segmentos é feita automaticamente pelo sistema.

Existe uma relação biunívoca e sobre entre os segmentos de estrutura e os segmentos de dados.

2.3 Procedimentos

Os procedimentos do sistema estão distribuídos em módulos e cada módulo com uma finalidade específica. Esta estrutura visa facilitar a utilização e manutenção do sistema.

Podemos classificar os módulos em:

- Primitivos: são módulos ativados como funções e destinam-se a atender necessidades específicas do usuário do Sistema de Banco de Dados. Estas funções são ativadas pelo usuário. O Sistema de Banco de Dados conta com vinte e quatro (24) primitivos.
- Subrotinas - executam funções comuns aos primitivos. O Sistema de Banco de Dados conta com três (3) subrotinas auxiliares. A ativação destas subrotinas é feita através dos primitivos. Existem também cinco (5) subrotinas que são ativadas pelos primitivos do Sistema de Banco de Dados, mas que tem como única finalidade, permitir que determinadas medições sejam feitas pelo Sistema de Avaliação de Eficiência.

2.3.1 Primitivos: Executam funções específicas do Sistema de Banco de Dados e operam sobre as estruturas de arquivos já descritas.

Os primitivos se classificam conforme a sua utilização em:

- primitivos para manuseio de segmentos;
- primitivos para manuseio de registros;
- primitivos de apoio.

2.3.1.1 Primitivos para manuseio de segmento - Destinam-se a criação, cancelamento, obtenção de informação ou troca de identificação dos arquivos.

Primitivos para criação de segmentos:

- CRIDAD - cria um segmento de dados;
- CRIELO - cria um segmento de elos;
- CRIINV - cria um segmento de inversões;
- CRIDUP - cria uma cópia de um segmento, tem entre outras a finalidade de criar uma cópia trocando o tipo do arquivo;
- CRICLS - cria arquivo de dados classificado;
- CRIPRJ - cria arquivo de dados a partir das classes de um arquivo de inversão. (este primitivo foi feito para ser usado com o interface relacional que está sendo desenvolvido).

Primitivos para cancelamento de segmentos:

- CANDAD - cancela um segmento de dados;
- CANELO - cancela um segmento de elos;
- CANINV - cancela um segmento de inversões.

Primitivos para obtenção de informações sobre um segmento:

- INFARQ - fornece informações constantes do mestre de um segmento de arquivos fundamentais ou informações sobre determinado campo de um segmento do arquivo de estruturas.

Primitivo para trocar a identificação de um segmento:

- TRIARQ - troca a identificação dada pelo usuário a um segmento.

2.3.1.2 Primitivos para manuseio de registros - Estes primitivos permitem selecionar, inserir, remover ou modificar registros em um arquivo.

Primitivos para selecionar um registro;

- SELDAD - seleciona um registro do segmento de dados;
- SELELO - seleciona um registro do segmento de elos;
- SELINV- seleciona um registro no segmento de inversão.

Estes primitivos possibilitam a seleção de determinado registro, do primeiro registro do segmento ou do registro seguinte ao último registro selecionado.

A seleção de registros de elos e de inversões permite também que seja selecionado o primeiro registro de uma classe e os registros seguintes ao selecionado antes, dentro da mesma classe.

Os três (3) tipos de seleção visam normalmente uma consulta a um segmento de dados, O SELDAD acessa diretamente o segmento de dados e o SELELO e SELINV produzem como resultado a identificação de um registro em um segmento de dados; esta identificação é então usada para se selecionar o registro desejado no segmento de dados.

Primitivos para inserção e remoção de dados:

- INSDAD - insere um registro no segmento de dados;
- INSELO - insere um registro no segmento de elos;
- REMDAD - remove um registro no segmento de dados;
- REMELO - remove um registro no segmento de elos.

A inserção e remoção em segmentos de inversão é sempre

automática, decorrendo da inserção e remoção de registros em segmentos de dados correspondentes. A remoção de registro em segmento de dados causa também a remoção automática em segmentos de elos.

Primitivos para modificação de registro:

- MODDAD - modifica o valor de um item no registro de segmento de dados. Esta modificação pode causar a inserção ou remoção automática em segmentos de inversões.

2.3.1.3 Primitivos de apoio - São primitivos com finalidade de dar apoio e segurança à utilização do Sistema de Banco de Dados.

São os seguintes os primitivos de apoio:

- ATIVDB - ativa o Sistema de Banco de Dados, para permitir a sua utilização;
- DATVDB - desativa o Sistema de Banco de Dados, guardando informações necessárias para nova utilização;
- COPIDB - copia os arquivos permanentes, este primitivo tem a finalidade de gerar um back up e de rearrumar os arquivos do Sistema de Banco de Dados;
- RESTDB - restaura os arquivos do Sistema a partir do back up;
- RESMDB - relatório contendo informações sobre todos os segmentos dos arquivos permanentes do Banco de Dados.

2.3.2 Subrotinas Auxiliares

São procedimentos comuns aos primitivos do Sistema de Banco de Dados.

O Sistema tem as seguintes subrotinas auxiliares:

- AUX1 - fornece informações sobre um segmento do Banco de Dados;
- AUX2 - atualiza informações sobre um segmento no Banco de Dados;
- AUX3 - abre e fecha os arquivos do Banco de Dados.

2.3.3 Subrotinas de Avaliação de Eficiência

Estes procedimentos embora sejam ativados pelos primitivos não fazem parte do Sistema de Banco de Dados mas sim do Sistema de Avaliação de Eficiência.

São as seguintes, as subrotinas de avaliação de eficiência:

- SUB1 - inicializa tabelas de medição de eficiência;
- SUB2 - guarda a hora de início do primitivo;
- SUB3 - guarda a hora do final do primitivo;
- SUB4 - atualiza tabela de medição do número de vezes que um arquivo é aberto;
- SUB5 - grava informações que serão usadas pelo Sistema de Avaliação de Eficiência.

Estas subrotinas serão descritas em detalhe no próximo capítulo.

3. DESENVOLVIMENTO DO SISTEMA E AVALIAÇÃO DE EFICIÊNCIA

3.1. O que entendemos por avaliação de eficiência de um Sistema de Banco de Dados.

3.1.1. Histórico e definição.

No princípio da utilização do Banco de Dados, a única preocupação dos profissionais da área, era, dado uma quantidade de dados, como arrumá-los, para ser possível fazer uma consulta. As limitações, a essa organização, eram impostas pela configuração de suas instalações e pelo tempo de resposta às consultas.

Em seguida, as pesquisas se encaminharam para tornar o Sistema de Banco de Dados, mais barato, mais seguro (hoje praticamente existe uma área independente pesquisando segurança em Banco de Dados), e com um desenho tal que permita atender as exigências do usuário.

A partir daí, verificou-se a necessidade de dotar um Sistema de Banco de Dados de medições que informem ao usuário o comportamento de várias funções, como tempo de acesso, área ocupada etc. O estudo dessas medições é o que chamamos de Avaliação da Eficiência de um Sistema de Banco de Dados.

3.1.2. Componentes da avaliação e eficiência

A atividade de pesquisa da Avaliação de Eficiência tem se concentrado em 2 (dois) componentes essenciais, a saber:

- elaboração dos parâmetros e dos métodos de medida
- análise dos resultados.

A análise dos sistemas de informações é determinada pela or

ganização e comportamento de dois elementos distintos:

- arquitetura do sistema; que é o conjunto de recursos materiais e lógicos (por exemplo na arquitetura do sistema não usamos PL/I optimizer nem o método de acesso VSAM, pois não eram recursos disponíveis na PUC/RJ, então tivemos que usar o PL/I-F e o sequencial indexado).
- desenho do sistema; que é o conjunto de tratamentos dados ao sistema.

3.1.3 - Medidas de Métodos

Diversas medidas de eficiência, são feitas simultaneamente sobre o mesmo sistema, para se poder comparar os resultados obtidos, em vários pontos de observação, sobre um mesmo fenômeno.

A primeira ferramenta de análise disponível para se entender e interpretar os resultados de medida, são as ferramentas estatísticas. Sua utilização é indispensável para reduzir a quantidade de dados recolhidos e apresentá-los sob uma forma utilizável.

Deve ser dado ao usuário um resumo dos resultados obtidos com a avaliação de eficiência, para que ele tenha os elementos necessários para tornar o seu sistema de Banco de Dados mais eficiente ou para desenvolver um outro sistema de Banco de Dados, mais voltado às suas necessidades.

Os métodos de avaliação da eficiência de Banco de Dados não se preocupam somente em avaliar a velocidade de acesso, hoje, os critérios de Avaliação de Eficiência grupam uma série de outras características, como preço, adaptabilidade, segurança e manutenção, necessários a utilização de um "software".

A avaliação da eficiência é hoje uma atividade técnica importante ao longo de todo o desenvolvimento e utilização de um sistema e os métodos existentes variam de sistema para sistema de acordo com as necessidades dos usuários.

Um método bastante utilizado é a apuração de médias dos tempos e de frequência estatística da utilização de instruções (no nosso sistema vamos calcular a frequência e o tempo médio dos primitivos).

Os métodos empíricos e o controle pelo bom senso ainda são muito utilizados.

No Sistema de Avaliação de Eficiência os parâmetros que vamos analisar estão definidos na seção 3.4, os tipos dos arquivos utilizados para fazer a análise constam da seção 3.5 e os programas da análise da seção 3.6.

3.2. Concepção

O Sistema de Avaliação de Eficiência foi desenvolvido com a finalidade de avaliar a eficiência do Sistema de Banco de Dados.

O sistema de Avaliação de Eficiência visa orientar o usuário quanto a modificações, que podem ser implementadas no sistema de Banco de Dados, tornando-o mais eficiente para suas aplicações, e ainda, com base nas análises feitas por diversos usuários, fornecer suporte a alterações que serão introduzidas em futuras versões do Sistema de Banco de Dados.

3.3. Considerações gerais

Os parâmetros que serão analisados pelo Sistema de Avaliação de Eficiência, foram sugeridos, com base na sua experiência

peçoal, pelo professor Kenneth Sevoik professor da Universidade de Toronto e como poderemos observar são pontos de medida bastante citados na literatura.

Como os elementos de medida da eficiência foram estabelecidos depois do Sistema de Banco de Dados ter sido definido, é evidente que as medições sofreram limitações impostas principalmente, pelo método de acesso utilizado nos arquivos do sistema de Banco de Dados.

Os Sistemas de Avaliação de Eficiência e de Banco de Dados são distintos, mas não independentes. Podemos classificar as medições em 3 (três) tipos em relação ao sistema de Banco de Dados.

- medições que necessitam de informações obtidas durante a execução do Sistema de Banco de Dados; se o usuário desejar efetuar alguma dessas avaliações deverá indicar este fato, em parâmetro previamente definido, quando ativar o Sistema de Banco de Dados.

p. ex. frequência de uso e tempo gasto por cada primitivo (função básica).

- medições que atuam sobre os arquivos de Banco de Dados, mas que podem ser solicitadas a qualquer momento pelo usuário.

p. ex. características dos dados armazenados no Banco de Dados.

- medições que orientam o usuário no dimensionamento da área a ser ocupada pelos seus arquivos, este tipo de medição pode ser utilizado pelo usuário antes mesmo de ter sido criado o Banco de Dados.

Os resultados destas medições serão discutidos de modo a orientar o usuário em que pontos o sistema deve ser alterado para que se obtenha uma maior eficiência e melhor se adapte às suas necessidades.

O Sistema de Avaliação de Eficiência está desenvolvido em PL/I-F e poderá ser usado em instalações que tenham esse software.

O sistema propriamente dito é composto de parâmetros a serem analisados, arquivos, programas e subrotinas e finalmente os resultados das análises (distribuição, frequência, alocação de espaço). Os parâmetros a serem analisados, os arquivos, os programas e as subrotinas são objeto dos próximos itens.

3.4. Parâmetros a analisar

O Sistema de Avaliação de Eficiência analisa os seguintes parâmetros referentes ao Sistema de Banco de Dados.

- características dos dados armazenados, envolvendo a distribuição do tamanho dos registros, do número de registros por arquivo, do tamanho dos itens, do número de registros por inversão, etc.
- características de utilização do Sistema de Banco de Dados, envolvendo a frequência de uso de cada primitivo.
- variáveis de observação do desempenho, distribuição do tempo de execução para cada primitivo.
- valores de equilíbrio da eficiência, tais como frequência de abertura dos arquivos, tamanho das chaves usadas para inversão.

3.5. Arquivos

O Sistema de Avaliação de Eficiência manipula 3 tipos de arquivos:

- arquivos do Sistema de Banco de Dados;
- arquivos do Sistema de Avaliação de Eficiência;
- arquivos de impressão.

Arquivos do Sistema de Banco de Dados utilizados pelo Sistema de Avaliação de Eficiência:

- arquivo de dados;
- arquivo de inversões;
- arquivo de estruturas.

Estes arquivos estão descritos no item arquivos do capítulo 2.

Os arquivos do Sistema de Avaliação de Eficiência são os seguintes:

- arquivo de distribuições;
- arquivo de frequência;
- arquivo de alocação de espaço.

Os arquivos de impressão do Sistema de Avaliação de Eficiência são os seguintes:

- relatórios de distribuição;
- relatórios de frequência;
- relatórios de alocação de espaço.

O quadro a seguir relaciona os arquivos às suas identificações no Sistema de Avaliação de Eficiência (SAE) e no Sistema Operacional.

NOME DO ARQUIVO	Identificação do arquivo	
	para o SAE (Filename)	para o Sistema Operacional (ddname)
Arquivo de Dados	ARQD	ARQDP
Arquivo de Inversões	ARQI	ARQIP
Arquivo de Estruturas	ARQS	ARQSP
Arquivo de Distribuições	ARQSAE1	ARQSAE1
Arquivo de Frequência	ARQSAE2	ARQSAE2
Arquivo de Alocação de Espaço	ARQSAE3	ARQSAE3
Relatórios de Distribuições	ARQIMP1	ARQIMP1
Relatórios de Frequência	ARQIMP2	ARQIMP2
Relatórios de Alocação de Espaço	ARQIMP3	ARQIMP3

3.5.1. Arquivos do Sistema de Avaliação de Eficiência

Todos os arquivos do Sistema de Avaliação de Eficiência são utilizados no Sistema como entrada. Tem a organização sequencial e o método utilizado é o QSAM.

A seguir descreveremos os arquivos do Sistema de Avaliação de Eficiência bem como os registros pertencentes a esses arquivos. Nesta descrição dos registros, a identificação do campo, representa efetivamente o nome desta variável nos programas do Sistema de Avaliação de Eficiência. O formato dos campos é "display" salvo quando for especificado.

3.5.1.1. - ARQSAE1 - Arquivo de Distribuições

Os registros deste arquivo especificam as distribuições que devem ser apuradas e quais as faixas da apuração.

O registro está definido em 80 posições, das quais apenas as 21 primeiras estão ocupadas, e tem o seguinte layout:

Identificação do campo	COD	LINFE	LSUPE	INCE	RESTO
Posição Início/Fim	1 3	4 9	10 15	16 21	22 80

Descrição dos campos:

COD - código da apuração a ser solicitada.
São válidos os seguintes códigos, com as suas significações:

DTR - tamanho dos registros de dados;

DNR - número de registros de dados por arquivos;

DTC - tamanho dos campos dos registros de dados;

IRC - número de registros de dados por classe de inversão;

ITC - tamanho da chave de inversão.

LINFE - limite inferior da distribuição.

LSUPE - limite superior da distribuição.

INCE - incremento da distribuição.

3.5.1.2 - ARQSDE2 - Arquivo de Frequência.

Este arquivo tem 2 tipos de registros de tamanho variável, a saber:

- registros de primitivos;
- registros de número de "opens".

Registros de Primitivos: Estes registros contêm informações sobre o tempo de execução dos pri

mitivos e a sua frequência de utilização. Existe no ARQSAE2 trinta e sete (37) registros deste tipo; sendo que a cada primitivo corresponde pelo menos 1 registro, e os primitivos que tem diversas finalidades (primitivo de ativação e de seleção) tem um registro associado a cada finalidade. Dizemos que existe um registro para cada Primitivo/Finalidade.

O registro tem o seguinte lay-out:

Identificação
do campo

Posição
Início/Fim

NUMT	CTOT	CMIN	CMAX	ETOT	EMIN	EMAX
1 3	4 9	10 13	14 17	18 23	24 27	28 31

Todos os campos do registro são números inteiros e estão em "decimal fixed".

Descrição dos campos:

- NUMT - número total de vezes que o primitivo foi executado;
- CTOT - tempo total de CPU gasto pelo primitivo em todas as corridas apuradas;
- CMIN - o menor tempo de CPU em que o primitivo foi executado;
- CMAX - o maior tempo de CPU em que o primitivo foi executado;
- ETOT - tempo de execução total gasto pelo primitivo em todas as corridas apuradas;
- EMIN - o menor tempo de execução do primitivo;
- EMAX - o maior tempo de execução do primitivo.

Registro de número de opens: Este registro contém o número de vezes que os arquivos permanentes de Sistema de Banco de Dados foram abertos, considerando os diversos métodos de acesso. Existe no ARQSAE2 apenas um (1) registro deste tipo, está localizado depois dos registros de primitivo, e tem o seguinte lay-out:

TABOPEN (I,J)

Este registro é formado de um array de 12 elementos com 2 níveis. A identificação de cada campo é feita pela variável TABOPEN (I,J).

Descrição dos campos:

TABOPEN(I,J) - indica o número de vezes que um arquivo foi aberto com determinado método de acesso.

- I - indica a natureza do arquivo, a saber:

1 - dados;		
2 - elos;		

3 - inversões;

4 - estruturas.

J - indica o método de acesso, a saber:

1 - sequencial update;

2 - sequencial output;

3 - direct update.

Os elementos do array estão em decimal fixed com tamanho de 3 bytes cada elemento.

3.5.1.3 - ARQSAE3 - Arquivo de Alocação de Espaço

Este arquivo tem como finalidade fornecer os elementos necessários a apuração de área suficiente para alocar os data-sets.

É composto por 2 tipos de registros:

- mestre de natureza;

- informações sobre o segmento.

Os registros aparecem na seguinte ordem: cada mestre de natureza é seguido pelos diversos registros de informações sobre

os segmentos daquela natureza.

Registro mestre de natureza: Este registro contém o número de segmentos que vão constar do Banco de Dados, para determinada natureza e o tamanho do bloco do arquivo.

O registro está definido com 80 posições embora apenas (9) não sejam utilizadas. Tem o seguinte lay-out:

Identificação do campo	NAT	NARQ	TAMBL
Posição Início/Fim	1 1	2 4	5 9

Descrição dos campos:

NAT - código da natureza do arquivo a ser alocado. São válidos os seguintes códigos com os respectivos significados:

- D - arquivos de dados;
- E - arquivos de elos ;
- I - arquivos de inversões;
- S - arquivos de estruturas.

NARQ - número de segmentos que serão criados em um arquivo.

TAMBL - tamanho do bloco, para o arquivo.

Registro de informações sobre o segmento: Este tipo de registro contém informações sobre cada um dos segmentos a serem criados. O número de registros deste tipo, colocados a seguir de um registro mestre de natureza, deve ser igual ao número especificado no campo NARQ deste mestre de natureza.

Este registro também está definido com 20 posições, embora apenas 17 sejam utilizadas.

Tem o seguinte lay-out:

Identificação
do campo

Posição
Início/Fim

NATS	IDENTA	NREG	TAMRD
1 1	2 8	9 14	15 17

Descrição dos campos:

NATS - código da natureza do segmento a ser criado. São válidos os seguintes códigos com os respectivos significados:

D - segmento de dados;

E - segmento de elos;

I - segmento de inversões;

S - segmento de estruturas.

IDENTA - identificação dada pelo usuário, ao segmento a ser criado.

NREG - número de registros previstos para o segmento.

TAMRD - tamanho do registro, só é necessário ser preenchido para registros de dados e representa o seu tamanho em bytes.

3.5.2 Arquivos de Impressão

3.5.2.1 - ARQIMP1 - Relatórios de Distribuições

Relatórios contendo as seguintes apurações de distribuição por faixa:

- Distribuição por tamanho do registro de dados, com os seguintes elementos:
 - a. faixa de tamanho do registro (comprimento do registro em bytes);
 - b. número de registros em determinada faixa;
 - c. tamanho médio dos registros da faixa;
 - d. percentual de registros nesta faixa.

- Distribuição por número de registros em cada arquivo de dados, com os seguintes elementos:
 - a. faixa de número de registros por segmento;
 - b. número de segmentos em determinada faixa;
 - c. número médio de registros por segmento da faixa;
 - d. percentual do número de segmentos nesta faixa.

- Distribuição por tamanho de campos dos registros de dados, com os seguintes elementos:
 - a. faixa de tamanho do campo dos registros de dados;
 - b. número de campos em determinada faixa;
 - c. tamanho médio dos campos da faixa;
 - d. percentual do número de campos na faixa.

- Distribuição por número de registros de dados em cada classe de inversão, com os seguintes elementos:
 - a. faixa de número de registros por classe de inversão;
 - b. número de classes de inversão em determinada faixa;
 - c. número médio de registros na classe de inversão;
 - d. percentual de classes em cada faixa.

- Distribuição por tamanho de chave de inversão, com os seguintes elementos:
 - a. faixa de tamanho de chave de inversão;
 - b. número de segmentos com tamanho de chave de inversão em determinada faixa;
 - c. tamanho médio da chave de inversão na faixa;
 - d. percentual do número de segmentos na faixa.

3.5.2.2 - ARQIMP2 - Relatórios de Frequência

Relatórios contendo as seguintes apurações de frequência:

- Frequência e tempo de uso por primitivo com os seguintes elementos:
 - a. Primitivo - código do primitivo, indicando a finalidade com que o primitivo foi utilizado, nos casos de primitivos com mais de uma finalidade;
 - b. Frequência de utilização - número de vezes que o primitivo foi executado. São gravados os primitivos utilizados;
 - c. Tempo total de CPU - utilizado pelo primitivo, em todas as corridas apuradas;
 - d. Tempo máximo de CPU - o maior tempo de CPU utilizado pelo primitivo;
 - e. Tempo mínimo de CPU - o menor tempo de CPU utilizado pelo primitivo;
 - f. Tempo médio de CPU - tempo total de CPU/freq.de util.
 - g. Tempo total de execução - utilizado pelo primitivo em todas as corridas apuradas;
 - h. Tempo máximo de execução + o maior tempo em que o primitivo foi executado;

- i. Tempo mínimo de execução - o menor tempo em que o primitivo foi executado;
- j. Tempo médio de execução - tempo total de execução /frequência de utilização.

- Array de frequência de opens; onde as linhas identificam o método de acesso em que o arquivo foi aberto e as colunas identificam o arquivo.

3.5.2.3 - ARQIMP3 - Relatórios de Alocação de Espaço

Relatórios onde são fornecidas informações sobre a área necesária a cada arquivo; temos um relatório para cada arquivo do Sistema de Banco de Dados, contendo as seguintes informações:

- nome do arquivo, onde temos o nome do segmento;
- número de trilhas necessárias ao segmento;
- ao final temos o número de cilindros necessários para alocar todos os segmentos de 1 (um) arquivo.

3.6 Programas

As identificações dos programas e dos arquivos, representa efetivamente o nome dessas variáveis nos programas do Sistema de Avaliação de Eficiência.

O Sistema de Avaliação de Eficiência é composto de 3 (três) programas a saber:

3.6.1 - PROG1 - Programa de Distribuições.

3.6.1.1 Objetivo: Destina-se a apurar distribuições com referência as características dos dados armazenados, envolvendo distribuições com referência a campos, registros e segmentos do arquivo de dados; número de registros de dados por classe de inversão; fornece informações sobre o equilíbrio dos valores do sistema com referência ao tamanho das chaves de inversão.

As apurações feitas por este programa atuam sobre os arquivos do Sistema de Banco de Dados, e podem ser solicitadas a qualquer momento pelo usuário.

3.6.1.2 - Arquivos de Entrada:

- ARQSAE1; onde o usuário especifica as distribuições que devem ser apuradas e quais as faixas de apuração;
- ARQD; arquivo de dados do sistema de Banco de Dados, necessário para fazer as apurações referentes ao tamanho dos registros de dados e do número de registros de dados por arquivos de dados.
- ARQS; arquivo de estruturas do Sistema de Banco de Dados, necessário para fazer a distribuição do tamanho dos campos dos registros de dados.

- ARQI; arquivo de inversões de Sistema de Banco de Dados necessário para fazer as distribuições referentes a número de registros de dados por classe de inversão e tamanho da chave de inversão.

3.6.1.3 - Arquivo de Saída:

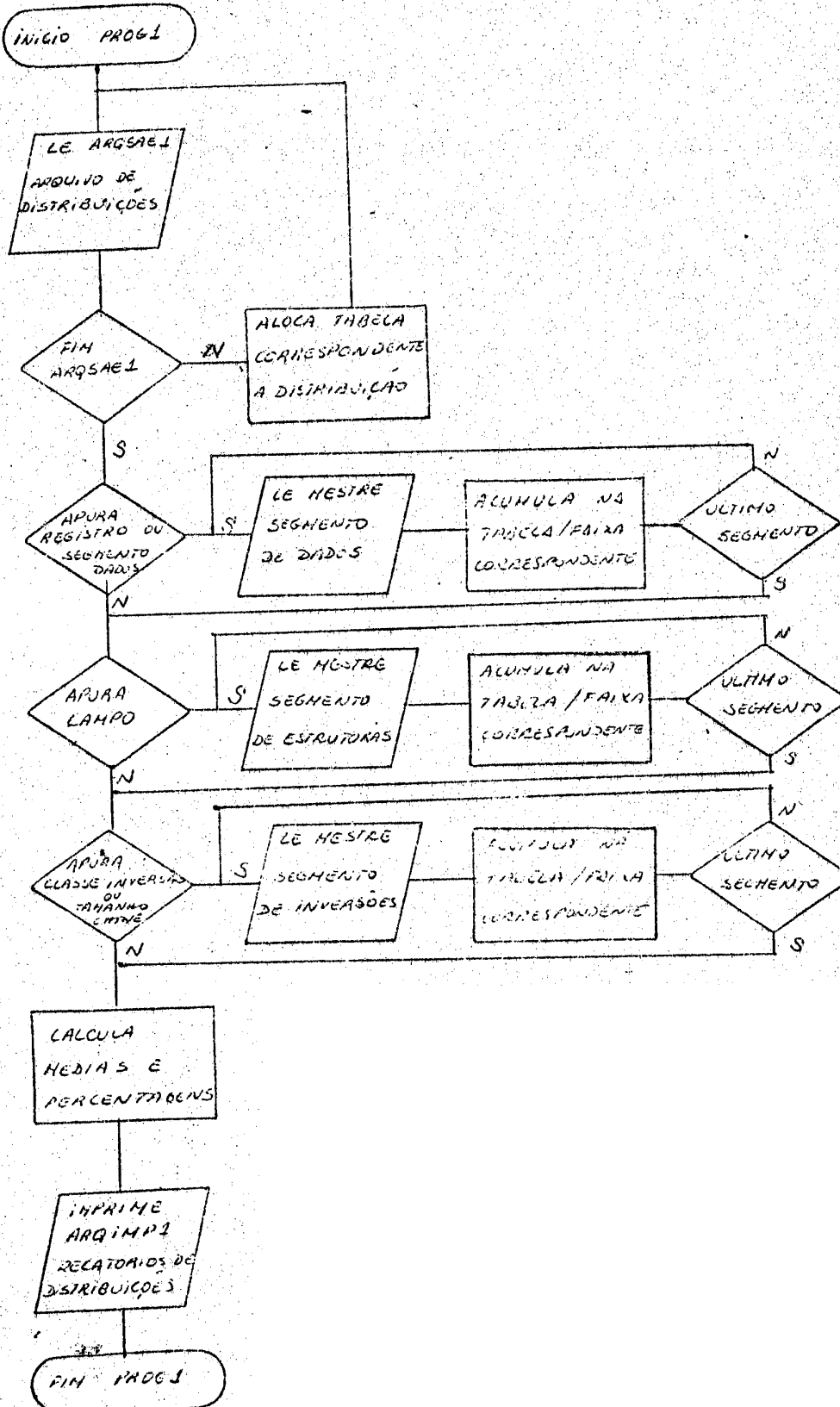
- ARQIMP1 ; relatório contendo as apurações das distribuições por faixa.

3.6.1.4 - Processamento:

- a. le o arquivo de distribuição - ARQSAE1 e para cada distribuição solicitada aloca tabela própria, onde serão armazenadas informações que permitirão fazer a apuração solicitada.
- b. Se tiver sido solicitada apuração referente a tamanho de registro ou número de registros por segmento, percorre os mestres de segmentos dos arquivos de dados, acumulando as informações necessárias a distribuição.
- c. Se for solicitada apuração referente a tamanho de campo, percorre os mestres de segmento do arquivo de estruturas, acumulando informações necessárias a distribuição.
- d. Se tiver sido solicitada apuração referente ao número de registros de dados por classe de inversão ou do tamanho da chave de inversão, percorre os mestres de segmento de arquivo de inversões, acumulando as informações necessárias a distribuição.
- e. As acumulações mencionadas nos itens b, c e d, são feitas em tabela própria da distribuição e na faixa específica.

- f. Para cada uma das tabelas em memória, faz os cálculos referentes a média e percentagem e imprime, as distribuições.

3.6.1.5 - Fluxograma:



3.6.2 - PROG2 - Programa de Frequências

3.6.2.1 - Objetivo:

Destina-se a apurar as características de utilização do Sistema de Banco de Dados, envolvendo a frequência de uso de cada primitivo, associando a esta apuração a observação de desempenho de cada primitivo, obtida através de uma análise dos tempos gastos na execução do primitivo. Para esta análise é apropriado o "CPU time" e o "elapsed time". Apura também valores do equilíbrio da eficiência do sistema no que se refere a frequência de abertura de arquivos.

As apurações neste programa necessitam de informações obtidas durante a execução do Sistema de Banco de Dados.

3.6.2.2 - Arquivo de Entrada:

- ARQSAE2; contendo informações sobre o tempo de CPU e de "elapsed time" dos primitivos e a sua frequência de utilização bem como do número de vezes que os arquivos permanentes foram abertos em cada método de acesso. Estas informações são de correntes de utilizações do Sistema de Banco de Dados, com a opção de avaliação de eficiência ativada.

3.6.2.3 - Arquivo de Saída:

- ARQIMP2; contendo os relatórios de frequência.

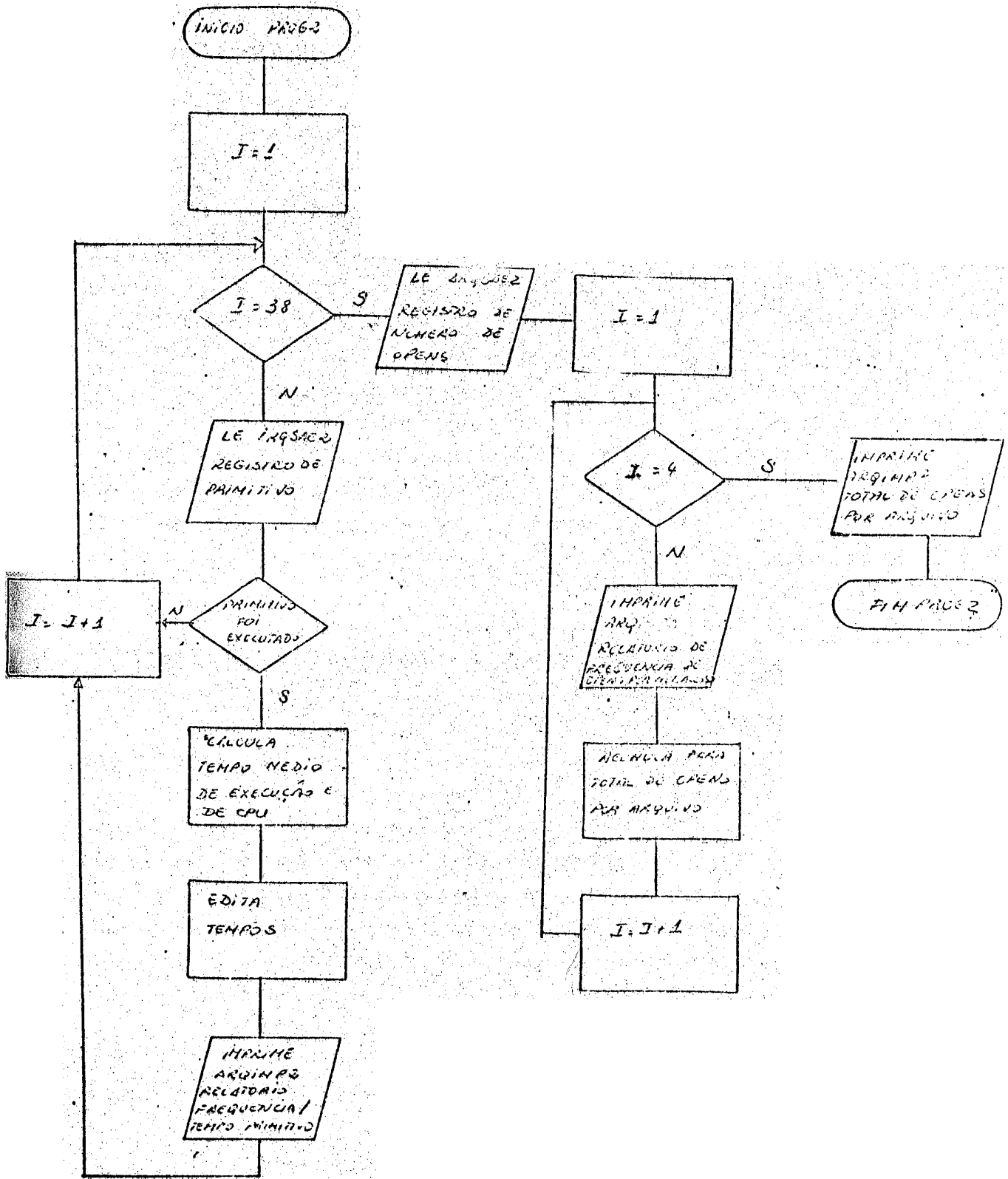
3.6.2.4 Processamento:

a. Lê os registros de primitivos do arquivo de frequência e para cada registro de primitivo temos os seguintes procedimentos:

- verifica se o primitivo foi executado uma vez. (Estamos chamando de primitivo ao conjunto primitivo/ finalidade) e caso ele tenha sido, calcula o tempo médio de CPU e de execução e edita os tempos total, máximo, mínimo e médio, transformando o tempo armazenado em milésimos de segundo em horas, minutos, segundos e centésimos de segundo. Imprime a frequência e os tempos do primitivo.

b. Lê o registro de número de opens e imprime array contendo frequência de opens por método de acesso e arquivo e imprime número total de opens por arquivo do Sistema de Banco de Dados.

3.6.2.5 - Fluxograma:



3.6.3 - PROG3 - Programa de Alocação de Espaço

3.6.3.1 Objetivo:

Destina-se a fazer uma previsão da área a ser alocada para os arquivos do Sistema de Banco de Dados.

Este programa visa orientar o usuário no dimensionamento da área a ser ocupada pelos seus arquivos e pode ser executado pelo usuário, antes mesmo de ter sido criado o Banco de Dados. O programa também pode ser utilizado depois do Banco de Dados existir e neste caso faz uma previsão da área alocada pelos seus arquivos, considerando-os rearrumados (após executar o primitivo RESTDB), o usuário pode também usar o programa para se saber para incluir determinados segmentos no sistema é necessário redimensioná-lo.

Os cálculos foram feitos considerando-se o equipamento da PUC/RJ, discos 3330, e baseados em fórmula constante do "Reference Manual for IBM3830, Storage Control Model 1 and IBM3330 Disk Storage".

Resaltamos que o programa fornece uma previsão de área a ser alocada e não a área mínima a ser alocada; pois a utilização da organização sequencial indexado faz com que a informação sobre a área real utilizada não seja disponível. Além disso devido a características do Sistema de Banco de Dados, esta área está dimensionada pelo máximo; assim sendo, cada registro de elo está multiplicado por 4 (quatro) considerando que ele pode dar origem a um registro de mestre de classe de elos, a um registro inverso e a um registro mestre de classe de elos correspondente ao registro inverso. Cada registro de inversão está multiplicado por 2 (dois), considerando que ele pode dar origem a um mestre de classe de inversão.

3.6.3.2- Arquivo de entrada:

- ARQSAE3; onde o usuário fornece informações sobre os arquivos do Sistema de Banco de Dados.

3.6.3.3- Arquivo de Saída:

- ARQIMP3; onde são fornecidas previsões de áreas para cada segmento do Sistema de Banco de Dados.

3.6.3.4- Processamento:

- a. Lê registro de mestre de natureza e estabelece constantes dependendo da natureza.
- b. Lê registro de informação sobre o segmento.
- c. Calcula:
 - número de blocos por trilha
$$13165 / (191 + \text{tamanho do bloco})$$

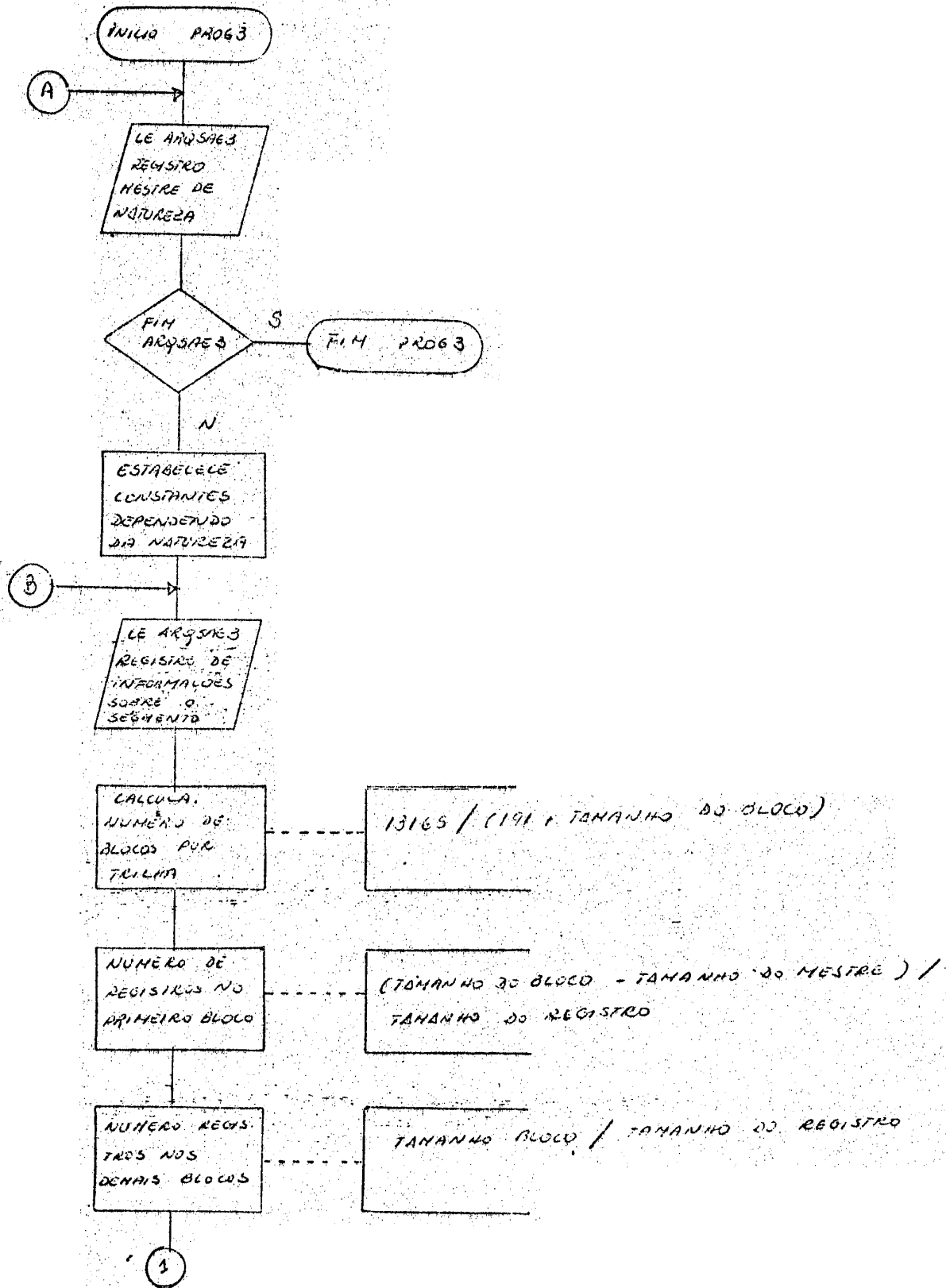
esta fórmula foi obtida no manual "Reference Manual for IBM 3830, Storage Control Model 1 and IBM 3330 Disk Storage".
 - número de registros no primeiro bloco
$$(\text{tamanho do bloco} - \text{tamanho do mestre de segmento}) / \text{tamanho do registro}.$$
 - número de registros nos demais blocos
$$\text{tamanho do bloco} / \text{tamanho do registro}.$$
 - número de registros máximo; se eles multiplicar número de registros por 4; se inversões multiplica por 2.
 - número de registros em todos os blocos exceto o primeiro
$$\text{número de registros máximo} - \text{número de registros no primeiro bloco}.$$

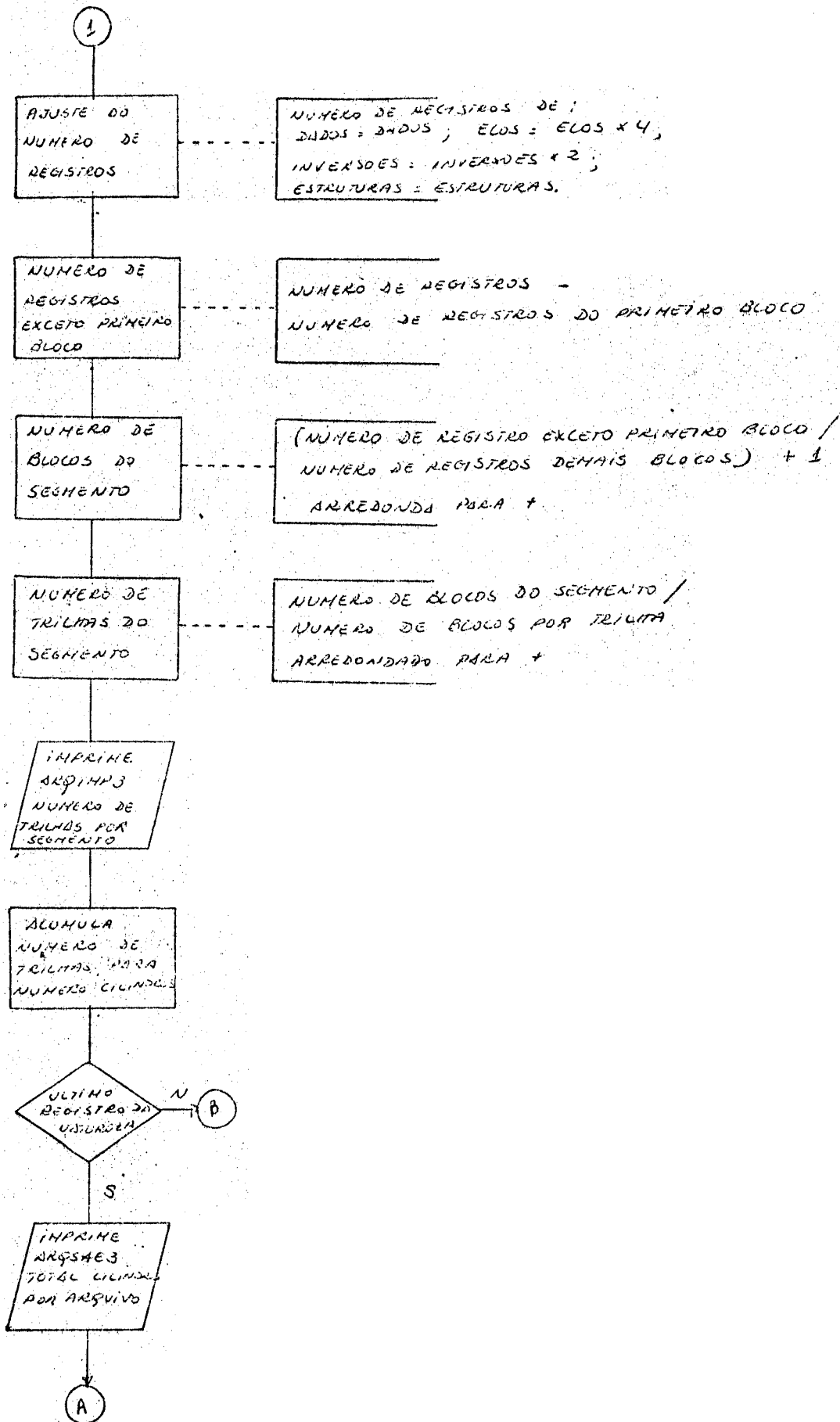
- número de blocos do segmento
(número de registros exceto o 1º bloco/nº de registros nos demais blocos) + 1
arredonda para +
- número de trilhas do segmento
número de blocos do segmento/número de blocos por trilha
arredonda para +

- d. imprime número de trilhas por segmento.
- e. acumula número de trilhas do segmento para cálculo do número de cilindros do arquivo.
- f. se este registro for o último desta natureza, imprime número de cilindros para o arquivo.
número total de trilhas/19
arredonda para +
vai para a.
- g. se não vai para b.

OBS.: Arredondar para + significa considerar o primeiro número inteiro imediatamente superior, se o número obtido tiver partes fracionárias.

3.6.3.5 - Fluxograma:





3.7 - SUBROTINAS

Subrotinas são procedimentos que embora sejam ativados pelos primitivos do Sistema de Banco de Dados, não fazem parte de le, mas sim do Sistema de Avaliação de Eficiência.

Elas são necessárias para obter as informações que são usadas no PROG2 - Programas de Frequência, que necessita de informações sobre a execução dos primitivos do sistema.

As subrotinas são executadas se determinado parâmetro do primitivo ATIVDB indicar que o usuário deseja guardar informações para a avaliação de eficiência.

Devido a sua maneira de funcionar, as subrotinas necessitam passar informações de uma para outra, assim sendo, foi necesário criar variáveis externas, no Sistema de Banco de Dados, mas que são utilizadas somente para avaliação da eficiência.

No Quadro 2, mostramos a utilização das variáveis externas pelas subrotinas.

3.7.1 - Variáveis externas do Sistema de Banco de Dados, utilizadas para avaliação de eficiência:

A identificação usada para as variáveis externas representa efetivamente o nome dessas variáveis nas subrotinas do Sistema de Avaliação de Eficiência e nos primitivos do Sistema de Banco de Dados. São as seguintes:

3.7.1.1 - TABTIME - array de 37 elementos contendo informações sobre o número de vezes que cada primitivo/indicador foi executado e o tempo gasto; chamaremos este array de "array de tempo por primitivo"; é composto dos seguintes elementos:

- NUNT	- número de vezes que o primitivo/finali <u>dade</u> foi executado;
- CTOT	- tempo total de CPU para o primitivo/finali <u>dade</u> ;
- CMIN	- menor tempo de CPU do primitivo/finali <u>dade</u> ;
- CMAX	- maior tempo de CPU do primitivo/finali <u>dade</u> ;
- ETOT	- tempo de execução total do primitivo/finali <u>dade</u> ;
- EMIN	- menor tempo de execução do primitivo/finali <u>dade</u> ;
- EMAX	- maior tempo de execução do primitivo/finali <u>dade</u> .

A tabela tem um elemento para cada primitivo/finalidade, ou seja, os primitivos que tem mais de uma finalidade, a cada finalidade é associada um elemento da tabela e identificamos estes elementos (diferentes de um mesmo primitivo) pelo termo primitivo/finalidade). Chamamos de tempo de execução o "elapsed time".

3.7.1.2 - TABOPEN - array com dimensão (4,3) contendo número de opens por arquivo e método de acesso; chamaremos este array de "array de opens"

3.7.1.3 - TIMEI - guarda a hora de início de execução de um primitivo/finalidade, na forma de horas, minutos, segundos e milésimos de segundos.

3.7.2 - A seguir relacionamos as subrotinas do Sistema de Avaliação de Eficiência; as identificações utilizadas para subrotinas, arquivos e parâmetros representam efetivamente o nome dessas variáveis no Sistema.

3.7.2.1 - SUB1 - Inicializa ou carrega em memória os arrays de tempo e de opens.

3.7.2.1.1 - Objetivo:

Inicializar os arrays de tempo e de opens ou carregar em memória estes arrays, a partir de informações apropriadas durante a execução dos primitivos.

3.7.2.1.2 - Ativação:

É ativada pelo primitivo ATIVDB.

3.7.2.1.3 - Parâmetro de entrada:

- S1 - PARM - indica a opção de manuseio dos arrays:
 - 1 - inicializa arrays com zeros.
 - 2 - carrega arrays, existentes na forma de arquivo, na memória.

3.7.2.1.4 - Arquivo de entrada:

- ARQSAE2 - Arquivo de Frequência, este arquivo é utilizado quando S1 - PARM = 2.

3.7.2.1.5 - Processamento:

a - testa o parâmetro S1-PARM que indica o tipo de procedimento a ser executado pela Subrotina:

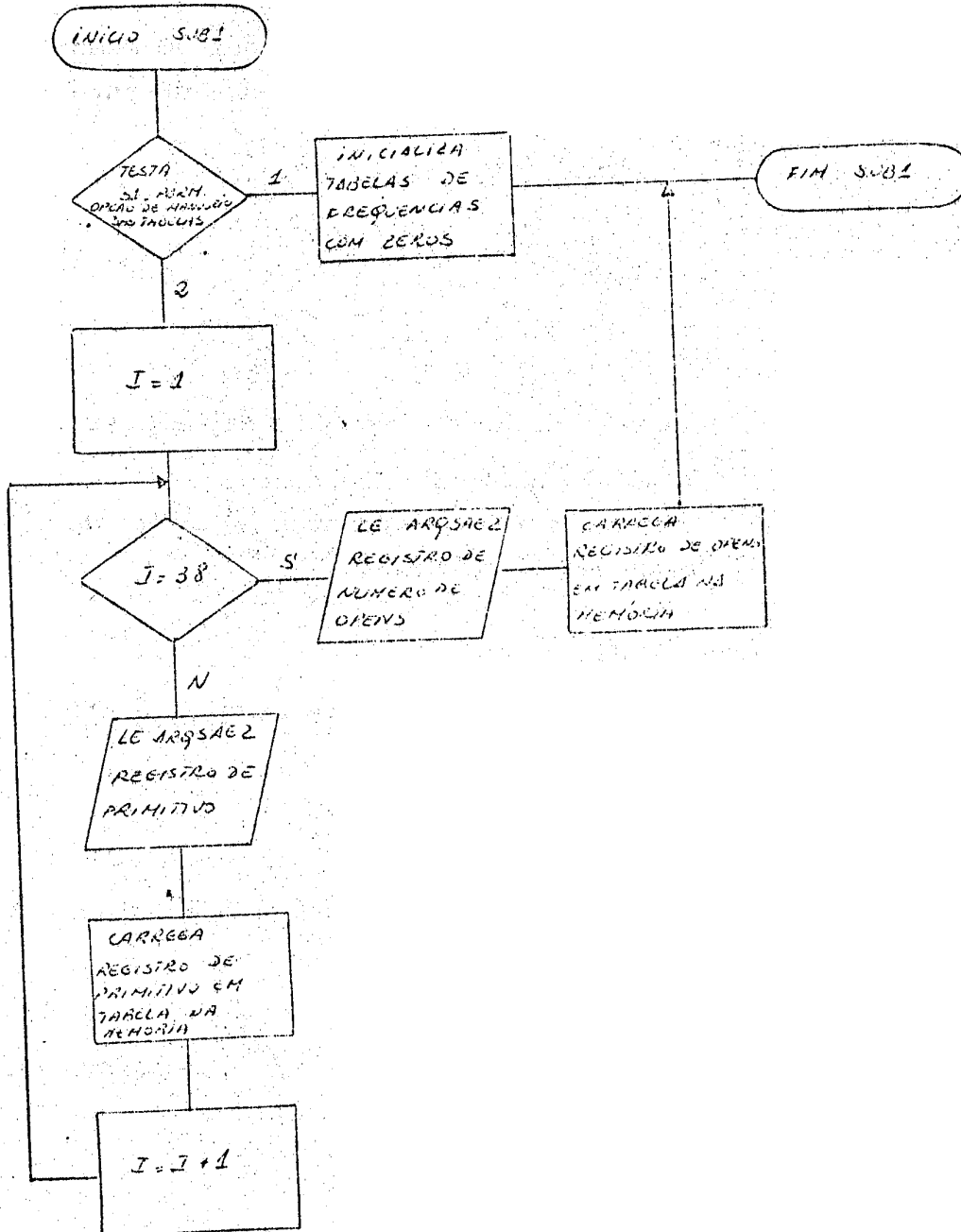
a₁ - S1-PARM = 1

Neste caso, não existe arquivo ARQSAE2, os arrays de tempo e opens devem ser inicializados com zero.

a₂ - S2-PARM = 2

Neste caso, já existe arquivo ARQSAE2, cujos registros são imagens dos arrays de tempo e opens; os registros são lidos e carregados em memória.

3.7.2.1.6 - Fluxograma:



3.7.2.2 - SUB2 - Hora de início do primitivo

3.7.2.2.1 - Objetivo:

Guardar a hora em que o primitivo começou a ser executado, para se obter o tempo de execução e ativar SUB21 para se obter tempo de CPU.

Esta subrotina não tem parâmetros nem arquivos.

3.7.2.2.2 - Ativação:

É ativada por todos os primitivos e precede todos os procedimentos referentes ao primitivo especificamente (procedimentos do Sistema de Banco de Dados).

3.7.2.2.3 - Processamento:

- guarda a hora na variável externa TIMEI;
- ativa SUB21.

3.7.2.2.4 - SUB21 - Ativa contagem de tempo de CPU.

3.7.2.2.4.1 - Objetivo:

Ativar a contagem do tempo de CPU através de uma subrotina em Assembler.

3.7.2.2.4.2 - Ativação:

É ativada pela subrotina SUB2.

3.7.2.2.4.3 - Processamento:

- Ativa STIMER iniciando contagem de tempo de CPU.

3.7.2.3 - SUB3 - Atualiza array de tempo

3.7.2.3.1 - Objetivo:

Atualizar o array de tempo, com base no tempo de CPU e de execução de um primitivo/finalidade.

3.7.2.3.2 - Ativação:

É ativada por todos os primitivos e segue o último procedimento referente ao primitivo especificamente.

3.7.2.3.3 - Parâmetro de entrada:

S3 - PARM - indica o primitivo/finalidade em execução.

3.7.2.3.4 - Processamento:

- apropria tempo de CPU do primitivo/finalidade através da ativação de SUB31;
- calcula tempo de execução do primitivo/finalidade;
- atualiza os elementos do array de tempo:
 - incrementa de 1 o número de vezes que o primitivo/ finalidade foi executado;
 - acumula nas variáveis de tempo total de CPU e de execução os tempos gastos;
 - verifica e atualiza se for o caso, se esta foi a execução mais rápida ou a mais longa deste primitivo/finalidade.

3.7.2.3.5 - SUB31 - Apropria tempo de CPU.

3.7.2.3.5.1 - Objetivo

Apropria o tempo de CPU através de uma subrotina em Assembler.

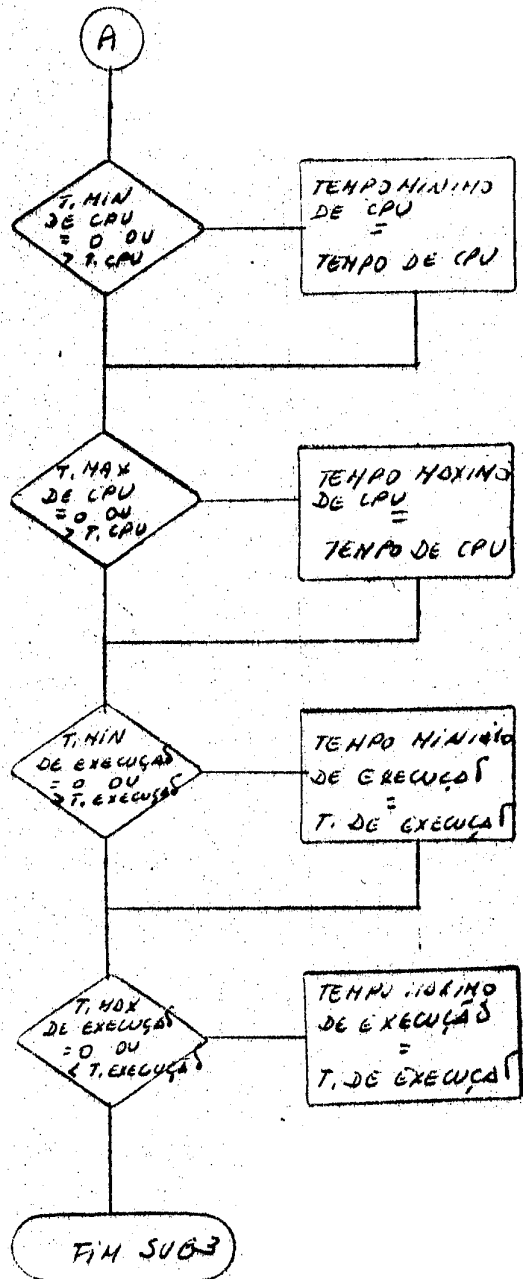
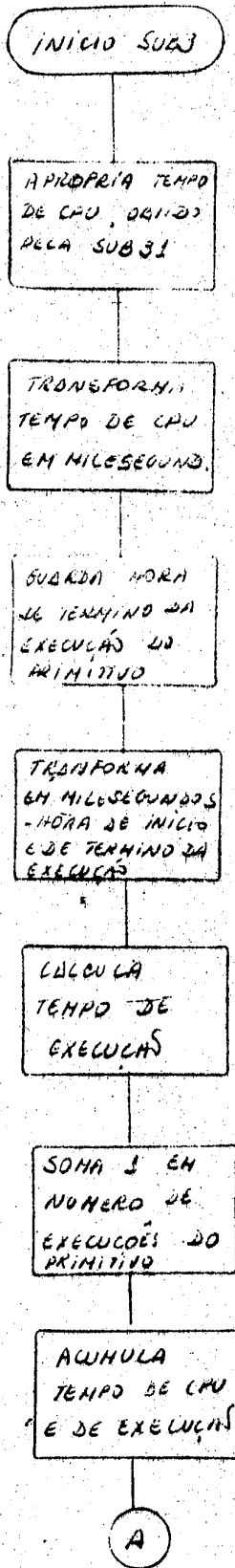
3.7.2.3.5.2 - Ativação:

É ativada pela subrotina SUB3.

3.7.2.3.5.3 - Processamento:

- executa TTIMER, apropriando o tempo de CPU.

3.7.2.3.5 - Fluxograma:



3.7.2.4 - SUB4 - Atualiza array de opens

3.7.2.4.1 - Objetivo:

Atualizar array de opens, sempre que um arquivo for aberto.

3.7.2.4.2 - Ativação:

É ativada pela subrotina auxiliar AUX3 do Sistema de Banco de Dados.

3.7.2.4.3 - Parâmetros de entrada:

S4 - NAT - identifica a natureza do arquivo aberto (se dados, elos, inversões ou estruturas).

S4 - MA - identifica o método de acesso com que o arquivo foi aberto (sequencial update, sequencial output, direct update)

3.7.2.4.4 - Processamento:

- Incrementa de um o elemento do array de opens correspondente ao arquivo e ao método de acesso com que ele foi aberto.

3.7.2.5 - SUB 5 - Grava array de tempo de opens

3.7.2.5.1 - Objetivo:

Gravar no arquivo ARQSAE2 os arrays de tempo e de opens.

3.7.2.5.2 Ativação:

É ativada pelo primitivo DATVDB.

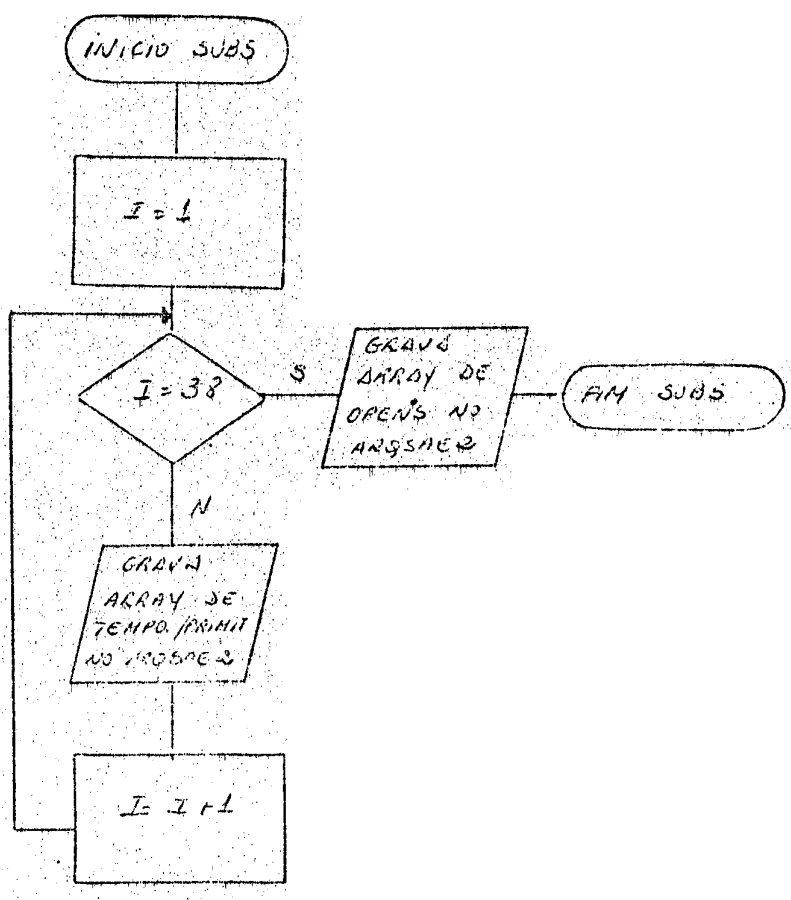
3.7.2.5.3 - Arquivo de saída:

- AROSAE2 - Arquivo de Frequência

3.7.2.5.4 - Processamento:

- para cada elemento do array de tempo, grava um registro de primitivos no arquivo de frequência.
- grava o array de opens no arquivo de frequência.

3.7.2.5.5 - Fluxograma:



Quadro 1 - Tabela de utilização de arquivos/Programas e Subrotinas

	PROG1	PROG2	PROG3	SUB1	SUB5
ARQD	X				
ARQI	X				
ARQS	X				
ARQSAE1	X				
ARQSAE2		X		X	X
ARQSAE3			X		
ARQIMP1	X				
ARQIMP2		X			
ARQIMP3			X		

Quadro 2: Tabela de utilização de variáveis externas/Subrotinas

Subrotinas variável ext.	SUB1	SUB2	SUB3	SUB4	SUB5
TABTIME	X		X		X
TABOPEN	X			X	X
TIMEI		X	X		

4. MANUAL DE UTILIZAÇÃO DO SISTEMA DE AVALIAÇÃO DE EFICIÊNCIA

Neste capítulo, vamos falar de como o usuário deve agir para se utilizar do Sistema de Avaliação de Eficiência, tanto no que se refere a executar os programas quanto a forma de analisar os resultados.

O manual de utilização, apresenta cada um dos programas do Sistema de Avaliação de Eficiência em separado, já que não existe nenhuma interdependência entre eles.

4.1 - PROG1 - Programa de Distribuição

4.1.1 - Pré-requisitos para utilização:

- o Sistema de Banco de Dados já deve existir e estar disponível para consulta.

4.1.2 - Arquivo de Entrada:

- ARQD, ARQS e ARQI do Sistema de Banco de Dados.

Os cartões DD referente a estes arquivos são semelhantes aos utilizados no Sistema de Banco de Dados;

- ARQSAE1 - arquivo de distribuição, descrito no item 3.5.1.1.

Este arquivo conterà as distribuições a serem apuradas e cabe ao usuário formatá-lo. Cada registro corresponde a uma apuração. Vamos supor para efeito dos exemplos que este arquivo está em cartão.

O cartão DD referente a este arquivo tem a forma:

```
//ARQSAE1 DD *
```

A seguir deste cartão estão os cartões de apuração perfurados pelo usuário.

4.1.3 Arquivo de Saída:

- ARQIMP1 - Relatórios de distribuição, descritos no item 3.5.2.1. São impressos 5 tipos de relatório, um para cada tipo de apuração.

O cartão DD referente a este arquivo:

```
//ARQIMP1 DD SYSOUT= A
```

4.1.4 - Exemplos:

A seguir daremos um exemplo da entrada e saída para cada apuração:

4.1.4.1 - Exemplo de apuração 1: Tamanho dos registros de dados.

Suponhamos que o usuário deseje obter uma distribuição de tamanho de registro, com as seguintes características:

- menor valor da distribuição - 10
- maior valor da distribuição - 99
- tamanho da faixa de distribuição - 10

Para solicitar esta apuração o usuário deve perfurar um cartão com as seguintes informações, começando na primeira coluna:

```
DTR000010000099000010
└──────────────────┘
CODLINFE LSUPE INCE
```

O relatório referente a esta apuração tem o seguinte formato:

DISTRIBUIÇÃO POR TAMANHO DE REGISTROS DE DADOS

TAMANHO DO REGISTRO NUMERO DE REGISTRO TAMANHO MEDIO PERCENTUAL

MENOR QUE 10		N1	M1	P1
10 a 19		N2	M2	P2
20 a 29		N3	M3	P3
30 a 39		N4	M4	P4
40 a 49		N5	M5	P5
MAIOR QUE 49		N6	M6	P6
TOTAL		N	M	100

4.1.4.2 - Exemplo de apuração 2: Número de registros de dados por segmento.

Suponhamos que o usuário deseje obter uma distribuição do número de registros de dados por segmento, com as seguintes características:

menor valor da distribuição - 100
maior valor da distribuição - 299
incremento da distribuição - 50

Deve perfurar um cartão contendo, a partir da coluna 1:

DNR000100000299000050
COD LINFE LSUPE INCE

O relatório referente a esta apuração tem o seguinte formato:

DISTRIBUIÇÃO POR NÚMERO DE REGISTROS EM CADA ARQUIVO DE DADOS				
NUMERO DE REGISTROS		NUMERO DE SEGMENTOS	NUMERO MEDIO	PERCENTUAL
MENOR QUE 100		N1	M1	P1
100 a 149		N2	M2	P2
150 a 199		N3	M3	P3
200 a 249		N4	M4	P4
249 a 299		N5	M5	P5
MAIOR QUE 199		N6	M6	P6
TOTAL		N	M	100

4.1.4.3 - Exemplo de apuração 3:

Tamanho dos campos dos registros de dados:

Vamos supor que o usuário deseje obter uma distribuição do tamanho dos campos dos registros de dados, com as seguintes características:

menor valor da distribuição - 5

maior valor da distribuição - 29

incremento da distribuição - 5

Deve perfurar um cartão, contendo a partir da primeira coluna o seguinte:

DTC000005000029000005
COD LINFE LSUPE INCE

O relatório referente a esta apuração tem o seguinte formato:

DISTRIBUIÇÃO POR TAMANHO DOS REGISTROS DE DADOS

TAMANHO DO CAMPO	NUMERO DE CAMPOS	TAMANHO MEDIO	PERCENTUAL
MENOR QUE 5	N1	M1	P1
5 a 9	N2	M2	P2
10 a 14	N3	M3	P3
15 a 19	N4	M4	P4
20 a 24	N5	M5	P5
25 a 29	N6	M6	P6
MAIOR QUE 29	N7	M7	P7
TOTAL	N	M	100

4.1.4.4 - Exemplo da apuração 4:

Número de registros de dados por classe de inversão:

Suponhamos que o usuário queira obter uma distribuição do número de registros de dados por classe de inversão, com as seguintes características:

menor valor da distribuição - 10

maior valor da distribuição - 39

incremento da distribuição - 10

Deve perfurar um cartão contendo as seguintes informações, a partir da coluna 1:

IRC000010000039000010

COD LINFE LSUPE INCE

O relatório referente a esta apuração tem o seguinte formato:

DISTRIBUIÇÃO POR NÚMERO DE REGISTROS DE DADOS EM CADA CLASSE DE INVERSÃO

NUMERO DE REGISTROS	NUMERO DE CLASSES	NUMERO MEDIO	PERCENTUAL
MENOR QUE 10	N1	M1	P1
10 a 19	N2	M2	P2
20 a 29	N3	M3	P3
30 a 39	N4	M4	P4
MAIOR QUE 39	N5	M5	P5
TOTAL	N	M	100

4.1.4.5 - Exemplo da apuração 5:

Tamanho da chave de inversão.

Vamos supor que o usuário deseje uma distribuição do tamanho da chave de inversão, com as seguintes características:

menor valor da distribuição - 10

maior valor da distribuição - 209

incremento da distribuição - 50

Deve perfurar num cartão, começando na coluna 1, o seguinte:

ITC000010000209000050

COD LINFE LSUPE INCE

O relatório referente a esta apuração tem o seguinte formato:

DISTRIBUIÇÃO POR TAMANHO DE CHAVE DE INVERSÃO

TAMANHO DA CHAVE	NUMERO DE SEGMENTOS	TAMANHO MEDIO	PERCENTUAL
MENOR QUE 10	N1	M1	P1
10 a 59	N2	M2	P2
60 a 109	N3	M3	P3
110 a 159	N4	M4	P4
159 a 209	N5	M5	P5
MAIOR QUE 209	N6	M6	P6
TOTAL	N	M	100

4.1.5 - Análise dos relatórios

Vamos fazer uma análise de cada um dos relatórios, visando fornecer ao usuário elementos que permitam levá-lo a concluir que tipos de modificações devem ser feitas no seu Sistema de Banco de Dados, com a finalidade de torná-lo mais eficiente para as suas aplicações.

4.1.5.1 - Análise de apuração 1

Este relatório nos dá a distribuição dos tamanhos dos registros nos segmentos de dados; a medida que o usuário verifica que através do estudo deste relatório, que existe uma grande concentração de tamanho de registros numa mesma faixa, o mais eficiente será ele passar a trabalhar no seu Sistema de Banco de Dados com registros de tamanho fixo no arquivo de dados, ao invés de tamanho variável, pois isso economizará espaço (4 bytes por registro e 4 bytes por bloco) e também tempo de processamento.

4.1.5.2 - Análise da Apuração 2

Este relatório nos dá o comportamento do número de registros dados por segmentos.

Se existe um arquivo pequeno muito acessado, deve ser colocado em memória.

Se tem arquivos grandes e voláteis, o usuário deve rearrumá-los periodicamente.

4.1.5.3 - Análise da Apuração 3

Este relatório nos dá uma visão da distribuição do comprimento dos campos nos registros de dados. Especial atenção deve ser dada aos campos grandes. Caso campos grandes sejam usados em inversões, o usuário deve verificar se podem ser codificados, para tornar a inversão mais rápida. No caso de campos numéricos grandes, deve ser verificada a possibilidade da compactação.

4.1.5.4 - Análise da Apuração 4

Este relatório deve ser observado com bastante cuidado, pois com base nele podemos obter informações úteis para fazer modificações no "lay-out" do arquivo de inversões. O arquivo de inversões tem registros de tamanho fixo com 336 bytes sendo que 256 bytes são ocupados por chaves e pelo carácter de controle do registro; e os outros 80 bytes para armazenar os identificadores referentes a até 20 registros de dados. Sendo assim, a medida que exista um grande número de classes com um número médio de registros por classe maior que 20, o usuário deve estudar o seu arquivo de inversão para saber o quanto deve aumentar o registro de inversões, de modo a diminuir, o número de registros por classe de inversão. Pois a cada registro de inversão que ele diminuir, estará na realidade eco

nomizando 256 bytes. Mas também não vale a pena fazer o registro de inversão com um tamanho tal que cada classe responderia a uma inversão, pois neste caso, as classes com um número pequeno de inversões, ficariam com uma grande quantidade de identificadores de registro de dados, desocupados.

A medida que o usuário chegue a conclusão que o seu registro de inversões deve ser bastante aumentado, ele deve pensar em transformar o arquivo em variável, pois a classe de inversão são necessários apenas 255 bytes (sendo que 249 bytes para o campo de inversão).

4.1.5.5 - Análise da apuração 5

O Sistema de Banco de Dados devido as suas características de ser um sistema geral e não voltado a um usuário específico, tem como chave, a maior permitida pelo sistema operacional, 255 bytes, sendo que 6 bytes são utilizados pelo sistema e 249 pelo campo de inversão. Com base nesta apuração o usuário deve estudar a possibilidade de diminuir o tamanho desta chave. Suponha, por exemplo, que o maior campo a ser classificado pelo usuário seja de 100 caracteres, deve modificar imediatamente o lay-out do arquivo de inversões, pois estará economizando 149 bytes por registro. Se realmente o usuário fizer classificações por variáveis grandes, deve verificar a possibilidade de codificar estas variáveis e classificar os códigos, tornando desta forma o processamento mais rápido e economizando área no arquivo de inversão, pois se a chave passa a ser um código ela certamente poderá ser diminuída.

Para fazer um bom estudo das modificações a serem implementadas no arquivo de inversões, deve se estudar os resultados deste relatório, conjuntamente com os do relatório anterior.

4.2 - PROG2 - Programa de Frequência

4.2.1 - Pré-requisito para utilização

Este programa necessita de informações que são obtidas durante a execução do Sistema de Banco de Dados.

Estas informações são obtidas na forma de um arquivo-ARQSAE2. Para que este arquivo seja gravado é preciso que seja dada a indicação no primitivo ATIVDB do Sistema de Banco de Dados.

O ATIVDB tem 2 parâmetros, a saber:

Q1_IND - indica se a base de dados existe ou não;

Q1_EFIC- indica se vão ser ativados os procedimentos para avaliação de eficiência e ainda se o arquivo ARQSAE2 já existe, a saber:

0 - não ativa procedimentos de eficiência, ARQSAE2 não é considerado;

1 - ativa os procedimentos de eficiência e cria ARQSAE2. Este código é utilizado quando ARQSAE2 não existe mas vai ser criado;

2 - ativa procedimentos de eficiência e atualiza ARQSAE2. Este código é utilizado quando já existe ARQSAE2, e vai ser atualizado.

Como podemos observar, o ARQSAE2 é gerado pelo sistema de Banco de Dados. Identificamos também 3 situações no Sistema de Banco de Dados. Situações estas que são identificadas pelo parâmetro Q1_EFIC do primitivo ATIVDB.

- ARQSAE2 não existe,
- ARQSAE2 é gravado,
- ARQSAE2 é lido e gravado.

Ressaltamos que, dependendo da forma como o Sistema de Banco de Dados é ativado, o ARQSAE2 tem um tratamento diferente. Como o ARQSAE2 pode ser de entrada e saída durante a mesma execução, ele tem duas identificações para o sistema ope

racional(ddname), uma ARQSAEI quando funciona como entrada, outra ARQSAEO quando funciona como saída.

No quadro abaixo, mostramos como deve ser o cartão DD para o ARQSAE2, durante a execução do Sistema de Banco de Dados.

ARQSAE2 Q1-EFIC	ENTRADA ARQSAEI	SAIDA ARQSAEO
0	DUMMY	DUMMY
1	DUMMY	especificado
2	especificado	especificado

4.2.2 - Arquivo de entrada:

- ARQSAE2 - Arquivo de frequência, obtido conforme especificado no item anterior.

O cartão DD referente a este arquivo, no PROG2, tem o seguinte formato:

//ARQSAE2 DD (especificação compatível com item anterior),

4.2.3 - Arquivo de saída

- ARQIMP2 - Relatórios de frequência, descrito no item 3.5.2.2.

O cartão DD referente a este arquivo

//ARQIMP2 DD SYSOUT = A

São impressos dois tipos de relatórios;

O primeiro relatório referente a este arquivo tem os seguintes elementos:

FREQUENCIA E TEMPO DE USO POR PRIMITIVO EM HOR.MIN.SEG.CTS.

PRIMITIVO	FREQUÊNCIA	TEMPO	DE CPU	-	(CPU TIME)
	DE	TOTAL	MÁXIMO	MÍNIMO	MÉDIO
	UTILIZAÇÃO				
			TEMPO DE EXECUÇÃO - (ELAPSED TIME)		
			TOTAL	MÁXIMO	MÍNIMO MÉDIO

O segundo relatório, tem o seguinte formato:

FREQUÊNCIA DE OPENS POR ARQUIVO

TIPO DE OPEN	DADOS	ELOS	INVERSÕES	ESTRUTURAS
SEQUENTIAL UPDATE	D1	E1	I1	S1
SEQUENTIAL OUTPUT	D2	E2	I2	S2
DIRECT UPDATE	D3	E3	I3	S3
TOTAL	D	E	I	S

4.2.4 - Análise das apurações

Podemos fazer as seguintes análises:

4.2.4.1 - Análise de Relatório de Frequência e Tempo de Uso por Primitivo.

Este relatório nos mostra o número de vezes que cada primitivo foi executado. Especial atenção deve ser dada aos primitivos com maior frequência de utilização. Os seus algoritmos devem ser otimizados. Se houver uma grande diferença entre o tempo de CPU máximo e o mínimo o usuário deve localizar os procedimentos que estão causando o aumento de tempo e procurar melhorar sua eficiência. O tempo de execução serve para mostrar o desempenho do equipamento. Uma grande diferença entre os tempos de execução máximos e mínimo ou tempos médios grandes em relação ao tempo de CPU indicam sobrecarga de processamento ou de operação.

4.2.5.1 - Análise do Relatório de Frequência de opens por arquivo.

Neste relatório aparecem o número de vezes que um arquivo foi aberto com determinado tipo de acesso. Lembramos que a cada open corresponde um close. A medida que este número for muito grande, ele pode prejudicar o sistema. O usuário deve modificar o sistema para poder permitir trabalhar com mais de um tipo de acesso simultâneo ao mesmo arquivo, sem ter que fechar o arquivo para método de acesso anterior.

4.3 - PROG3 - Programa de Alocação de Área

4.3.1 - Pré-requisito para utilização.

- nenhum, o Banco de Dados pode não existir ainda.

4.3.2 - Arquivo de Entrada:

- ARQSAE3 - Arquivo de Alocação de Área, descrito no item 3.5.1.3.

Este arquivo contém informações sobre os segmentos que farão parte do Banco de Dados. Cabe ao usuário formatá-lo. Vamos supor para efeito dos exemplos, que este arquivo está em cartão.

O cartão DD referente a este arquivo tem a forma:

//ARQSAE3 DD

Os cartões preparados pelo usuário, devem seguir este cartão.

4.3.3. - Arquivo de Saída:

- ARQIMP3 - Relatórios de Alocação de Espaço, descritos no item 3.5.2.3. São impressos até 4 relatórios, um para cada natureza de arquivo (dados, elos, inversões e estruturas), cada relatório contendo informações sobre os segmentos do arquivo.

4.3.4 - Exemplo:

Suponhamos que o usuário deseje dimensionar a área a ser alocada para um arquivo de dados, com 4 segmentos com as seguintes características:

	<u>identificação do usuário</u>	<u>nº de registros</u>	<u>tamanho dos registros</u>
1º Segmento	SEG1	200	90
2º Segmento	SEG2	500	300
3º Segmento	SEG3	10500	200
4º Segmento	SEG4	2500	150
5º Segmento	SEG5	5900	220

Para solicitar estas previsões de área, o usuário deve providenciar os seguintes cartões, todos começando na coluna 1.

1º cartão - cartão mestre de natureza:

D 0 4 5 1 6 4
NAT NARQ TAMBL

O tamanho do bloco, é o mesmo indicado para o arquivo de dados no Sistema de Banco de Dados.

2º ao 5º cartão - cartões de informações sobre o segmento:

D S E G 1	0 0 0 2 0 0 0 9 0
D S E G 2	0 0 0 2 0 0 5 1 2
D S E G 3	0 1 0 5 0 0 2 0 0
D S E G 4	0 0 2 5 0 0 1 5 0
D S E G 5	0 0 5 9 0 0 2 2 0
NAT IDENTA	NREG TAMRD

Se o usuário desejar fazer previsões para outros arquivos (elos, inversões, estruturas), deverá preparar conjuntos semelhantes a este e colocá-los em sequencia; neste caso não é necessário especificar o tamanho do registro.

A saída referente a esta apuração terá o seguinte formato:

PREVISÃO PARA ALOCAÇÃO DE ÁREA EM DISCO

ARQUIVO DE DADOS

NOME DO ARQUIVO	NUMERO DE TRILHAS
SEG1	2
SEG2	11
SEG3	108
SEG4	38
SEG5	129

NUMERO DE CILINDROS NECESSÁRIOS - 20

OBS.: - Neste caso não há análise a ser feita, este é um programa de previsão de espaço, o Banco de Dados pode nem existir.

5. ESTUDO DE UM CASO

Para darmos um exemplo do sistema, criamos e utilizamos um Banco de Dados. Passaremos, agora a fazer a análise da sua eficiência, com base nos relatórios emitidos pelo Sistema de Avaliação de Eficiência.

O conteúdo do Banco de Dados (a criação e utilização), que vamos analisar, consta do exemplo descrito no Capítulo 4 - Manual de Operação da tese de Implementação e Documentação de um Suporte para manipulação de Banco de Dados de Roberto Pires Vasques.

5.1 - Relatórios de Distribuições

Os relatórios a seguir, foram obtidos, com base nos arquivos constantes do exemplo citado acima.

DISTRIBUICAO POR TAMANHO DE REGISTROS DE DADOS

TAMANHO DO REGISTRO	NUMERO DE REGISTROS	TAMANHO MEDIO	PERCENTUAL
MENOR QUE 21	0	0	0,00
21 A 30	15	28	20,00
31 A 40	14	40	18,66
41 A 50	0	0	0,00
51 A 60	0	0	0,00
61 A 70	22	70	29,33
MAIOR QUE 70	24	90	32,00
TOTAL	75	62	100,00

Análise:

Com base neste relatório, podemos observar que o Banco de Dados contém registros com tamanhos bem diferentes. Não é aconselhável, neste caso, trabalhar com registros de tamanho fixo.

DISTRIBUICAO POR NUMERO DE REGISTROS EM CADA ARQUIVO DE DADOS

NUMERO DE REGISTROS	NUMERO DE SEGMENTOS	NUMERO MEDIO	PERCENTUAL
MENOR QUE 11	2	7	40,00
11 A 20	1	14	20,00
21 A 30	2	23	40,00
31 A 40	0	0	0,00
41 A 50	0	0	0,00
MAIOR QUE 50	0	0	0,00
TOTAL	5	15	100,00

Análise:

Observamos no relatório acima, que os arquivos do sistema contém poucos registros e podemos verificar nos exemplos de execução do Sistema de Banco de Dados, constantes da tese de ROBERTO PIRES VASQUES, que nenhum dos arquivos apresenta um número de acessos tal que justifique uma modificação no sistema, de acordo com a análise descrita no item 4.1.5.2 deste trabalho.

DISTRIBUICAO POR TAMANHO DE CAMPO DOS REGISTROS DE DADOS

TAMANHO	DO	CAMPO	NUMERO	DE	CAMPOS	TAMANHO MEDIO	PERCENTUAL
MENOR	QUE	6			0	0	0,00
	6	A	10		1	7	7,69
	11	A	15		2	15	15,38
	16	A	20		3	20	23,07
	21	A	25		0	0	0,00
	26	A	30		2	30	15,38
	31	A	35		1	33	7,69
	36	A	40		2	40	15,38
	41	A	45		0	0	0,00
	46	A	50		0	0	0,00
	51	A	55		0	0	0,00
	56	A	60		1	60	7,69
	61	A	65		0	0	0,00
	66	A	70		0	0	0,00
	71	A	75		0	0	0,00
	76	A	80		1	80	7,69
MAIOR	QUE	80			0	0	0,00
TOTAL					13	31	100,00

Análise:

Os campos dos registros deste Banco de Dados são de maneira geral grandes. Verificamos através Relatórios Resumo dos Arquivos sistema referente a este caso (vide Anexo III), que são feitas duas inversões com base em campos de 40 posições. O usuário deve criar códigos para estes campos e fazer a classificação através do código, tornando desta forma, a inversão mais rápida.

DISTRIBUICAO POR NUMERO DE REGISTROS DE DADOS EM CADA CLASSE DE INVERSAO

NUMERO DE REGISTRCS	NUMERO DE CLASSES	NUMERO MEDIO	PERCENTUAL
MENOR QUE 2	41	1	93,18
2 A 3	3	2	6,81
4 A 5	0	0	0,00
6 A 7	0	0	0,00
8 A 9	0	0	0,00
10 A 11	0	0	0,00
12 A 13	0	0	0,00
14 A 15	0	0	0,00
16 A 17	0	0	0,00
18 A 19	0	0	0,00
MAIOR QUE 19	0	0	0,00
TOTAL	44	1	100,00

Análise:

As classes de inversão são bastante pequenas, o arquivo de inversões não precisa ter a parte referente a identificadores de registro aumentada.

DISTRIBUICAO POR TAMANHO DE CHAVE DE INVERSAO

TAMANHO	DA	CHAVE	NUMERO	DE	SEGMENTOS	TAMANHO	MEDIO	PERCENTUAL
MENOR	QUE	6			0		0	0,00
	6	A	10		0		0	0,00
	11	A	15		0		0	0,00
	16	A	20		0		0	0,00
	21	A	25		0		0	0,00
	26	A	30		0		0	0,00
	31	A	35		0		0	0,00
	36	A	40		2		40	100,00
MAIOR	QUE	40			0		0	0,00
TOTAL					2		40	100,00

Análise:

Como vemos, as duas chaves de inversão do sistema, tem um tamanho de 40 bytes. O sistema prevê a inversão de campos com até 249 bytes ao todo. Neste caso, poderíamos reduzir a parte da chave referente ao campo invertido para 40 bytes. O registro de inversão poderia ser reduzido, desta forma em 209 bytes.

Análise geral das distribuições do sistema, sugestões de mo
dificações.

Sugerimos que sejam criados itens de código para os campos classificados. Como ambos os campos classificados são de "nome", teríamos, por exemplo no arquivo EMP os campos de "código" e "nome" e no arquivo PROJPEs teríamos apenas o campo de "código" (ver Relatórios Resumo dos Arquivos do Sistema de Banco de Dados no Anexo III). Considerando-se a dimensão do Banco de Dados, o campo de "código" pode ter 2 bytes de tamanho. Desta forma o arquivo EMP teria o seu tamanho de registro aumentado de 2 bytes e o registros do PROJPEs seriam diminuídos de 38 bytes.

Considerando ainda que o sistema tem apenas 2 inversões, e que feita a alteração sugerida no parágrafo anterior, as inversões seriam feitas para campos de 2 bytes, poderíamos reduzir a parte da chave referente ao campo invertido de 249 bytes para 2 bytes. Embora esta chave seja suficiente para esta aplicação, não considero interessante esta modificação pois restringe a utilização do sistema. O usuário deve pesquisar o tamanho de chave ideal para suas aplicações.

5.2 Relatórios de Frequência

Os relatórios a seguir, refletem informações colhidas durante a execução dos Job's EXECUTA 1, EXECUTA2 e EXECUTA3 descritos na Capítulo 4 da tese de Implementação e Documentação de um Suporte de Banco de Dados.

FREQUENCIA E TEMPO DE USE POR PRIMITIVO EM MCS-MINISEG.SYS

PRIMITIVO	PREJENCIA DE UTILIZACAO	TEMPO TOTAL	MAXIMO CPU	MINIMO CPU	MEIO CPU	TEMPO TOTAL	MAXIMO	MEIO
ATIVO - IB JA EXISTE	2	3.31	1.71	1.59	1.65	2.25.99	1.34.10	1.31.17
ATIVO - IB NAO EXISTE	1	1.09	1.09	1.09	1.09	1.32.70	1.32.70	1.31.17
CRIDAC	5	7.30	1.74	0.91	1.04	1.31.75	1.31.75	1.31.17
CRIELO	4	1.66	0.99	0.71	0.84	1.31.75	1.31.75	1.31.17
CRITAV	7	6.48	2.29	0.81	0.92	1.31.75	1.31.75	1.31.17
DELIV - PRIMEIROS ARQUIVO	4	0.47	0.47	0.47	0.47	1.31.75	1.31.75	1.31.17
INSCAC	1	1.52	1.52	1.52	1.52	1.31.75	1.31.75	1.31.17
INSELL	126	41.15	0.77	0.27	0.42	1.31.75	1.31.75	1.31.17
REMOCD	1	2.29	2.29	2.29	2.29	1.31.75	1.31.75	1.31.17
RESFLE	3	0.29	0.29	0.29	0.29	1.31.75	1.31.75	1.31.17
SATVCE	3	1.41	0.47	0.45	0.45	1.31.75	1.31.75	1.31.17

Análise:

Podemos observar que o primitivo mais utilizado, é o INSELO. (insere elos) ele é bastante rápido, mas mesmo assim detem o maior tempo total de CPU, o usuário deve procurar otimizar o seu algoritmo.

A diferença entre os tempos máximo e mínimo de CPU nos primitivos CRIDAD (cria um arquivo de dados) deve-se ao número de registros manipulados pelo arquivo; no caso do CRIELO (cria um arquivo de elos) esta diferença de tempo deve-se ao número de registros pré-formatados para cada arquivo. O número inicial de registros de cada arquivo seja de dados, elos ou inversões consta dos Relatórios Resumo dos Arquivos do Sistema de Banco de Dados no Anexo III.

Grandes diferenças entre os tempos máximo e mínimo de execução (não proporcionais as de CPU) servem para detectar sobrecarga de processamento ou de operação.

FREQUENCIA DE OPENS POR ARQUIVO

TIPO DE OPEN	DADOS	ELOS	INVERSOES	ESTRUTURAS
SEQUENTIAL UPDATE	4	2	3	3
SEQUENTIAL CUTPUT	5	4	2	5
DIRECT UPDATE	11	3	5	3
TOTAL	20	9	10	11

Análise:

O número de "opens" ocorridos durante as três execuções de "job's" sobre o Sistema de Banco de Dados não justificam uma modificação do sistema por parte do usuário.

5.3 Relatórios de Alocação de Espaço

Estes relatórios foram utilizados, para se estimar a área a ser alocada para os arquivos que fazem parte do exemplo descrito no Capítulo 4 da tese Implementação e Documentação de um Suporte de Banco de Dados.

PREVISAO PARA ALOCAÇÃO DE AREA EM DISCO
ARQUIVO DE DADOS

NOME DO ARQUIVO	NUMERO DE TRILHAS
EMP	1
PROJETO	1
TAREFA	1
HABILIT	1
PROJPES	1
NUMERO DE CILINDROS NECESSARIOS-	1

PREVISAO PARA ACOGACAO DE APET EN DISCO

ARQUIVO DE LIPS

Nome do Arquivo	Numero de Trilhas
MNG	1
REQ	1
CAP	1
ASN	1
Numero de Cilindros Necessarios-	1

PREVISAO PARA ALOCACAO DE AREA EM DISCO	
ARQUIVO DE INVERSOES	
NOME DO ARQUIVO	NUMERO DE TRILHAS
EMPINV	4
PRJINV	3
NUMERO DE CILINDROS NECESSARIOS-	1

PREVISAO PARA ALOCACAO DE AREA EM DISCO
ARQUIVO DE ESTRUTURAS

NOME DO ARQUIVO	NUMERO DE TRILHAS
PROJETO	1
TAREFA	1
HABILIT	1
PROJPES	1
NUMERO DE CILINDROS NECESSARIOS-	1

REFERÊNCIAS BIBLIOGRÁFICAS

- FURTADO, J.T. - Computer Database Organization. Prentice Hall, 1975
- DATE, C.J. - An Introduction to Database Systems. Addison Wesley, 1975
- FURTADO, A. L. - Uma Introdução a Banco de Dados. PUC/RJ, 1976
- FURTADO, A. L. - Towards the Design of Database Interfaces for non Programmers. PUC/RJ, 1976
- CHALLIS, M.F. - The Jackdaw Database Package University of Cambridge, Computer Laboratory, 1974
- STONEBRAKER, M. - The Design and Implementation of Ingres. ACM Transactions on Database Systems, vol 1, num 3, 1976
- TSICHRITZIS, D.C. & LOCHOVSKY, F.H. - Database Management Systems. Academic Press, 1976
- FURTADO, A. L., PASSOS, S. & VASQUES, R. - Um Suporte para Implementação de Sistemas de Gerência de Banco de Dados. PUC/RJ, 1977
- KLEINROCK-Queuing Systems. Volume I Teory. Wiley-Intercience, 1975
- Informatique et Gestion, nº 83, Decembre 1976
- ACM Transactions on Database Systems, volume number 3, September 1976.