



Manual do IFLOAD*

Carga de arquivo invertido (índice)

www.naves.com.br/arte

elysio@naves.com.br

*Texto preparado para o Curso de Utilitários Cisis via Internet
por Elysio Mira Soares de Oliveira
com base em textos publicados pela Bireme*

Versão

20/04/00 11:27:57

* O IFLOAD é um "software" desenvolvido e distribuído pela Bireme/OPAS/OMS
* Este manual é um componente da metodologia utilizada nos cursos de Utilitários Cisis via Internet
* Veja outros instrumentos de suporte disponíveis neste endereço: <http://www.naves.com.br/arte>

TABELA DE CONTEÚDO

<u>INTRODUÇÃO</u>	3
<u>APRESENTAÇÃO</u>	3
<u>GERANDO O ARQUIVO DE LIGACÕES</u>	4
<u>ORDENANDO O ARQUIVO DE CHAVES CURTAS</u>	4
<u>ORDENANDO O ARQUIVO DE CHAVES LONGAS</u>	4
<u>CARREGANDO O ARQUIVO INVERTIDO</u>	4
<u>SINTAXE</u>	4
<u>PARÂMETROS OBRIGATÓRIOS</u>	4
<nome do arquivo>	4
<arquivo lk1>.....	5
<arquivo lk2>.....	5
<u>PARÂMETROS OPCIONAIS <OPÇÕES></u>	5
<i>master=<nome da base></i>	5
<i>-reset</i>	5
<i>-balan</i>	5
<i>-posts</i>	5
<i>tell=<n></i>	6
<i>+fix</i>	6
<i>+fix/m</i>	7

Introdução

O IFLOAD (Inverted File Load) é um utilitário desenvolvido em CISIS e se encontra disponíveis para as plataformas MS-DOS e Linux. Sua principal utilidade está relacionada com a carga de arquivos invertidos de grandes bases de dados, face a sua rapidez no processo. Chega a ser até 100 vezes mais rápido que o CDS/ISIS MS-DOS ou Windows. Além desta característica, o IFLOAD permite a carga de arquivo invertido a partir de outros formatos diferentes do padrão utilizado pelo CDS/ISIS. Pode ser utilizado para criar e reinicializar um arquivo invertido.

O IFLOAD requer como arquivo de entrada, arquivos de ligações contendo chaves que irão ser carregadas no arquivo invertido. Estes arquivos de ligações podem ser gerados pelo CDS/ISIS ou pelo utilitário MX.

Neste documento, em todos os exemplos, utilizaremos o MX para extrair as chaves da base de dados e gerar o arquivo de ligações. Utilizaremos outro utilitário, denominado de MYS para ordenar o arquivo de chaves que será utilizado como arquivos de entrada para o IFLOAD.

Apresentação

O comando:

```
ifload \cisis\base \isis\work\base.lk1 \isis\work\base.lk2  
tell=99
```

Carrega o arquivo invertido base localizado no diretório **\cisis\bases** usando os arquivos de ligações de chaves curtas e chaves longas localizados no diretório **\isis\work**. A opção **tell=99** produz uma mensagem progressiva a cada 99 ligações processadas, mostrando a chave atual que está sendo carregada.

Supõe-se que os arquivos de ligações estejam no formato de arquivo de ligações padrão do MicroISIS, conforme exemplo abaixo:

```
102 24 1 1 ABOUT  
42 24 1 9 ABSENCE  
6 24 1 10 ABSORPTION  
87 24 1 5 ACCOUNT  
136 69 1 1 ACCOUNTING  
40 24 1 6 ACID  
101 24 1 5 ACTION  
49 24 1 6 ACTIVITIES  
130 24 1 7 ACTIVITIES  
23 24 1 5 ACTUAL
```

Este arquivo tem linhas de tamanho variável que contém campos que identificam a origem da chave e a chave propriamente dita. Os primeiros 4 campos são os componentes do *posting*.

MFN	<i>Número de registro na base de dados (arquivo mestre).</i>
TAG	<i>Identificador de campo, atribuído pela FST.</i>
OCC	<i>Número de ocorrência do campo</i>
CNT	<i>Número sequencial do termo no campo</i>

Gerando o arquivo de ligações

Exemplo de procedimento para gerar um arquivo de ligações utilizando o MX:

```
mx base fst=@base.fst stw=@base.stw ln1=base.ln1 ln2=base.ln2 +fix  
tell=100
```

Onde:

mx	<i>Comando de execução do MX</i>
base	<i>Nome da base de dados que será processada pelo MX</i>
fst=	<i>Parâmetro que recebe as instruções para a extração de chaves</i>
@	<i>Sinal que indica que as instruções estão em um arquivo</i>
base.fst	<i>Nome do arquivo de instruções para a extração de chaves</i>
stw=	<i>Parâmetro que recebe as chaves a serem ignoradas</i>
@	<i>Sinal que indica que as instruções estão em um arquivo</i>
base.stw	<i>Nome do arquivo com as chaves que devem ser ignoradas</i>
ln1=	<i>Parâmetro para indicação do arquivo de chaves curtas</i>
base.ln1	<i>Nome do arquivo de chaves curtas</i>
ln2=	<i>Parâmetro para indicação do arquivo de chaves longas</i>
base.ln2	<i>Nome do arquivo de chaves longas</i>
+fix	<i>Parâmetro para geração de arquivos de tamanho fixo</i>
tell=	<i>Parâmetro para indicação da mensagem de processamento</i>
100	<i>Argumento da mensagem de processamento</i>

Ordenando o arquivo de chaves curtas

```
mys 36 base.ln1 base.lk1
```

Ordenando o arquivo de chaves longas

```
mys 56 base.ln2 base.lk2
```

Carregando o arquivo invertido

```
ifload base base.lk1 base.lk2 +fix tell=1000
```

Sintaxe

```
ifload <nome_do_arquivo> <arquivo_lk1> <arquivo_lk2> [<opções> [...]]
```

Parâmetros obrigatórios

<nome_do_arquivo>

Nome do arquivo invertido (índice) a ser criado durante o processo de carga. Normalmente recebe o mesmo nome da base de dados de onde se originaram as chaves dos arquivos de ligação. Pode receber nomes diferentes. Caso já exista um arquivo com o nome fornecido, este será reinicializado.

<arquivo_lk1>

O **lk1** é um arquivo de ligação contendo as chaves curtas, ordenadas. Neste arquivo se encontram todas os termos com até 10 caracteres, extraídos da base de dados através das instruções contidas na FST utilizada no processo de extração.

<arquivo_lk2>

O **lk2** é um arquivo de ligação contendo as chaves longas, ordenadas. Neste arquivo se encontram todas os termos com 11 até 30 caracteres, extraídos da base de dados através das instruções contidas na FST utilizada no processo de extração.

Parâmetros opcionais <opções>

master=<nome_da_base>

Apaga o sinal de pendência para atualização do arquivo invertido da base de dados <nome_da_base>; se este parâmetro for omitido, será apagado o sinal de pendência da base de dados que coincide com o <nome_do_arquivo>.

Sempre que se carrega um arquivo invertido, por padrão, o sinal de pendência existente na base de dados é apagado de todos os registros associados ao arquivo invertido.

Para apagar o sinal de pendência de uma base de dados com nome distinto do nome do arquivo invertido, há que se usar a opção **master=**. Veja o exemplo:

```
ifload kw kw.lk1 kw.lk2 master=\cisis\bases\lilacs
```

-reset

Não apaga o sinal de pendência para atualização do arquivo invertido da base de dados <nome_da_base> ou <nome_do_arquivo>

-balan

Não efetua o balanceamento da árvore (B-Tree) do índice.

-posts

Carrega somente o índice. Não carrega os *postings* no arquivo .ifp

Este parâmetro é bastante útil quando se pretende criar arquivos a serem utilizados com o parâmetro **jchk** do MX, reduzindo sensivelmente o tamanho do índice que será utilizado como tabela de verificação ou controle.

Exemplo:

```
ifload au au.lk1 au.lk2 -posts
```

Quando for necessária a criação de vários arquivos invertidos de uma mesma base de dados, há necessidade de indicar que o sinal de pendência do arquivo invertido deve ser mantido. Essa indicação é feita com o parâmetro **reset=0**, Exemplo:

```
ifload au au.lk1 au.lk2 reset=0
```

```
ifload ti ti.lk1 ti.lk2 reset=0
```

tell=<n>

Produz uma mensagem progressiva sobre o andamento do processamento a cada <n> ligações carregadas.

+fix

Carrega o arquivo de ligações que contém registros de tamanho fixo no formato:

```
KEY MFN TAG OCC CNT
```

Nota: A grande vantagem da utilização de arquivos fixos para o processamento está relacionado com a velocidade do processamento que passa a ser sensivelmente maior.

Exemplo de um arquivo de tamanho fixo utilizado como entrada para o IFLOAD:

ABOUT	102	24	1	1
ABSENCE	42	24	1	9
ABSORPTION	6	24	1	10
ACCOUNT	87	24	1	5
ACCOUNTING	136	69	1	1
ACID	40	24	1	6
ACTION	101	24	1	5
ACTIVITIES	49	24	1	6
ACTIVITIES	130	24	1	7
ACTUAL	23	24	1	5

O comando seguinte, carrega o arquivo invertido teste localizado no diretório \cisis\bases usando um arquivo de ligações de tamanho fixo teste.lk1 e teste.lk2 localizado no diretório \isis\work:

```
ifload \cisis\bases\teste \isis\work\teste.lk1 \isis\work\teste.lk2 +fix
```

Pode-se criar um procedimento BAT para a geração completa de um arquivo invertido usando as potencialidades do IFLOAD. Segue abaixo um exemplo de um arquivo BAT com parâmetros DOS fornecidos em linha de comando:

Arquivo INV.BAT:

```
@echo off
mx %1 fst=@%1 stw=@%1 ln1=%1.ln1 ln2=%1.ln2 +fix tell=100
del *.$$$
mys 36 %1.ln1 %1.lk1
del %1.ln1
del *.$$$
mys 56 %1.ln2 %1.lk2
del %1.ln2
ifload %1 %1.lk1 %1.lk2 +fix tell=1000
del %1.lk1
del %1.lk2
```

A sintaxe a ser utilizada neste BAT para geração do arquivo invertido da base de dados TESTE, cuja FST seja TESTE.FST e arquivo de chaves a serem omitidas TESTE.STW seria:

inv TESTE

+fix/m

Carrega o arquivo de ligações que contém registros de tamanho fixo em um formato reduzido:

KEY MFN

De utilidade quando se deseja criar vários arquivos invertidos (índices) para uma mesma base de dados, e não seja necessário a operação de busca por proximidade, sendo somente necessário o MFN da chave. Exemplo:

ABOUT	102
ABSENCE	42
ABSORPTION	6
ACCOUNT	87
ACCOUNTING	136
ACID	40
ACTION	101
ACTIVITIES	49
ACTIVITIES	130
ACTUAL	23

Este arquivo seria carregado com o nome BASE, com o seguinte comando:

```
ifload base base.lk1 base.lk2 +fix/m
```